Prediction of conversion from Mild Cognitive Impairment to Alzheimer's disease

Hossein Tabatabaei Jafari (MD)

December 2019

A thesis submitted for the degree of Doctor of Philosophy

of

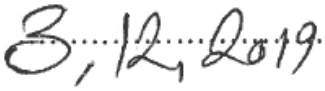
The Australian National University

Statement of originality and candidate contribution

I hereby declare that this submission is my original work using the Alzheimer's Disease Neuroimaging Initiative database. I analysed all data independently and drafted all chapters of this thesis. It contains no materials previously written by another person except it is cited by their name. Any contribution made to the studies of this project is acknowledged in the acknowledgement section.

Hossein Tabatabaei Jafari

Signed 

Date 

Acknowledgements

This thesis becomes a reality with the kind support and help of many people. I would like to express my gratitude and appreciation to all of them. Foremost, I would like to thank my supervisors, Professor Nicolas Cherbuin, Dr. Marnie Shaw and Dr. Erin Walsh for their support, guidance and contributions in all my studies. Especially, I owe a deep sense of thanks to Professor Nicolas Cherbuin for his constant supports and his outstanding and inspiring mentorship. His timely advice, meticulous scrutiny, and scientific approach have made the process as smooth as possible and helped me to a very great extent to accomplish this task.

I need to thank all the participants and researchers of the Alzheimer's Disease Neuroimaging Initiative (ADNI) project for establishing such an invaluable longitudinal data. Additionally, I would like to thank all staffs and colleagues at the Centre for Research on Ageing, Health and Wellbeing and the Research School of Population Health for providing friendly environment that helped me to be focused on my studies.

It is also my privilege to thank my children, Parsa and Kiana for their love and support.

Abstract

Predicting the conversion from mild cognitive impairment (MCI) to Alzheimer's disease (AD) and of the time to conversion, remain some of the most important clinical challenges despite having been investigated for many years. To this day the available evidence has not identified any reliable methods that can be applied in clinical settings mostly because of the complexity of the most effective methods. Taking feasibility into account, this thesis aimed to use simple MRI markers such as brain volumes to predict the risk and the time of conversion from MCI to AD. This thesis is built upon five step-by-step studies, which demonstrate that hippocampal volume is a practical, reliable measure for MCI prognosis.

The first three studies aimed to develop a novel brain MRI volumetric measure to identify individuals with MCI who progress to AD within five years. The last two studies aimed to explore the contribution of MRI measures in predicting time to conversion and to investigate their interaction with cognitive performance. Data used in this project were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database.

It was hypothesized that the volumetric ratio of the brain region with the greatest atrophy rate in MCI to that of a region with no or a substantially lower atrophy rate in MCI could be a reliable and sensitive index in the prediction of AD conversion. The first study, a systematic review, revealed that the hippocampus and entorhinal cortex were the brain regions with the greatest atrophy rates in MCI, with atrophy rates about two-fold greater in MCI than cognitively normal (CN) people. The second study revealed that the cerebellum does not shrink faster in MCI than in CN individuals. Based on these findings, the hippocampal volume to cerebellar

volume ratio was investigated as a predictor of conversion from MCI to AD within five years. The results revealed that the measure was predictive of conversion, and when combined with a Mini Mental Examination Score (MMSE), could effectively identify individuals at risk of AD from those individuals with MCI who remained stable for at least five years or reverted to CN.

Further comprehensive investigation in the last two studies revealed that brain volumes - the whole brain, ventricles, hippocampal and entorhinal cortex volumes - were predictive of time to conversion from MCI to AD. Additionally, although individual cognitive/functional performance was predictive of time to conversion, its predictive values was dependent on hippocampal volume. The same conclusions were drawn from analyses investigating atrophy rates of these regions. That is, the rates of atrophy in whole brain, ventricles, hippocampus, and entorhinal cortex were predictive of time to AD conversion but dependent on their baseline volumes. Moreover, individuals with MCI, who had hippocampal or entorhinal cortex volumes smaller than 5500 mm^3 and 2800 mm^3 (respectively), progressed to AD more quickly regardless of the ensuing atrophy rate.

Taking all these findings into consideration, this thesis suggests that hippocampal volume is a reliable biological marker for the identification of individuals with MCI at demonstrable risk of conversion to AD. Additionally, it is a reliable biomarker of time to conversion from MCI to AD. Indeed, at volumes less than a defined threshold it is highly prognostic of early conversion. Importantly, the prediction accuracy of a simple volumetric measure of the hippocampus is comparable to that of highly complex and sophisticated methods, such as machine learning, but with the advantage of being practical and easier to use in clinical or

research settings. In clinical practice, early identification of those at risk can assist with early intervention and lifestyle modification, which subsequently can decrease the burden of the disease on the patients, their caregivers, and the health systems.

List of publications

1. Hossein Tabatabaei-Jafari, Marnie E. Shaw, Nicolas Cherbuin, “Cerebral atrophy in mild cognitive impairment: a systematic review with meta-analysis.” Alzheimer’s & Dementia: Diagnosis, Assessment & Disease Monitoring, 1 (2015) 487-504

2. Hossein Tabatabaei-Jafari, Erin Walsh, Marnie E. Shaw, Nicolas Cherbuin, “The cerebellum shrinks faster than normal ageing in Alzheimer’s disease but not in mild cognitive impairment”, Human brain mapping, 2017;38:3141-3150.

3. Hossein Tabatabaei-Jafari, Erin Walsh, Marnie E. Shaw, Nicolas Cherbuin, “A simple and clinically relevant combination of neuroimaging and functional indexes for the identification of those at highest risk of Alzheimer’s disease”, Neurobiology of Aging, 69 (2018) 102-110

4. Hossein Tabatabaei-Jafari, Marnie E. Shaw, Erin Walsh, Nicolas Cherbuin, “Cognitive/functional measures predict Alzheimer’s disease, dependent on hippocampal volume”, The journals of Gerontology: Series B, 2019 Jan 21. doi: 10.1093/geronb/gbz011. [Epub ahead of print]

5. Hossein Tabatabaei-Jafari, Marnie E. Shaw, Erin Walsh, Nicolas Cherbuin, “Regional brain atrophy predicts time to conversion to Alzheimer's disease, dependent on baseline volume”, Neurobiology of Aging, 83 (2019) 86 – 94

Table of Contents

INTRODUCTION

"By the time you are eighty years old you have learned everything. You only have to remember it". George Burns (1896 – 1996)

George Burns was an active actor until he was ninety-eight years old and was mentally sharp until his death two months after his one-hundredth birthday. His satire is a gentle expression of the fear that most people may experience in old age. Indeed, it can be a real concern for anyone who wishes longevity particularly because it is not yet clear whether one can ever be safe from future mental decline. Identifying those at risk of this decline is the aim of this thesis.

Research context

Acquired mental decline is known as dementia and it is characterized by cognitive decline, severe enough to interfere with activities of daily living and is not better explained by other neurological or psychiatric disorders. Although a number of different causes may lead to dementia (Alzheimer's Association, 2018), Alzheimer' disease (AD) is its leading cause (about 70%) especially among people over 65 years of age (Fiest et al., 2016).

Clinical symptoms of AD typically begin with difficulty in recent memory, apathy, depression and progresses to impaired communication, disorientation, confusion, poor judgment, behavior changes and, ultimately, difficulty in speaking and motor function including walking and swallowing (Forstl & Kurz, 1999). The emergence of clinical

symptoms is underpinned by progressive neural damage and degeneration (Falahati, Westman, & Simmons, 2014). Pathologically, AD is characterized by the accumulation of extracellular amyloid β protein and the aggregation of intracellular abnormal phosphorylated tau protein (Braak & Braak, 1991; Dickson, 1997). The severity of accumulation of these extra/intracellular proteins increases and gradually spreads to new brain regions during the course of the disease. This accumulation is neurotoxic and leads to a decreased number of synapses, neural connections, and eventually neural death (Falahati et al., 2014). Neurodegeneration begins typically in the medial temporal lobes and spreads to the parietal and frontal lobes, and eventually to the posterior parts of the brain (Braak & Braak, 1991; Thal, Rub, Orantes, & Braak, 2002). The aetiology of AD is not yet understood but several risk factors have been identified. Some of the well established risk factors for AD are APOE e4 genotype, family history of AD, female sex, lower education, sedentary lifestyle, psychiatric comorbidity, unhealthy diet, smoking, illicit drug use and medical diseases such as diabetes and obesity, however, age is the strongest AD risk (Livingston et al., 2017).

The prevalence of AD rises dramatically with age, markedly after 65. There is a 15-fold increase in the prevalence of AD between the ages of 60 and 85 years (Mayeux & Stern, 2012). In the United State of America (USA), about 3% of people aged 65-74, 17% of people aged 75-84, and 32% of people aged 85 and older suffer from AD (Hebert, Weuve, Paul A. Scherr, & Evans, 2013). AD's annual incidence rate is about 0.5% (500 cases per 100,000 person-year) between 65-70, which steadily increases with age, to about 6% to 8% for individuals over the age 85 (Mayeux & Stern, 2012). In 2013, 44.4 million people were estimated to be suffering from AD worldwide. This number is expected to reach 75.7 million in 2030 (Alzheimer's Disease International, 2013). These findings clearly demonstrate why this disease is one of the biggest public health challenges that society faces.

Although the prevalence of AD only become substantial after the age of 65 years, the brain changes attributable to AD pathology precede the onset of clinical symptoms by several decades (Alzheimer's Association, 2018). The deposition of amyloid plaque and the development of tau neuro-fibrillary tangles may start as early as young adulthood in limited brain regions and gradually increases in number and spreads to most parts of the brain (Braak & Braak, 1997; Vlassenko, Benzinger, & Morris, 2012), resulting in neural damage/loss. Initially, people may cope with the impact of the neural damage/loss and maintain their cognitive performance at normal levels while remaining asymptomatic. However, gradually the capacity of the brain to cope with the damage is overwhelmed by the increasing pathology and cognitive symptoms (particularly memory deficits) emerge (Steffener, Reuben, Rakitin, & Stern, 2011; Stern, 2009). Further neural pathology progressively interferes with the activities of daily living that are dependent on cognitive integrity. The emergence of these symptoms represents the onset of clinical dementia due to AD. Over time, the neural degeneration extends to most parts of the brain and impairs behavioral function and personality, and eventually basic bodily functions such as walking and swallowing (Forstl & Kurz, 1999).

Since AD is currently an irreversible and incurable condition, primary prevention of the disease by controlling its modifiable risk factors has been a particular focus of risk reduction studies. However, because AD is a slowly progressing condition, the hope for secondary prevention (intervention at the early stages of the disease) has been rising in recent years. This raises two important questions: "who is at the early stage of the disease and at risk of progression to AD dementia?" and "how can we predict the time to conversion to AD dementia in those at risk?" These questions are currently among the most important challenges that researchers and clinicians in this field have been attempting to resolve, and thus the focus of this thesis.

In the following sections findings from previous studies on this topic are first reviewed. Gaps in the literature are then outlined. Finally, the aims of this thesis are presented as well as a brief review of the conducted studies.

Mild cognitive impairment precursor of Alzheimer's disease

Since AD pathology is irreversible, management is mainly palliative by the time the neural damage is severe enough to fulfill the dementia criteria, (Alzheimer's Association, 2018; Canu, Sarasso, Filippi, & Agosta, 2018). Therefore, early intervention aimed at modifying the course, before it reaches advanced clinical stages, may lead to better outcomes. However, early intervention requires identifying those in the early stages of the disease. Identifying individuals at the early symptomatic stage has been a focus of interest for researchers investigated the course of the disease in the last few decades. The concept of an early symptomatic stage of the disease has gradually developed into a clinical entity that is currently known as Mild Cognitive Impairment (MCI). MCI has been widely accepted as the earliest clinical presentation of AD pathology that is prodromal to the dementia stage (Crook, Bahar, & Sudilovsky, 1987; Hugo & Ganguli, 2014; Jack et al., 2018; Levy, 1994; O'Brien, 1999). Cognitive decline in MCI is more than that expected in normal ageing but it is not severe enough to interfere with activities of daily living, which is characteristic of the dementia stage of AD (R. C. Petersen, 2004b).

MCI is a clinical stage that occurs between normal cognition and AD, and it is currently recognised as of the AD clinical continuum. This continuum consists of three stages: (1) the asymptomatic preclinical stage of AD, in which tau neurofibrillary tangles and amyloid β protein deposition is abnormally high but not yet clinically symptomatic, (2) the MCI stage, in which at the presence of protein deposition and neurodegeneration cognitive impairment is clinically noticeable but not severe enough to interfere with activities of daily living, and (3)

the dementia stage, in which at the presence of protein deposition and neurodegeneration cognitive impairment is severe enough to interfere with activities of daily living (Albert et al., 2011; Jack et al., 2018). However, despite the importance of verification of the pathology a definite in vivo verification is not yet established and MCI and AD are still clinically diagnosed based on clinical symptoms and neuropsychological assessments. Therefore, MCI is a clinical syndrome with different underpinning pathology and as a result its progression to AD is not inevitable. Moreover, different MCI subtypes have been identified based on the type of cognitive abnormalities observed e.g. impairment in memory, language, executive function, or visuospatial skills (R. C. Petersen, 2004a). Although those with MCI, who predominantly suffer from memory deficits, are most at risk of progression to AD (Albert et al., 2011; R. C. Petersen, 2004a; Ronald C. Petersen et al., 1999; Winblad et al., 2004), progression is not inevitable.

MCI, the precursor to but not always predictive of Alzheimer's disease

Reviews of the literature show that only half of those clinically diagnosed with MCI have been observed to progress to AD within five years (Falahati et al., 2014; Mitchell & Shiri-Feshki, 2009; Pandya, Clem, Silva, & Woon, 2016) with the rest remaining stable or reversing to Cognitively Normal (CN) status (Falahati et al., 2014; Mitchell & Shiri-Feshki, 2009; Pandya et al., 2016). Surprisingly, the rate of conversion remains largely stable even after 10 years of follow-up (Mitchell & Shiri-Feshki, 2009). This variability in progression from MCI to AD is not completely consistent with the concept of an AD clinical continuum. It may be due to different factors including variability in pathology (i.e. presence of other pathologies such as vascular dementia), the rate of progression of AD pathology (i.e. different pace of progression across different people), or uncertainty in diagnosis due to variability in

clinical expertise (i.e. false positive diagnosis). These potential explanations are discussed in turn in the following sections.

Variability in pathology: It is possible that in the absence of pathology load measurement those who remain stable do not carry significant levels of AD pathology and their clinical presentation is a variation of normal ageing. Additionally, some MCI may suffer from a different type of pathology and thus they are not in the AD continuum and they have been misclassified because of their clinical presentation. Moreover, previous studies reveal that about half of individuals with AD have pathological evidence of a second cause of (most commonly vascular) dementia (Schneider, Arvanitakis, Bang, & Bennett, 2007; Schneider, Arvanitakis, Leurgans, & Bennett, 2009). Therefore, the contribution of secondary pathology may determine the MCI outcome.

Variability of the rate of progression of AD pathology: It is also plausible that AD pathology is the underlying pathology but that it follows a very slow progression. Indeed, because people are not homogeneous in regard to the number of risk factors and vulnerability to the pathology, it is likely that some individuals with MCI due to AD pathology may have slower progression rate compared with others. Thus, in this context individuals may remain clinically stable for a while despite the slow, gradual progression in the pathology. However it is important to note that even at the presence of similar pathological progression, progression in clinical presentation may differ due to brain/cognitive differences (Stern, 2009).

Variability in expertise: Another possible reason for variability in MCI outcomes may be attributable to the uncertainty in diagnosis due to variability in clinicians/researchers' expertise and the context of assessment. This may be particularly the case for those who are diagnosed with MCI but who revert to normal cognition after a while (Park, Han, &

Initiative, 2015). This explanation is supported by the fact that there is a higher rate of reversion from MCI to CN in community-based studies, which tend to have somewhat less stringent clinical assessments, compared to clinical studies (Mitchell & Shiri-Feshki, 2009). In clinical studies, clinicians have access to patients' medical history and they can personalize further investigations to better detect and characterise cognitive decline. In contrast, researchers are restricted to screening batteries and predefined assessments, and a limited number of tools for cognitive assessment in community-based studies, which may decrease the accuracy of diagnosis. This may lead to less robust diagnoses in community-based studies which in turn leads to a greater rate of diagnosis reversal despite the use of the same diagnostic criteria (Mitchell & Shiri-Feshki, 2009).

Whether variability in MCI outcome is due to variability in pathology or accuracy of diagnosis, the use of biomarkers, which can confirm the presence of AD pathology, may help better explain variability in disease progression. For this reason, the National Institute of Aging and Alzheimer's Association (NIA-AA) suggested incorporating biological marker of AD in the diagnostic guideline of MCI and AD, and also suggested using AD biomarkers for prognostic purpose (Dubois et al., 2014; Jack et al., 2018). The NIA-AA guideline system groups the AD biomarkers into those of amyloid β deposition, pathologic tau, and neurodegeneration. The guideline recommends – at this stage for research frameworks only-- the markers of amyloid β, pathologic tau and neurodegeneration for diagnostic purpose and neurodegeneration markers and clinical assessment for prognostic purpose. Numerous studies have investigated different biological markers for prognostic purposes and to predict conversion from MCI to AD. In the following section, these studies are briefly reviewed and discussed in relation to their practical implementation in research and clinical practice.

Prediction of conversion and time to conversion

Newly developed technologies have provided an opportunity to investigate in vivo biological markers reflecting pathological changes underpinning disease progression and to predict conversion from MCI to AD (Falahati et al., 2014; Rathore, Habes, Iftikhar, Shacklett, & Davatzikos, 2017). Positron-Emission Tomography (PET) imaging has been used for the evaluation of amyloid β deposition and pathologic tau aggregation using specific ligands such as Pittsburgh Compound-B PET (PIB-PET) for amyloid β and flortaucipir-PET for phosphorylated tau (Klunk et al., 2004; Pontecorvo et al., 2017). Additionally, the level of amyloid β and tau proteins' fragments in cerebrospinal fluid (CSF) (Lista et al., 2017; Ritchie et al., 2017) and recently blood plasma (Deters et al., 2017; Fan et al., 2018; Nakamura et al., 2018) have been assayed to investigate abnormal cortical levels of these proteins. Other studies have also explored markers of neural injury and degeneration related to the protein deposition by tracking brain structural and functional changes using neuroimaging methods such as Magnetic Resonance Imaging (MRI), Diffusion Tensor Imaging (DTI), functional MRI and Fluorodeoxyglucose-PET (FDG-PET) (Caso, Agosta, & Filippi, 2016; Dennis & Thompson, 2014; Dona, Thompson, & Druchok, 2016; Mosconi, 2005). Some of these methods have been used to predict conversion from MCI to AD and to a lesser extent time to conversion. They are reviewed in the following two sections.

Predicting MCI to AD conversion

Studies investigating conversion from MCI to AD can be categorized into those that used biomarkers of amyloid β and tau deposition, biomarkers of neurodegeneration, or a combination of these.

Amyloid β and tau biomarkers

Systematic review of PIB-PET studies shows that cerebral amyloid β deposition predicts conversion to AD with an average sensitivity of 94.7% (ranged from 83.3% to 100%) and 57.2% specificity (ranged from 41.1% to 100%) (Ma et al., 2014). Similarly, systematic reviews revealed that CSF amyloid β and tau biomarkers are predictive of AD conversion with 75% sensitivity (ranged from 51% to 91%) and 72% specificity (ranged from 48% to 88%) (Diniz, Pinto Junior, & Forlenza, 2008; Ritchie et al., 2017). These findings suggest that confirming the presence of abnormal amyloid β deposition and pathologic tau in those who are clinically diagnosed with MCI should indicate that conversion to AD is highly likely although far from perfect.

An important issue relating to the use of amyloid β and tau biomarkers is that these methods (i.e. PET scan and CSF) are invasive and have limited availability in practice. This is particularly the case for PET, which is a highly invasive procedure that involves the injection of a radiotracer. CSF measures are also highly invasive they require a lumbar puncture for which complications are common. Indeed, CSF leakage and post-dural puncture headache have been reported in up to one-third of those receiving lumbar puncture (Wang et al., 2015). Moreover, it requires inpatient facilities for its administration and post-intervention monitoring. Therefore, PET and CSF are invasive and expensive procedures and relatively impractical in clinical practice.

Biomarkers of neurodegeneration

While different biomarkers of neurodegeneration have been identified, only brain metabolic activity using FDG-PET and brain volumetry using structural MRI have been investigated for prediction of conversion from MCI to AD.

A meta-analytic study suggests that metabolic hypoactivity detected by FDG–PET imaging is predictive of conversion from MCI to AD with pooled sensitivity of 88.8% and specificity of 84.9% (Yuan, Gu, & Wei, 2009). It is likely that this metabolic hypoactivity is subsequent to neural injury and death in the brain regions involved in AD pathology, and thus representative of the severity of neurodegeneration and predictive of clinical progression.

Besides measures of brain metabolic activity, MRI imaging investigating brain structure has been used to evaluate the severity of neurodegeneration. Indeed, structural MRI is the most common approach for investigating neurodegeneration to date (Falahati et al., 2014; Rathore et al., 2017). One of the best predictive values for prediction of conversion from MCI to AD has been demonstrated for certain patterns of cortical atrophy with an 84.8% sensitivity and 97.22% specificity (Guo, Lai, Wu, Cen, & Alzheimer's Disease Neuroimaging, 2017). The volume of the hippocampus has also been investigated and has shown to be a moderate predictor of AD conversion with a sensitivity of 67%, and specificity of 72% (Chupin et al., 2009). Comparing the results of topological studies and prediction using hippocampal volume suggests that the predictive values of topological features are better than a regional brain measure.

In contrast to PET imaging and spinal taps, structural MRI is minimally invasive, affordable and a more widely available technique for investigating AD neurodegeneration. However, machine-learning approaches have been used to detect the pattern of neurodegeneration in structural MRI, while their implementation in clinical practice or some research settings such as clinical trials is not practical. Machine learning is a method, which involves two steps. In a first training step a proportion of the data is used to identify the topological pattern characteristic of individuals who convert to AD. In a second step, the rest of the data is used to test the accuracy of the learned topological pattern (Deo, 2015). This is

a sophisticated but complex method that requires large data sets for training and validation, which may not always be available. Moreover, in clinical practice or in clinical trials having data for the training step is not practical because the outcomes of the participants' diagnoses are yet to be determined. It is possible to use different large databases for the training step and it may be possible to establish a normalized topological pattern applicable in clinical settings and clinical trials but based on current evidence such an approach has not yet been validated. In addition, because of the "black box" nature of the training (i.e. the ability to automatically learn and improve from experience without being explicitly programmed), the approach is vulnerable to "adversarial attacks" (intentional inputs that have been crafted to force the model to misclassify) especially in real world medical practice (Finlayson, Chung, Kohane, & Beam, 2018). Therefore, we are still in the initial stages of utilising this method, with significant obstacles that need to be addressed (additional to the effectiveness of the approach in a research framework) before machine learning is utilised in clinical practice.

Multi-modal biomarkers

Using multi-domain measures such as a combination of neuroimaging and CSF biomarkers or a combination of different neuroimaging modalities has been shown to have higher predictive accuracy than individual measures (Falahati et al., 2014; Rathore et al., 2017). However, this improvement comes at the expense of increasing complexity and using invasive methods. For example, the combination of MRI, PET, and CSF shows more than 91% accuracy in prediction of conversion from MCI to AD (Zhang et al., 2011), but this is hardly practical since as discussed above CSF and PET are invasive methods with limited availability in daily practice.

Additionally, a combination of cognitive assessments and biological markers such as MRI, CSF, and PET imaging has shown improvement in predicting the conversion from MCI

to AD (Barnes, Cenzer, Yaffe, Ritchie, & Lee, 2014; Devanand et al., 2007; Ewers et al., 2012; Falahati et al., 2014; Moradi et al., 2015; Zhang, Shen, & Alzheimer's Disease Neuroimaging, 2012) when compared to cognitive assessments alone (Rabin et al., 2012; Silva et al., 2012; Silva et al., 2013). Despite this improvement the combinations have the same practical limitations discussed.

In general, despite being informative, biological markers that have been investigated to date are mostly too complex or are not widely available to be implemented in clinical settings/trials to predict conversion from MCI to AD. Therefore, identifying an approach practical enough to be widely used in clinical and research settings would be highly desirable. In addition to predicting conversion from MCI to AD, prediction of time to AD conversion is also important and this question will be discussed in the next section.

Predicting time to conversion

In contrast to studies predicting conversion from MCI to AD, those that have aimed to predict the timeframe within which conversion occurs are limited in number. Moreover, available studies have mainly investigated the predictive value of biomarkers of neurodegeneration. These studies have generally suggested an association between longitudinal brain volume loss and the time to AD conversion (Falahati et al., 2017; Jack et al., 2005; Liu, Chen, Yao, & Guo, 2017; Teipel, Kurth, Krause, Grothe, & Alzheimer's Disease Neuroimaging, 2015). While few of these studies showed that regional brain volume loss is predictive of time to AD conversion (Falahati et al., 2017; Jack et al., 2005), the majority of them demonstrated that the pattern of longitudinal volume loss (using machine learning approaches) is predictive of time to AD conversion (Gavidia-Bovadilla, Kanaan-Izquierdo, Mataro-Serrat, Perera-Lluna, & Alzheimer's Disease Neuroimaging, 2017; Li et al., 2012; Risacher et al., 2010; Teipel et al., 2015; Thung et al., 2018). Additionally, few

studies have used a combination of neuroimaging and clinical performance to predict time to AD conversion, such as the combination of hippocampal and entorhinal volume and Mini Mental State Examination (MMSE) and Selective Reminding Test (SRT) delayed recall (episodic memory) and Wechsler Adult Intelligence Scale–Revised (WAIS-R) Digit Symbol (attention/psychomotor/executive function) (Devanand et al., 2007), the average gray matter volume of several brain networks and MMSE scores, Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-cog) scores and Clinical Dementia Rating (CDR) scores (Liu et al., 2017), and the gray matter volume (from 93 region of interest) and MMSE, ADAS-cog, and CDR revealed prediction of time to conversion (Thung et al., 2018).

The aforementioned studies mostly used machine-learning approaches, which as noted above are not yet practical for individual assessment such as in clinical settings and clinical trials. In the few studies that used brain volumes or longitudinal volume loss the results have not been consistent due to different methodology. Additionally, these studies mostly had short follow-up periods of three years or less. Hence, they are only informative in relation to early conversion. The predictive values of these MRI measures in a longer follow-up are not yet clear.

Gaps in the literature

Despite the large number of studies, which have been conducted to investigate AD progression in the past few decades, some key gaps in the literature remain to be investigated/addressed. They are summarised below.

1. Practicality of the predictive methods at individual level

To date, biomarkers that are most predictive of conversion from MCI to AD are relatively impractical in clinical settings or in clinical trials. This is because the focus of

previous studies has been mostly on using complex approaches, while the predictive values of some simple measures such as regional volume have not been thoroughly investigated.

2. The earliest practical prediction of conversion from MCI to AD

The conceptualisation of the progressive nature of AD pathology suggests that all individuals with MCI should eventually progress to AD. However, the average time to conversion is not yet determined. In order to provide sufficient time for intervention we must be able to predict conversion earlier in the disease process. To date, most studies have tried to identify individuals with MCI who converted to AD within one to two years or, at most, up to three years (Falahati et al., 2014; Rathore et al., 2017). A longer prediction interval, which can provide longer time for potential interventions or lifestyle modification, is currently lacking.

3. Interaction between biomarkers and cognitive performance in timing of conversion

While the association between severity of AD neurodegeneration and cognitive impairment has been well documented, this association is not straightforward (Neuropathology Group of the Medical Research Council Cognitive Function and Aging Study, 2001). Some people can cope with same level of degeneration better than other people (Medaglia, Pasqualetti, Hamilton, Thompson-Schill, & Bassett, 2017; Steffener & Stern, 2012; Stern, 2009). Investigating the interaction between neurodegenerative markers and cognitive function in predicting the time to conversion may not only help disentangle the predictive value of each of these measures, but will also help to better understand how they relate to one-another.

4. Brain volume, atrophy and time to conversion

Spread of neurodegeneration in the brain is characteristic of AD pathology progression and the pattern of this expansion is predictive of time to conversion from MCI to AD (Liu et al., 2017; Teipel et al., 2015). However, it is not clear if regional brain volumes and atrophy rates (besides the pattern of neurodegeneration) are also predictive of conversion, and if so, whether there is any interaction between them. This issue is particularly relevant because atrophy rate is the main outcome measure of clinical trials of disease-modifying treatments.

Thesis scope

The main aim of this thesis is to establish a practical approach to identify those with MCI at the highest risk of conversion to AD dementia and to predict the timing of this conversion.

Taking practicality into consideration, simple structural MRI markers of neurodegeneration are the target of this thesis to track AD progression. Structural MRI is accurate, cost-effective, relatively simple to interpret, and the safest in vivo investigation of the nervous system that is widely available in the clinic. The purpose is to avoid complex image processing techniques such as machine learning and keep the approach simple enough for implementation in clinical practice. Thus, simple volumetric measures of the brain are examined, separately or in combination with cognitive performance, to predict the risk and the timing of conversion from MCI to AD.

While neurodegeneration is not limited to AD, some characteristic patterns of neurodegeneration are. Thus, the intention is to identify a volumetric measure that can serve as a proxy of the pattern of AD neurodegeneration in addition to being simple and practical. To fulfil this goal, the plan is to normalize the volume of the brain region with the earliest and greatest involvement in AD pathology by the brain region with the least and latest involvement. Therefore, before conducting any study to predict conversion it is necessary to

determine which brain regions are affected the most by neurodegeneration in a longitudinal follow-up, and which regions are least affected.

In addition to having a practical measure for predicting AD conversion, it is import to predict the conversion early enough to provide sufficient time for either intervene to change the course of the disease or at least implement lifestyle modification to better cope with the situation. Although the ideal scenario is to identify those at high risk of AD conversion as early as it is possible, the goal of this thesis is to identify those at risk of conversion within five years. Five-year is an arbitrary timeframe but it is consistent with what is usually reported in medicine in regard to prognosis of chronic progressive diseases.

In addition to prediction of AD conversion, the scope of this thesis is also to shed light on the prediction of the timing of AD conversion. The intention is to explore if simple brain volume measure and cognitive performance can predict time to AD conversion, and if so, characterise the nature of their interaction? Furthermore, it is also important to determine if brain atrophy rate can predict time to conversion, and if so, better understand the interaction between brain volume and brain atrophy in predicting time to AD conversion. This is especially important for interpreting the results of clinical trials aiming to delay AD progression.

The scope of the thesis is developed into five aims that are outlined below.

Thesis aims

The overarching focus of this thesis is to demonstrate whether simple structural MRI markers can predict conversion and the time to conversion from MCI to AD. To explore this question, five specific aims were investigated through five separate studies

1. To combine, contrast, and integrate the findings from different studies to produce normative information on regional atrophy rates, and to identify the most sensitive anatomic biomarkers characteristic for MCI.
2. To evaluate cross- sectional and longitudinal structural differences in a brain region (that is less likely to be involved in AD pathology) across cognitively different populations i.e. CN, MCI, and AD.
3. To determine if a simple MRI measure, separately and in combination with a standard cognitive test, can predict conversion from MCI to AD within five years.
4. To investigate the value of brain MRI measures and cognitive performance in predicting time to conversion.
5. To investigate the value of atrophy rates in predicting time to AD conversion from MCI.

Thesis outline and summary

Five studies have been conducted to fulfil the five aims of the thesis. The first study is a systematic review of differences in structural brain changes in MCI. The review revealed that the whole brain, ventricles, hippocampus and entorhinal cortex volumetric measures were the most studied measures. A meta-analysis was conducted to produce normative information on global/regional atrophy rates.

The second study aimed to evaluate cerebellar atrophy across different cognitive status i.e. CN, MCI, and AD. The cerebellum was hypothesized to be the area that shrinks least in the disease process since cross-sectional studies have shown that this region experiences the least amount of amyloid β deposition, tau aggregation or histological changes. The result revealed that the rate of atrophy in the cerebellum in MCI is essentially the same as in normal ageing. This suggests that the cerebellum can be considered as a reference area for

normalizing more strongly implicated brain regions in the disease process such as the hippocampus.

The third study was a conceptual development of the first two studies. It was hypothesized that normalising the hippocampus with the cerebellum may reliably predict conversion from MCI to AD. Therefore, the aim of the study was to evaluate the predictive value of this ratio, either alone or in combination with a sensitive measure of cognitive decline e.g. the Mini Mental State Examination (MMSE) to identify those MCI at highest risk of progression to AD within five years. Findings of this study revealed that the combination of hippocampal to cerebellar volume and MMSE was a sensitive and reliable predictor of conversion.

The purpose of the fourth study was to explore the value of a combination of hippocampal volume and some widely used cognitive and functional tests in predicting time to conversion from MCI to AD. Findings showed that the combination of hippocampal volume and cognitive/functional measures was better at predicting time to conversion than each of these measure in isolation. However, effectiveness of cognitive measures in predicting conversion from MCI to AD was dependent on hippocampal volume.

The fifth study aimed to investigate the value of brain volume at first consultation and its ensuing atrophy rate in predicting the time to conversion from MCI to AD. Based on findings from the first study the whole brain, ventricles, hippocampus and entorhinal cortex were selected as region of interest (ROI). The results were consistent with the findings of those of the fourth study and suggested that the atrophy rate was predictive of the time of conversion from MCI to AD but its predictive value was dependent on the initial brain volume.

The conducted studies are presented in the following chapters and the findings and their implications are discussed in the final chapters. Printed copies of the papers in their published format are provided in the appendices.

References

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., . . . Phelps, C. H. (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. *Alzheimers Dement, 7*(3), 270-279. doi:10.1016/j.jalz.2011.03.008

Alzheimer's Disease International. (2013). *Policy Brief for Heads of Government. The Global Impact of Dementia 2013–2050*. Retrieved from London:

Alzheimer's Association. (2018). 2018 Alzheimer's disease facts and figures. *Alzheimer's & Dementia, 14*(3), 367-429. doi:10.1016/j.jalz.2018.02.001

Barnes, D. E., Cenzer, I. S., Yaffe, K., Ritchie, C. S., & Lee, S. J. (2014). A point-based tool to predict conversion from mild cognitive impairment to probable Alzheimer's disease. *Alzheimer's & Dementia, 10*(6), 646-655. doi:10.1016/j.jalz.2013.12.014

Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol, 82*, 239-259.

Braak, H., & Braak, E. (1997). Frequency of stages of Alzheimer-related lesions in different age categories. *Neurobiol Aging, 18*(4), 351-357.

Canu, E., Sarasso, E., Filippi, M., & Agosta, F. (2018). Effects of pharmacological and nonpharmacological treatments on brain functional magnetic resonance imaging in Alzheimer's disease and mild cognitive impairment: a critical review. *Alzheimers Res Ther, 10*(1), 21. doi:10.1186/s13195-018-0347-1

Caso, F., Agosta, F., & Filippi, M. (2016). Insights into White Matter Damage in Alzheimer's Disease: From Postmortem to in vivo Diffusion Tensor MRI Studies. *Neurodegener Dis, 16*(1-2), 26-33. doi:10.1159/000441422

Chupin, M., Gerardin, E., Cuingnet, R., Boutet, C., Lemieux, L., Lehericy, S., . . . Alzheimer's Disease Neuroimaging, I. (2009). Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. *Hippocampus, 19*(6), 579-587. doi:10.1002/hipo.20626

Crook, T., Bahar, H., & Sudilovsky, A. (1987). Age-associated memory impairment: diagnostic criteria and treatment strategies. *Int J Neurol, 21-22*, 73-82.

Dennis, E. L., & Thompson, P. M. (2014). Functional brain connectivity using fMRI in aging and Alzheimer's disease. *Neuropsychol Rev, 24*(1), 49-62. doi:10.1007/s11065-014-9249-6

Deo, R. C. (2015). Machine Learning in Medicine. *Circulation, 132*(20), 1920-1930. doi:10.1161/CIRCULATIONAHA.115.001593

Deters, K. D., Risacher, S. L., Kim, S., Nho, K., West, J. D., Blennow, K., . . . Alzheimer Disease Neuroimaging, I. (2017). Plasma Tau Association with Brain Atrophy in Mild Cognitive Impairment and Alzheimer's Disease. *J Alzheimers Dis*. doi:10.3233/JAD-161114

Devanand, D. P., Pradhaban, G., Liu, X., Khandji, A., Santi, S. D., Segal, S., . . . Leon, M. J. d. (2007). Hippocampal and entorhinal atrophy in mild cognitive impairmentPrediction of Alzheimer disease. *Neurology, 68*, 828-836.

Dickson, D. W. (1997). The pathogenesis of senile plaques. *J Neuropathol Exp Neurol, 56*(4), 321-339.

Diniz, B. S., Pinto Junior, J. A., & Forlenza, O. V. (2008). Do CSF total tau, phosphorylated tau, and beta-amyloid 42 help to predict progression of mild cognitive impairment to

Alzheimer's disease? A systematic review and meta-analysis of the literature. *World J Biol Psychiatry, 9*(3), 172-182. doi:10.1080/15622970701535502

Dona, O., Thompson, J., & Druchok, C. (2016). Comprehensive Review on Magnetic Resonance Imaging in Alzheimer's Disease. *Crit Rev Biomed Eng, 44*(3), 213-225. doi:10.1615/CritRevBiomedEng.2016019544

Dubois, B., Feldman, H. H., Jacova, C., Hampel, H., Molinuevo, J. L., Blennow, K., . . . Cummings, J. L. (2014). Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *The Lancet Neurology, 13*(6), 614-629. doi:10.1016/s1474-4422(14)70090-0

Ewers, M., Walsh, C., Trojanowski, J. Q., Shaw, L. M., Petersen, R. C., Jack, C. R., . . . Neuroimagin, N. A. A. s. D. (2012). Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. *Neurobiol Aging, 33*(7), 1203-+. doi:10.1016/j.neurobiolaging.2010.10.019

Falahati, F., Ferreira, D., Muehlboeck, J. S., Eriksdotter, M., Simmons, A., Wahlund, L. O., & Westman, E. (2017). Monitoring disease progression in mild cognitive impairment: Associations between atrophy patterns, cognition, APOE and amyloid. *Neuroimage Clin, 16*, 418-428. doi:10.1016/j.nicl.2017.08.014

Falahati, F., Westman, E., & Simmons, A. (2014). Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging. *J Alzheimers Dis, 41*(3), 685-708. doi:10.3233/JAD-131928

Fan, L. Y., Tzen, K. Y., Chen, Y. F., Chen, T. F., Lai, Y. M., Yen, R. F., . . . Chiu, M. J. (2018). The Relation Between Brain Amyloid Deposition, Cortical Atrophy, and Plasma Biomarkers in Amnesic Mild Cognitive Impairment and Alzheimer's Disease. *Front Aging Neurosci, 10*, 175. doi:10.3389/fnagi.2018.00175

Fiest, K. M., Roberts, J. I., Maxwell, C. J., Hogan, D. B., Smith, E. E., Frolkis, A., . . . Jette, N. (2016). The Prevalence and Incidence of Dementia Due to Alzheimer's Disease: a Systematic Review and Meta-Analysis. *Can J Neurol Sci, 43 Suppl 1*, S51-82. doi:10.1017/cjn.2016.36

Finlayson, S. G., Chung, H. W., Kohane, I. S., & Beam, A. L. (2018). Adversarial Attacks Against Medical Deep Learning Systems. *arXiv.org > cs > arXiv:1804.05296*.

Forstl, H., & Kurz, A. (1999). Clinical features of Alzheimer's disease. *Eur Arch Psychiatry Clin Neurosci, 249*(6), 288-290.

Gavidia-Bovadilla, G., Kanaan-Izquierdo, S., Mataro-Serrat, M., Perera-Lluna, A., & Alzheimer's Disease Neuroimaging, I. (2017). Early Prediction of Alzheimer's Disease Using Null Longitudinal Model-Based Classifiers. *PLoS One, 12*(1), e0168011. doi:10.1371/journal.pone.0168011

Guo, S., Lai, C., Wu, C., Cen, G., & Alzheimer's Disease Neuroimaging, I. (2017). Conversion Discriminative Analysis on Mild Cognitive Impairment Using Multiple Cortical Features from MR Images. *Front Aging Neurosci, 9*, 146. doi:10.3389/fnagi.2017.00146

Hebert, L. E., Weuve, J., Paul A. Scherr, & Evans, D. A. (2013). Alzheimer disease in the United States (2010–2050) estimated using the 2010 census.

Hugo, J., & Ganguli, M. (2014). Dementia and cognitive impairment: epidemiology, diagnosis, and treatment. *Clin Geriatr Med, 30*(3), 421-442. doi:10.1016/j.cger.2014.04.001

Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., . . . Contributors. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement, 14*(4), 535-562. doi:10.1016/j.jalz.2018.02.018

Jack, C. R., Jr., Shiung, M. M., Weigand, S. D., O'Brien, P. C., Gunter, J. L., Boeve, B. F., . . . Petersen, R. C. (2005). Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic MCI. *Neurology, 65*(8), 1227-1231. doi:10.1212/01.wnl.0000180958.22678.91

Klunk, W. E., Engler, H., Nordberg, A., Wang, Y., Blomqvist, G., Holt, D. P., . . . Langstrom, B. (2004). Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. *Ann Neurol, 55*(3), 306-319. doi:10.1002/ana.20009

Levy, R. (1994). Aging-associated cognitive decline. Working Party of the International Psychogeriatric Association in collaboration with the World Health Organization. *Int Psychogeriatr, 6*(1), 63-68. doi:10.1017/S1041610294001626

Li, Y., Wang, Y., Wu, G., Shi, F., Zhou, L., Lin, W., & Shen, D. (2012). Discriminant analysis of longitudinal cortical thickness changes in Alzheimer's disease using dynamic and network features. *Neurobiol Aging, 33*(2), 427 e415-430. doi:10.1016/j.neurobiolaging.2010.11.008

Lista, S., Toschi, N., Baldacci, F., Zetterberg, H., Blennow, K., Kilimann, I., . . . Alzheimer Precision Medicine, I. (2017). Diagnostic accuracy of CSF neurofilament light chain protein in the biomarker-guided classification system for Alzheimer's disease. *Neurochem Int*. doi:10.1016/j.neuint.2017.05.010

Liu, K., Chen, K., Yao, L., & Guo, X. (2017). Prediction of Mild Cognitive Impairment Conversion Using a Combination of Independent Component Analysis and the Cox Model. *Front Hum Neurosci, 11*, 33. doi:10.3389/fnhum.2017.00033

Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., . . . Mukadam, N. (2017). Dementia prevention, intervention, and care. *The Lancet, 390*(10113), 2673-2734. doi:10.1016/s0140-6736(17)31363-6

Ma, Y., Zhang, S., Li, J., Zheng, D. M., Guo, Y., Feng, J., & Ren, W. D. (2014). Predictive accuracy of amyloid imaging for progression from mild cognitive impairment to Alzheimer disease with different lengths of follow-up: a meta-analysis. [Corrected]. *Medicine (Baltimore), 93*(27), e150. doi:10.1097/MD.0000000000000150

Mayeux, R., & Stern, Y. (2012). Epidemiology of Alzheimer Disease. *Cold Spring Harbor Perspectives in Medicine, 2*(8), a006239-a006239. doi:10.1101/cshperspect.a006239

Medaglia, J. D., Pasqualetti, F., Hamilton, R. H., Thompson-Schill, S. L., & Bassett, D. S. (2017). Brain and cognitive reserve: Translation via network control theory. *Neurosci Biobehav Rev, 75*, 53-64. doi:10.1016/j.neubiorev.2017.01.016

Mitchell, A. J., & Shiri-Feshki, M. (2009). Rate of progression of mild cognitive impairment to dementia - meta-analysis of 41 robust inception cohort studies. *Acta Psychiatrica Scandinavica, 119*(4), 252-265. doi:10.1111/j.1600-0447.2008.01326.x

Moradi, E., Pepe, A., Gaser, C., Huttunen, H., Tohka, J., & Alzheimer's Disease Neuroimaging, I. (2015). Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *Neuroimage, 104*, 398-412. doi:10.1016/j.neuroimage.2014.10.002

Mosconi, L. (2005). Brain glucose metabolism in the early and specific diagnosis of Alzheimer's disease. FDG-PET studies in MCI and AD. *Eur J Nucl Med Mol Imaging, 32*(4), 486-510. doi:10.1007/s00259-005-1762-7

Nakamura, A., Kaneko, N., Villemagne, V. L., Kato, T., Doecke, J., Dore, V., . . . Yanagisawa, K. (2018). High performance plasma amyloid-beta biomarkers for Alzheimer's disease. *Nature, 554*(7691), 249-254. doi:10.1038/nature25456

Neuropathology Group of the Medical Research Council Cognitive Function and Aging Study. (2001). Pathological correlates of late-onset dementia in a multicentre, community-based population in England and Wales. Neuropathology Group of the

Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). *Lancet, 357*(9251), 169-175. doi:10.1016/s0140-6736(00)03589-3

O'Brien, J. T. (1999). Age-associated memory impairment and related disorders. *Advance in psychiatric treatment*.

Pandya, S. Y., Clem, M. A., Silva, L. M., & Woon, F. L. (2016). Does mild cognitive impairment always lead to dementia? A review. *J Neurol Sci, 369*, 57-62. doi:10.1016/j.jns.2016.07.055

Park, M. H., Han, C., & Initiative, t. A. s. D. N. (2015). Is there an MCI reversion to cognitively normal? Analysis of Alzheimer's disease biomarkers profiles. *international Psychiatrics*

, 27(3), 429-437. doi:10.1017/S1041610214002129

Petersen, R. C. (2004a). Mild cognitive impairment as a diagnostic entity. *J Intern Med, 256*(3), 183-194. doi:DOI 10.1111/j.1365-2796.2004.01388.x

Petersen, R. C. (2004b). Mild cognitive impairment as a diagnostic entity. *J Intern Med, 256*(3), 183-194. doi:10.1111/j.1365-2796.2004.01388.x

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol, 56*(3), 303-308. doi:10.1001/archneur.56.3.303

Pontecorvo, M. J., Devous, M. D., Sr., Navitsky, M., Lu, M., Salloway, S., Schaerf, F. W., . . . investigators, F. A.-A. (2017). Relationships between flortaucipir PET tau binding and amyloid burden, clinical diagnosis, age and cognition. *Brain, 140*(3), 748-763. doi:10.1093/brain/aww334

Rabin, L. A., Wang, C., Katz, M. J., Derby, C. A., Buschke, H., & Lipton, R. B. (2012). Predicting Alzheimer's disease: neuropsychological tests, self-reports, and informant reports of cognitive difficulties. *J Am Geriatr Soc, 60*(6), 1128-1134. doi:10.1111/j.1532-5415.2012.03956.x

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, 155*, 530-548. doi:10.1016/j.neuroimage.2017.03.057

Risacher, S. L., Shen, L., West, J. D., Kim, S., McDonald, B. C., Beckett, L. A., . . . Alzheimer's Disease Neuroimaging, I. (2010). Longitudinal MRI atrophy biomarkers: relationship to conversion in the ADNI cohort. *Neurobiol Aging, 31*(8), 1401-1418. doi:10.1016/j.neurobiolaging.2010.04.029

Ritchie, C., Smailagic, N., Noel-Storr, A. H., Ukoumunne, O., Ladds, E. C., & Martin, S. (2017). CSF tau and the CSF tau/ABeta ratio for the diagnosis of Alzheimer's disease dementia and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database Syst Rev, 3*, CD010803. doi:10.1002/14651858.CD010803.pub2

Schneider, J. A., Arvanitakis, Z., Bang, W., & Bennett, D. A. (2007). Mixed brain pathologies account for most dementia cases in community-dwelling older persons. *Neurology, 69*(24), 2197-2204. doi:10.1212/01.wnl.0000271090.28148.24

Schneider, J. A., Arvanitakis, Z., Leurgans, S. E., & Bennett, D. A. (2009). The neuropathology of probable Alzheimer disease and mild cognitive impairment. *Ann Neurol, 66*(2), 200-208. doi:10.1002/ana.21706

Silva, D., Guerreiro, M., Maroco, J., Santana, I., Rodrigues, A., Bravo Marques, J., & de Mendonca, A. (2012). Comparison of four verbal memory tests for the diagnosis and predictive value of mild cognitive impairment. *Dement Geriatr Cogn Dis Extra, 2*, 120-131. doi:10.1159/000336224

Silva, D., Guerreiro, M., Santana, I., Rodrigues, A., Cardoso, S., Maroco, J., & de Mendonca, A. (2013). Prediction of long-term (5 years) conversion to dementia using neuropsychological tests in a memory clinic setting. *J Alzheimers Dis, 34*(3), 681-689. doi:10.3233/JAD-122098

Steffener, J., Reuben, A., Rakitin, B. C., & Stern, Y. (2011). Supporting performance in the face of age-related neural changes: testing mechanistic roles of cognitive reserve. *Brain Imaging Behav, 5*(3), 212-221. doi:10.1007/s11682-011-9125-4

Steffener, J., & Stern, Y. (2012). Exploring the neural basis of cognitive reserve in aging. *Biochim Biophys Acta, 1822*(3), 467-473. doi:10.1016/j.bbadis.2011.09.012

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia, 47*(10), 2015-2028. doi:10.1016/j.neuropsychologia.2009.03.004

Teipel, S. J., Kurth, J., Krause, B., Grothe, M. J., & Alzheimer's Disease Neuroimaging, I. (2015). The relative importance of imaging markers for the prediction of Alzheimer's disease dementia in mild cognitive impairment - Beyond classical regression. *Neuroimage Clin, 8*, 583-593. doi:10.1016/j.nicl.2015.05.006

Thal, D. R., Rub, U., Orantes, M., & Braak, H. (2002). Phases of A beta-deposition in the human brain and its relevance for the development of AD. *Neurology, 58*(12), 1791-1800.

Thung, K. H., Yap, P. T., Adeli, E., Lee, S. W., Shen, D., & Alzheimer's Disease Neuroimaging, I. (2018). Conversion and time-to-conversion predictions of mild cognitive impairment using low-rank affinity pursuit denoising and matrix completion. *Med Image Anal, 45*, 68-82. doi:10.1016/j.media.2018.01.002

Vlassenko, A. G., Benzinger, T. L., & Morris, J. C. (2012). PET amyloid-beta imaging in preclinical Alzheimer's disease. *Biochim Biophys Acta, 1822*(3), 370-379. doi:10.1016/j.bbadis.2011.11.005

Wang, Y. F., Fuh, J. L., Lirng, J. F., Chen, S. P., Hseu, S. S., Wu, J. C., & Wang, S. J. (2015). Cerebrospinal fluid leakage and headache after lumbar puncture: a prospective non-invasive imaging study. *Brain, 138*(Pt 6), 1492-1498. doi:10.1093/brain/awv016

Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L. O., . . . Petersen, R. C. (2004). Mild cognitive impairment--beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *J Intern Med, 256*(3), 240-246. doi:10.1111/j.1365-2796.2004.01380.x

Yuan, Y., Gu, Z. X., & Wei, W. S. (2009). Fluorodeoxyglucose-positron-emission tomography, single-photon emission tomography, and structural MR imaging for prediction of rapid conversion to Alzheimer disease in patients with mild cognitive impairment: a meta-analysis. *AJNR Am J Neuroradiol, 30*(2), 404-410. doi:10.3174/ajnr.A1357

Zhang, D., Shen, D., & Alzheimer's Disease Neuroimaging, I. (2012). Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers. *PLoS One, 7*(3), e33182. doi:10.1371/journal.pone.0033182

Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D., & Alzheimer's Disease Neuroimaging, I. (2011). Multimodal classification of Alzheimer's disease and mild cognitive impairment. *Neuroimage, 55*(3), 856-867. doi:10.1016/j.neuroimage.2011.01.008

STUDY 1

CEREBRAL ATROPHY IN MILD COGNITIVE IMPAIRMENT: A SYSTEMATIC REVIEW WITH META-ANALYSIS

Hossein Tabatabaei-Jafari [1,] Marnie E. Shaw [1] Nicolas Cherbuin [1]

[1]Centre for Research on Ageing, Health and Wellbeing, Australian National University, Canberra, ACT 2601, Australia

Published in Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring,

1 (2015) 487-504

DOI: 10.1016/j.dadm.2015.11.002

Abstract

Introduction: Although mild cognitive impairment (MCI) diagnosis is mainly based on cognitive assessment, reliable estimates of structural changes in specific brain regions - that could be contrasted against normal brain ageing and inform diagnosis - are lacking. This study aimed to systematically review the literature reporting on MCI-related brain changes.

Methods: The MEDLINE database was searched for studies investigating longitudinal structural changes in MCI. Studies with compatible data were included in the meta-analyses. A qualitative review was conducted for studies excluded from meta-analyses.

Results: The analyses revealed a 2.2-fold higher volume loss in the hippocampus, 1.8-fold in the whole brain and 1.5-fold in the entorhinal cortex in MCI participants.

Discussion: Although the medial temporal lobe is likely to be more vulnerable to MCI pathology, atrophy in this brain region represents a relatively small proportion of whole brain loss, suggesting that future investigations are needed to identify the source of unaccounted volume loss in MCI.

Introduction

Although Alzheimer's disease (AD) was first characterized more than 100 years ago, little concrete progress has been made towards an effective cure of this progressive disorder. Identification of mild cognitive impairment (MCI) as a prodromal phase of AD has raised hopes of the possibility of preventing or modifying progressive neurodegeneration in AD. Indeed, initial attempts at early therapeutic interventions have reported some successes in the early phase of MCI (Douaud et al., 2013; Smith et al., 2010).

Clinically, MCI is defined based on the detection of cognitive decline greater than that expected at any given age and less than those observed in dementia in the context of preserved activities of daily living and the absence of other neurological disorders. However, clinical evaluation is complicated by heterogeneity in cognitive reserve and diversity in daily function. Considering that each cognitive measure is designed to target a particular brain function, selecting which cognitive measures are appropriate to assess functional decline in the MCI trajectory is a matter of concern not only for diagnostic purposes but also in the evaluation of clinical trials (Snyder et al., 2014). Besides higher uncertainty in characterizing MCI based on functional impairment (Park, Han, & Initiative, 2015), cognitive evaluation is not currently informative enough for demonstrating patterns of deterioration that will accurately discriminate those who will remain stable from those who will convert to AD or other dementias. Therefore, without a better understanding of the neurological basis of the disorder, as well as the identification of structural biomarkers, reliable detection of MCI and estimation of future risk of dementia remains elusive.

Assuming that impairment in cognitive function is the result of neurodegeneration, monitoring structural brain changes may be beneficial in understanding the pathophysiology of MCI. Recent development in neuroimaging technologies has provided an opportunity to

investigate structural biomarkers in living subjects. In the past two decades, the use of magnetic resonance imaging to assess cerebral structure has become widespread. Most early studies have used a cross-sectional design and have suggested that, although the presence of structural differences in any particular brain region is not specific to MCI or AD (i.e. it can also be observed in ‘’normal’’ ageing), the pattern of regional atrophy rates and the topological progression of atrophy are quite characteristic, particularly in AD (Braak & Braak, 1991). Moreover, these studies also revealed that regional atrophy rates are different in MCI and AD (Sluimer et al., 2009). Consequently, identification of regionally specific atrophy rates in MCI may be beneficial for detecting the early stage of AD development, as well as evaluating the magnitude of expected structural changes in clinical trials.

Available longitudinal studies have identified a subset of brain regions that may be involved in MCI pathology. An important next step is to combine, contrast and integrate the findings from different studies to produce normative information on regional atrophy rates, and to identify the most sensitive anatomical biomarkers characteristic for MCI. As far as we are aware, no study has systematically summarised these findings to date. Therefore, the aim of this study was to systematically review the literature concerning MCI-related structural brain changes.

Methodology

This systematic review was conducted based on an established methodology (Fraser, Shaw, & Cherbuin, 2015) using pre-specified search terms and inclusion and exclusion criteria, and was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (Moher, Liberati, Tetzlaff, Altman, & Group, 2009).

To retrieve all references relating to longitudinal brain structural changes in MCI published in the MEDLINE database, a literature search was conducted through the PubMed

portal in two stages, (1) at the beginning of the study, (2) at the end of February 2015 to update pooled data with the most recent published studies. The following search string was used for both searches; *(Brain or Cerebral or Cortical)* ***And (****Mild Cognitive Impairment Or MCI Or Cognitive disorder Or Neurocognitive disorder Or Cognitive decline Or Cognition)* ***And*** *(Structur* Or Volum* Or Thickness Or MRI Or Neuroimaging)* ***And*** *(Atrophy Or Change Or Longitudinal Or shrinkage)*. Both literal and Medical Subject Heading searches were performed. Searches were limited to studies published in English and focusing on human subjects.

Selection criteria and selection process

To be selected, studies were required to use a longitudinal methodology with two or more structural MRI scans conducted over a follow-up of 12 months or more. As MCI status defined the group being compared with Cognitively Normal (CN), cognitive status of CN and MCI was required to be stable between all time points. Studies were required to use Peterson or Winblad criteria for MCI diagnosis. Cross-sectional, experimental and review articles were excluded. Studies were also excluded if they had a combined total of less than 30 CN and MCI participants. All retrieved articles were first screened by title and abstract and irrelevant studies were excluded. The full text of all remaining articles was double screened by two reviewers (H.T-J and M.E.S) against selection criteria.

Data extraction and structural measures

Two reviewers (H.T-J and M.E.S) extracted data from all included articles and any disagreement was resolved by consensus. Data extracted consisted of: (1) study design including sample source, number of participants in each group, type of structural measurement, and follow up period, (2) participants' demographics including age, sex ratio, APOEe4 ratio, years of education, dropout rate, MCI subtype for MCI groups, subjective

memory complaint for CN and handedness, (3) measurement details including number of scans, scan intervals, follow up period, MRI parameters, segmentation method, and method of analysis, (4) study results including areas of interest (left and right) and effect sizes (left, right and total).

All structural measures were evaluated and studies were categorized according to the following structural measurements; voxel-based morphometry (VBM), volumetry, tensor-based morphometry (TBM), cortical thickness, sulcal morphometry, diffusion tensor imaging (DTI), white matter hyperintensities (WMH), susceptibility weighted imaging (SWI), and other structural measures.

Studies meeting the selection criteria were assessed for quality using the Newcastle-Ottawa scale (Wells et al., 2011). The Newcastle-Ottawa scale is an instrument for assessing the quality of studies included in a systematic review. Each study was evaluated on eight items classified into three categories including the selection of the study groups, the comparability of the groups, and the ascertainment of outcome of interest. Each quality item was awarded by a star (except two for comparability) and for each study up to nine stars in total.

Multiple reports

In the case of multiple reports for the same cohort, or any overlap of participants, an annual change rate estimate from only one publication was used in any single analysis. The most appropriate reports were selected based on recency, availability of effect size and moderators, sample size and methodology. Studies that reported effect sizes (or provided them after contact) were the first priority and from those the most recent study with the largest sample size was selected. If there was more than one study similar in sample size and recency, the one with the highest quality rating was selected. When different studies on the

same cohort reported on different brain regions, estimates from the same cohort but from different studies might be used in different analyses.

Statistical analysis

The R statistical software (Version 3.1.1) was used for the statistical analysis and the Metafor package (version 1.9-4) was used for meta-analysis. The annual percentage mean atrophy rate was considered as the effect size and calculation of required standard error (SE) for meta-analysis was based on the standard deviation and number of participants in each group for each individual study. Availability of mean annual atrophy rate (%/year), either reported or computed based on other reported results, was the essential requirement for the meta-analysis. Where insufficient data were available for inclusion in the meta-analysis, authors were contacted directly to seek additional information.

Meta-Analysis

It was assumed that the heterogeneity in the atrophy rates across reviewed studies was the impact of the between-study and within-study heterogeneities, and the random-effects for between and within-studies were normally distributed. A random-effects model using the restricted maximum likelihood estimator was applied for all analyses. Random-effects model was chosen based on the assumption that cerebral atrophy rates (effect size) are not similar in population with different characteristics and there is no single effect size representative of all population but an array of effect sizes. Therefore, each included study was assumed to represent a random sample of a particular effect size and a random-effects model estimates a mean of the distribution of these effect sizes (Borenstein, Hedges, Higgins, & Rothstein, 2011). Separate meta-analyses were performed for healthy and MCI atrophy rate and also for the mean difference in atrophy rate between MCI and CN (MCI-CN) across each cerebral region.

Heterogeneity across studies was assessed with the Q and I^2 statistics. P-value < 0.01 considered as significant heterogeneity in the Q test and in the I^2 statistic values of 25%, 50% and 75% were suggestive of low, moderate and high heterogeneity respectively. Heterogeneity in the atrophy rates was also assumed to be in part the result of disparities in age, sex ratio, ApoeE4 ratio, and education levels in the studies' participants as well as scan intervals and different segmentation approaches. Therefore, these variables were investigated as possible moderators for subgroup and meta-regression analyses. Subgroup analyses were conducted to investigate the impact of manual versus automated segmentations. Meta-regression analyses using a mixed-effect model were conducted to determine the influence of moderators.

To identify studies contributing excessively to heterogeneity, sensitivity analyses were conducted using the-leave-one-out method. Visual evaluation of asymmetry of the funnel plots was used to assess the bias in the meta-analyses results towards publication of studies with significant outcomes. The trim and fill method was used to estimate the number of missing studies (representative of unreported effect sizes) in the meta-analysis to estimate adjusted effect sizes (Duval & Tweedie, 2000).

Results

Literature Search and Studies Included in the Review

The search strategy identified 5220 unique citations. After exclusion of irrelevant studies based on title and abstracts, 219 publications remained for full text assessment. A further 151 studies did not meet the inclusion criteria and were excluded leaving 68 studies for further analysis (Figure 1).

Of the studies included, 45 assessed brain structure with volumetry, nine with cortical thickness and 18 with a wide variety of structural measurements including sulcal morphometry, VBM, TBM, DTI, WMH, SWI, and quantitative scaling methods such as the medial temporal atrophy scale (MTAS) (Song, Mitnitski, Zhang, Chen, & Rockwood, 2013) and the brain atrophy and legion index (BALI) (Zhang, Song, & Zhang, 2012) (Table 1).

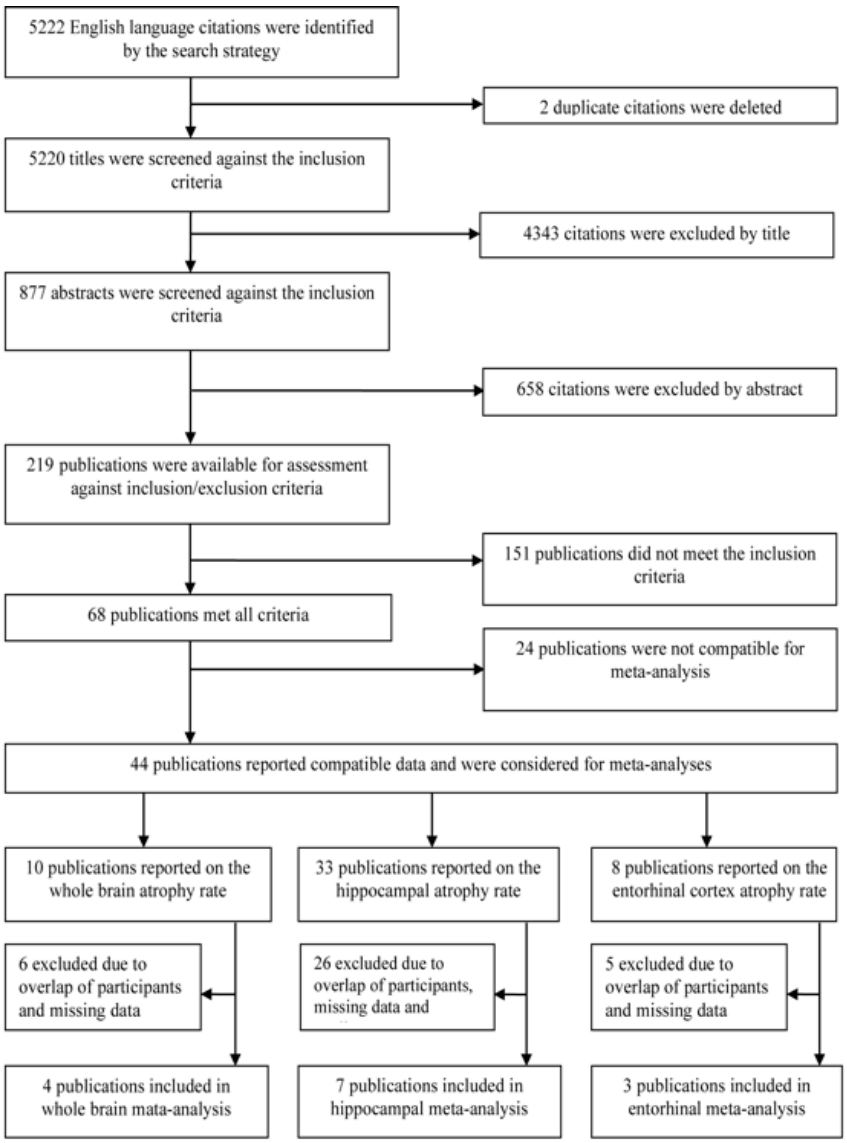


Figure 1: Screening and selection process for studies included in the systematic review and the meta-analyses.

Table 1: Studies included in the review

#	Study	Measurement	Cohort	Newcastle-Ottawa Quality Assessment Scale[1]								Compatible for Meta-analysis		
				Selection				Comparability	Outcome					
				Q 1	Q 2	Q 3	Q 4	Q 5	Q 6	Q 7	Q 8	Yes/ No	Meta-analysis	Details
1	(Madsen et al., 2015)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Ventricle meta-analysis due to overlap of participants
2	(Lorenzi, Pennec, Frisoni, & Ayache, 2015)	SVF	ADNI	*	*	*	*	**	*	*	*	No	No	No quantitative structural measure
3	(Toledo et al., 2014)	???	ADNI	*	*	*	*	**	-	*	*	No	No	Incomplete report of structural data
4	(Teipel et al., 2014)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
5	(Mulder et al., 2014)	Volumetry	ADNI	*	*	*	*	*	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
6	(Marshall et al., 2014)	Cortical thickness	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible brain region with other studies
7	(Manning et al., 2014)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
8	(Lillemark, Sorensen, Pai, Dam, & Nielsen, 2014)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	WB and Hip meta-analyses due to overlap of participants
9	(Kljajevic, Grothe, Ewers, & Teipel, 2014)	Volumetry, Cortical Thickness	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible brain region with other studies
10	(Insel et al., 2014)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip and ERC meta-analyses due to overlap of participants
11	(H. Guo et al., 2014)	BALI	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible with independent study
12	(Aguilar et al., 2014)	Volumetry, Cortical Thickness	AddNeuroMed	*	*	*	*	**	*	*	*	Yes	N0	Missing and mismatch data
13	(Franko & Joly, 2013)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
14	(Nowrangi et al., 2013)	DTI	Community-dwelling Volunteers	*	*	*	*	*	*	*	*	No	No	Incompatible brain region with other studies

15	(L. H. Guo, Alexopoulos, Wagenpfeil, Kurz, & Perneczky, 2013)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	WB meta-analysis due to overlap of participants
16	(Adaszewski, Dukart, Kherif, Frackowiak, & Draganski, 2013)	VBM, SVM	ADNI	*	*	*	*	*	*	*	*	No	No	No quantitative structural measures
17	(Villemagne et al., 2013)	Volumetry	AIBL	*	*	*	*	**	*	*	*	Yes	Yes	Hip meta-analysis
18	(Song et al., 2013)	MTAS, BALI	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible with independent study
19	(Selnes et al., 2013)	DTI	Memory Clinics	*	*	*	*	**	*	*	*	No	No	Incompatible brain region with other studies
20	(Liu et al., 2013)	Sulcal Morphology, Cortical Thickness	MAS	*	*	*	*	*	*	*	*	No	No	Incompatible brain region with other studies
21	(Gutman et al., 2013)	Volumetry	ADNI	*	*	*	*	*	*	*	*	Yes	No	Ventricles meta-analysis due to overlap of participants
22	(Zhang et al., 2012)	MTAS, BALI	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible with independent study
23	(Yao, Hu, Liang, Zhao, & Jackson, 2012)	Cortical Thickness	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible brain region with other studies
24	(Schuff et al., 2012)	Volumetry	ADNI	*	*	*	*	*	*	*	*	Yes	Yes	ERC meta-analysis
25	(McDonald et al., 2012)	Volumetry	ADNI	*	*	*	*	**	*	*	*	No	No	Due to mismatch of brain regions
26	(Li et al., 2012)	Cortical Thickness	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible brain region with other studies
27	(Leung, Ridgway, Ourselin, & Fox, 2012)	Volumetry	ADNI	*	*	*	*	*	*	*	*	No	No	Mismatch data
28	(Andrawis et al., 2012)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
29	(Zhang et al., 2011)	BALI	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible with independent study
30	(Tosun, Schuff, Shaw, Trojanowski, & Weiner, 2011)	Cortical Thickness	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible brain region with other studies
31	(Skup et al., 2011)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip and ERC meta-analyses due to overlap of participants

32	(Mouiha & Duchesne, 2011)	Volumetry	ADNI	*	*	*	*	*	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
33	(Lo et al., 2011)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
34	(Desikan et al., 2011)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	ERC meta-analysis due to overlap of participants
35	(Chiang et al., 2011)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
36	(Villain et al., 2010)	VBM	Memory Clinic	*	*	*	*	**	*	*	*	No	N0	No quantitative structural measure
37	(Vemuri et al., 2010)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	WB meta-analysis due to overlap of participants
38	(Tosun et al., 2010)	Volumetry, cortical Thickness	ADNI	*	*	*	*	**	*	*	*	Yes	No	Ventricle, Hip and ERC meta-analyses due to overlap of participants
39	(Stoub, Rogalski, Leurgans, Bennett, & deToledo-Morrell, 2010)	Volumetry	RADC & ROS and MAP	*	*	*	*	**	*	*	*	Yes	No	Hip and ERC meta-analyses due to overlap of participants
40	(Schott et al., 2010)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	Yes	WB and Hip meta-analyses
41	(Prestia et al., 2010)	VBM	TOMC	*	*	*	*	**	*	*	*	No	No	No quantitative structural measure
42	(Leung et al., 2010)	Volumetry	ADNI	*	*	*	*	*	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
43	(Hua et al., 2010)	TBM	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible brain region with other studies
44	(Ho et al., 2010)	TBM	ADNI	*	*	*	*	*	*	*	*	No	No	Incompatible brain region with other studies
45	(Evans et al., 2010)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	WB and Ventricle meta-analyses due to overlap of participants
46	(Desikan et al., 2010)	Volumetry, Cortical Thickness	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
47	(Carmichael et al., 2010)	WMH	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible with independent study
48	(Beckett et al., 2010)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip and Ventricle meta-analyses due to overlap of participants

49	(Ayaz, Boikov, Haacke, Kido, & Kirsch, 2010)	SWI	???	-	-	*	*	*	*	*	*	No	No	Incompatible with independent study
50	(Archer et al., 2010)	Volumetry	Hospital & Memory Clinic	*	*	*	*	**	*	*	*	Yes	Yes	WB and Hip meta-analyses
51	(Apostolova et al., 2010)	Volumetry	ADNI	*	*	*	*	**	*	*	*	No	No	Hip meta-analysis due to overlap of participants
52	(Wang, Liu, Lirng, Lin, & Wu, 2009)	Volumetry	Neurological clinic	*	*	*	*	**	*	*	*	Yes	Yes	Hip meta-analysis
53	(Sluimer et al., 2009)	Volumetry	Memory Clinic	*	*	*	*	**	*	*	*	No	No	Incompatible brain region with other studies
54	(Schuff et al., 2009)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
55	(Morra et al., 2009)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
56	(Leow et al., 2009)	TBM	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible brain region with other studies
57	(Jack et al., 2009)	Volumetry	Mayo, ADNI	*	*	*	*	**	*	*	*	Yes	No	Ventricle meta-analysis due to overlap of participants
58	(Hua et al., 2009)	TBM	ADNI	*	*	*	*	*	*	*	*	No	No	Incompatible brain region with other studies
59	(Holland, Brewer, Hagler, Fennema-Notestine, & Dale, 2009)	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	WB, Ventricle, Hip and ERC meta-analyses due to overlap of participants
60	(Henneman, Vrenken, et al., 2009)	Volumetry	Memory Clinic	*	*	*	*	**	*	*	*	Yes	Yes	Hip meta-analysis
61	(Henneman, Sluimer, et al., 2009)	Volumetry	Memory Clinic	*	*	*	*	**	*	*	*	Yes	Yes	WB meta-analysis
62	(Brys et al., 2009)	VBM, MTL-rBS	AD research centre	*	*	*	*	**	*	*	*	No	No	Incompatible with independent study
63	(Jack et al., 2008)	Volumetry	Mayo	*	*	*	*	**	*	*	*	Yes	No	WB and Ventricle meta-analyses due to overlap of participants
64	(Eckerstrom et al., 2008)	Volumetry	Goteborg MCI study	-	*	*	*	**	*	*	*	Yes	Yes	Hip meta-analysis

65	(Desikan et al., 2008)	Volumetry	Community-dwelling Volunteers	*	*	*	*	**	*	*	*	Yes	Yes	Hip and ERC meta-analyses
66	(Jack et al., 2005)	Volumetry	Mayo	*	*	*	*	**	*	*	*	Yes	Yes	WB, Hip and ERC meta-analyses
67	(Jack et al., 2004)	Volumetry	Mayo	*	*	*	*	**	*	*	*	Yes	No	WB, Ventricle, Hip and ERC meta-analyses due to overlap of participants
68	(Jack et al., 2000)	Volumetry	Mayo	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants

Keys: ADNI; Alzheimer's disease Neuroimaging Initiative, AIBL; Australian Imaging, Biomarker and Lifestyle, RADC; Rush Alzheimer's Disease Centre, ROS and MAP; Religious Order Study and Rush Memory and Aging Project, TOMC; The Transitional Outpatient Memory Clinic, MAS; Sydney memory Aging Study, AddNeuroMed; six European sites compatible with the US ADNI study, DTI; Diffusion Tensor Imaging, VBM; Voxel-Based Morphometry, TBM; Tensor-Based Morphometry, MTL-rBS; Medial Temporal Lobe atrophy using regional Boundary Shift, SVF; Stationary Velocity Field, BALI; Brain Atrophy and Lesion Index, SVM; Support Vector Machine, MTAS; The Medial Temporal Atrophy Scale, WMH; White Matter Hyperintensities, SWI; Susceptibility Weighted Imaging, AD; Alzheimer's Disease, Hip; Hippocampus, ERC; Entorhinal Cortex, WB; Whole Brain.

(1): Q-1; Representativeness of the exposed cohort, Q-2; Selection of the non-exposed cohort, Q-3; Ascertainment of exposure, Q-4; Demonstration of interest was not present at start of study, Q-5; Comparability of cohorts on the basis of the design or analysis, Q-6; Assessment of outcome, Q-7; Was follow-up long enough for outcomes to occur, Q-8; Adequacy of follow up of cohorts

Study Quality

All studies except one, which was rated 6 (Ayaz et al., 2010), were rated as high quality (8 or 9 stars) based on the Newcastle-Ottawa scale (Table 1). Fifty-four of 68 studies fulfilled the maximum of nine stars, two studies were rated as not representative of the population due to a higher rate of medical diseases in the participants, and one study did not describe the derivation of the CN. Twelve studies only controlled for age to establish comparability between controls and MCI participants.

Multiple Reports

A number of multiple reports were identified. Forty-six studies reported on participants taking part in the Alzheimer's Disease Neuroimaging Initiative (ADNI) (to date up to 229 CN and 395 MCI), four studies used Mayo AD research centre and AD patient registry data

(up to 91 CN and 72 MCI), and one study used a mixture of ADNI and Mayo data. There was also an overlap of participants in 2 studies reported by Henneman et al. (Henneman, Sluimer, et al., 2009; Henneman, Vrenken, et al., 2009). A total of 15 publications reported on separate independent cohorts including in total 629 CN and 571 MCI participants from 10 countries across four continents (eight in Europe, five in North America, one in Asia, and one in Australia).

Compatible Studies for Meta-analysis

A sufficient number of compatible studies were only available for meta-analysis of volumetric measurements. Quantitative report of structural measures in VBM and TBM studies were not comparable. Brain regions investigated by cortical thickness or DTI studies were not anatomically compatible. There was only one study in each given category of sulcal morphometry, WMH and SWI. Finally, studies using MTAS and BALI scales were all based on the same cohort except for one study (Table 1). Therefore, of the 68 studies that met the selection criteria, 24 studies could not be included in the meta-analyses, leaving 44 volumetric studies for inclusion. Because too few sporadic reports of laterality were available this factor could not be investigated. There was also no report of handedness.

Volumetric studies evaluated a wide variety of brain regions including the whole brain, hippocampus, entorhinal cortex, ventricles, parahippocampal gyrus, amygdala, fusiform gyrus, superior temporal, medial lateral and inferior temporal lobes, medial and lateral orbitofrontal cortex, superior frontal cortex, cingulate cortex, parietal and occipital lobes. Beside the first four measures, other brain regions were investigated sporadically. Three of 44 studies evaluated brain regions incompatible with other studies and were not considered for meta-analysis. Forty-one studies were identified as potentially compatible and were included in meta-analyses. These studies evaluated annual atrophy rate of the whole brain

(n=10), the hippocampus (n=33) and the entorhinal cortex (n=10), as well as annual expansion rate of the ventricles (n=14).

Of 41 studies, 29 were excluded due to overlap in participants and 1 due to missing data which could not be obtained from authors (Table 1). Although three studies were available for ventricle expansion analysis, reported expansion rates did not use the same units (ml/year vs. %/year) and requests for more information from authors was not successful. Therefore, meta-analysis could not be conducted for this region. Final numbers of studies included in the meta-analyses were four for whole brain, eight for hippocampal and three for entorhinal cortex atrophy (Table 2).

Whole Brain Atrophy

Four studies (Archer et al., 2010; Henneman, Sluimer, et al., 2009; Jack et al., 2005; Schott et al., 2010), which were included for whole brain analysis (Figure 2), surveyed 351 control and 466 MCI participants over an average follow-up of 1.30 years (range 1.00-1.80). Estimated mean atrophy rates were 1.02%/year (SE=0.13) for MCI and 0.57%/year (SE=0.03) for controls. Thus, the additional annual total brain atrophy attributable to MCI above the effect of "normal" ageing was 0.46%/year (SE=0.10). There was no significant heterogeneity (based on the Q test) for whole brain atrophy rates in CN and MCI after removing the effect attributable to normal ageing (MCI-CN). The proportion of real observed variance (not related to random error) between studies (I^2) was moderate in MCI-CN and high in MCI (Table 3).

Table 2: Studies included in Meta-analyses

First author, year	Measurement				Recruit	Participants		Age		Female %		APOEe4 %		Change rate	
	WB	Hip	ERC	Vent		CN	MCI	CN	MCI	CN	MCI	CN	MCI	CN	MCI
Villemagne, 2013		✔			AIBL	112	32	71.2 (7.2)	74.2 (6.6)	48.21	43.75	46	65	-0.911 (1.15) %/y	-2.15 (1.33) %/y
Schuff, 2012			✔		ADNI	147	164	76 (5)	75 (7)	49.66	37.8	22	45	-1.6 (0.4) %/y	-2.4 (0.4)
Schott, 2010	✔				ADNI	199	334	76 (5.1)	74.9 (7.2)	46.73	36.53	28.64	53.3	-0.592 (0.581) %/y	-1.08 (0.84) %/y
		✔												-1.01 (1.72) %/y	-2.63 (2.35) %/y
				✔										**-1.43 (1.63) ml/y**	**-2.85 (2.75) ml/y**
Archer, 2010	✔				Clinic	27	16	62.3 (8.3)	67.1 (6.9)	51.85	31.25	18.5	75	-0.47 (0.67) %/y	-1 (0.81) %/y
		✔												-0.78 (0.91) %/y	-2.8 (1.68) %/y
				✔										**-1.14 (1.73) ml/y**	**-3.62 (2.33) ml/y**
Wang, 2009		✔			Clinic	20	39	75.1 (3.7)	75.6 (3.6)	45	20.51	20	26.5	-1 (0.7) %/y	-2.1 (1.5) %/y
Hennema, 2009a		✔			Clinic	19	25	66 (9)	71 (6)	42.11	56	47	71	-2 (1.5) %/y	-3.7 (1.2) %/y
Hennema, 2009b	✔				Clinic	34	44	67 (9)	71 (6)	47.06	47.72	_	_	-0.6(0.6) %/y	-1.3(0.9) %/y
Eckerstrom[2], 2008		✔			GMS	19	15	?	?	?	?	?	?	**-0.168 (0.464) ml/y**	**+0.082 (0.329) ml/y**
Desikan, 2008		✔			Media	19	22	69.7 (3.7)	70.1 (4.4)	63.16	59.1	31.6	31.8	-0.71 (0.88) %/y	-1.13 (1.01) %/y
			✔											-0.68 (1.4) %/y	-1.92 (2.12) %/y
Jack Jr, 2005	✔				MAYO	91	72	80.5 (?)	78.7 (?)	60.44	43.06	_	_	-0.5 (0.7) %/y	-0.7 (1) %/y
		✔												-1.7 (1.4) %/y	-3.3 (2.7) %/y
			✔											-5 (3.6) %/y	-7 (4.3) %/y
				✔										**-2.4(2) %/y**	**-3.3(2.3) %/y**

Measures provided as mean (standard deviation). 1: Studies in ventricular meta-analysis were not matched in atrophy rate unit. 2. This study was an outlier and excluded from final hippocampal meta-analysis.

Keys: WB; Whole Brain, Hip; Hippocampus, ERC; Entorhinal cortex, Vent; Ventricles, CN; Cognitively Normal, MCI; Mild cognitive impairment, AIBL; Australian Imaging, Biomarker and Lifestyle, ADNI; Alzheimer's Disease Neuroimaging Initiative, GMS; Goteborg MCI study, MAYO; Mayo AD research centre and AD patient registry

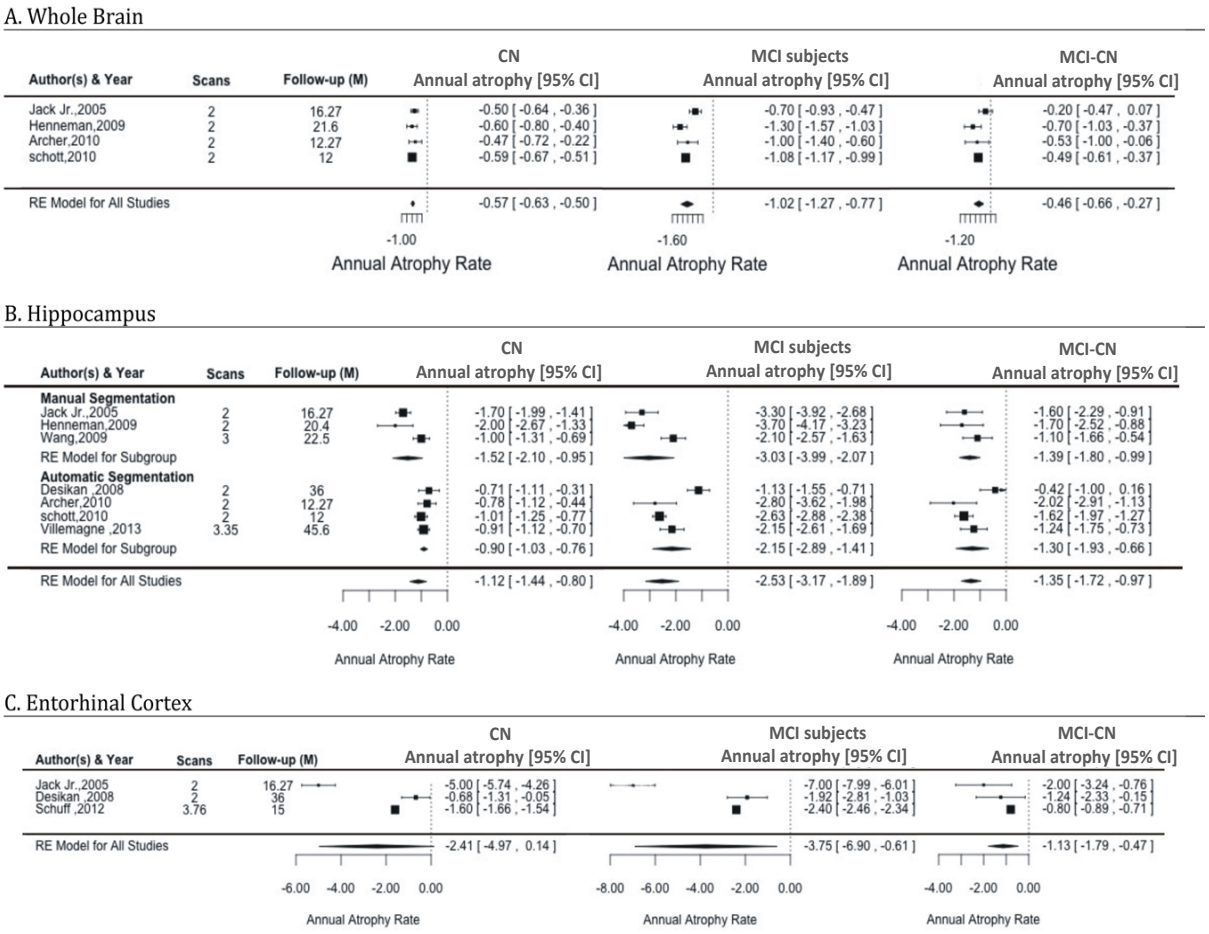


Figure 2: Forest plots of atrophy rates for (A) whole brain, (B) hippocampus, and (C) entorhinal cortex in CN, MCI and the difference in atrophy rate between MCI and CN (MCI-CN). Studies are ordered by year of publication.

Hippocampal Atrophy

Of eight studies (Archer et al., 2010; Desikan et al., 2008; Eckerstrom et al., 2008; Henneman, Vrenken, et al., 2009; Jack et al., 2005; Schott et al., 2010; Villemagne et al., 2013; Wang et al., 2009) which were included for hippocampal meta-analysis, one study (Eckerstrom et al., 2008) reported an increase in hippocampal volume in MCI and a decrease

in volume in CN as well as standard deviations larger than twice the mean atrophy rates. These characteristics were interpreted as being potentially methodologically problematic and after further investigation, the study was excluded from the meta-analysis because it was remarkably different in quality and design compared to other studies in the group, including sex proportion misbalance and high level of medical illness in the participants.

The remaining seven studies estimated hippocampal atrophy rates for 487 CN and 540 MCI participants with an average follow up of 1.97 years (range 1-3.8) (Figure 2). The estimated mean atrophy rates were 2.53 %/year (SE=0.33) for MCI, 1.12%/year (SE=0.16) for controls and 1.35%/year (SE=0.19) for MCI after removing the effect attributable to normal aging. Significant heterogeneity was found for hippocampal atrophy rates in MCI and MCI-CN but not in CN. The proportion of real observed variance (not related to random error) between studies (I^2) was moderate to high in all groups (Table 3).

Entorhinal Cortex Annual Atrophy

Three studies [(Desikan et al., 2008; Jack et al., 2005; Schuff et al., 2012)], which were included for entorhinal cortex meta-analysis (Figure 2), surveyed 257 controls and 258 MCI participants, followed up for 2.28 (range 1.25-3.00) years. Estimated mean atrophy rates were 3.75%/year (SE=1.60) for MCI, 2.41%/year (SE=1.30) for CN. After removing the effect attributable to normal aging, the mean atrophy rate exclusively associated with MCI was 1.13 %/year (SE=0.33). Significant heterogeneity was identified in entorhinal cortex atrophy rates in MCI and CN, but not MCI-CN. The proportion of real observed variance (not related to random error) between studies (I^2) was moderate to high in all groups (Table 3).

Table 3: Random effect models of whole brain, hippocampus and entorhinal cortex and mixed effect models of hippocampal atrophy rate in CN and MCI and in MCI after removing the effect attributable to normal ageing.

Random Effects Model											
Brain Regions	k	age	estimate (se) *	95% Cl	z-value	p-value	T^2	T	I^2 %	H^2	Heterogeneity**
Whole Brain											
CN	351	71.45	-0.5665 (0.0328)	-0.6308, -0.5023	-17.2757	<0.0001	0	0	0	1.0	1.8707(3) 0.5997
MCI	466	72.92	-1.0203 (0.1263)	-1.2679, -0.7727	-8.0772	<0.0001	0.0477	0.2185	79.98	4.99	12.6691(3) 0.0053
MCI-CN	-	-	-0.4634 (0.0987)	-0.6569, -0.2699	-4.6944	<0.0001	0.0194	0.1393	51.86	2.08	5.7540(3) 0.1242
Entorhinal Cortex											
CN	257	75.40	-2.4146 (1.3036)	-4.9696, 0.1505	-1.8522	0.0640	5.0168	2.2398	98.81	83.72	89.1356(2) <0.0001
MCI	258	74.60	-3.754 (1.6065)	-6.9028, -0.6052	-2.3367	0.0195	7.5905	2.7551	98.51	67.29	83.2905(2) <0.0001
MCI-CN	-	-	-1.1301 (0.3373)	-1.7911, -0.4691	-3.3509	0.0008	0.1936	0.4400	52.49	2.10	4.1965(2) 0.1227
Hippocampus											
CN	487	71.54	-1.1197 (0.1622)	-1.4376, -0.8019	-6.9048	<0.0001	0.1513	0.3890	86.22	7.26	34.2283(6) <0.0001
MCI	540	73.09	-2.5303 (0.3261)	-3.1694, -1.8912	-7.7598	<0.0001	0.6741	0.8211	92.87	14.02	78.1854(6) <0.0001
MCI-CN	-	-	-1.3450 (0.1906)	-1.7186, -0.9715	-7.0571	<0.0001	0.1556	0.3945	64.69	2.83	16.5628(6) 0.0110
Subgroup and meta-regression analyses (Hippocampus; MCI-CN)											
	K	age	Coef (se)	95%Cl	z-value	p-value	T^2	T	I^2 %	H^2	R^2	Hetrogeneity**
Model 1												
Automatic Segmentation	4	71.57	-1.2900 (0.2682)	-1.8156, -0.7644	-4.8106	<0.0001	0.2019	0.4494	69.90	3.32	-	16.5244(5) 0.0055
Manual Segmentation	3	75.1	-1.4383 (0.3289)	-2.0829, -0.7936	-4.3730	<0.0001						
Model 2												
aMCI	2	77.15	-1.3337 (0.3939)	-2.1057, -0.5618	-3.3863	0.0007	0.2091	0.4572	70.73	3.42	-	16.4832(5) 0.0056
MCI	5	71.46	-1.3562 (0.2488)	-1.8438, -0.8686	-5.4510	<0.0001						
Model 3												
Intercept	-	-	-0.9973 (5.0703)	-10.9349, 8.9403	-0.1967	0.8441	0.0384	0.1960	36.24	1.57	79.71	2 (2.9841) 0.2249
Age	-	-	0.0006 (0.0640)	-0.1249, 0.1261	0.0093	0.9926						
Female rate	-	-	0.0209 (0.0132)	-0.0050, 0.0467	1.5821	0.1136						
APOEe4 rate	-	-	-0.0233 (0.0088)	-0.0406, -0.0061	-2.6477	0.0081						

CN=Cognitively Normal; MCI= Mild Cognitive Impairment; aMCI= amnestic MCI; Coef= Coefficient; se=Standard Error; T = Standard deviation of true effects; r^2 =proportion of observed dispersion accounted for by the model; H^2= total variability / sampling variability; R^2= heterogeneity accounted for the moderator(s); df= degrees of freedom, * %/Y, ** Q (df) P-value

Sensitivity Analyses

The influence of single studies was investigated with leave-one-out analyses. Globally, the analysis revealed no particularly influential study and showed consistency in reported estimates.

Publication Bias

Some evidence of publication bias was detected based on the funnel plot asymmetry diagnostic and the trim-and-fill test. The funnel plots revealed some degree of asymmetry for all three groups of analyses (the whole brain, hippocampus and entorhinal) and the trim-and-fill method identified one or two missing studies in each analysis group. One missing study was identified in the whole brain and hippocampal analyses and two studies in entorhinal analysis, representing 20%, 12.5% and 40% of included studies respectively. Although asymmetry and presence of missing studies suggest some publication bias towards studies reporting higher atrophy rates, the differences between actual and reported atrophy rates were generally small, particularly for the hippocampus (Figure 3).

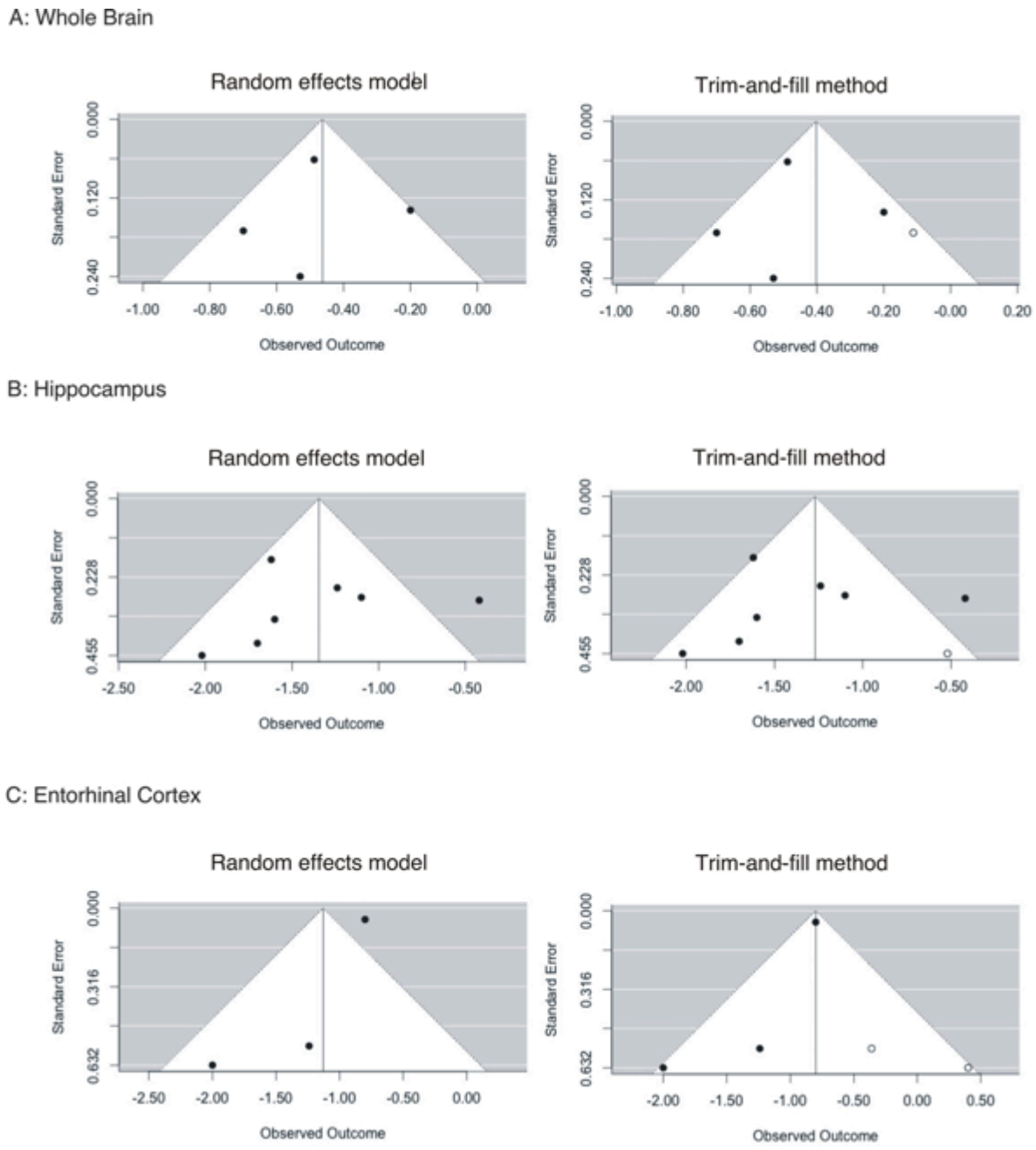


Figure 3: Funnel plots of (A) Whole brain, (B) Hippocampus, and (C) entorhinal cortex using random effects model (left column) and trim and fill method (right column). Filled circles represent included studies in the meta-analyses and open circles represent possible missing studies.

Subgroup and meta-regression Analyses

The influence of segmentation methods (automatic vs. manual), MCI subtype (amnestic MCI vs. MCI), female proportion, APOEe4 genotype and sample mean age on pooled estimates was investigated by subgroup meta-analyses and meta-regression on hippocampal volumetry only, as too few studies were available for other regions of interest (Table 3). Subgroup analyses showed that the estimated mean hippocampal atrophy rates in studies (Henneman, Vrenken, et al., 2009; Jack et al., 2005; Wang et al., 2009) using manual segmentation were significantly higher than studies (Archer et al., 2010; Desikan et al., 2008; Schott et al., 2010; Villemagne et al., 2013) using automatic segmentation (Figure 2 and table 3) by 68% in CN, 40% in MCI, and 7% in MCI-CN. Additionally, subgroup analysis of MCI subtypes (amnestic MCI vs. MCI) showed significantly higher hippocampal atrophy rate in amnestic MCI (Jack et al., 2005; Wang et al., 2009) compared to MCI (all subtypes) (Archer et al., 2010; Desikan et al., 2008; Henneman, Vrenken, et al., 2009; Schott et al., 2010; Villemagne et al., 2013) (2.68 %/year (SE=0.66) vs 2.47 %/year (SE=0.42)) in MCI participants. After removing the effect attributable to normal ageing, the hippocampal atrophy rate was significantly higher in analyses including all generic/unspecified MCI (1.35 %/year, SE=0.25) compared to those including amnestic MCI only (1.33 %/year, SE=0.39). However, the atrophy rate difference was relatively small especially in MCI-CN analyses, and also numbers of studies in each subgroup were limited. In addition, (as it is notified in the discussion) studies, which were not specific in detecting MCI subtype, generally utilized cognitive measures that commonly used for detecting amnestic MCI in other studies.

The influence of age, female sex and APOEe4 rate on hippocampal atrophy was separately investigated in CN, MCI and MCI-CN. Except for APOEe4, which significantly predicted the unexplained variance (55.38%) in annual atrophy rate, age and female sex did not contribute substantially to the heterogeneity detected between studies. A mixed effects

model using age, female sex and APOEe4 rate as moderators accounted for 79.7% of heterogeneity in hippocampal atrophy rate in MCI-CN, however, only APOEe4 rate was a significant moderator of atrophy rate (Table 3).

Incompatible Studies

Ventricular Expansion

Although it was not possible to produce a pooled estimate of ventricular expansion rate due to insufficient reports of separate cohorts, the remaining studies reported very similar estimates (Archer et al., 2010; Jack et al., 2009; Schott et al., 2010) of, on average, two-fold (3.30%/year vs 2.40%/year in one report and 2.85ml/year vs 1.43ml/year and 3.62ml/year vs 1.14 ml/year in two other reports) increase in expansion rate in MCI compared to CN. When considering that whole brain volume is about 1200-1500 ml, reported ventricular expansion rate are approximately 0.1%/year of the whole brain volume in CN and 0.2%/year of the whole brain volume in MCI.

Grey mater atrophy

Besides the hippocampus and entorhinal cortex, which were the focus of most volumetric studies, there were also sporadic reports of volume loss for other parts of the brain including; the parahippocampus, amygdala and fusiform gyrus (Desikan et al., 2008), lateral temporal lobe (McDonald et al., 2012), cingulate (Desikan et al., 2008; McDonald et al., 2012), insula (Sluimer et al., 2009), parietal lobe (Desikan et al., 2008; McDonald et al., 2012; Sluimer et al., 2009), frontal and occipital lobes(McDonald et al., 2012; Sluimer et al., 2009). Atrophy rates in these regions were less than the average hippocampal atrophy rate and also differed based on the clinical outcome. Volume loss in the temporal and parietal lobes was higher for MCI subjects who had converted to AD within 4-5 years compared to

stable MCI (lowest Cohen d for the inferior parietal lobe =0.53 and largest for the hippocampus =1.39) (Desikan et al., 2008). However, in clinically diagnosed AD, the atrophy rate in the medial temporal lobe was less than in MCI, whereas volume loss in frontal, parietal and occipital regions was greater in MCI than AD (Sluimer et al., 2009).

Cortical Thickness and Sulcal Morphometry

Cortical thickness was the second most commonly reported structural measure. Reports covered almost all parts of the brain but without quantitative estimates amenable to meta-analysis. Overall, studies revealed that, controls and MCI participants demonstrated a similar spatial distribution of cortical loss, specifically in the parahippocampal cortex, middle/inferior temporal gyrus, supramarginal gyrus, angular gyrus and superior frontal gyrus (Li et al., 2012). However, these studies suggested that atrophy rates were higher (no report of effect size) in MCI than controls, mainly in the temporal, superolateral parietal, and frontal lobes (Marshall et al., 2014; Yao et al., 2012). The only available longitudinal sulcal morphometry study showed an almost two-fold higher rate of superior frontal and superior temporal sulcal widening in MCI compared to controls (Liu et al., 2013).

White Matter

A minority of studies evaluated longitudinal changes in white matter. Recent DTI studies demonstrated a loss of integrity (increase in mean diffusivity) in the white matter fibre tracts (Selnes et al., 2013) particularly in the fornix (fitted mean changes in mean diffusivity over 12 months of 0.003 in controls vs. 0.051 in MCI), inferior and anterior cingulum (fitted mean changes in mean diffusivity over 6 months of -0.003 in controls vs. 0.013 in MCI), (Nowrangi et al., 2013) in MCI compared to controls. DTI studies were limited in number and restricted to regions of interest evaluation.

Discussion

This study aimed to systematically review the literature on longitudinal structural brain changes specific to stable MCI. The main findings of this review were that: (1) atrophy rates were 1.5 to 2.2 times larger in MCI participants than CN, (2) atrophy rate estimates were greater when assessed with manual than automatic segmentation, (3) Age, sex, and APOEe4 were the most important moderators and together explained almost 80% of the between-study heterogeneity.

Global and local atrophy

Whole brain annual atrophy rate in MCI was twice that observed in controls. After removing the effect of normal ageing, MCI related shrinkage was estimated at 0.46%/year or almost 5 ml per year. This finding was consistent with studies reporting approximately 0.1%/year ventricular expansion in MCI in addition to that observed in normal ageing (Archer et al., 2010; Jack et al., 2009; Schott et al., 2010), when considering that 20% to 25% of the whole brain shrinkage is accounted for ventricular expansion (Standring, 2008).

Shrinkage in the whole brain is not necessarily the result of homogenous atrophy in all parts of the brain. Studies using measurement of cortical thickness and grey/white matter density in different parts of the brain demonstrated that atrophy rates in different brain regions were different and that some areas were more susceptible to neurodegeneration in normal ageing as well as MCI related degeneration (Liu et al., 2013; Sluimer et al., 2009; Wang et al., 2009; Yao et al., 2012). Studies suggested that in MCI, noticeable atrophy was restricted to the medial temporal lobe, while frontal lobe and sensory motor cortices remained less atrophic until late in AD (Hua et al., 2009; Leow et al., 2009). Additionally, previous evidence suggested that medial temporal lobe atrophy was higher in MCI

participants who converted to Alzheimer's disease compared to those with stable MCI (Evans et al., 2010; Leow et al., 2009).

It is important to consider that most reviewed studies used general diagnostic criteria to recruit MCI participants and did not investigate MCI subtypes. However, study design and cognitive tests, which were used in these studies, suggested that there was probably a higher prevalence of amnestic MCI in MCI participants. Therefore, reported findings are likely to be more representative of amnestic MCI than other MCI subtypes.

The hippocampus and entorhinal cortex were two of the most commonly investigated subregions of the medial temporal lobe, and direct evaluation of the medial temporal lobe volume change was not an issue in volumetric studies. Therefore, there is no estimation of the whole medial temporal lobe atrophy rate in the literature. However, overall atrophy rates in these medial temporal lobe subregions were similar to the whole brain atrophy rate, i.e. approximately twice in MCI compared to CN. Although, to our knowledge there is no other systematic review of brain regions atrophy rates in MCI, a systematic review estimating annual hippocampal atrophy rate in healthy ageing across the lifespan revealed hippocampal annual atrophy rate of 1.12%/year in healthy ageing over the age of 70 years (Fraser et al., 2015), which is consistent with the present findings. The roles of the hippocampus and entorhinal cortex in memory function have been known for a long time and the association between atrophy rates in these regions and cognitive decline have been well documented in MCI. However, the mean estimates of annual atrophy rates in these regions do not explain a 5 ml annual reduction in the whole brain volume. The cerebral atrophy observed in MCI above that detected in normal ageing was 1.35%/year in the hippocampus and 0.94%/year in entorhinal cortex. This indicates a total annual volume loss of about 0.068 ml in these areas (Standring, 2008), which covers less than 1.5% (of 5 ml) of the whole brain annual volume

loss. This suggests that volume loss in areas well known for memory and cognition may only be the tip of the iceberg. In summary, although most available evidence has suggested that high rates of atrophy are mostly restricted to the medial temporal lobe in stable MCI, this conclusion might be due to under-investigation of other cerebral regions.

Grey matter and White matter

Apart from medial temporal lobe atrophy, decrease in grey matter volume was reported in the lateral temporal, parietal and frontal lobes (Villain et al., 2010). These findings are consistent with reports demonstrating cortical thinning in the superolateral parietal lobe and some regions of the frontal cortex (Yao et al., 2012) as well as sulcal widening in the superior temporal and superior frontal sulci (Liu et al., 2013). There are also sporadic reports suggesting decrease in the volume of the parahippocampal gyrus, amygdala, fusiform gyrus, superior temporal lobe (Desikan et al., 2008), lateral temporal lobe (McDonald et al., 2012), inferior temporal lobe (Desikan et al., 2008), frontal lobe (McDonald et al., 2012; Sluimer et al., 2009), cingulate (McDonald et al., 2012) and parietal and occipital lobes (McDonald et al., 2012; Sluimer et al., 2009), and insula (Sluimer et al., 2009). Therefore, although higher atrophy rates have been prominently reported in the medial temporal lobe and the atrophy rate in this region was positively associated with cognitive decline, brain atrophy is also widely distributed to other parts of the temporal, parietal and frontal lobes. Nonetheless, in spite of the widespread grey matter atrophy, estimated atrophy rates in these areas alone cannot explain the whole brain atrophy rate. Indeed, the grey matter forms less than half of the brain tissue and atrophy rates as high as the atrophy rate in the hippocampus are needed in all parts of the grey matter to explain the total brain volume loss.

Therefore, atrophy of white matter is likely to significantly contribute to whole brain atrophy, especially because axonal integrity depends on cell body viability in the grey matter and theoretically cell loss in grey matter atrophy should have an impact on white matter integrity. Loss of integrity in the white matter fibre tracts, particularly in the fornix and anterior and inferior cingulum, has been detected by DTI studies (Nowrangi et al., 2013; Selnes et al., 2013). These studies are limited in number and restricted in the selection of regions of interest. A relationship between hippocampal grey matter atrophy and subsequent disruption in the uncinate fasciculus and the cingulum bundle has also been reported (Villain et al., 2010).

Although too few studies investigating white matter atrophy were available for review and for reliable assessment of their magnitude, they suggest that white matter is not spared from MCI pathology. However, the rate of atrophy in white matter and its association with grey matter and whole brain volume loss are some important unanswered questions. White matter forms the dominant proportion of brain structure, which reflects the importance of connection and networks in neural structure and consequently brain function. Therefore, it is essential that more investigations focus on these questions.

Furthermore, while neuroimaging studies largely interpret their results in relation to neural tissue, the brain also consists of connective tissue forming the brain's structural frame, supporting neural content and providing nutrients to neural tissue. This structural frame has an important role in preserving neural integrity and brain function. Therefore, any change in brain connective tissue may affect the structure and function of neural system. The effect of ageing on connective tissues in other parts of the body, including the skin, have been well documented but the involvement of brain connective tissue in ageing and age related disorders needs to be evaluated in more detail. In summary, further longitudinal investigation

of non-grey matter (e.g. white matter and connective tissue) atrophy might be informative and may help explain gaps in our understanding of pathological processes associated with MCI and dementia.

Segmentation method

We investigated the impact of segmentation methodologies (manual vs automated) through meta-regression analyses and found that manual segmentation of the hippocampus resulted in larger atrophy rate estimates compared to automatic segmentation using FreeSurfer. While previous studies suggested that automatic segmentation with FreeSurfer resulted in a larger estimation of hippocampal volume in comparison with manual segmentation of the same images (Cherbuin, Anstey, ́glade-Meslin, & Sachdev, 2009; Wenger et al., 2014), atrophy rates have been reported to be lower in investigations using automatic segmentation (Mulder et al., 2014). As detailed in figure 2, differences between manual and automatic estimations of hippocampal atrophy are bigger in CN than MCI participants (68% compared to 40%), and in MCI (after removing the effect of normal ageing) the difference is remarkably less than CN (7% compared to 68%).

As suggested by Wenger (Wenger et al., 2014), automatic segmentation may classify some non-hippocampal tissue -- with lower atrophy rate -- as hippocampal tissue. This would explain how the automatic approach could result in higher volume estimates but lower atrophy rate. A systematic review by Fraser (Fraser et al., 2015), estimating annual hippocampal atrophy rate in healthy ageing across the lifespan, also detected a similar difference between manual tracing and automatic FreeSurfer segmentation and suggested that most studies using manual tracing excluded the tail of hippocampus and estimate the atrophy rate based on the atrophy of the head of the hippocampus. They concluded that hippocampal atrophy in CN was mostly restricted to the head of the hippocampus, rather

than the tail, therefore manual approaches, which excluded the tail were likely to estimate a lower atrophy rate compared with automatic FreeSurfer approaches, which included the tail. In summary, although manual tracing is traditionally considered as the gold standard method of hippocampal volume estimation, the difference between manual tracing and automatic approaches appear to be largely related to the subregions included in each method, rather than accuracy of estimation.

Moderators

An important question is whether study-specific factors such as age, female sex and APOEe4 influenced the reported estimates of brain atrophy in MCI. To investigate this question we performed a mixed-effects model analysis for hippocampal atrophy rates (the largest analysis group). The results showed that these moderators accounted for almost 80% of the observed heterogeneity between studies, with APOEe4 showing the largest moderating effect.

Moderating effects of age on brain atrophy have been well documented, although the pattern of association needs more investigation. It seems that this association is nonlinear and that the atrophy rate in stable MCI is larger at younger than older ages (Hua et al., 2010) although this was not confirmed in our meta-regression, possibly be due to a narrow age range as well as small number of studies in the meta-regression. Indeed, research consistent with this finding suggests that a higher whole brain atrophy rate is present in female compared to male individuals with MCI as well as in CN (Hua et al., 2010). However, although this appears to be the case across the brain, it may not apply at regional levels. This is the likely reason we did not find a sex effect in our hippocampal meta-regression. Previous evidence revealed that in different brain regions are different in male and female not only in MCI but also in CN. For example, atrophy rates for the thalamus, caudate nucleus and right

middle temporal gyrus are higher in male MCI, compared to female, and atrophy rates in the left middle temporal gyrus and precuneus are higher in female MCI than male (Skup et al., 2011). Our finding that APOEe4 genotype is a significant moderator and is associated with a higher rate of hippocampal atrophy in MCI is consistent with reviewed longitudinal studies that were not included in the meta-analysis. Moreover the effect appears to become more salient across the disease process with MCI and AD showing that APOEe4 genotype is associated with faster atrophy rates, (Aguilar et al., 2014; Vemuri et al., 2010) particularly in the hippocampus (Andrawis et al., 2012; Manning et al., 2014). Association between APOEe4 genotype and greater atrophy rate has been reported previously in CN (Morra et al., 2009). Thus, all parts of the brain do not seem to have a similar vulnerability to the effect of APOEe4 genotype and brain regions primarily involved in AD pathology, i.e. medial temporal lobe and particularly the hippocampus, are more affected, although the pattern of vulnerability is disease-stage specific (Tosun et al., 2011; Tosun et al., 2010). APOEe4 genotype is also associated with lower level of beta-amyloid (Schuff et al., 2009) and higher level of total and phosphorylated tau proteins (Tosun et al., 2010) in cerebrospinal fluid. All these biomarkers are shown to be associated with faster regional brain atrophy (particularly the hippocampus) together and separately (Chiang et al., 2011; Schuff et al., 2009; Toledo et al., 2014; Tosun et al., 2011; Tosun et al., 2010).

Strength and Limitations of the study

A broad search of the literature (e.g. using a wide range of search terms) and inclusion of all available studies (using all sorts of structural measurements) were major strengths of this review. Special care was taken to combine studies with compatible measurements -- to investigate pooled estimates of atrophy rates – and an attempt was made to comprehensively integrate incompatible findings and to summarise available knowledge about structural changes in MCI pathology. However, the review was limited by a relatively small number of

available studies that could be included in meta-analyses, particularly where whole brain and entorhinal cortex analyses are concerned. Additionally, many brain regions (such as the cerebellum) could not be analysed due to lack of evidence and should be the focus of future studies. Also due to the small number of studies in the meta-analysis, in relation to the number of moderators, it was recognized that estimates of moderator effects might be imprecise. The review was limited to comparing stable CN and prevalent MCI and data related to healthy participants converting to MCI and MCI participants converting to AD were insufficient to consider them in the present investigation.

Conclusion

To our knowledge this is the first systematic review of longitudinal studies investigating MCI related brain structural changes. The analyses revealed that the whole brain shrinks approximately two times faster in MCI participants compared to matched healthy people of the same age. Additionally, the medial temporal lobe regions -- particularly the entorhinal cortex and hippocampus -- are remarkably affected in AD pathology and associated with risk factors including APOEe4 genotype and female sex. These regions demonstrate an atrophy rate of 1.5 to 2.2%/year times for MCI compared to CN. Although the medial temporal lobe was reported as the region highly involved in AD related neurodegeneration, estimated atrophy rates in this region do not convincingly explain the amount of annual whole brain volume loss observed in MCI. Further investigation of other components of neural tissue, including white matter and non-neural brain tissue (e.g. connective tissue) are needed.

Acknowledgments

This study was funded by Australian Research Council project grant number 120101705. The funding sources were not involved in the design, collection, analysis or interpretation of data; or in the writing of the report or the decision to submit.

References:

Adaszewski, S., Dukart, J., Kherif, F., Frackowiak, R., & Draganski, B. (2013). How early can we predict Alzheimer's disease using computational anatomy? *Neurobiol Aging, 34*(12), 2815-2826. doi:10.1016/j.neurobiolaging.2013.06.015

Aguilar, C., Muehlboeck, J. S., Mecocci, P., Vellas, B., Tsolaki, M., Kloszewska, I., . . . Westman, E. (2014). Application of a MRI based index to longitudinal atrophy change in Alzheimer disease, mild cognitive impairment and healthy older individuals in the AddNeuroMed cohort. *Front Aging Neurosci, 6*, 145. doi:10.3389/fnagi.2014.00145

Andrawis, J. P., Hwang, K. S., Green, A. E., Kotlerman, J., Elashoff, D., Morra, J. H., . . . Apostolova, L. G. (2012). Effects of ApoE4 and maternal history of dementia on hippocampal atrophy. *Neurobiol Aging, 33*(5), 856-866. doi:10.1016/j.neurobiolaging.2010.07.020

Apostolova, L. G., Morra, J. H., Green, A. E., Hwang, K. S., Avedissian, C., Woo, E., . . . Thompson, P. M. (2010). Automated 3D mapping of baseline and 12-month associations between three verbal memory measures and hippocampal atrophy in 490 ADNI subjects. *Neuroimage, 51*(1), 488-499. doi:10.1016/j.neuroimage.2009.12.125

Archer, H. A., Kennedy, J., Barnes, J., Pepple, T., Boyes, R., Randlesome, K., . . . Fox, N. C. (2010). Memory complaints and increased rates of brain atrophy: risk factors for mild cognitive impairment and Alzheimer's disease. *Int J Geriatr Psychiatry, 25*(11), 1119-1126. doi:10.1002/gps.2440

Ayaz, M., Boikov, A. S., Haacke, E. M., Kido, D. K., & Kirsch, W. M. (2010). Imaging cerebral microbleeds using susceptibility weighted imaging: one step toward detecting vascular dementia. *J Magn Reson Imaging, 31*(1), 142-148. doi:10.1002/jmri.22001

Beckett, L. A., Harvey, D. J., Gamst, A., Donohue, M., Kornak, J., Zhang, H., & Kuo, J. H. (2010). The Alzheimer's Disease Neuroimaging Initiative: Annual change in biomarkers and clinical outcomes. *Alzheimers Dement, 6*(3), 257-264. doi:10.1016/j.jalz.2010.03.002

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). Introduction to Meta. *John Wiley & Sons.*

Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol, 82*, 239-259.

Brys, M., Glodzik, L., Mosconi, L., Switalski, R., De Santi, S., Pirraglia, E., . . . de Leon, M. J. (2009). Magnetic resonance imaging improves cerebrospinal fluid biomarkers in the early detection of Alzheimer's disease. *J Alzheimers Dis, 16*(2), 351-362. doi:10.3233/jad-2009-0968

Carmichael, O., Schwarz, C., Drucker, D., Fletcher, E., Harvey, D., Beckett, L., . . . DeCarli, C. (2010). Longitudinal changes in white matter disease and cognition in the first year of the Alzheimer disease neuroimaging initiative. *Arch Neurol, 67*(11), 1370-1378. doi:10.1001/archneurol.2010.284

Cherbuin, N., Anstey, K. J., glade-Meslin, C. R., & Sachdev, P. S. (2009). In Vivo Hippocampal Measurement and Memory: A Comparison of Manual Tracing and Automated Segmentation in a Large Community-Based Sample. *PLoS One, 4*(4), e5265. doi:10.1371/journal.pone.0005265

Chiang, G. C., Insel, P. S., Tosun, D., Schuff, N., Truran-Sacrey, D., Raptentsetsang, S. T., . . . Weiner, M. W. (2011). Impact of apolipoprotein E4-cerebrospinal fluid beta-amyloid interaction on hippocampal volume loss over 1 year in mild cognitive impairment. *Alzheimer's & Dementia, 7*(5), 514-520. doi:10.1016/j.jalz.2010.12.010

Desikan, R. S., Fischl, B., Cabral, H. J., Kemper, T. L., Guttmann, C. R., Blacker, D., . . . Killiany, R. J. (2008). MRI measures of temporoparietal regions show differential rates of atrophy during prodromal AD. *Neurology, 71*(11), 819-825. doi:10.1212/01.wnl.0000320055.57329.34

Desikan, R. S., McEvoy, L. K., Thompson, W. K., Holland, D., Roddey, J. C., Blennow, K., . . . Dale, A. M. (2011). Amyloid-beta associated volume loss occurs only in the presence of phospho-tau. *Ann Neurol, 70*(4), 657-661. doi:10.1002/ana.22509

Desikan, R. S., Sabuncu, M. R., Schmansky, N. J., Reuter, M., Cabral, H. J., Hess, C. P., . . . Fischl, B. (2010). Selective disruption of the cerebral neocortex in Alzheimer's disease. *PLoS One, 5*(9), e12853. doi:10.1371/journal.pone.0012853

Douaud, G. l., Refsum, H., Jager, C. A. d., Jacoby, R., Nichols, T. E., Smith, S. M., & Smith, A. D. (2013). Preventing Alzheimer's disease-related gray matter atrophy by B-vitamin treatment. *PANS, 110*(23), 9523-9528.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455-463.

Eckerstrom, C., Olsson, E., Borga, M., Ekholm, S., Ribbelin, S., Rolstad, S., . . . Malmgren, H. (2008). Small baseline volume of left hippocampus is associated with subsequent conversion of MCI into dementia: the Goteborg MCI study. *J Neurol Sci, 272*(1-2), 48-59. doi:10.1016/j.jns.2008.04.024

Evans, M. C., Barnes, J., Nielsen, C., Kim, L. G., Clegg, S. L., Blair, M., . . . Fox, N. C. (2010). Volume changes in Alzheimer's disease and mild cognitive impairment: cognitive associations. *Eur Radiol, 20*(3), 674-682. doi:10.1007/s00330-009-1581-5

Franko, E., & Joly, O. (2013). Evaluating Alzheimer's disease progression using rate of regional hippocampal atrophy. *PLoS One, 8*(8), e71354. doi:10.1371/journal.pone.0071354

Fraser, M. A., Shaw, M. E., & Cherbuin, N. (2015). A systematic review and meta-analysis of longitudinal hippocampal atrophy in healthy human ageing. *Neuroimage,* http://dx.doi.org/10.1016/j.neuroimage.2015.03.035. doi:10.1016/j.neuroimage.2015.03.035

Guo, H., Song, X., Schmidt, M. H., Vandorpe, R., Yang, Z., LeBlanc, E., . . . Rockwood, K. (2014). Evaluation of whole brain health in aging and Alzheimer's disease: a standard procedure for scoring an MRI-based brain atrophy and lesion index. *J Alzheimers Dis, 42*(2), 691-703. doi:10.3233/jad-140333

Guo, L. H., Alexopoulos, P., Wagenpfeil, S., Kurz, A., & Perneczky, R. (2013). Brain size and the compensation of Alzheimer's disease symptoms: a longitudinal cohort study. *Alzheimers Dement, 9*(5), 580-586. doi:10.1016/j.jalz.2012.10.002

Gutman, B. A., Hua, X., Rajagopalan, P., Chou, Y. Y., Wang, Y., Yanovsky, I., . . . Thompson, P. M. (2013). Maximizing power to track Alzheimer's disease and MCI progression by LDA-based weighting of longitudinal ventricular surface features. *Neuroimage, 70*, 386-401. doi:10.1016/j.neuroimage.2012.12.052

Henneman, W. J., Sluimer, J. D., Barnes, J., van der Flier, W. M., Sluimer, I. C., Fox, N. C., . . . Barkhof, F. (2009). Hippocampal atrophy rates in Alzheimer disease: added value over whole brain volume measures. *Neurology, 72*(11), 999-1007. doi:10.1212/01.wnl.0000344568.09360.31

Henneman, W. J., Vrenken, H., Barnes, J., Sluimer, I. C., Verwey, N. A., Blankenstein, M. A., . . . van der Flier, W. M. (2009). Baseline CSF p-tau levels independently predict progression of hippocampal atrophy in Alzheimer disease. *Neurology, 73*(12), 935-940. doi:10.1212/WNL.0b013e3181b879ac

Ho, A. J., Hua, X., Lee, S., Leow, A. D., Yanovsky, I., Gutman, B., . . . Thompson, P. M. (2010). Comparing 3 T and 1.5 T MRI for tracking Alzheimer's disease progression

with tensor-based morphometry. *Hum Brain Mapp, 31*(4), 499-514. doi:10.1002/hbm.20882

Holland, D., Brewer, J. B., Hagler, D. J., Fennema-Notestine, C., & Dale, A. M. (2009). Subregional neuroanatomical change as a biomarker for Alzheimer's disease. *Proc Natl Acad Sci U S A, 106*(49), 20954-20959.

Hua, X., Hibar, D. P., Lee, S., Toga, A. W., Jack, C. R., Jr., Weiner, M. W., & Thompson, P. M. (2010). Sex and age differences in atrophic rates: an ADNI study with n=1368 MRI scans. *Neurobiol Aging, 31*(8), 1463-1480. doi:10.1016/j.neurobiolaging.2010.04.033

Hua, X., Lee, S., Yanovsky, I., Leow, A. D., Chou, Y. Y., Ho, A. J., . . . Thompson, P. M. (2009). Optimizing power to track brain degeneration in Alzheimer's disease and mild cognitive impairment with tensor-based morphometry: an ADNI study of 515 subjects. *Neuroimage, 48*(4), 668-681. doi:10.1016/j.neuroimage.2009.07.011

Insel, P. S., Mattsson, N., Donohue, M. C., Mackin, R. S., Aisen, P. S., Jack, C. R., Jr., . . . Weiner, M. W. (2014). The transitional association between beta-amyloid pathology and regional brain atrophy. *Alzheimers Dement*. doi:10.1016/j.jalz.2014.11.002

Jack, C. R., Jr., Lowe, V. J., Weigand, S. D., Wiste, H. J., Senjem, M. L., Knopman, D. S., . . . Petersen, R. C. (2009). Serial PIB and MRI in normal, mild cognitive impairment and Alzheimer's disease: implications for sequence of pathological events in Alzheimer's disease. *Brain, 132*(Pt 5), 1355-1365. doi:10.1093/brain/awp062

Jack, C. R., Jr., Petersen, R. C., Xu, Y., O'Brien, P. C., Smith, G. E., Ivnik, R. J., . . . Kokmen, E. (2000). Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. *Neurology, 55*(4), 484-489.

Jack, C. R., Jr., Shiung, M. M., Gunter, J. L., O'Brien, P. C., Weigand, S. D., Knopman, D. S., . . . Petersen, R. C. (2004). Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD. *Neurology, 62*(4), 591-600.

Jack, C. R., Jr., Shiung, M. M., Weigand, S. D., O'Brien, P. C., Gunter, J. L., Boeve, B. F., . . . Petersen, R. C. (2005). Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic MCI. *Neurology, 65*(8), 1227-1231. doi:10.1212/01.wnl.0000180958.22678.91

Jack, C. R., Jr., Weigand, S. D., Shiung, M. M., Przybelski, S. A., O'Brien, P. C., Gunter, J. L., . . . Petersen, R. C. (2008). Atrophy rates accelerate in amnestic mild cognitive impairment. *Neurology, 70*(19 Pt 2), 1740-1752. doi:10.1212/01.wnl.0000281688.77598.35

Kljajevic, V., Grothe, M. J., Ewers, M., & Teipel, S. (2014). Distinct pattern of hypometabolism and atrophy in preclinical and predementia Alzheimer's disease. *Neurobiol Aging, 35*(9), 1973-1981. doi:10.1016/j.neurobiolaging.2014.04.006

Leow, A. D., Yanovsky, I., Parikshak, N., Hua, X., Lee, S., Toga, A. W., . . . Thompson, P. M. (2009). Alzheimer's disease neuroimaging initiative: a one-year follow up study using tensor-based morphometry correlating degenerative rates, biomarkers and cognition. *Neuroimage, 45*(3), 645-655.

Leung, K. K., Barnes, J., Ridgway, G. R., Bartlett, J. W., Clarkson, M. J., Macdonald, K., . . . Ourselin, S. (2010). Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease. *Neuroimage, 51*(4), 1345-1359. doi:10.1016/j.neuroimage.2010.03.018

Leung, K. K., Ridgway, G. R., Ourselin, S., & Fox, N. C. (2012). Consistent multi-time-point brain atrophy estimation from the boundary shift integral. *Neuroimage, 59*(4), 3995-4005. doi:10.1016/j.neuroimage.2011.10.068

Li, Y., Wang, Y., Wu, G., Shi, F., Zhou, L., Lin, W., & Shen, D. (2012). Discriminant analysis of longitudinal cortical thickness changes in Alzheimer's disease using

dynamic and network features. *Neurobiol Aging, 33*(2), 427 e415-430. doi:10.1016/j.neurobiolaging.2010.11.008

Lillemark, L., Sorensen, L., Pai, A., Dam, E. B., & Nielsen, M. (2014). Brain region's relative proximity as marker for Alzheimer's disease based on structural MRI. *BMC Med Imaging, 14*, 21. doi:10.1186/1471-2342-14-21

Liu, T., Sachdev, P. S., Lipnicki, D. M., Jiang, J., Cui, Y., Kochan, N. A., . . . Wen, W. (2013). Longitudinal changes in sulcal morphology associated with late-life aging and MCI. *Neuroimage, 74*, 337-342. doi:10.1016/j.neuroimage.2013.02.047

Lo, R. Y., Hubbard, A. E., Shaw, L. M., Trojanowski, J. Q., Petersen, R. C., Aisen, P. S., . . . Jagust, W. J. (2011). Longitudinal change of biomarkers in cognitive decline. *Arch Neurol, 68*(10), 1257-1266. doi:10.1001/archneurol.2011.123

Lorenzi, M., Pennec, X., Frisoni, G. B., & Ayache, N. (2015). Disentangling normal aging from Alzheimer's disease in structural magnetic resonance images. *Neurobiol Aging, 36 Suppl 1*, S42-52. doi:10.1016/j.neurobiolaging.2014.07.046

Madsen, S. K., Gutman, B. A., Joshi, S. H., Toga, A. W., Jack, C. R., Jr., Weiner, M. W., & Thompson, P. M. (2015). Mapping ventricular expansion onto cortical gray matter in older adults. *Neurobiol Aging, 36 Suppl 1*, S32-41. doi:10.1016/j.neurobiolaging.2014.03.044

Manning, E. N., Barnes, J., Cash, D. M., Bartlett, J. W., Leung, K. K., Ourselin, S., & Fox, N. C. (2014). APOE epsilon4 is associated with disproportionate progressive hippocampal atrophy in AD. *PLoS One, 9*(5), e97608. doi:10.1371/journal.pone.0097608

Marshall, G. A., Lorius, N., Locascio, J. J., Hyman, B. T., Rentz, D. M., Johnson, K. A., & Sperling, R. A. (2014). Regional cortical thinning and cerebrospinal biomarkers predict worsening daily functioning across the Alzheimer's disease spectrum. *J Alzheimers Dis, 41*(3), 719-728. doi:10.3233/jad-132768

McDonald, C. R., Gharapetian, L., McEvoy, L. K., Fennema-Notestine, C., Hagler, D. J., Jr., Holland, D., & Dale, A. M. (2012). Relationship between regional atrophy rates and cognitive decline in mild cognitive impairment. *Neurobiol Aging, 33*(2), 242-253. doi:10.1016/j.neurobiolaging.2010.03.015

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, t. P. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses:

The PRISMA Statement. *Annuals of Internal Medicine, 151*(4), 264-270.

Morra, J. H., Tu, Z., Apostolova, L. G., Green, A. E., Avedissian, C., Madsen, S. K., . . . Thompson, P. M. (2009). Automated mapping of hippocampal atrophy in 1-year repeat MRI data from 490 subjects with Alzheimer's disease, mild cognitive impairment, and elderly controls. *Neuroimage, 45*(1 Suppl), S3-15. doi:10.1016/j.neuroimage.2008.10.043

Mouiha, A., & Duchesne, S. (2011). Hippocampal atrophy rates in Alzheimer's disease: automated segmentation variability analysis. *Neurosci Lett, 495*(1), 6-10. doi:10.1016/j.neulet.2011.02.065

Mulder, E. R., de Jong, R. A., Knol, D. L., van Schijndel, R. A., Cover, K. S., Visser, P. J., . . . Vrenken, H. (2014). Hippocampal volume change measurement: quantitative assessment of the reproducibility of expert manual outlining and the automated methods FreeSurfer and FIRST. *Neuroimage, 92*, 169-181. doi:10.1016/j.neuroimage.2014.01.058

Nowrangi, M. A., Lyketsos, C. G., Leoutsakos, J. M., Oishi, K., Albert, M., Mori, S., & Mielke, M. M. (2013). Longitudinal, region-specific course of diffusion tensor

imaging measures in mild cognitive impairment and Alzheimer's disease. *Alzheimer's & Dementia, 9*(5), 519-528. doi:10.1016/j.jalz.2012.05.2186

Park, M. H., Han, C., & Initiative, t. A. s. D. N. (2015). Is there an MCI reversion to cognitively normal? Analysis of Alzheimer's disease biomarkers profiles. *international Psychiatrics*

, 27(3), 429-437. doi:10.1017/S1041610214002129

Prestia, A., Drago, V., Rasser, P. E., Bonetti, M., Thompson, P. M., & Frisoni, G. B. (2010). Cortical changes in incipient Alzheimer's disease. *J Alzheimers Dis, 22*(4), 1339-1349. doi:10.3233/jad-2010-101191

Schott, J. M., Bartlett, J. W., Barnes, J., Leung, K. K., Ourselin, S., & Fox, N. C. (2010). Reduced sample sizes for atrophy outcomes in Alzheimer's disease trials: baseline adjustment. *Neurobiol Aging, 31*(8), 1452-1462, 1462 e1451-1452. doi:10.1016/j.neurobiolaging.2010.04.011

Schuff, N., Tosun, D., Insel, P. S., Chiang, G. C., Truran, D., Aisen, P. S., . . . Weiner, M. W. (2012). Nonlinear time course of brain volume loss in cognitively normal and impaired elders. *Neurobiol Aging, 33*(5), 845-855. doi:10.1016/j.neurobiolaging.2010.07.012

Schuff, N., Woerner, N., Boreta, L., Kornfield, T., Shaw, L. M., Trojanowski, J. Q., . . . Weiner, M. W. (2009). MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. *Brain, 132*(Pt 4), 1067-1077. doi:10.1093/brain/awp007

Selnes, P., Aarsland, D., Bjornerud, A., Gjerstad, L., Wallin, A., Hessen, E., . . . Fladby, T. (2013). Diffusion tensor imaging surpasses cerebrospinal fluid as predictor of cognitive decline and medial temporal lobe atrophy in subjective cognitive impairment and mild cognitive impairment. *J Alzheimers Dis, 33*(3), 723-736. doi:10.3233/jad-2012-121603

Skup, M., Zhu, H., Wang, Y., Giovanello, K. S., Lin, J. A., Shen, D., . . . Zhang, H. (2011). Sex differences in grey matter atrophy patterns among AD and aMCI patients: results from ADNI. *Neuroimage, 56*(3), 890-906. doi:10.1016/j.neuroimage.2011.02.060

Sluimer, J. D., van der Flier, W. M., Karas, G. B., van Schijndel, R., Barnes, J., Boyes, R. G., . . . Barkhof, F. (2009). Accelerating regional atrophy rates in the progression from normal aging to Alzheimer's disease. *Eur Radiol, 19*(12), 2826-2833. doi:10.1007/s00330-009-1512-5

Smith, A. D., Smith, S. M., Jager, C. A. d., Whitbread, P., Johnston, C., Agacinski, G., . . . Refsum, H. (2010). Homocysteine-Lowering by B Vitamins Slows the Rate of Accelerated Brain Atrophy in Mild Cognitive Impairment: A Randomized Controlled Trial. *PLoS One, 5*(9). doi:10.1371/journal.pone.0012244

Snyder, P. J., Kahle-Wrobleski, K., Brannan, S., Miller, D. S., Schindler, R. J., DeSanti, S., . . . Carrillo, M. C. (2014). Assessing cognition and function in Alzheimer's disease clinical trials: Do we have the right tools? *Alzheimer's & Dementia, 10*, 853-860. doi:10.1016/j.jalz.2014.07.158

Song, X., Mitnitski, A., Zhang, N., Chen, W., & Rockwood, K. (2013). Dynamics of brain structure and cognitive function in the Alzheimer's disease neuroimaging initiative. *J Neurol Neurosurg Psychiatry, 84*(1), 71-78. doi:10.1136/jnnp-2012-303579

Standring, S. (2008). Gray's Anatomy The Anatomical Basis of Clinical Practice, 40th Ed *40th*.

Stoub, T. R., Rogalski, E. J., Leurgans, S., Bennett, D. A., & deToledo-Morrell, L. (2010). Rate of entorhinal and hippocampal atrophy in incipient and mild AD: relation to

memory function. *Neurobiol Aging, 31*(7), 1089-1098. doi:10.1016/j.neurobiolaging.2008.08.003

Teipel, S., Heinsen, H., Amaro, E., Jr., Grinberg, L. T., Krause, B., & Grothe, M. (2014). Cholinergic basal forebrain atrophy predicts amyloid burden in Alzheimer's disease. *Neurobiol Aging, 35*(3), 482-491. doi:10.1016/j.neurobiolaging.2013.09.029

Toledo, J. B., Da, X., Weiner, M. W., Wolk, D. A., Xie, S. X., Arnold, S. E., . . . Trojanowski, J. Q. (2014). CSF Apo-E levels associate with cognitive decline and MRI changes. *Acta Neuropathol, 127*(5), 621-632. doi:10.1007/s00401-013-1236-0

Tosun, D., Schuff, N., Shaw, L. M., Trojanowski, J. Q., & Weiner, M. W. (2011). Relationship between CSF biomarkers of Alzheimer's disease and rates of regional cortical thinning in ADNI data. *J Alzheimers Dis, 26 Suppl 3*, 77-90. doi:10.3233/jad-2011-0006

Tosun, D., Schuff, N., Truran-Sacrey, D., Shaw, L. M., Trojanowski, J. Q., Aisen, P., . . . Weiner, M. W. (2010). Relations between brain tissue loss, CSF biomarkers, and the ApoE genetic profile: a longitudinal MRI study. *Neurobiol Aging, 31*(8), 1340-1354. doi:10.1016/j.neurobiolaging.2010.04.030

Vemuri, P., Wiste, H. J., Weigand, S. D., Knopman, D. S., Trojanowski, J. Q., Shaw, L. M., . . . Jack, C. R., Jr. (2010). Serial MRI and CSF biomarkers in normal aging, MCI, and AD. *Neurology, 75*(2), 143-151. doi:10.1212/WNL.0b013e3181e7ca82

Villain, N., Fouquet, M., Baron, J. C., Mezenge, F., Landeau, B., de La Sayette, V., . . . Chetelat, G. (2010). Sequential relationships between grey matter and white matter atrophy and brain metabolic abnormalities in early Alzheimer's disease. *Brain, 133*(11), 3301-3314. doi:10.1093/brain/awq203

Villemagne, V. L., Burnham, S., Bourgeat, P., Brown, B., Ellis, K. A., Salvado, O., . . . Masters, C. L. (2013). Amyloid beta deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol, 12*(4), 357-367. doi:10.1016/s1474-4422(13)70044-9

Wang, P. N., Liu, H. C., Lirng, J. F., Lin, K. N., & Wu, Z. A. (2009). Accelerated hippocampal atrophy rates in stable and progressive amnestic mild cognitive impairment. *Psychiatry Res, 171*(3), 221-231. doi:10.1016/j.pscychresns.2008.05.002

Wells, G., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2011). The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analysis. *Available:* http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

Wenger, E., Martensson, J., Noack, H., Bodammer, N. C., Kuhn, S., Schaefer, S., . . . Lovden, M. (2014). Comparing manual and automatic segmentation of hippocampal volumes: reliability and validity issues in younger and older brains. *Hum Brain Mapp, 35*(8), 4236-4248. doi:10.1002/hbm.22473

Yao, Z., Hu, B., Liang, C., Zhao, L., & Jackson, M. (2012). A longitudinal study of atrophy in amnestic mild cognitive impairment and normal aging revealed by cortical thickness. *PLoS One, 7*(11), e48973. doi:10.1371/journal.pone.0048973

Zhang, N., Song, X., & Zhang, Y. (2012). Combining structural brain changes improves the prediction of Alzheimer's disease and mild cognitive impairment. *Dement Geriatr Cogn Disord, 33*(5), 318-326. doi:10.1159/000339364

Zhang, N., Song, X., Zhang, Y., Chen, W., D'Arcy, R. C., Darvesh, S., . . . Rockwood, K. (2011). An MRI brain atrophy and lesion index to assess the progression of structural changes in Alzheimer's disease, mild cognitive impairment, and normal aging: a follow-up study. *J Alzheimers Dis, 26 Suppl 3*, 359-367. doi:10.3233/jad-2011-0048

STUDY 2

THE CEREBELLUM SHRINKS FASTER THAN NORMAL AGEING IN ALZHEIMER'S DISEASE BUT NOT IN MILD COGNITIVE IMPAIRMENT

Hossein Tabatabaei-Jafari[1], Erin Walsh, Marnie E. Shaw[1], Nicolas Cherbuin[1], PhD;
For the Alzheimer's Disease Neuroimaging Initiative (ADNI)

[1]Centre for Research on Ageing, Health and Wellbeing, Australian National University, Canberra, ACT 2601, Australia

Published in Human Brain Mapping, 38 (2017): 3141–3150

DOI: 10.1002/hbm.23580

Abstract

Background: While acceleration in age-related cerebral atrophy has been well documented in Alzheimer's disease, the cerebellar contributions to this effect have not been thoroughly investigated.

Objective: This study investigated cerebellar volume and atrophy rate using magnetic resonance imaging in individuals with normal cognition (CN), mild cognitive impairment (MCI) and Alzheimer's disease (AD).

Method: Two hundred and twenty nine CN, 398 MCI and 191 AD participants of stage one ADNI database with screening scans were evaluated for cerebellar volume. Of those, 758 individuals with two or more follow up scans were categorized into stable, converted and reverted CN, MCI and AD and evaluated for cerebellar atrophy rate.

Results: Cerebellar volume was 2.5% larger in CN than in those with AD but there were no differences between CN and MCI and MCI and AD in cross-sectional analysis. Similarly, the atrophy rate was 49% larger in AD and 64% larger in MCI who converted to AD but no difference was detected between CN and MCI. There were no associations between education and APOEe4 and cerebellar volume or cerebellar atrophy across the diagnostic groups.

Conclusion: Cerebellar atrophy contributes to Alzheimer's clinical progression but mostly at the late stage of the disease. However, even in the late stage shrinkage rate is less than the average of the shrinkage in the cerebrum and is not associated with AD moderators. This suggests that cerebellar involvement is secondary to cerebral involvement and can be due to network connection spread regardless of the primary pathology.

Introduction

The human cerebellum is a brain structure well known for its role in motor function and recently has drawn attention for its implication in cognitive functions (Schmahmann & Sherman, 1998; Stoodley, 2012; Weier et al., 2014; Wolf, Rapoport, & Schweizer, 2009). It is connected to almost all parts of the nervous system, comprises more than 50% of the total brain neurons, but surprisingly contributes to only 10% of the whole brain volume (B. B. Andersen, Korbo, & Pakkenberg, 1992). This mismatch is a reflection of the difference in neural architecture. Gray matter makes up 80% of the cerebellar volume (compared with less than half for the cerebrum) (Hoogendam et al., 2012) and consists of densely packed small granular neurons tightly folded which are less diverse compared to those of the cerebral cortex. In contrast to the variety of cytoarchitectonic organisation observed in different regions of the cerebral cortex, all regions of the cerebellar cortex appear similar in histological sections (Standring, 2008). Specific histological architecture in addition to rich connections to the other parts of the brain makes the cerebellum an important region to investigate in the context of neurodegenerative disorders.

Pathologically, Alzheimer's disease (AD) is characterized by abnormal intra and extra cellular protein aggregations, i.e. intracellular tau phosphorylation and extracellular β-amyloid deposition. Studies using positron emission tomography (PET) revealed significant correlations between post-mortem and in vivo presence and density of amyloid plaques and phosphorylated tau: ^{11}C-labeled Pittsburgh compound B (^{11}C-PiB) (Driscoll et al., 2012) and Florbetapir-PET imaging (Clark et al., 2011) for β-amyloid deposition and labelled THK5117-PET (Lemoine et al., 2015) for aggregated hyper phosphorylated tau. PET studies suggested no difference in the cerebellar uptake in AD and cognitively normal (CN) participants (Jack, Lowe, et al., 2008; Jonasson et al., 2016; Rowe et al., 2007) and therefore

it has been adopted as a normalizing area for standardized uptake values (SUVs) (Jonasson et al., 2016; Lopresti et al., 2005).

Although AD related shrinkage and neuronal death are thought to be associated with and possibly due to β-amyloid deposition and tau aggregation (Wang, D'Andrea, & Nagele, 2002), their topological patterns and progression are different (Braak & Braak, 1991; Thal, Rub, Orantes, & Braak, 2002). Moreover, the pattern of regional brain atrophy in AD does not follow precisely either β-amyloid or tau topological patterns (Sluimer et al., 2009). Therefore, normal level of β-amyloid deposition and tau aggregation may not rule out the presence of neuronal loss or shrinkage in the cerebellum. A recent post-mortem stereological study suggested no significant differences in the cerebellar total Purkinje and granular cell number nor in the volume of the granular layer between people with severe AD and normal individuals (K. Andersen, Andersen, & Pakkenberg, 2012). However, this finding is inconsistent with a previous study that showed a significant reduction in the granular layer in AD (Wegiel et al., 1999) although both studies reported significant reduction in whole cerebellar volume. These somewhat inconsistent findings may be due to the fact that these studies were post-mortem (cross-sectional) with low sample sizes (20 and 16 subjects respectively) in qualitatively different cohorts and thus afforded low statistical power.

To bypass the inevitable limitations of post mortem studies (single measurement occasion and small sample size), structural neuroimaging techniques including magnetic imaging are the best available option for longitudinal examination of brain volume change over time. Our recent published systematic review (Tabatabaei-Jafari, Shaw, & Cherbuin, 2015) revealed that there is no morphological longitudinal study aimed at comparing cerebellar structural change in normal ageing and cognitively impaired populations including mild cognitive impairment (MCI) and Alzheimer's disease. Therefore, the main aim of this

study is to evaluate cross-sectional and longitudinal structural differences in the cerebellum across cognitively different populations including CN, MCI and AD.

Methodology

Study Participants

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 MCI and early AD.

All individuals participating in ADNI1 study who underwent MRI screening and diagnostic evaluations were included in the cross-sectional analysis and categorized into 3 diagnostic groups: CN, MCI and AD. Participants with additional scans in follow-up assessments were included in the longitudinal analysis and categorised into more specific diagnostic groups according to the diagnosis at the first and last scanning time points. Details of the diagnostic criteria can be found on the ADNI web site (http://www.adni-info.org/Scientists/AboutADNI.aspx). Briefly, participants were categorized as CN if they had a Mini Mental State Examination (MMSE) score higher than 24, a Clinical Dementia Rating (CDR) of 0 and were not diagnosed with MCI, dementia or depression. Participants were categorized as MCI if they had a MMSE score higher than 24, a subjective report of memory concern, a measured objective memory loss, a CDR of 0.5, absence of dementia and preserved daily living activities. Participants were categorized as AD if they had a MMSE

score lower than 26, a CDR of 0.5 or 1.0, and fulfilled criteria for clinically probable AD according to the Institute of Neurological and Communicative Diseases and Stroke/ Alzheimer's Disease and Related Disorders Association. Participants with follow-up evaluation were categorized into stable, converted or reverted CN, MCI and AD according to the first and last time points diagnoses: stable if the first and last evaluation were similar, converted if the last evaluation progressed to declined cognitive diagnosis and reverted if the last evaluation was improved.

Image Acquisition

Participants underwent a high-resolution MRI scans of the brain on 1.5 T scanners from General Electric, Siemens, or Philips (Milwaukee, WI, USA; Germany; the Netherlands respectively) across multiple scanners using a standardized MRI protocol for 3D MP-RAGE sequences (Jack, Bernstein, et al., 2008) and following parameters: TR= 2400 ms, minimum full TE, TI=1000 ms, flip angle= 8°, 24 cm field of view, acquisition matrix of 192 ×192 ×166 and yielding 1.25 ×1.25 ×1.2 mm^3.

Segmentation and Image analysis

Volumetric segmentation were conducted by the ADNI team at the University of California, San Francisco using FreeSurfer version 5.1 for longitudinal analyses (Reutera, Schmansky, Rosasa, & Fischl, 2012). The cerebellum was automatically segmented into gray matter and white matter. Sum values of the gray and white matter were considered as hemisphere volume and total of left and right were considered as cerebellar volumes.

Statistical analysis

The R statistical software (version 3.1.1) was used for the cross-sectional and longitudinal analyses. The intra-class correlation coefficient (ICC) for the repeated

longitudinal cerebellar volumes measurements was 0.98 (95%CI 0.9803 - .9843), which indicates that most of the variance (~96%) occurs between participants while only 4% occurs within participants.

Non-parametric locally weighted scatterplot smoothing (LOWESS) was used to visually inspect the data to determine whether linear models were appropriate. The LOWESS approach uses weighted least squares (giving more weight to points near the point whose response is being estimated) to estimate the mean response value at each time point and provide a smooth line representing the relationship between dependent and explanatory variables, when there are no assumptions about the relationship. The LOWESS plots for cerebellar volume versus age suggested that linear modelling of the relationship between cerebellar volume and age was appropriate for cross-sectional and longitudinal analyses since little departure from linearity was observed across groups except for CNc, which assumed to be due to low sample size i.e. 19 participants (Fig 1).

The lme4 package (version 1.1-7) was used to conduct linear regressions analyses. In cross-sectional analyses, multiple linear regressions were conducted to investigate the cross-sectional relationship between cerebellar volume and clinical diagnosis status. Cerebellar volume was applied as dependent variable and age (centred on 55, the youngest participants at baseline), sex, education, APOE e4, diagnosis and intracranial volume (ICV) were considered as explanatory variables. In longitudinal analyses, mixed effects models were applied with the same explanatory variables for linear regressions in addition to a random effect by scanner and two random effects by subjects: a random intercept and a random slope for age at each time point. The random slope of time (centred age at each time point) was tested in a minimally controlled model and if statistically significant was included in the model as random effect (Bernal-Rusiel et al., 2013). A time by clinical diagnosis group

interaction effect was tested to determine whether the rate of change in cerebellar volume differed between groups. Fixed effect of age on cerebellar volume for each diagnostic group was considered as cerebellar atrophy rate.

The final models were visually checked for any obvious deviations from homoscedasticity, normality of residuals and linearity. Likelihood ratio test of the model with the effect in question against the model without was used to determine statistical significance.

Results

Demography

Cross-sectional: Eight hundred and eighteen participants were categorized into CN, MCI and AD. There were no significant differences in age across the groups, but significant differences in sex and APOE e4 distributions among the diagnostic groups. The male ratio was higher in MCI and, as expected, APOEe4 frequencies were significantly higher in MCI and AD. AD participants were significantly less educated than CN (Table 1).

Longitudinal: Of 818 participants with screening scans 758, who had one or more follow-up scans and cognitive tests, were included in the longitudinal part. They were categorized into different diagnostic groups according to the first and last time points diagnoses: stable CN (CNs), CN converted to MCI (CNc), stable MCI (MCIs), MCI converted to AD (MCIc), stable AD (ADs), CN converted to AD, MCI reverted to CN (MCIr) and AD reverted to MCI (ADr). There were no significant differences in age and education across the diagnostic groups except for education between CNs and ADs. Pearson chi-squared test revealed no significant difference in sex distribution but a significant difference in APOE e4 distributions between diagnostic groups. APOEe4 distributions were

higher in MCIs than CNs and in ADs than CNs. The mean follow up period across the groups was 2.54 (1.20) years, which was shorter in MCIs and ADs compared with CNs.

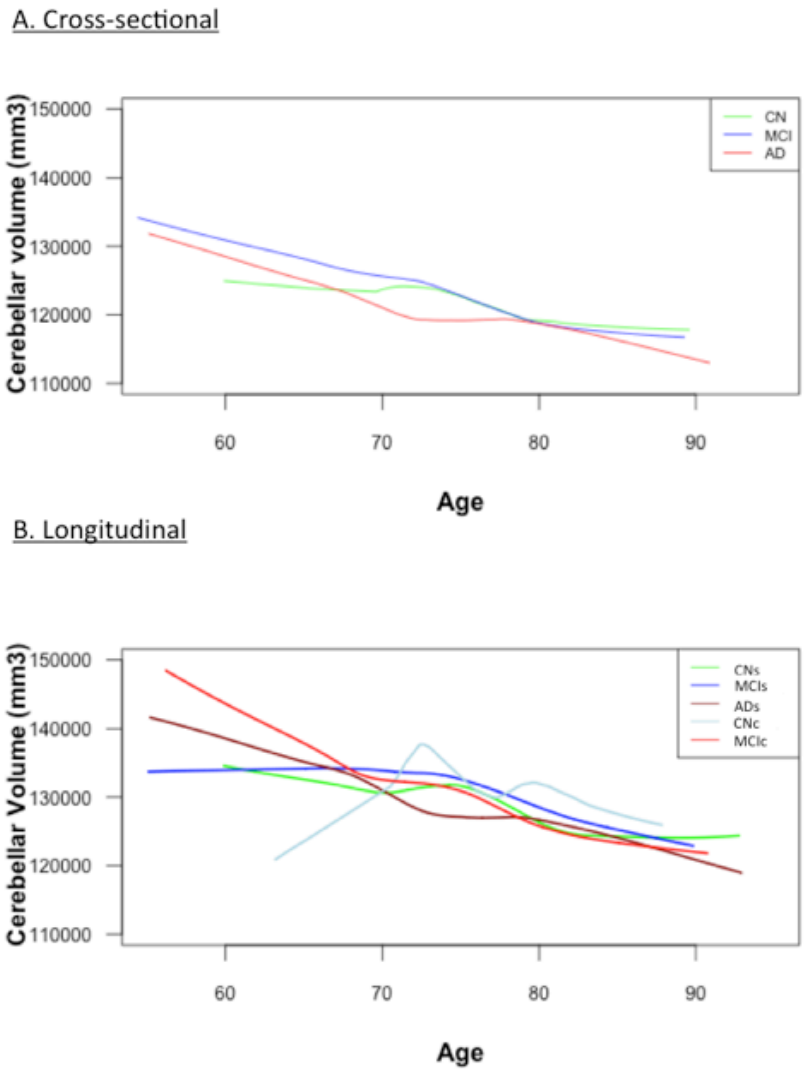


Figure 1: Locally weighted smoothed mean measurement trajectory (LOWESS plot) cerebellar volumes vs. age. (A) Three clinical groups including cognitively normal (CN), mild cognitive impairment (MCI), and Alzheimer's disease (AD) in cross-sectional level. (B) Five clinical groups including stable cognitively normal (CNs), cognitively normal converted to mild cognitive impairment (CNc), stable mild cognitive impairment (MCIs), mild cognitive impairment converted to AD (MCIc), and stable Alzheimer's disease (ADs) in serial scans.

Table 1: Demographic Information

	Cross-sectional (N=818)			Longitudinal (N=758)							
	CN	MCI	AD	CNs	CNc	MCIs	MCIc	CN to AD	ADs	MCIr	ADr
No. of participants	229	398	191	196	19	193	161	2	172	13	2
Age at baseline, year (SD)	75.87 (5.02)	74.74 (7.39)	75.27 (7.46)	75.76 (5.03)	77.45 (5.22)	75.00 (7.42)	74.73 (6.71)	80.55 (3.61)	75.12 (7.61)	73.43 (9.96)	79.50 (4.38)
Male sex, n (%)	119 (52)	257 (65)	100 (52)	103 (53)	11 (58)	124 (64)	101 (63)	0 (0)	91 (53)	9 (69)	2 (100)
Education, year (SD)	16.07 (2.86)	15.64 (3.03)	14.70 (3.15)	16.13 (2.88)	15.95 (2.39)	15.45 (3.15)	15.84 (2.81)	15.00 (0.00)	14.77 (3.14)	16.00 (2.42)	16.00 (0.00)
APOEe4, n (%)	61 (27)	212 (53)	127 (67)	49 (25)	8 (42)	92 (48)	104 (65)	1 (50)	117 (68)	4 (31)	1 (50)
MMSE at baseline (SD)	29.11 (0.99)	27.03 (1.78)	23.31 (2.04)	29.08 (1.06)	29.32 (0.75)	27.22 (1.77)	26.71 (1.72)	29.5 (0.71)	23.40 (1.97)	27.85 (1.77)	26.00 (0.00)

CN; cognitively normal, MCI; mild cognitive impairment, AD; Alzheimer's disease, CNs, stable cognitively normal; cognitively normal converted to mild cognitive impairment; MCIs, stable mild cognitive impairment; MCIc, mild cognitive impairment converted to Alzheimer's disease; Ads, stable Alzheimer's disease; MCIr, mild cognitive impairment reverted to cognitively normal; ADr, Alzheimer's disease reverted to mild cognitive impairment; APOEe4, Apolipoprotein alleles e 4 genotype; MMSE; mini-mental state examination

Cross-sectional results

A significant association between cognitive diagnosis and cerebellar volume ($F(2,811)=3.95$, $P<0.01$) was detected.

Pairwise comparisons demonstrated ($3400mm^3$; ~ 2.5%) larger cerebellar volume in CN compared to AD ($F(1,413)= 9.82$, $P<0.001$), but no differences between CN and MCI ($F(1, 620)=3.40$ $P>0.1$), and MCI and AD ($F(1, 582)=1.62$, $P>0.1$). Table 2 presents the mean ICV-adjusted cerebellar volumes and the fixed effect of age for the three diagnostic groups. Although, the average cerebellar volume was significantly smaller in AD compared to CN and MCI, the slope of decrease in cerebellar volume for each year increase in age was only 0.41% (CN; 0.34%, MCI; 0.42%, AD; 0.38%) and was not significantly different across groups ($F(2, 809)= 0.28$, $p> 0.5$) and in pair-wise comparisons ($F < 0.5$, $p> 0.1$). When all explanatory variables were included, the linear regression model explained 44.7% of the variance in cerebellar volume ($F(8, 809)= 83.61$, $P< 0.0001$) mostly explained by ICV (37.9%) with 7.7% explained by age alone, and 0.7% by clinical group.

The scatters plot presenting the association between age and cerebellar volume for each group also revealed an initial overlap of CN and MCI regression lines followed by deviation of MCI regression line to AD line suggesting that cerebellar volumes are highly similar in CN and MCI at younger ages but lower in MCI in older individuals (Figure 2A). In contrast the AD regression line while following a similar slope had a clearly different intercept suggesting a constant smaller cerebellar volume in AD across the age span investigated. Similar patterns were demonstrated for the left and right cerebellar volumes (Table 2).

Table 2: Cerebellar measures

	Cross-sectional (N=818)			Longitudinal (N=758)							
	CN	MCI	AD	CNs	CNc	MCIs	MCIc	CN to AD	ADs	MCIr	ADr
No. of participants	229	398	191	196	19	193	161	2	172	13	2
Baseline volume, mm^3 (SD)*											
Total	123249.80 (10018.69)	122394.90 (10632.40)	120706.90 (10591.60)	131276.60 (11163.98)	132820.00 (9450.25)	130319.50 (11317.70)	130154.90 (11680.98)	131354.80 (778.99)	129524.10 (10838.42)	131603.50 (10703.89)	114157.80 (17439.08)
Left	61739.12 (6539.80)	60420.98 (6431.27)	60193.35 (7254.14)	65382.32 (5724.23)	66270.67 (4846.23)	64842.59 (5688.15)	64934.41 (5939.78)	63636.93 (154.53)	64600.50 (5454.41)	65556.00 (5127.40)	57525.16 (9387.68)
Right	62498.44 (6404.02)	61213.30 (6584.98)	60910.64 (7329.76)	65894.26 (5570.00)	66549.30 (4843.58)	65476.95 (5784.85)	65220.46 (5943.95)	65717.89 (624.47)	64923.62 (5525.67)	66047.48 (5681.92)	56632.64 (8051.40)
Follow up period, year (SD)	-	-	-	3.12 (1.18)	3.52 (0.73)	2.41 (1.24)	2.84 (1.08)	2.99 (0.00)	1.61 (0.62)	2.68 (1.09)	3.45 (0.76)
Number of scan	1	1	1	4.81 (1.15)	5.21 (0.86)	4.76 (1.57)	5.38 (1.37)	5.00 (0.00)	3.48 (0.38)	4.31 (1.44)	4.00 (0.00)
Last scan volume, mm^3 (SD)†											
Total	-	-	-	129686.50 (11392.32)	130245.70 (9597.18)	129016.00 (11286.36)	127633.00 (11669.41)	130699.20 (3139.66)	128100.80 (10866.31)	131491.30 (10366.15)	118023.80 (7945.15)
Left	-	-	-	64607.51 (5804.00)	64839.90 (5017.74)	64181.13 (5647.45)	63673.82 (5896.68)	65943.72 (1230.14)	63894.04 (5467.48)	65287.25 (5149.04)	59466.88 (4355.83)
Right	-	-	-	65078.97 (5704.68)	65405.81 (4766.31)	64834.86 (5813.87)	63959.14 (5960.80)	64755.50 (1909.52)	64206.73 (5560.26)	66204.10 (5307.67)	58556.96 (3589.32)
Coef. Of age (CS)/atrophy rate (longi)‡, mm^3/year (SE)											
Total,	-417.90 (128.80)	-531.60 (70.33)	-463.50 (100.70)	511.98 (53.27)	615.84 (120.36)	498.71 (68.52)	833.04 (79.52)	-§	747.84 (80.79)	373.60 (201.40)	- §
Left	-192.90 (66.25)	-256.30 (35.39)	-238.30 (50.17)	249.79 (28.39)	347.77 (62.36)	249.66 (37.03)	412.00 (41.59)	-	362.69 (41.92)	153.72 (83.20)	
Right	-225.00 (64.35)	-275.30 (36.09)	-225.30 (51.90)	262.54 (26.66)	271.53 (66.57)	248.74 (33.28)	419.13 (39.83)	-	374.76 (40.90)	200.00 (95.16)	

CN; cognitively normal, MCI; mild cognitive impairment, AD; Alzheimer's disease, CNs, stable cognitively normal; cognitively normal converted to mild cognitive impairment; MCIs, stable mild cognitive impairment; MCIc, mild cognitive impairment converted to Alzheimer's disease; Ads, stable Alzheimer's disease; MCIr, mild cognitive impairment reverted to cognitively normal; ADr, Alzheimer's disease reverted to mild cognitive impairment; APOEe4, Apo lipoprotein alleles e 4 genotype; MMSE; mini-mental state examination

* The mean of cerebellar volume adjusted by the intra cranial volume for the cross-sectional part and baseline volume adjusted by the intra cranial volume for the longitudinal part

† Adjusted by the intra cranial volume

‡ Fixed effects of age for the cross-sectional data; extracted from the linear regression adjusted by intra cranial volume, gender, education and APOEe4, and atrophy rate for the longitudinal data; extracted from the linear mixed effects model adjusted by intra cranial volume, gender, education and APOEe4.

§ Insufficient data for calculation

Longitudinal results

The linear mixed model achieved a good fit and fixed factors in the model explained 43% (marginal R^2) while fixed and random factors together explained 99% (conditional R^2) of variance in cerebellar atrophy. A significant negative fixed effect of age was detected (χ^2 (1, 9)= 586.99, P< 0.0001); each year beyond age 55 was associated with a 0.47% lower cerebellar volume compared to baseline. Additionally, a significant random effect of age on cerebellar volume (χ^2 (2, 18)= 227.92, P< 0.0001) and interaction between age and diagnosis (χ^2 (7, 25)= 22.72, P< 0.01) were detected. The model revealed no differences in cerebellar volume across the diagnostic groups (χ^2 (7,18)= 11.31, p>0.1) i.e. the average of cerebellar volumes in CNs, CNc, MCIs, MCIc and ADs were not significantly different. However, a significant effect of cognitive diagnosis on cerebellar atrophy rates was detected (χ^2 (7, 25)= 22.71, P< 0.001). There was also a significant effect of sex on cerebellar volume (1, 18)= 14.12, P< 0.001) with less shrinkage in male.

An annual shrinkage of 0.36% (SE= 0.04) was detected in CNs individuals. A pairwise comparison revealed that it was not significantly different in MCIs (0.36%/year, SE=0.05) and CNc (0.42%/year, SE=0.08) however it was about 49% larger in ADs (0.53%/year, SE=0.06). Similarly, the atrophy rate was about 64% larger in MCIc (0.62%/year, SE=0.06) compared to CNs (Table 2, Table 3). The annual atrophy was also about 53% larger in ADs than MCIs (χ^2 (2, 13)= 8.67 p<0.01) and 68% larger in MCIc than MCIs (χ^2 (2, 13)= 12.57, P<0.001; Table 2). CN who converted to AD, MCI who reverted to CN and AD who reverted to MCI were excluded from pairwise comparison due to small samples sizes. Atrophy trajectories across groups are presented in Figure 2B.

Similar patterns of findings were observed for the left and right cerebellar volumes (Table 2), as well as left and right cerebellar gray matter and white matter volumes.

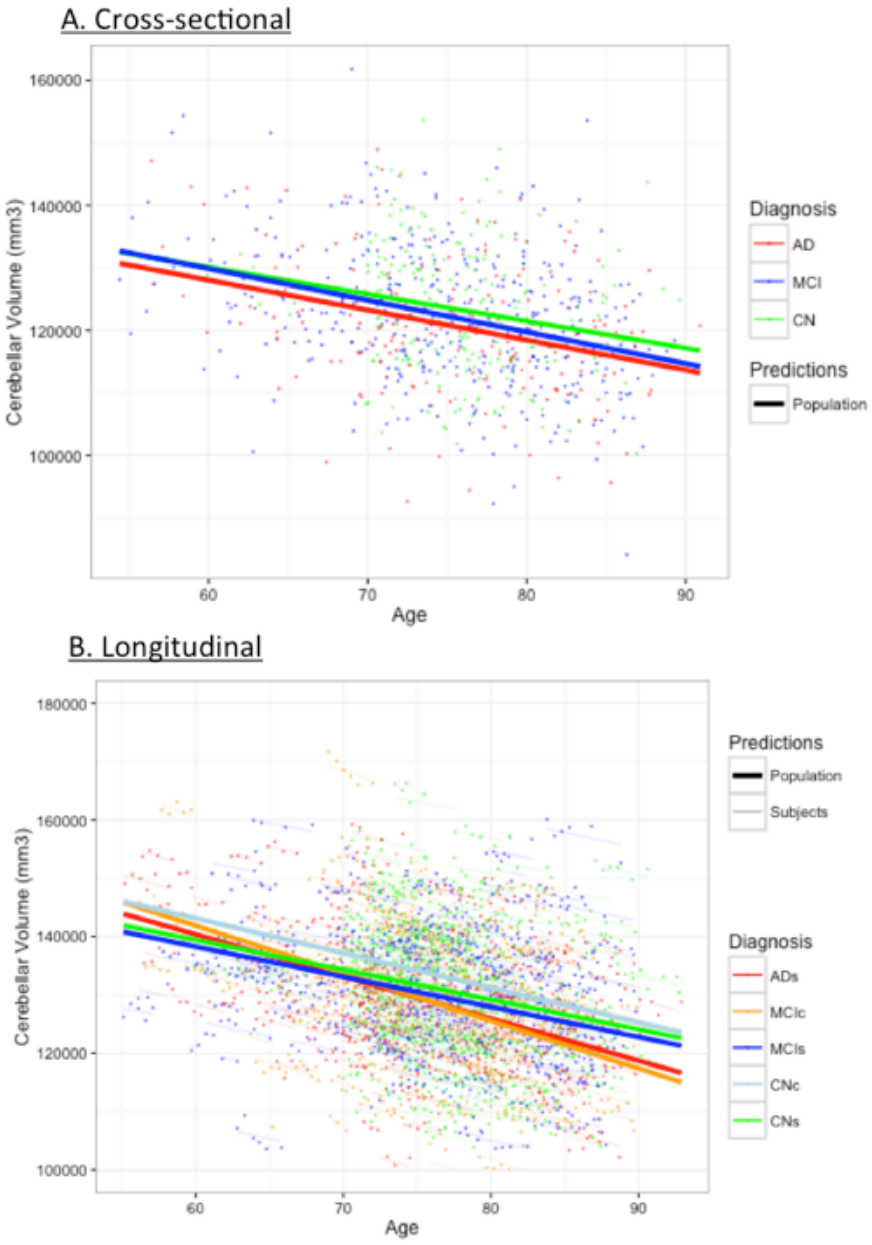


Figure 2; Linear Prediction of the cerebellar volumes for age at time points. (A) Prediction of the cerebellar volumes in three clinical groups including cognitively normal (CN), mild cognitive impairment (MCI) and Alzheimer's disease (AD) in cross-sectional level. (B) Prediction in subject and group (population) levels in five diagnostic groups including stable cognitively normal (CNs), cognitively normal converted to mild cognitive impairment (CNc), stable mild cognitive impairment (MCIs), and mild cognitive impairment converted to AD (MCIc) illustrating different slopes for the diagnostic groups.

Table 3: Pair-wise comparison of group diagnosis

	CNs vs CNc		CNs vs MCIs		CNs vs MCIc		CNs vs ADs	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Intercept	142212.26 (4460.56)	146127.2 (6078.7)	139769.58 (3247.02)	139091.58 (3445.00)	137510.50 (3521.38)	139282.56 (3755.78)	138821.74 (3172.06)	141244.50 (3523.62)
Volume slope in CNs, mm^3/ year (SE)	-	-4286.9 (4450.6)	-	1745.53 (2164.33)	-	-3236.69 (2257.59)	-	-1504.24 (2371.78)
Pr(>|t|)	-	0.3366	-	0.421	-	0.1525	-	0.5263
Shrinkage slope in CNs, mm^3/ year (SE)	-	100.9 (163.6)	-	-19.72 (84.54)	-	312.57 (91.80)	-	214.37 (94.66)
Pr(>|t|)	-	0.5382	-	0.816	-	**0.0007 *** **	-	**0.0240 ***
Loglik	-9793	-9793	-17438	-17438	-16960	-16951	-14515	-14509
Chisq	1.0195		1.3075		17.126		10.884	
Chi Df	2		2		2		2	
Pr(> Chisq)	0.6006		0.5201		**0.0001911 *** **		**0.004332 ** **	

Significant codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '-' 1;

Key: CN; cognitively normal, CNs; stable cognitively normal, CNc; cognitively normal converted to mild cognitive impairment; MCIs, stable mild cognitive impairment; MCIc, mild cognitive impairment converted to Alzheimer's disease; ADs, stable Alzheimer's disease

Discussion

This study aimed to investigate cerebellar shrinkage in normal ageing and prodromal (MCI) and clinical phases of AD. It revealed that cerebellar shrinkage occurs mostly in the late stages of the disease. The main findings were that (1) in cross-sectional analyses cerebellar volume was larger in CN compared to AD but not compared to MCI, (2) in longitudinal analyses cerebellar atrophy was higher in ADs and MCIc compared to CNs but not in CNc and MCIs, and (3) APOEe4 was not a significant predictor of baseline cerebellar volume nor of cerebellar atrophy across clinical groups.

Cross-sectional

The smaller cerebellar volume observed in AD compared to CN and no difference between MCI and CN are in agreement with available cross-sectional studies reporting

smaller cerebellar volume in AD (Kusbeci et al., 2009; Moller et al., 2013) but normal volume in MCI (Thomann et al., 2008; Yoon et al., 2013). This discrepancy is consistent with the documented progression of AD pathology. However, the cerebellum can be parsed functionally and morphologically into different subdivisions and it is likely that AD pathology targets each subdivision differently. Previous voxel-based morphometric studies showed bilateral lower gray matter density in lobule VI (Colloby, O'Brien, & Taylor, 2014) and Crus I/II (Guo et al., 2016) in AD compared with CN, suggesting that network-selective vulnerability underlies the cerebellar neurodegeneration(Guo et al., 2016). Regardless of selective or non-selective volume loss in the cerebellum and its subregions, cross-sectional approach needs to be affirmed by tracking atrophy in a longitudinal approach.

Longitudinal

The negative association between age and cerebellar volume is consistent with that demonstrated in the cross-sectional analysis (0.41%/ year in cross-sectional and 0.47% in longitudinal). Pairwise analyses demonstrated significantly larger cerebellar atrophy rates in ADs and MCIc but not in CNc and MCIs compared to CNs. This pattern of results is suggestive of an increasing rate of cerebellar atrophy with progression of AD pathology. It is also consistent with the chronological development of AD pathology with progressive spreading of tau fibrillary tangles (Braak stages), amyloid deposition, and subsequently gradual decline in cognitive function (Murray et al., 2015). As Thal et al. demonstrated, clinically diagnosed AD occurs in the amyloid phase 3-5 while the cerebellar involvement mostly occurs in the 5th phase (Thal et al., 2002). Thus, the available evidence suggests that the cerebellum is relatively spared of neurodegeneration in the preclinical stages of the disease and gradually becomes affected as the clinical presentation fully develops. However, it remains unclear whether association of the cerebellum with AD clinical progression is due

to spreading of fibrillary tangle and/or amyloid deposition, or secondary to cerebral neurodegeneration.

Although the findings suggest shrinkage in the cerebellum with ageing and larger cerebellar atrophy in ADs compared with CNs and MCIs, it is worthy to consider that cerebellar atrophy in the diagnostic groups were less than that reported for whole brain atrophy (CNs: 0.36%/year vs 0.57%/year; MCIs: 0.36%/year vs 1.02%/year; ADs: 0.53%/year vs 1.90%/year) (Henneman et al., 2009; Tabatabaei-Jafari et al., 2015). This is in contrast to brain regions characteristics for AD pathology, including hippocampus and entorhinal cortex, for which atrophy rates are roughly 200% higher for MCI and 300% higher for AD compared to normal ageing (Desikan et al., 2008; Tabatabaei-Jafari et al., 2015), further emphasising the relative resistance of the cerebellum to AD related degeneration. However, despite the small effect size and partial resistance, the cerebellum is not intact in AD pathology and future investigation is needed to elucidate the impact of cerebellar atrophy on uptake measurement when using the cerebellum to standardise FDG uptake in PET studies.

Covariates and correlates

Age is a common predictor for CN and AD-related brain atrophy and all cognitive groups in the current study were matched for age. However, they were differences in sex distribution, education and APOEe4 alleles-- the most well known risk factors of AD pathology-- as were expected. An effect of sex on cerebellar volume was detected such that males showed less cerebellar atrophy than females. However, no significant association between education or APOEe4 alleles and cerebellar volume were detected. APOEe4 is a known moderator of hippocampal atrophy in AD pathology (Tabatabaei-Jafari et al., 2015), therefore it might have been expected that carrying the APOEe4 allele would be associated

with increased cerebellar atrophy. However, this was not the case in our findings. It may indicate that while neurodegeneration in the cerebrum is directly related to the development of neurofibrillary tangles and β-amyloid deposition which occurs at higher rates in APOEe4 carriers, cerebellar atrophy is the product of secondary processes associated with cerebral neuronal loss, Wallerian degeneration, and widespread disconnection. To clarify this question future investigations need to further elucidate the impact of risk factors in different AD clinical stages.

Strengths and limitation:

This study is unique in using in vivo evaluation of the cerebellum with a reasonable follow up period in a relatively large sample while computing both cross-sectional and longitudinal estimates and using advanced and well-controlled mixed-effects models. Most AD related cerebellar studies conducted to date have been post-mortem or if in vivo, cross-sectional in design, thus raising questions as to the precision and generalizability of their estimates. Consequently, the present study fills an important gap. However, it should be noted that this investigation was restricted to the gray and white matter volumes in the left and right cerebellum and therefore do not provide information on the cerebellar subregions.

Conclusion

The cerebellum is often thought to be spared from neurodegenerative processes but the present findings indicate that this is not the case. The present findings demonstrate that although the cerebellum is not significantly affected in the preclinical phase of AD (i.e. MCI), it is affected in the clinical phase. However, acceleration in atrophy rate is less than the average of the atrophy in the cerebrum and it is not associated with AD moderators (education and APOEe4 status). These findings in addition to previous evidence of network-selective vulnerability of the cerebellum suggest that AD related cerebellar atrophy might be

secondary to the development of AD pathology in the cerebrum rather than the cerebellum itself. Therefore, modifying interventions targeting the non-specific network progression is a potential therapeutic option additional to interventions targeting the specific pathological process.

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: 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. Additionally, Nicolas Cherbuin is funded by ARC fellowship grant number 120101705.

References:

Andersen, B. B., Korbo, L., & Pakkenberg, B. (1992). A Quantitative Study of the Human Cerebellum With Unbiased Stereological Techniques. *J Comp Neurol, 326*, 549-560.

Andersen, K., Andersen, B. B., & Pakkenberg, B. (2012). Stereological quantification of the cerebellum in patients with Alzheimer's disease. *Neurobiol Aging, 33*(1), 197.e111-120. doi:10.1016/j.neurobiolaging.2010.06.013

Bernal-Rusiel, J. L., Greve, D. N., Reuter, M., Fischl, B., Sabuncu, M. R., & Initiative, t. A. s. D. N. (2013). Statistical Analysis of Longitudinal Neuroimage Data with Linear Mixed Effects Models. *Neuroimage, 1*, 249-260. doi:10.1016/j.neuroimage.2012.10.065

Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol, 82*, 239-259.

Clark, C. M., Schneider, J. A., Bedell, B. J., Beach, T. G., Bilker, W. B., Mintun, M. A., . . . Skovronsky, D. M. (2011). Use of florbetapir-PET for imaging beta-amyloid pathology. *Jama, 305*(3), 275-283. doi:10.1001/jama.2010.2008

Colloby, S. J., O'Brien, J. T., & Taylor, J. P. (2014). Patterns of cerebellar volume loss in dementia with Lewy bodies and Alzheimers disease: A VBM-DARTEL study. *Psychiatry Res, 223*(3), 187-191. doi:10.1016/j.pscychresns.2014.06.006

Desikan, R. S., Fischl, B., Cabral, H. J., Kemper, T. L., Guttmann, C. R., Blacker, D., . . . Killiany, R. J. (2008). MRI measures of temporoparietal regions show differential rates of atrophy during prodromal AD. *Neurology, 71*(11), 819-825. doi:10.1212/01.wnl.0000320055.57329.34

Driscoll, I., Troncoso, J. C., Rudow, G., Sojkova, J., Pletnikova, O., Zhou, Y., . . . Resnick, S. M. (2012). Correspondence between in vivo (11)C-PiB-PET amyloid imaging and postmortem, region-matched assessment of plaques. *Acta Neuropathol, 124*(6), 823-831. doi:10.1007/s00401-012-1025-1

Guo, C. C., Tan, R., Hodges, J. R., Hu, X., Sami, S., & Hornberger, M. (2016). Network-selective vulnerability of the human cerebellum to Alzheimer's disease and frontotemporal dementia. *Brain, 139*(Pt 5), 1527-1538. doi:10.1093/brain/aww003

Henneman, W. J., Sluimer, J. D., Barnes, J., van der Flier, W. M., Sluimer, I. C., Fox, N. C., . . . Barkhof, F. (2009). Hippocampal atrophy rates in Alzheimer disease: added value over whole brain volume measures. *Neurology, 72*(11), 999-1007. doi:10.1212/01.wnl.0000344568.09360.31

Hoogendam, Y. Y., Geest, J. N. v. d., Lijn, F. v. d., Lugt, A. v. d., Niessen, W. J., Krestin, G. P., . . . Ikram, M. A. (2012). Determinants of cerebellar and cerebral volume in the general elderly population. *Neurobiology of Aging, 33*, 2774-2781.

Jack, C. R., Jr., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., . . . Weiner, M. W. (2008). The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging, 27*(4), 685-691. doi:10.1002/jmri.21049

Jack, C. R., Jr., Lowe, V. J., Senjem, M. L., Weigand, S. D., Kemp, B. J., Shiung, M. M., . . . Petersen, R. C. (2008). 11C PiB and structural MRI provide complementary information in imaging of Alzheimer's disease and amnestic mild cognitive impairment. *Brain, 131*(Pt 3), 665-680. doi:10.1093/brain/awm336

Jonasson, M., Wall, A., Chiotis, K., Saint-Aubert, L., Wilking, H., Sprycha, M., . . . Lubberink, M. (2016). Tracer kinetic analysis of (S)-18F-THK5117 as a PET tracer for assessing tau pathology. *J Nucl Med*. doi:10.2967/jnumed.115.158519

Kusbeci, O. Y., Bas, O., Gocmen-Mas, N., Karabekir, H. S., Yucel, A., Ertekin, T., & Yazici, A. C. (2009). Evaluation of cerebellar asymmetry in Alzheimer's disease: a stereological study. *Dement Geriatr Cogn Disord, 28*(1), 1-5. doi:10.1159/000228544

Lemoine, L., Saint-Aubert, L., Marutle, A., Antoni, G., Eriksson, J. P., Ghetti, B., . . . Nordberg, A. (2015). Visualization of regional tau deposits using (3)H-THK5117 in Alzheimer brain tissue. *Acta Neuropathol Commun, 3*, 40. doi:10.1186/s40478-015-0220-4

Lopresti, B. J., Klunk, W. E., Mathis, C. A., Hoge, J. A., Ziolko, S. K., Lu, X., . . . Price, J. C. (2005). Simplified quantification of Pittsburgh Compound B amyloid imaging PET studies: a comparative analysis. *J Nucl Med, 46*(12), 1959-1972.

Moller, C., Vrenken, H., Jiskoot, L., Versteeg, A., Barkhof, F., Scheltens, P., & van der Flier, W. M. (2013). Different patterns of gray matter atrophy in early- and late-onset Alzheimer's disease. *Neurobiol Aging, 34*(8), 2014-2022. doi:10.1016/j.neurobiolaging.2013.02.013

Murray, M. E., Lowe, V. J., Graff-Radford, N. R., Liesinger, A. M., Cannon, A., Przybelski, S. A., . . . Dickson, D. W. (2015). Clinicopathologic and 11C-Pittsburgh compound B implications of Thal amyloid phase across the Alzheimer's disease spectrum. *Brain, 138*(Pt 5), 1370-1381. doi:10.1093/brain/awv050

Reutera, M., Schmansky, N. J., Rosasa, D., & Fischl, B. (2012). Within-Subject Template Estimation for Unbiased Longitudinal Image Analysis. *Neuroimage, 16;61*(4), 1402-1418.

Rowe, C. C., Ng, S., Ackermann, U., Gong, S. J., Pike, K., Savage, G., . . . Villemagne, V. L. (2007). Imaging beta-amyloid burden in aging and dementia. *Neurology, 68*(20), 1718-1725. doi:10.1212/01.wnl.0000261919.22630.ea

Schmahmann, J. D., & Sherman, J. C. (1998). The cerebellar cognitive affective syndrome. *Brain, 121*, 561-579.

Sluimer, J. D., van der Flier, W. M., Karas, G. B., van Schijndel, R., Barnes, J., Boyes, R. G., . . . Barkhof, F. (2009). Accelerating regional atrophy rates in the progression from normal aging to Alzheimer's disease. *Eur Radiol, 19*(12), 2826-2833. doi:10.1007/s00330-009-1512-5

Standring, S. (2008). Gray's Anatomy The Anatomical Basis of Clinical Practice, 40th Ed *40th*.

Stoodley, C. J. (2012). The cerebellum and cognition: evidence from functional imaging studies. *Cerebellum, 11*(2), 352-365. doi:10.1007/s12311-011-0260-7

Tabatabaei-Jafari, H., Shaw, M. E., & Cherbuin, N. (2015). Cerebral atrophy in mild cognitive impairment: A systematic review with meta-analysis. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 1*(4), 487-504. doi:10.1016/j.dadm.2015.11.002

Thal, D. R., Rub, U., Orantes, M., & Braak, H. (2002). Phases of A beta-deposition in the human brain and its relevance for the development of AD. *Neurology, 58*(12), 1791-1800.

Thomann, P. A., Schlafer, C., Seidl, U., Santos, V. D., Essig, M., & Schroder, J. (2008). The cerebellum in mild cognitive impairment and Alzheimer's disease - a structural MRI study. *J Psychiatr Res, 42*(14), 1198-1202. doi:10.1016/j.jpsychires.2007.12.002

Wang, H.-Y., D'Andrea, M. R., & Nagele, R. G. (2002). Cerebellar diffuse amyloid plaques are derived from dendritic A-beta 42 accumulations in Purkinje cells. *Neurobiol Aging, 23*, 213-223.

Wegiel, J., Wisniewski, H. M., Dziewiatkowski, J., Badmajew, E., Tarnawski, M., Reisberg, B., . . . Miller, D. C. (1999). Cerebellar atrophy in Alzheimer's disease-clinicopathological correlations. *Brain Res, 818*(1), 41-50.

Weier, K., Penner, I. K., Magon, S., Amann, M., Naegelin, Y., Andelova, M., . . . Sprenger, T. (2014). Cerebellar Abnormalities Contribute to Disability Including Cognitive Impairment in Multiple Sclerosis. *PLoS One, 9*(1), e86916. doi:10.1371/journal.pone.0086916

Wolf, U., Rapoport, M. J., & Schweizer, T. A. (2009). Evaluating the Affective Component of the Cerebellar Cognitive Affective Syndrome. *J Neuropsychiatry Clin Neurosci., 21*(3), 245-253.

Yoon, C. W., Seo, S. W., Park, J. S., Kwak, K. C., Yoon, U., Suh, M. K., . . . Na, D. L. (2013). Cerebellar atrophy in patients with subcortical-type vascular cognitive impairment. *Cerebellum, 12*(1), 35-42. doi:10.1007/s12311-012-0388-0

STUDY 3

A SIMPLE AND CLINICALLY RELEVANT COMBINATION OF NEUROIMAGING AND FUNCTIONAL INDEXES FOR THE IDENTIFICATION OF THOSE AT HIGHEST RISK OF ALZHEIMER'S DISEASE

Hossein Tabatabaei-Jafari MD[1], Erin Walsh PhD[1], Marnie E. Shaw PhD[1], Nicolas Cherbuin PhD[1], For the Alzheimer's Disease Neuroimaging Initiative (ADNI)

[1]Centre for Research on Ageing, Health and Wellbeing, The Australian National University, Canberra, Australia

Published in Neurobiology of Aging, 69 (2018) 102-110

DOI: 10.1016/j.neurobiolaging.2018.05.006

Abstract

The current challenge in clinical practice is to identify those with mild cognitive impairment (MCI) who are at greater risk of Alzheimer's disease (AD) conversion in the near future. The aim of this study was to assess a clinically practical new hippocampal index – hippocampal volume normalized by cerebellar volume (Hippocampus to Cerebellum volume Ratio; HCCR) used alone or in combination with scores on the mini mental state examination (MMSE), as a predictor of conversion from MCI to AD. The predictive value of the HCCR was also contrasted to that of the hippocampal volume to intracranial volume (ICV) ratio. The findings revealed that the performance of the combination of measures was significantly better than that of each measure used individually. The combination of MMSE and hippocampal volume, normalized by the cerebellum or by intracranial volume, accurately discriminated individuals with MCI who progress to AD within five years from other MCI types (stable, reverters) and those with intact cognition (area under receiver operating curve 0.88 and 0.89 respectively). Normalization by cerebellar volume was as accurate as normalization by ICV with the advantage of being more practical, particularly for serial assessments.

Introduction

Mild cognitive impairment (MCI) refers to modest cognitive decline along with preserved daily activities (Association, 2013). While many people with MCI live largely normal lives, they are at higher risk of developing Alzheimer's disease (AD) than those without MCI (Forlenza, Diniz, Stella, Teixeira, & Gattaz, 2013). The available evidence suggests that less than half of patients diagnosed with MCI may progress to AD in a five-year period while the rest remain stable or reverse to cognitively normal (CN) status (Falahati, Westman, & Simmons, 2014; Pandya, Clem, Silva, & Woon, 2016). Generally, there is an expectation of eventual conversion from MCI to AD due to the progressive nature of the neurodegenerative processes involved, and MCI stability can depend on the duration of follow-up (Ganguli, 2013). Reversion to CN is still an unresolved question but may relate to the relatively unspecific nature of diagnostic criteria, interaction with co-morbid conditions and/or variability in the pathological process (Park, Han, & Initiative, 2015). Thus, the current clinical challenge is to discriminate individuals with MCI who are more likely to convert to AD.

In their revised position, the National Institute on Aging and the Alzheimer's Association (NIA-AA) consider MCI and AD as different stages of the AD continuum rather than two distinct clinical entities (Albert et al., 2011; Jack et al., 2018). In 2011, NIA-AA reviewed diagnostic guidelines and suggested that, owing to greater diagnostic uncertainty earlier in the AD continuum, MCI diagnosis should be supported by biological markers reflecting AD pathology (Albert et al., 2011). In 2018, the NIA-AA work group further qualified this position and recommended that biological markers should reflect neuropathological processes that define the disease instead of simply supporting the diagnosis (Jack et al., 2018). Based on this expert consensus, the work group recommended

that AD biomarkers should be incorporated in MCI/AD diagnostic criteria. The NIA-AA work group identified three types of AD biomarkers directly related to the underlying pathological processes. The biomarkers include: (1) amyloid β deposition including cortical amyloid PET ligand bonding (F^{18}-flutemetamol PET) and low CSF $A\beta_{42}$, (2) aggregated tau including cortical tau PET ligand bonding (flortaucipir-PET) and elevated CSF phosphorylated tau (P-tau), and (3) neurodegeneration or neural injury including PET detected hypometabolism (Fluorodeoxyglucose-PET), CSF total tau (T-tau), and cortical/volume atrophy on MRI scan (Jack et al., 2018).

Much research has been conducted to evaluate amyloid β deposition, tau aggregation and hypometabolism using PET scans and CSF biomarkers --separately or in combination-- to classify MCI at risk of AD conversion, with some promising performance (Mitchell, 2009; Ritchie et al., 2017; Vandenberghe et al., 2013; Yuan, Gu, & Wei, 2009). However, these methods are invasive and, especially for PET imaging, have limited availability in clinical practice. Ideally, a practical biomarker should be widely available, accurate, cost effective, relatively simple to interpret, easy to use, and be acceptable to patients while not imposing an excessive burden. It is important that -- prior to assessing a new biomarker -- clear criteria for selection be established, and the likelihood of meeting them be considered. As a minimum, the proposed new biomarker should perform at least as well as simple, non-invasive and currently available biomarkers.

A type of non-invasive and more widely available biomarker is provided by structural brain measurement obtained using MRI. Cerebral cortical thickness and hippocampal measures are the most predictive and practical MRI methods to date (Falahati et al., 2014; Rathore, Habes, Iftikhar, Shacklett, & Davatzikos, 2017). Although cerebral cortical thickness has been shown to be more predictive compared to volumetric measures based on

single brain regions, it requires agreement on a standard pattern of cerebral cortical thickness in AD to be adoptable in clinical practice. Hippocampal volume, which has been shown to be a moderate predictor of AD conversion with a sensitivity of 67% and specificity of 72%, has the advantage of being less invasive compared to a CSF biomarker, less costly than a PET scan, and more widely available and clinically easier to use compared to cortical atrophy measures (Chupin et al., 2009). However, using hippocampal volume in the clinical setting is less straightforward compared to the use of this measure in a research setting.

Hippocampal volume needs to be normalized by or adjusted for intracranial volume (ICV) (Whitwell, Crum, Watt, & Fox, 2001) to control for inter-subject (Barnes et al., 2010) and sex (Pintzka, Hansen, Evensmoen, & Haberg, 2015) variations in head size, as well as variation in head size estimations in serial scans (Whitwell et al., 2001). The most widely used method in neuroimaging research is adjustment for ICV using its inclusion as a covariate in regression analyses. A less commonly used normalization approach is to divide the hippocampal volume by another volume that can be accurately measured and which is not significantly impacted by neurodegenerative processes, typically ICV. In this study, we investigate normalization by cerebellar volume (hippocampus to cerebellar volume ratio) as an alternative approach, to correct for head size/pre-morbid brain volume as the cerebellum has been shown to be little affected by age-related atrophy in the absence of clinical dementia. Neurodegeneration in AD gradually progresses from the medial temporal lobe to the parietal and frontal lobes and then to the posterior parts of the brain. The cerebellum is among the last brain regions affected by AD pathology (Thal, Rub, Orantes, & Braak, 2002). We have recently shown that cerebellar atrophy is not different in MCI compared to normal ageing (Tabatabaei-Jafari, Walsh, Shaw, Cherbuin, & Alzheimer's Disease Neuroimaging, 2017). Furthermore, while cerebellar atrophy increases in AD it remains lower than in other regions and particularly in the medial temporal lobe (Tabatabaei-Jafari et al., 2017). Thus,

using the cerebellum as a reference area should be both methodologically robust and practical in a clinical context. Importantly, regional brain volume is more accurately measured than ICV using semi-automated methods such as FreeSurfer (Heinen et al., 2016), and unlike ICV also less affected by field strength (Heinen et al., 2016; Nordenskjold et al., 2013) and segmentation method (Hansen, Brezova, Eikenes, Haberg, & Vangberg, 2015; Keihaninejad et al., 2010; Malone et al., 2015).

Although hippocampal volume is not sufficiently accurate to be clinically useful as a single predictor of MCI who progress to AD, it is a useful benchmark. If other measures sufficiently improve on the predictive value of hippocampal volume, they may be worth further consideration. The mini mental state examination (MMSE) may be a good candidate. A recent Cochrane review indicated that the weighted sensitivity and specificity of the MMSE for conversion from MCI to AD are 54% and 80% in a limited number of available studies (Arevalo-Rodriguez et al., 2015). Moreover, evidence suggests that a combination of cognitive measures with hippocampal volume can improve the predictive value of hippocampal volume for predicting AD conversion in MCI (Devanand et al., 2008). Therefore, such a combination is also likely to improve on the classification performance of hippocampal volume for identifying MCI who convert to AD in short term from all those who do not convert.

In the present study, we investigated the classification performance of MMSE and hippocampal volume normalized by cerebellar volume or ICV both individually and in combination, to identify individuals with MCI who will convert to AD within five years. We expected that these combinations of measures would have classification accuracies high enough to be useful in clinical practice.

Methodology

Study Participants

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 MCI and early AD.

Total numbers of 1289 participants with MCI (n=872) or CN (n=417) at baseline were considered for inclusion. All MCI participants who were stable for at least 6 months after baseline and converted to AD or reverted to CN within five years (confirmed with two consecutive stable diagnosis) or were stable for at least five years were included. Participants who were CN at baseline and were stable throughout the study were also included.

Based on diagnosis and diagnostic change, participants were categorized into four groups; (1) MCIc (N=187); MCI who converted to AD in less than five years, (2) MCIs (N=112); MCI who were stable for five years or more, (3) MCIr (N=39); MCI who reverted to CN in less than five years, and (4) CN (N=322); who remained cognitively healthy for the whole follow-up period.

Details of the diagnostic criteria can be found at the ADNI web site (http://www.adni-info.org/Scientists/AboutADNI.aspx). Briefly, participants were classified as CN if they had an MMSE greater than 24, a Clinical Dementia Rating (CDR) of 0 and did not meet diagnostic criteria for MCI, dementia or depression. Participants were classified as MCI if

they had an MMSE greater than 24, a CDR of 0.5, a subjective report of memory concern, an objective memory loss, preserved daily living activity and did not meet diagnostic criteria for dementia. AD participants have MMSE scores less than 26, CDR 0.5 or 1.0 and fulfill criteria for clinically probable AD according to the Institute of Neurological and Communicative Diseases and Stroke/ Alzheimer's Disease and Related Disorders Association.

Neuroimaging acquisition and processing

Participants underwent high-resolution MRI brain scans on 1.5 (N=335) or 3 T (N=325) scanners from General Electric, Siemens, or Philips (Milwaukee, WI, USA; Germany; the Netherlands respectively) using a standardized ADNI acquisition protocol for 3D MP-RAGE sequence(Jack et al., 2008). Baseline images which had undergone specific ADNI preprocessing correction steps to standardize images from different sites and platforms, were obtained for this study: (1) Grad wrap; a specific correction of image geometry distortion due to non-linearity, (2) B1 non-uniformity; B1 calibration to correct the image intensity non-uniformity that results when RF transmission is performed with a more uniform body coil while reception is performed with a less uniform head coil, (3) N3 correction; a histogram peak sharpening algorithm applied after grad wrap and B1 correction. We conducted automatic volumetric segmentation using FreeSurfer (version 5.3, http://surfer.nmr.mgh.harvard.edu/) and the output images were visually checked for the hippocampal and cerebellar segmentations. The criterion was a clear segmentation error as assessed by an experienced neuroscientist. Scans with segmentation errors were re-processed and would only be excluded if the error could not be corrected. In this sample no image was excluded.

Measurements

ICV was computed by the sum of the whole brain gray and white matter and CSF volume. Total cerebellar volume was computed by summing the left and right cerebellar gray and white matter. Total hippocampal volume was the sum of the volumes of the left and right hippocampus. Hippocampus to intracranial volume ratio (HCICV) was the ratio of total hippocampal volume to intracranial volume adjusted for age and field strength. Hippocampus to cerebellar volume ratio (HCCR) was the ratio of total hippocampal volume to total cerebellar volume adjusted for age and field strength. No significant correlation was detected between HCICV (correlation= -0.09) or HCCR (correlation= -0.09) and ICV. There was a moderate correlation between hippocampal volume and MMSE (r=0.35, Figure 1). The residual method was used for all adjustments implemented by running a regression line between raw ratios and the variables using the whole data (Pintzka et al., 2015).

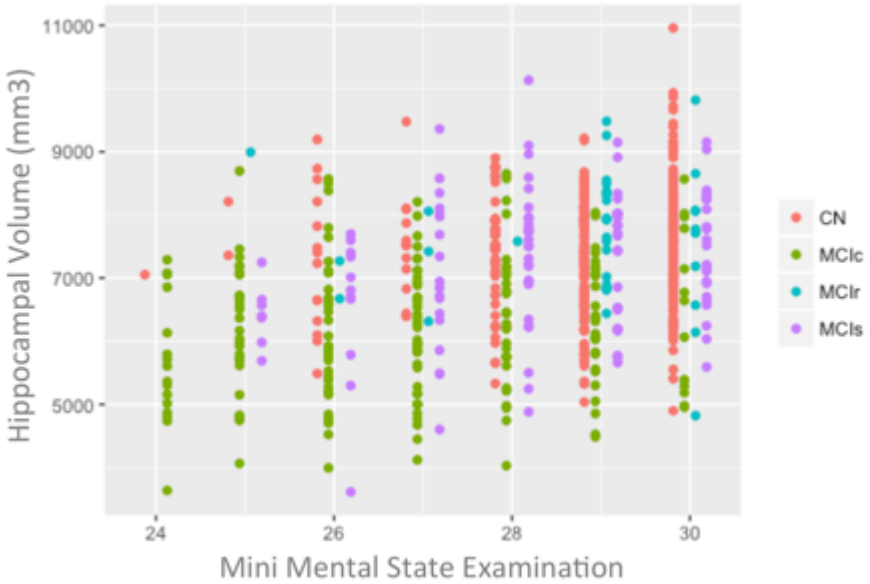


Figure 1: Total hippocampal volume (mm^3) at different mini mental state examination scores in four diagnostic groups

Statistical analysis

Statistical analyses were performed using the R statistical software (version 3.3.2). Data were checked for missing values and univariate and multivariate outliers using Mahalanobis distance. Discriminant analysis was used to estimate the predictive value of HCICV, HCCR, MMSE and their combination for clinical status. The DiscriMiner package (version; 0.01-29, https://CRAN.R-project.org/package=DiscriMiner) was used for descriptive discrimination and the MASS (version; 7.3-45, http://www.stats.ox.ac.uk/pub/MASS4) and Caret package (version; 6.3-73, https://CRAN.R-project.org/package=caret) for predictive discrimination (classification). Data were evaluated for normality of all measures (Q-Q plot), linearity, and multicollinearity and singularity (variation inflation factor) assumptions of discriminant analysis, which were all satisfied. Statistically significant heterogeneity of variance-covariance matrices was observed (Box's M-test; $\chi^2 > 51.19$, $p<0.001$) and therefore a quadratic classification procedure was used, because linear discriminant analysis is known to perform poorly in the presence of heterogeneous covariance matrices(Tabachnick & Fidell, 2013).

For binary classification analyses using quadratic classification procedure, MCIc was contrasted with (1) CN, MCIs and MCIr pooled together, (2) CN alone, and (3) MCIs and MCIr pooled together and CN was contrasted with MCIs and MCIr pooled together. The stability of the classification procedure was checked by a 10-fold cross-validation. The sample randomly partitioned into 10 equal size subsamples. Nine subsamples (combined) were used as training data and the remaining single subsample was retained as the validation data to evaluate predictive model. The process was repeated 10 times, with each of the 10 subsamples was used only once as the validation data. The average of the results was provided with confidence interval. We measured reliability using the Kappa coefficient, a

chance-corrected measure of agreement between the reference classification (categorized by long-term clinical follow-up) and predictive classification (classifications based on study measures) (Fritz & Wainner, 2001). The receiver operating characteristic (ROC) curve (package pROC version 1.9.1, http://www.biomedcentral.com/1471-2105/12/77/) and the area under the curve (AUC) was used to estimate the discriminant capacity of each model and DeLong's test was used to compare different models (Tabachnick & Fidell, 2013).

Results

Demography and brain measures

The average age of all participants together was aged 73.76 (SD=6.80). Participants within the four diagnostic groups were similar in age, except for MCIr who were three to five years younger. APOEe4 genotype was significantly higher and MMSE scores lower in the MCI sub-groups compared to CN. The average time for MCIc to convert to AD and MCIr to revert to CN was similar at about two years. Baseline imaging measures showed that there was a trend of ascending hippocampal volume (adjusted for age, field strength and ICV), HCICV and HCCR values in MCIc, MCIs, MCIr and CN. No such trend was detected for cerebellar volume (Table 1).

Discriminant analyses; Descriptive statistics

Discriminant analyses were conducted to evaluate discriminative performance of the HCICV-MMSE and HCCR-MMSE models. Two discriminant functions were calculated for each model separately. The first function significantly distinguished among the diagnostic groups (HCICV-MMSE: F [6, 1310]=74.556, HCCR-MMSE: F [6, 1310]= 70.096) and accounted for 99.6% of prediction of MCIc from CN, MCIs and MCIr (first function's eigenvalue/ sum of all eigenvalues *100) in both models, whereas the second function was

not effective in distinguishing between CN, MCIs and MCIr. Predictive values of the combination of HCICV and MMSE or HCCR and MMSE were almost equal (equal standardised coefficient correlation of predictors and discriminant functions) in the first discriminant functions for distinguishing among the groups (Table 2).

The binary classification analyses revealed that HCICV, HCCR and MMSE were equally predictive of MCIc with loadings of more than 0.5 on the discriminant functions (standardised coefficient correlation) with large effect sizes (canonical R^2 and eigenvalue) in all contrasts. In comparison, the standardized coefficients in CN contrasted with MCIs and MCIr groups were more than 0.5, but because the effect sizes were very low the discriminant functions were not effective in separating the groups (Table 2).

Discriminant Analysis; Classification

Individual predictor classification

HCICV, HCCR and MMSE performed similarly in identifying diagnostic groups when tested individually, and classified participants of the four diagnostic groups into two groups, CN and MCIc. A high proportion of CN and MCIc were correctly classified, whereas the majority of MCIs and MCIr were classified as CN and the remainder as MCIc (Table 3).

Table 1: Characteristics: Demographic information, MMSE and brain measures. Trends of decrease in the average of MMSE and hippocampal measures are noticeable across the groups.

	CN	MCIr	MCIs	MCIc	Test of significance (p < 0.05)	
Sample size	322	39	112	187	Across groups	Significant pairs
Age; year, Mean (SD)	74.55 (5.80)	69.33 (8.32)	72.08 (7.65)	74.31 (7.02)	F (3)=10.09 *	CN vs. MCIr CN vs. MCIs MCIc vs. MCIr
Age range, year	59 - 90	55 - 87	57 - 88	55 - 89		
Male sex; N (%)	158 (49)	17 (44)	72 (64)	113 (60)	χ^2 (3) 12.68	All pairs are different
Education; year, Mean (SD)	16.38 (2.74)	16.87 (2.38)	15.75 (3.03)	16.09 (2.73)	F (3) 2.285	No difference in pairs
APOE e4; N (%)	82 (25)	19 (49)	40 (36)	127 (68)	χ^2 (3) 90.63*	All pairs are different
One allele	75 (23)	18 (46)	32 (29)	96 (51)		
Two alleles	7 (2)	1 (3)	8 (7)	31 (17)		
Age at DX change; year, Mean (SD)	-	71.38 (8.31)	-	76.74 (7.15)	-	MCIc vs MCIr
Time to DX change; year, Mean (SD)	-	2.06 (1.14)	-	2.43 (0.91)	-	-
Measures						
MMSE; Mean (SD)	29.08 (1.14)	28.85 (1.23)	28.11 (1.48)	26.95 (1.72)	F (3) 95.22*	MCIc vs. CN MCIs vs CN MCIr vs. MCIc MCIs vs MCIc
Hippocampus, mm^3, mean (SD) †	7510.06 (807.29)	7210.85 (756.46)	7052.82 (909.03)	6240.78 (888.32)	F (3) 89.32*	MCIc vs. CN, MCIc vs. MCIr MCIc vs. MCIs MCIs vs. CN
Cerebellum, mm^3, mean (SD) †	121937.60 (9539.73)	120522.40 (9840.47)	121318.00 (10337.83)	122673.50 (10510.29)	F (3) 0.458	No difference in pairs
HCICV, mean (SD)	0.50 (0.06)	0.47 (0.05)	0.46 (0.07)	0.41 (0.06)	F (3) 87.86*	MCIc vs. CN, MCIc vs. MCIr, MCIc vs.MCIs MCIs vs. CN
HCCR, mean (SD)	6.21(0.73)	5.99 (0.68)	5.85 (0.94)	5.09 (0.79)	F (3) 79.83*	MCIc vs. CN, MCIc vs. MCIr, MCIc vs. MCIs MCIs vs. CN

CN; cognitively normal, MCIr; mild cognitive impairment reverted to normal, MCIc; mild cognitive impairment converted to Alzheimer's disease in five years, MCIs; mild cognitive impairment stable for five years or more, APOE e4; Apolipoprotein E allele 4, MMSE; mini mental status examination, DX; diagnosis, HCICV; Hippocampus to intracranial volume ratio ×100 adjusted by age and field strength, HCCR; Hippocampus to cerebellum volume ratio ×100 adjusted by age and field strength.

* Indicates significance at p<0.0001

† Adjusted by age, field strength and intracranial volume

Table 2: Descriptive discriminants analysis of predictors: Predictors are generally significant with loadings more than 0.5 on the discriminant functions (standardised coefficient correlation) with large effect sizes (canonical R^2 and eigenvalue) across all groups as well as in all contrasts except for CN contrasting pooled of MCIs and MCIr, which the effect sizes are small and not effective in separating the groups.

		Correlation of predictors with discriminant functions (std. coef.)		Pooled within-group correlation among predictors	Univariate significance
		1	2	MMSE	F [DF]
All groups					
HCICV + MMSE					
	HCICV	-0.675	-0.740	0.024	87.86[3, 656]
	MMSE	-0.705	-0.711		95.22[3, 656]
	Canonical R^2	0.443	0.002		
	Eigenvalue	0.793	0.002		
HCCR + MMSE					
	HCCR	-0.644	0.768	0.035	79.83 [3, 656]
	MMSE	-0.715	-0.703		95.22 [3, 656]
	Canonical R^2	0.425	0.003		
	Eigenvalue	0.738	0.003		
MCIc vs. [CN, MCIs and MCIr]					
HCICV + MMSE					
	HCICV	0.666		0.108	221.73 [1,658]
	MMSE, std. coef	0.683			231.08 [1,658]
	Canonical R^2	0.385			
	Eigenvalue	0.627			
HCCR + MMSE					
	HCCR	-0.655		0.119	215.79 [1,658]
	MMSE	-0.683			231.08 [1,658]
	Canonical R^2	0.379			
	Eigenvalue	0.609			
MCIc vs. CN					
HCICV + MMSE					
	HCICV	-0.700		0.013	275.73 [1,507]
	MMSE	-0.707			281.25 [1,507]
	Canonical R^2	0.521			
	Eigenvalue	1.088			
HCCR + MMSE					
	HCCR	-0.676		0.032	256.09 [1,507]
	MMSE	-0.710			281.25 [1,507]
	Canonical R^2	0.505			
	Eigenvalue	1.019			
MCIc vs [MCIs and MCIr]					
HCICV + MMSE					
	HCICV	-0.717		0.082	69.94[1, 336]
	MMSE	-0.647			58.46 [1, 336]

	Canonical R^2	0.263			
	Eigenvalue	0.356			
HCCR + MMSE					
	HCCR	-0.717		0.142	75.00 [1, 336]
	MMSE	-0.610			58.46 [1, 336]
	Canonical R^2	0.260			
	Eigenvalue	0.352			
CN vs. [MCIs and MCIr]					
HCICV + MMSE					
	HCICV	-0.617		0.080	30.45 [1, 471]
	MMSE	-0.734			40.72 [1, 471]
	Canonical R^2	0.122			
	Eigenvalue	0.139			
HCCR + MMSE					
	HCCR	-0.498		0.069	17.47 [1, 471]
	MMSE	-0.826			40.72[1, 471]
	Canonical R^2	0.103			
	Eigenvalue	0.115			

CN; cognitively normal, MCIr; mild cognitive impairment reverted to normal, MCIc; mild cognitive impairment converted to Alzheimer's disease in five years, MCIs; mild cognitive impairment stable for five years or more, APOE e4; Apolipoprotein E allele 4, MMSE; mini mental status examination, HCICV; Hippocampus to intracranial volume ratio × 100 adjusted for age and field strength, HCCR; Hippocampus to Cerebellum volume ratio ×100 adjusted for age and field strength, std.coef; standardized coefficient

In binary classifications (Table 4), classification performance of MMSE, HCICV and HCCR was generally comparable and more specific than sensitive for detecting MCIc from the other three groups: classification accuracy from 77.6% to 78.9%, specificity from 90.9% to 92%, and sensitivity from 41.2% to 47.1%. Similar trends were demonstrated in all other contrasts. ROC analyses demonstrated no statistically significant difference between AUC for MMSE, HCICV and HCCR based on Delong's test in all contrasts (Table 5 and Figure 2).

Importantly, using ICV ratio to normalize the hippocampus or using regression to adjust for ICV was separately assessed, which was found to have little impact on the classification results (Figure 3).

Table 3: Group classification performance: Predictors separate MCIc from CN but cannot separate MCIs and MCIr from others and majority of them were classified as CN and minority as MCIc.

		MMSE				HCICV				HCCR				HCICV + MMSE				HCCR + MMSE			
References		CN	MCIc	MCIs	MCIr	CN	MCIc	MCIs	MCIr	CN	MCIc	MCIs	MCIr	CN	MCIc	MCIs	MCIr	CN	MCIc	ICIs	MCIr
Prediction	CN	293	71	76	33	272	70	78	29	283	69	83	34	290	42	75	34	293	43	73	34
	MCIc	29	116	36	6	50	117	34	10	39	118	29	5	27	144	37	5	25	142	36	5
	MCIs	0	0	0	0	0	0	0	0	0	0	0	0	5	1	0	0	4	2	3	0
	MCIr	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sensitivity %		91.0	62.0	-	-	84.5	62.57	-	-	87.9	63.1	-	-	90.1	77.0	-	-	91.0	75.9	-	-
Specificity %		46.8	85.0	-	-	47.6	80.13	-	-	45.0	84.6	-	-	55.3	85.4	-	-	55.6	86.1	-	-
Pos Pred Value %		62.0	62.0	-	-	60.6	55.45	-	-	60.3	61.8	-	-	65.8	67.6	-	-	66.1	68.3	-	-
Neg Pred Value %		84.5	85.0	-	-	76.3	84.41	-	-	79.6	85.3	-	-	85.4	90.4	-	-	86.6	90.1	-	-
Prevalence %		48.8	28.3	17.0	5.9	48.8	28.33	17.0	5.9	48.8	28.3	17.0	5.9	48.8	28.3	17.0	5.9	48.8	28.3	17.0	5.9
Accuracy (95%CI)		62.0 (58.1 – 65.7)				58.94 (55.1– 62.7)				60.8 (56.9 – 64.5)				65.8 (62.0 – 69.4)				66.4 (62.6 – 67.0)			
Kappa %		33.3				28.90				31.3				41.1				42.1			

CN; cognitively normal, MCIr; mild cognitive impairment reverted to normal, MCIc; mild cognitive impairment converted to Alzheimer's disease in five years, MCIs; mild cognitive impairment stable for five years or more, MMSE; mini mental status examination, HCICV; Hippocampus to intracranial volume ratio adjusted for age and field strength, HCCR; Hippocampus to Cerebellum volume ratio adjusted for age and field strength, Pos Pred Value; positive predictive value, Neg Pred value; negative predictive value, 95%CI; 95% confident interval.

Combined predictors classification

The combination of predictors (hippocampal and MMSE) improved almost all aspects of classification performance, but as for individual predictor models, classification was optimal in classifying participants into two groups, CN and MCIc. A high proportion of CN and MCIc were correctly classified, whereas a majority of MCIs and MCIr were misclassified as CN and a minority as MCIc (Table 3).

Almost all aspects of classification performance in all binary classifications that identified MCIc from other groups (i.e. MCIc vs. pooled of others, MCIc vs. CN, and MCIc vs. pooled of MCIs and MCIr) were improved with the combination of HCICV or HCCR and MMSE, when compared with the individual predictor. In contrast, combination models did not show improvement in discriminating CN from pooled MCIs and MCIr groups. Combination models showed significant improvements over those of individual predictors (Table 4).

Table 4: Contrast classification performance: MCIc contrasted separately with all three groups together, other two MCI groups and CN alone. CN also contrasted with MCIs and MCIr together. In MCIc contrasts (with all groups or CN alone), predictors were mostly specific than sensitive when they were not in combinations while combinations improved all classification performances.

	Classification Accuracy % (95% CI)	Kappa %	McNemar Test P-value	Sensitivity %	Specificity %	Positive predictive value %	Negative predictive value %	R^+	R^-	AUROC (95% CI)
MCIc vs. [CN + MCIs + MCIr]										
MMSE	77.6 (74.2 – 80.7)	37.5	<0.0001	41.2	92.0	67.0	69.8	.2	.6	0.80 (0.76 – 0.84)
HCICV	78.9(75.6 - 82.0)	44.0	<0.0001	50.3	90.3	67.1	82.1	.2	.6	0.82 (0.79 – 0.86)
HCCR	78.5 (75.2 – 81.6)	41.8	<0.0001	47.1	90.9	67.2	81.3	.2	.6	0.82 (0.78 – 0.85)
HCICV + MMSE	83.2 (80.1 – 86.0)	56.6	0.008	62.6	91.3	74.1	86.1	.2	.4	0.89 (0.86 – 0.91)
HCCR + MMSE	83.5 (80.4 – 86.2)	57.9	0.0554	65.2	90.7	73.5	86.8	.0	.4	0.88 (0.85 – 0.91)
MCIc vs. CN										
MMSE	80.4 (76.6 - 83.7)	55.7	<0.0001	62.1	91.0	80.0	80.5	.9	.4	0.84 (0.81- 0.88)
HCICV	76.4 (72.5 – 80.1)	48.12	0.0828	62.6	84.5	70.1	79.5	.0	.4	0.86 (0.82 – 0.89)
HCCR	78.8 (75.0 – 82.3)	52.8	0.0053	63.1	87.9	75.2	80.4	.2	.4	0.85 (0.81 – 0.88)

HCICV + MMSE	85.5 (82.1 – 88.4)	68.3	0.2010	77.0	90.4	82.3	87.1	.0	.3	0.93 (0.90 – 0.95)
HCCR + MMSE	86.1 (82.7 – 88.9)	69.4	0.0576	76.5	91.6	84.1	87.0	.1	.3	0.92 (0.89 – 0.94)
MCIc vs. [MCIs + MCIr]										
MMSE	66.6 (61.3 – 71.6)	33.6	0.0084	62.0	72.2	73.4	60.6	.2	.5	0.72 (0.67 – 0.77)
HCICV	69.2 (64.0 – 74.1)	36.8	0.0241	78.6	57.6	69.7	68.5	.9	.4	0.75 (0.69 – 0.80)
HCCR	69.8 (64.6 - 74.7)	38.1	0.0376	78.6	58.9	70.3	69.0	.9	.4	0.75 (0.70 – 0.81)
HCICV + MMSE	74.6 (69.6 – 79.1)	48.3	0.5898	78.6	69.5	76.2	72.4	.6	.3	0.81 (0.76 - 0.85)
HCCR + MMSE	72.8 (67.7 – 77.5)	44.9	0.9170	75.9	68.9	75.1	69.8	.4	.4	0.80 (0.75 – 0.85)
CN vs. [MCIs + MCIr]										
MMSE	70.8 (66.5 – 74.9)	22.2	<0.0001	73.0	58.9	90.7	28.5	.8	.5	0.66 (0.61 – 0.72)
HCICV	69.1 (64.8 – 73.3)	12.7	<0.0001	93.8	16.6	70.6	55.6	.1	.4	0.65 (0.60 – 0.70)
HCCR	69.3 (65.0 – 73.5)	11.2	<0.0001	95.7	13.3	70.2	58.8	.1	.3	0.61 (0.55 – 0.66)
HCICV + MMSE	70.4 (66.01– 74.5)	20.4	<0.0001	91.0	26.5	72.5	62.0	.2	.3	0.70 (0.65 – 0.75)
HCCR + MMSE	71.7 (67.4 – 75.7)	23.8	<0.0001	91.9	28.5	73.3	62.3	.3	.3	0.68 (0.63 – 0.73)

CN; cognitively normal, MCIr; mild cognitive impairment reverted to normal, MCIc; mild cognitive impairment converted to Alzheimer's disease in five years, MCIs; mild cognitive impairment stable for five years or more, MMSE; mini mental status examination, HCICV; Hippocampus to intracranial volume ratio adjusted for age and field strength, HCCR; Hippocampus to Cerebellum volume ratio adjusted for age and field strength, LR^+; positive likelihood ratio, LR^-; negative likelihood ratio, AUROC; area under operating characteristic curve, 95%CI; 95% confident interval.

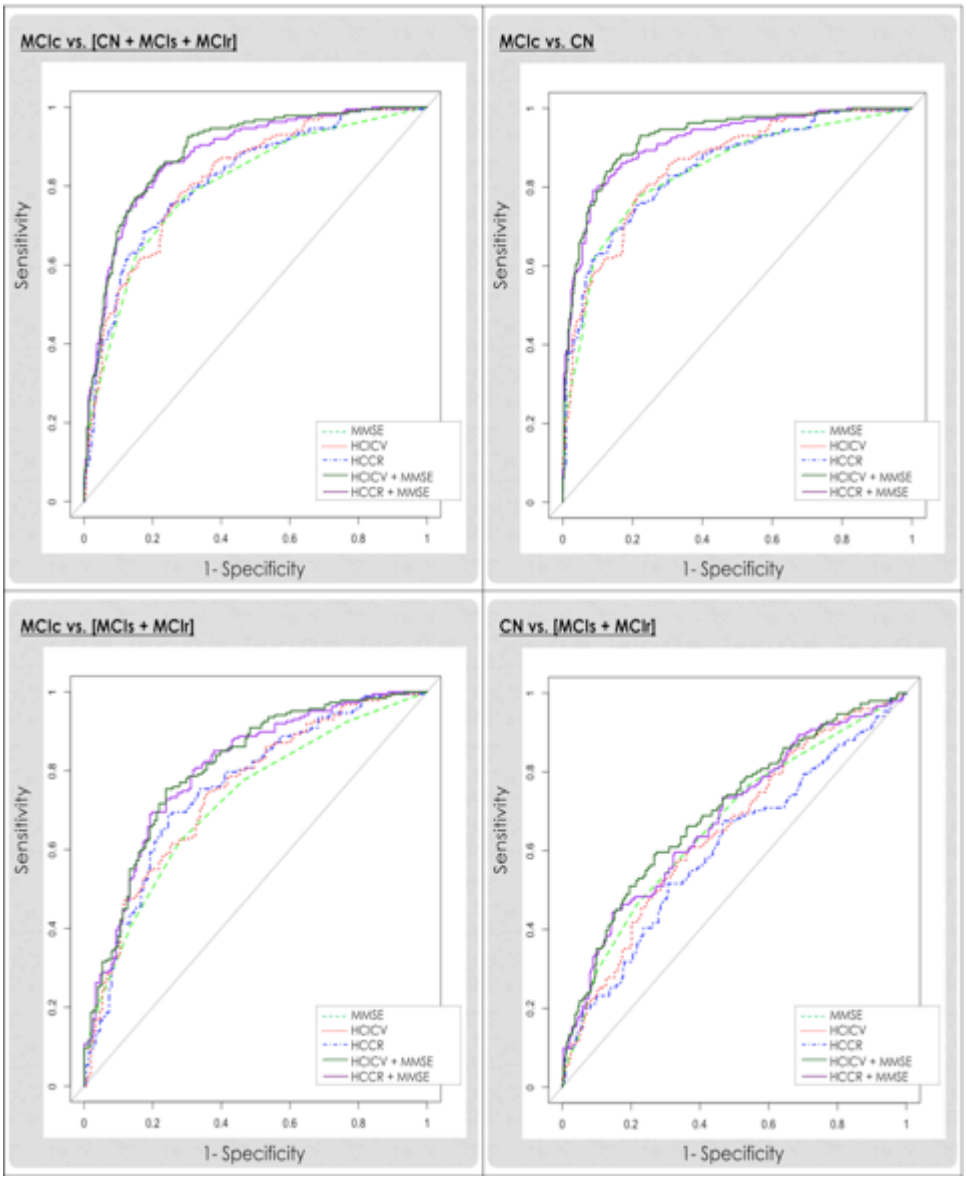


Figure 2: Receiver Operating Characteristic (ROC) curve for group membership: Area under the curve (AUC) revealed that in mild cognitive impairment converted to Alzheimer (MCIc) contrasted with pooled of other groups (upper left) or cognitively normal (CN) alone (upper right) combination of MMSE and hippocampus to intracranial volume ratio (HCICV) or hippocampus to cerebellum volume ratio (HCCR) were better than each predictor separately. This was partially true for MCIc contrasted pooled of other MCI groups (lower left), while not true for CN contrasted other MCI groups (lower right).

The discrimination ability (AUC of ROC analyses) of combinations of HCICV or HCCR and MMSE were significantly better than each predictor individually (Delong's test;

z<-4, p<0.001), while there was no significant difference between the HCICV-MMSE and HCCR-MMSE models. Additionally, analyses suggested that there was no difference in discrimination ability between the combination models and MMSE alone in separating CN from MCIs and MCIr groups. In contrast, the combination of hippocampal ratios (to ICV or cerebellar volume) and MMSE was significantly better in discriminating MCIc from pooled MCIs and MCIr (Table 4 and figure 2). Additional analyses investigating the ability to discriminate MCI who convert within specified time periods (1-5 years) revealed that performance was better in the first three years of follow-up compared to the final two years (Table 5).

Table 5: Area under receiver operating characteristic curve of mild cognitive impairment convert to Alzheimer's disease in one up to five years vs. pooled of mild cognitive impairment remain stable for five years or more and those who revert to cognitively normal

MCIc vs. [MCIs & MCIr]	AUROC (95% CI)	
	HCICV-MMSE	HCCR-MMSE
MCI convert in year 1	0.89 (0.74 – 0.99)	0.93 (0.85 – 0.99)
MCI convert in year 2	0.75 (0.67 – 0.82)	0.75 (0.67 - 0.82)
MCI convert in year 3	0.78 (0.72 – 0.85)	0.79 (0.72 – 0.85)
MCI convert in year 4	0.72 (0.63 – 0.80)	0.75 (0.67 – 0.83)
MCI convert in year 5	0.66 (0.54 – 0.79)	0.68 (0.56 – 0.81)

MCIc; mild cognitive impairment convert to Alzheimer's disease, MCIs; mild cognitive impairment remain stable for five years or more, MCIr; mild cognitive impairment revert to cognitively normal, AUROC; area under receiver operating characteristic curve, HCICV; hippocampus to intracranial volume ratio adjusted for age and field strength, HCCR; hippocampus to cerebellar volume ratio adjusted for age and field strength, MMSE; mini mental state examination.

Classification performance of the predictors in combination (HCCR-MMSE and HCICV-MMSE), for discriminating between MCIc from other groups in all contrasts was generally substantial: classification accuracy for MCIc vs. all other groups was more than 83% with sensitivity between 65.2% - 62.6%, with a specificity of 90.7% - 91.3% and an AUC of 0.88 – 0.89. The performance was even better when discriminating MCIc from CN (Table 4).

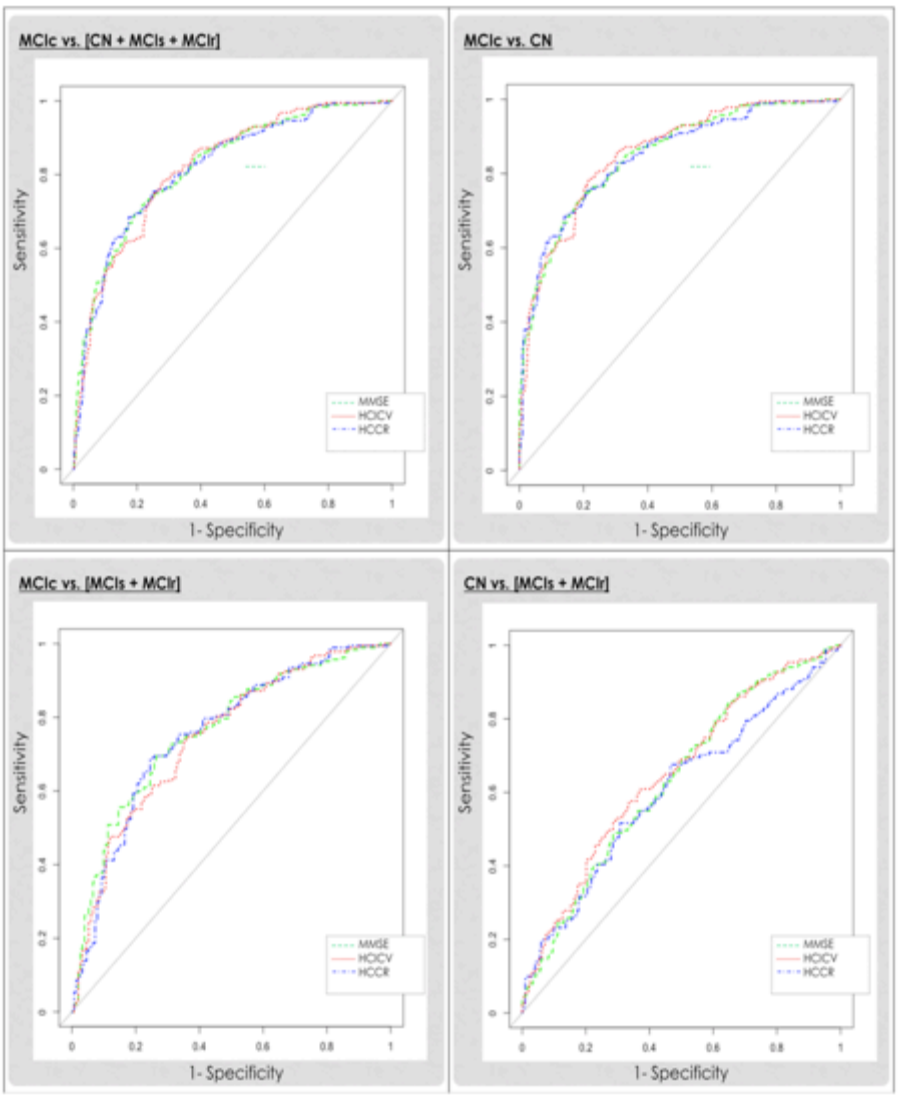


Figure 3: Receiver Operating Characteristic (ROC) curve for the three predictors: Area under curve reveals similar discrimination for the predictors in all group contrasts.

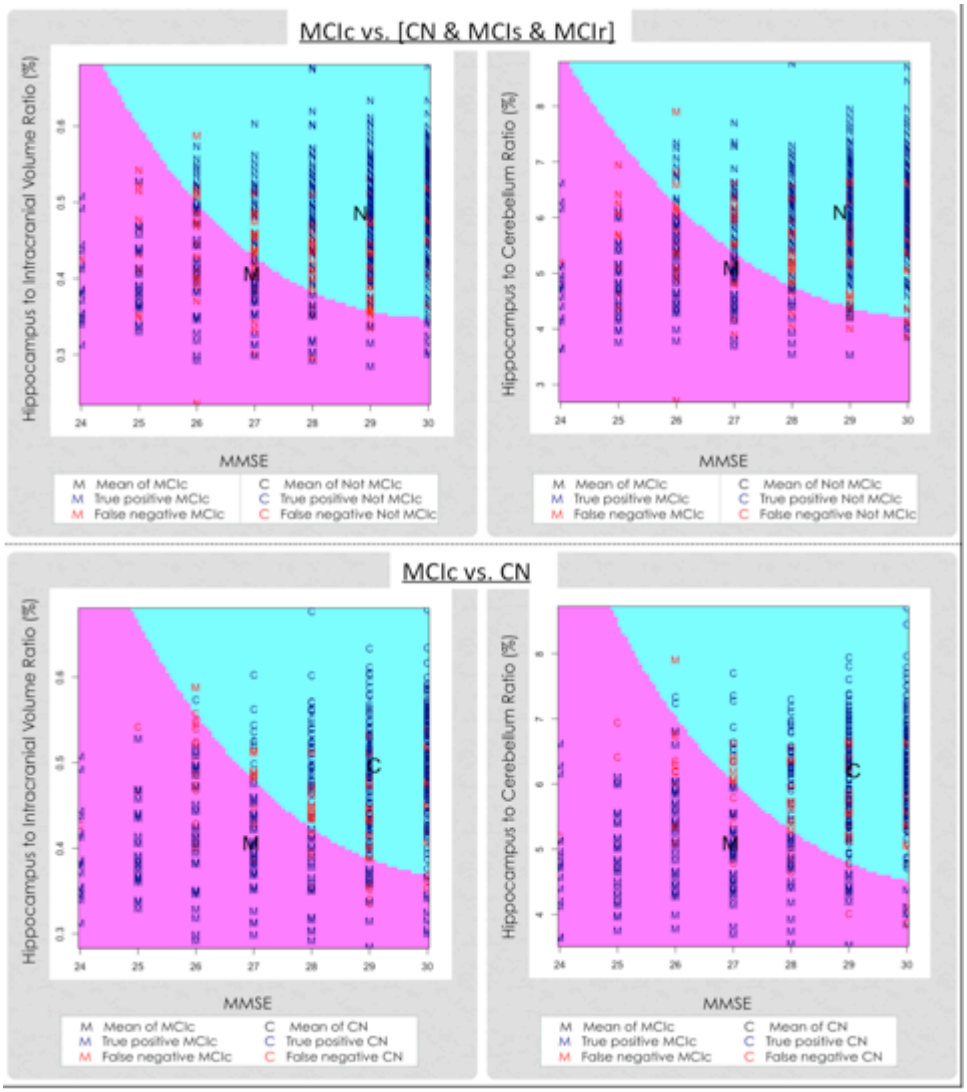


Figure 4: Partition plots: Thresholds of different hippocampus to intracranial volume (HCICV, right) or hippocampus to cerebellum ratios (HCCR, left) based on different MMSE scores, which separate mild cognitive impairment converted to Alzheimer (MCIc) from the pooled of cognitively normal (CN) and other MCI groups (upper) and from CN alone (lower).

Based on the partition plots in Figure 4, individuals with MMSE scores of less than 25 were mostly classified as MCIc regardless of the HCICV and HCCR values. For individuals with higher MMSE values, lower hippocampal volumes were observed in those who were classified as MCIc. For example, for a MMSE equal to 25, HCICV needed to be less than 0.6% or HCCR less than 7.5%, to be classified as MCIc. The thresholds for HCICV or HCCR were 0.5% and 6.3% for a MMSE of 26, 0.42% and 5.3% for 27, 0.38% and 4.8% for

MMSE for 28. HCICV or HCCR needed to be less than 0.35% and 4.3% respectively for MCIc diagnosis, when MMSE scores were 29-30. These thresholds were slightly smaller for discriminating MCIc from CN.

Discussion

This study aimed to investigate the performance of hippocampal volume normalized to cerebellar volume as a new measure for the clinical discrimination of MCI individuals at risk of AD conversion within five years. A combination of HCCR and MMSE was most effective in identifying MCI at risk of conversion. The main findings were that (1) the combination of HCCR or HCICV and MMSE and MMSE performed better in classifying MCI at risk of AD conversion than each measure individually, (2) the classification performance of HCCR and MMSE was similar to that of HCICV and MMSE, and (3) CN and MCI who did not convert to AD within five years did not differ statistically in their normalised hippocampal measures at a particular MMSE score.

Among the brain regions implicated in AD neuropathology, hippocampal shrinkage is most predictive of AD related cognitive dysfunction (Jack et al., 2000) and MMSE is the most widely used screening instrument for AD/dementia. We found that HCCR, a new normalized hippocampal measure, performed as well as HCICV in classification performance. Although, none of HCICV, HCCR or MMSE reliably identified MCI individuals who progressed to AD alone, we confirmed that HCICV or HCCR in combination with MMSE were effective in differentiating MCI who progressed to AD from CN and MCI who did not progress.

Both combinations were similar in performance and revealed a high level of classification accuracy, particularly for discriminating between MCIc and CN. However, classification accuracy only reflects the proportion of true results (positive or negative) in the

sample. In order to be practical and useful, a test needs to be sensitive and specific. Our results revealed that of those with MCIc, 65.2% - 62.6% were correctly identified (satisfactory sensitivity) by the combination models (HCCR + MMSE or HCICV + MMSE), while 91.3% - 90.7% of non-converters (CN, MCIs and MCIr) were correctly identified (high specificity). Furthermore, in those who were positively identified as MCIc, the likelihood of being truly MCIc was about nine-fold that of those who were falsely identified as MCIc (high positive likelihood ratio). For those who were positively identified as not being MCIc, the likelihood of being MCIc was close to a third that of those who were correctly identified as not being MCIc (low negative likelihood ratio). Therefore, not only was the overall accuracy of the combinations high, but the probabilities of false positive/negative results were also acceptable. Altogether, the combinations of hippocampal measures and MMSE are likely to be better than any single measure in identifying individuals with MCI at risk of AD conversion, but also effective in ruling out those individuals unlikely to convert within 5 years.

Interestingly, using either a combination of HCICV and MMSE, or HCCR and MMSE resulted in similar performance. This is important because it indicates that normalisation of hippocampal volume by ICV or cerebellar volume is equally effective and thus validates our approach. ICV estimation is more sensitive to scanning parameters and segmentation methods than cerebellar volume. This is probably because ICV segmentation relies on the correct identification of the boundary between the subarachnoid space and CSF fluid whose contrast is more variable to that between cerebellar gray matter and CSF. Thus, cross-sectional comparison between patients (or longitudinal within patients) assessed with different scanning parameters may be more difficult when using the ICV ratio. Consequently, in these contexts normalization by cerebellar volume may be more reliable and preferable.

The classification performance of HCICV and MMSE were in agreement with previous studies (in spite of different study parameters) that revealed a sensitivity of 67% and specificity of 72% for ICV adjusted hippocampal volume and a sensitivity of 54% and specificity of 80% for MMSE in identifying MCIc from CN (Arevalo-Rodriguez et al., 2015; Chupin et al., 2009). Better performance for the combination models was consistent and comparable with a previous study that showed better prediction of a combination of hippocampal volume, entorhinal cortex volume, MMSE, informant report of functioning questionnaire, the University of Pennsylvania Smell Identification Test, and Selective Reminding Test immediate recall score with a sensitivity of 70% and a specificity of 90% (Devanand et al., 2008). Additionally the models' performances were comparable with other studies with combination of multiple modalities (including MRI and cognitive measures), which mostly had many predictors in each modality (Costafreda et al., 2011; Ferrarini et al., 2009; Moradi et al., 2015; Zhang et al., 2011). This suggests that adding more predictors into a model may not necessarily improve classification performance when the predictors are from a single domain. Therefore, as well as the comparability of the current findings with previous studies which used complex combinations of predictors, the combination of HCCR and MMSE have the advantage of being easily implementable and interpretable, and thus may facilitate clinical adoption.

It is interesting to note that MCIs and MCIr did not differ from CN based on the combination of HCICV or HCCR and MMSE while they differed from MCIc. This suggests that those who are not at actual risk of short term AD conversion are not substantially different from CN. A measure of concurrent decline in function and structure is likely to be a better predictor of AD conversion in short term.

Most classification studies conducted to date were predominantly based on multi-domain/ multivariate predictors, and thus too complex to be easily adoptable in clinical practice. This study stands out in its use of a combination of simple structural (HCCR) and functional (MMSE) measures with a potential diagnostic value for identifying MCI subjects at risk of converting to AD in 5 years easily applicable in clinical practice.

Conclusion

The need to evaluate AD-related biological markers for identifying those at risk of AD conversion and to include them in MCI diagnosis has been well documented. However, there is no agreement on a biomarker that can be effectively applied in clinical practice. In the present study, we show that a combination of one brain biomarker, either HCCR or HCICV, and MMSE can accurately identify individuals at risk of AD conversion within five years. Moreover, normalization by cerebellar volume is as precise as normalization by intracranial volume with the advantage of being more practical in a clinical setting.

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). 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 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.; Neu- roRx 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 co- ordinated 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 partly undertaken on the National Computational Infrastructure (NCI) facility in Canberra, Australia, which is supported by the Australian Commonwealth Government and was supported by an Australian Government Research Training Program (RTP) Scholarship.

References:

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., . . . Phelps, C. H. (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. *Alzheimers Dement, 7*(3), 270-279. doi:10.1016/j.jalz.2011.03.008

Arevalo-Rodriguez, I., Smailagic, N., Roque, I. F. M., Ciapponi, A., Sanchez-Perez, E., Giannakou, A., . . . Cullum, S. (2015). Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database Syst Rev*(3), CD010783. doi:10.1002/14651858.CD010783.pub2

Association, T. A. P. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC.

Barnes, J., Ridgway, G. R., Bartlett, J., Henley, S. M., Lehmann, M., Hobbs, N., . . . Fox, N. C. (2010). Head size, age and gender adjustment in MRI studies: a necessary nuisance? *Neuroimage, 53*(4), 1244-1255. doi:10.1016/j.neuroimage.2010.06.025

Chupin, M., Gerardin, E., Cuingnet, R., Boutet, C., Lemieux, L., Lehericy, S., . . . Alzheimer's Disease Neuroimaging, I. (2009). Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. *Hippocampus, 19*(6), 579-587. doi:10.1002/hipo.20626

Costafreda, S. G., Dinov, I. D., Tu, Z., Shi, Y., Liu, C. Y., Kloszewska, I., . . . Simmons, A. (2011). Automated hippocampal shape analysis predicts the onset of dementia in mild cognitive impairment. *Neuroimage, 56*(1), 212-219. doi:10.1016/j.neuroimage.2011.01.050

Devanand, D. P., Liu, X., Tabert, M. H., Pradhaban, G., Cuasay, K., Bell, K., . . . Pelton, G. H. (2008). Combining early markers strongly predicts conversion from mild cognitive impairment to Alzheimer's disease. *Biol Psychiatry, 64*(10), 871-879. doi:10.1016/j.biopsych.2008.06.020

Falahati, F., Westman, E., & Simmons, A. (2014). Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging. *J Alzheimers Dis, 41*(3), 685-708. doi:10.3233/JAD-131928

Ferrarini, L., Frisoni, G. B., Pievani, M., Reiber, J. H., Ganzola, R., & Milles, J. (2009). Morphological hippocampal markers for automated detection of Alzheimer's disease and mild cognitive impairment converters in magnetic resonance images. *J Alzheimers Dis, 17*(3), 643-659. doi:10.3233/JAD-2009-1082

Forlenza, O. V., Diniz, B. S., Stella, F., Teixeira, A. L., & Gattaz, W. F. (2013). Mild cognitive impairment (part 1): clinical characteristics and predictors of dementia. *Revista Brasileira de Psiquiatria, 35*(2), 178-185. doi:10.1590/1516-4446-2012-3503

Fritz, J. M., & Wainner, R. S. (2001). Examining diagnostic tests: an evidence-based perspective. *Phys Ther, 81*(9), 1546-1564.

Ganguli, M. (2013). Can the DSM-5 framework enhance the diagnosis of MCI? *Neurology, 81*, 2045- 2050.

Hansen, T. I., Brezova, V., Eikenes, L., Haberg, A., & Vangberg, T. R. (2015). How Does the Accuracy of Intracranial Volume Measurements Affect Normalized Brain Volumes? Sample Size Estimates Based on 966 Subjects from the HUNT MRI Cohort. *AJNR Am J Neuroradiol, 36*(8), 1450-1456. doi:10.3174/ajnr.A4299

Heinen, R., Bouvy, W. H., Mendrik, A. M., Viergever, M. A., Biessels, G. J., & de Bresser, J. (2016). Robustness of Automated Methods for Brain Volume Measurements across Different MRI Field Strengths. *PLoS One, 11*(10), e0165719. doi:10.1371/journal.pone.0165719

Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., . . . Contributors. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement, 14*(4), 535-562. doi:10.1016/j.jalz.2018.02.018

Jack, C. R., Jr., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., . . . Weiner, M. W. (2008). The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging, 27*(4), 685-691. doi:10.1002/jmri.21049

Jack, C. R., Jr., Petersen, R. C., Xu, Y., O'Brien, P. C., Smith, G. E., Ivnik, R. J., . . . Kokmen, E. (2000). Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. *Neurology, 55*(4), 484-489.

Keihaninejad, S., Heckemann, R. A., Fagiolo, G., Symms, M. R., Hajnal, J. V., Hammers, A., & Alzheimer's Disease Neuroimaging, I. (2010). A robust method to estimate the intracranial volume across MRI field strengths (1.5T and 3T). *Neuroimage, 50*(4), 1427-1437. doi:10.1016/j.neuroimage.2010.01.064

Malone, I. B., Leung, K. K., Clegg, S., Barnes, J., Whitwell, J. L., Ashburner, J., . . . Ridgway, G. R. (2015). Accurate automatic estimation of total intracranial volume: a nuisance variable with less nuisance. *Neuroimage, 104*, 366-372. doi:10.1016/j.neuroimage.2014.09.034

Mitchell, A. J. (2009). CSF phosphorylated tau in the diagnosis and prognosis of mild cognitive impairment and Alzheimer's disease: a meta-analysis of 51 studies. *J Neurol Neurosurg Psychiatry, 80*(9), 966-975. doi:10.1136/jnnp.2008.167791

Moradi, E., Pepe, A., Gaser, C., Huttunen, H., Tohka, J., & Alzheimer's Disease Neuroimaging, I. (2015). Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *Neuroimage, 104*, 398-412. doi:10.1016/j.neuroimage.2014.10.002

Nordenskjold, R., Malmberg, F., Larsson, E. M., Simmons, A., Brooks, S. J., Lind, L., . . . Kullberg, J. (2013). Intracranial volume estimated with commonly used methods could introduce bias in studies including brain volume measurements. *Neuroimage, 83*, 355-360. doi:10.1016/j.neuroimage.2013.06.068

Pandya, S. Y., Clem, M. A., Silva, L. M., & Woon, F. L. (2016). Does mild cognitive impairment always lead to dementia? A review. *J Neurol Sci, 369*, 57-62. doi:10.1016/j.jns.2016.07.055

Park, M. H., Han, C., & Initiative, t. A. s. D. N. (2015). Is there an MCI reversion to cognitively normal? Analysis of Alzheimer's disease biomarkers profiles. *international Psychiatrics*

, 27(3), 429-437. doi:10.1017/S1041610214002129

Pintzka, C. W., Hansen, T. I., Evensmoen, H. R., & Haberg, A. K. (2015). Marked effects of intracranial volume correction methods on sex differences in neuroanatomical structures: a HUNT MRI study. *Front Neurosci, 9*, 238. doi:10.3389/fnins.2015.00238

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, 155*, 530-548. doi:10.1016/j.neuroimage.2017.03.057

Ritchie, C., Smailagic, N., Noel-Storr, A. H., Ukoumunne, O., Ladds, E. C., & Martin, S. (2017). CSF tau and the CSF tau/ABeta ratio for the diagnosis of Alzheimer's disease dementia and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database Syst Rev, 3*, CD010803. doi:10.1002/14651858.CD010803.pub2

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.): Pearson.

Tabatabaei-Jafari, H., Walsh, E., Shaw, M. E., Cherbuin, N., & Alzheimer's Disease Neuroimaging, I. (2017). The cerebellum shrinks faster than normal ageing in Alzheimer's disease but not in mild cognitive impairment. *Hum Brain Mapp, 38*(6), 3141-3150. doi:10.1002/hbm.23580

Thal, D. R., Rub, U., Orantes, M., & Braak, H. (2002). Phases of A beta-deposition in the human brain and its relevance for the development of AD. *Neurology, 58*(12), 1791-1800.

Vandenberghe, R., Nelissen, N., Salmon, E., Ivanoiu, A., Hasselbalch, S., Andersen, A., . . . Dupont, P. (2013). Binary classification of (1)(8)F-flutemetamol PET using machine learning: comparison with visual reads and structural MRI. *Neuroimage, 64*, 517-525. doi:10.1016/j.neuroimage.2012.09.015

Whitwell, J. L., Crum, W. R., Watt, H. C., & Fox, N. C. (2001). Normalization of cerebral volumes by use of intracranial volume: Implications for longitudinal quantitative MR imaging. *American Journal of Neuroradiology, 22*(8), 1483-1489.

Yuan, Y., Gu, Z. X., & Wei, W. S. (2009). Fluorodeoxyglucose-positron-emission tomography, single-photon emission tomography, and structural MR imaging for prediction of rapid conversion to Alzheimer disease in patients with mild cognitive impairment: a meta-analysis. *AJNR Am J Neuroradiol, 30*(2), 404-410. doi:10.3174/ajnr.A1357

Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D., & Alzheimer's Disease Neuroimaging, I. (2011). Multimodal classification of Alzheimer's disease and mild cognitive impairment. *Neuroimage, 55*(3), 856-867. doi:10.1016/j.neuroimage.2011.01.008

STUDY 4

COGNITIVE/FUNCTIONAL MEASURES PREDICT ALZHEIMER'S DISEASE, DEPENDENT ON HIPPOCAMPAL VOLUME

Hossein Tabatabaei-Jafari[1], Marnie E. Shaw[2], Erin Walsh[1], Nicolas Cherbuin[1], For the Alzheimer's Disease Neuroimaging Initiative (ADNI)

[1]Centre for Research on Ageing, Health and Wellbeing, The Australian National University, Canberra, Australia

[2]College of Engineering and Computer Science, The Australian National University, Canberra, Australia

Published in The journals of Gerontology: Series B, 2019 Jan 21 [Epub ahead of print]

DOI: 10.1093/geronb/gbz011

Abstract

Objective: This study aimed to investigate the predictive value of cognitive/functional measures in combination with hippocampal volume on the probability of conversion from mild cognitive impairment (MCI) to Alzheimer's disease (AD).

Method: The Rey Auditory Verbal Learning Test for immediate memory, Mini Mental State Examination, a functional assessment for independent daily activities and Alzheimer's Disease Assessment Scale were used as cognitive/functional measures and hippocampal volume as neuroimaging measure. Logistic regression and Cox proportional hazard analyses were used to explore the measures' predictive values for AD conversion and time to conversion.

Results: The probability of conversion from MCI to AD was associated with cognitive function, but this was moderated by hippocampal volume: higher at lower hippocampal volume and lower at higher hippocampal volume. General cognitive/functional measures were less predictive than immediate memory in predicting time to conversion to AD at small hippocampal volumes.

Conclusion: Effectiveness of cognitive measures and subtle functional abnormality in predicting conversion from MCI to AD is dependent on hippocampal volume, thus combined evaluation should be considered. A combination of hippocampal volume and immediate memory appear to perform best in predicting time to conversion.

Introduction

Alzheimer's disease (AD) is a progressive degenerative disorder that involves cognitive decline severe enough to substantially impair daily activities. Cognitive decline accompanied by preserved daily activities has been specified as mild cognitive impairment (MCI), and is commonly known to be the prodromal phase of AD (Petersen et al., 1999). Approximately half of those with MCI progress to AD within five years (Pandya, Clem, Silva, & Woon, 2016). Identifying those who will progress to AD and predicting time to conversion remains an important clinical challenge.

Cognitive and functional performance is the central component of AD/MCI diagnostic. Thus, it is to be expected that cognitive performance is a sensitive predictor of conversion from MCI to AD (Belleville et al., 2017). A combination of measures from a range of domains typically provides a better predictor of disease progression (Belleville et al., 2017). Additionally, although intact daily function is the main clinical differentiator of MCI and AD diagnosis, subtle decline in daily function, while it remains in the normal range, is still predictive of conversion from MCI to AD (Gomar et al., 2011; Li et al., 2017). Furthermore, a combination of cognitive/functional measures with neuroimaging measures has been reported to produce significantly higher predictive accuracy (Devanand et al., 2008; Falahati, Westman, & Simmons, 2014; Moradi et al., 2015). Our recent study showed that the combination of a new hippocampal index – hippocampus to cerebellum volume ratio, HCCR - and Mini Mental State Examination (MMSE) could reliably identify those who progress from MCI to AD within five years with an area under receiver operating characteristic curve of 0.9 (Tabatabaei-Jafari, Walsh, Shaw, & Cherbuin, 2018).

Cognitive/functional impairment is positively associated with neurodegeneration, but this association is not straightforward and there is a mismatch between the extent of neural

pathology and the severity of cognitive/functional impairment (Steffener & Stern, 2012). Although a combination of cognitive performance and neuroimaging measures has been previously shown to have a higher predictive value compared to either measure alone, the relative contribution of these measures to each other across their range is not well understood. To answer these important questions, this study aimed to investigate the predictive value of cognitive/functional measures across the range of hippocampal volumes, in those who have a diagnosis of MCI and convert to AD within five years. Hippocampal volumes and cognitive/functional measures were selected on the basis of established associations with MCI and AD (Jack et al., 2005; Li et al., 2017; Tabatabaei-Jafari, Shaw, & Cherbuin, 2015; Tabatabaei-Jafari et al., 2018). We hypothesized that hippocampal volume would moderate the predictive value of cognitive/functional performance. Additionally, we aimed to investigate how well a combination of these measures would predict time to conversion from MCI to AD.

Methodology

Study Participants

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 MCI and early AD.

All participants of ADNI 1/GO/2 who were diagnosed with MCI at baseline were considered for inclusion. Those who were stable for at least 6 months after baseline diagnosis were included if they converted to AD within five years (MCIc; n=183) or remained stable for more than five years (MCIs; n=112).

Details of the diagnostic criteria can be found at the ADNI web site (http://www.adni-info.org/Scientists/AboutADNI.aspx). Briefly, Participants were classified as MCI if they had an MMSE greater than 24, a CDR of 0.5, a subjective report of memory concern, an objective memory loss, preserved daily living activity and did not meet diagnostic criteria for dementia. Participants were classified as having AD if they had MMSE scores less than 26, CDR 0.5 or 1.0 and fulfill criteria for clinically probable AD according to the Institute of Neurological and Communicative Diseases and Stroke/ Alzheimer's Disease and Related Disorders Association.

Neuroimaging acquisition and processing

Participants underwent high-resolution MRI brain scans on 1.5 (N=165) or 3 T (N=130) scanners from General Electric, Siemens, or Philips (Milwaukee, WI, USA; Germany; the Netherlands respectively) using a standardized ADNI acquisition protocol for 3D MP-RAGE sequence (Jack et al., 2008). Images which had undergone specific ADNI preprocessing correction steps to standardize images from different sites and platforms, were obtained for this study: (1) Grad wrap; a specific correction of image geometry distortion due to non-linearity, (2) B1 non-uniformity; B1 calibration to correct the image intensity non-uniformity that results when RF transmission is performed with a more uniform body coil while reception is performed with a less uniform head coil, (3) N3 correction; a histogram peak sharpening algorithm applied after grad wrap and B1 correction.

FreeSurfer version 5.3 (http://surfer.nmr.mgh.harvard.edu/) was used for automatic volumetric segmentation. The output images were visually checked for accurate segmentation.

Measures

One neuroimaging, three cognitive and one functional measure that have been extensively used for diagnostic purposes and cognitive and functional evaluation in clinical trials (Estevez-Gonzalez, Kulisevsky, Boltes, Otermin, & Garcia-Sanchez, 2003; Ito, Hutmacher, & Corrigan, 2012; Petersen et al., 2005) and with established associations with AD and predictive of MCI conversion (Ito et al., 2012; Li et al., 2017) were considered.

Hippocampal volume (HCV): The hippocampus is one of the first brain regions to be impacted by AD pathology, and one of the areas with greatest shrinkage over the course of the disease (Tabatabaei-Jafari et al., 2015). It is also the most sensitive structural predictor of AD conversion in MCI individuals (Eckerstrom et al., 2008). Therefore, hippocampal volume (HCV), the total volume of the left and right hippocampi adjusted for age, field strength, and ICV using the residual regression method described elsewhere (Pintzka, Hansen, Evensmoen, & Haberg, 2015) was investigated as neuroimaging predictor.

Mini Mental State Examination (MMSE): The MMSE (Folstein, Folstein, & McHugh, 1975) is the most widely used screening instrument for AD/ dementia (Arevalo-Rodriguez et al., 2015). It consists of 11 items with total scores ranging between 0 and 30, which lower scores reflecting more severe cognitive impairment. The items evaluate orientation in time and space (10 points), immediate recall (3 points), attention and calculation (5 points), delayed recall (3 points), language naming (2 points), following command (3 points), repetition (1 point), reading (1 point), writing (1 point), and visuospatial (1 point).

The Alzheimer's Disease Assessment Scale (ADAS): The modified 13-items ADAS-cog version (Petersen et al., 2005) was used here to assess general cognitive function. The modified ADAS consists of word recall (10 items), commands (5 items), construction (5 items), naming (5 items), ideational praxis (5 items), orientation (8 items), word recognition (12 items), recall instruction (5 items), spoken language (5 items), word finding (5 items), comprehension (5 items), delayed word recall (10 items), and number cancellation (5 items) in total 85 scores, which the higher score the severest the cognitive impairment.

Rey Auditory Verbal Learning Test (RAVLT): The RAVLT was used to evaluate episodic memory (Rey, 1941, 1964). It involves free recall of a list of 15 words in any order over five sequential trials. It is followed by recall of a second list of 15 words. Finally, the participant is asked to remember as many words as possible from the first list immediately following the second list recall and after 30 minutes. The scoring system of the RAVLT based on the correct number of words in each trial (5 in total) and evaluates a wide diversity of learning and memory functions including immediate memory, learning, and forgetting. The immediate recall score, RAVLT immediate, was considered for this study based on our introductory analyses that showed better predictive value for immediate memory compared with RAVLT learning and percentage of forgetting. The RAVLT immediate was computed as the total scores of trials one to five.

The Functional Assessment Questionnaire (FAQ): The FAQ assesses abilities of daily living with total scores ranging from 0 to 30. A score of 0 indicates "no impairment" and 30 "severely impaired" (Ito et al., 2012; Pfeffer, Kurosaki, Harrah, Chance, & Filos, 1982). The total FAQ score is the sum of 10 daily activities, with each activity being rated from 0 to 3 (0 = normal, 1 = has difficulty but does by self, 2 = requires assistance, 3 = dependent). Evaluated activities are (1) writing checks, paying bills, or balancing a

checkbook, (2) assembling tax records, business affairs, or other papers, (3) shopping alone for clothes, household necessities, or groceries, (4) playing a game of skill such as bridge or chess or working on a hobby, (5) heating water, making a cup of coffee, turning off the stove, (6) preparing a balanced meal, (7) keeping track of current events, (8) paying attention to and understanding a TV program, book, or magazine, (9) remembering appointments, family occasions, holidays, medications, and (10) traveling out of the neighborhood, driving, or arranging to take public transportation.

Statistical analysis

Statistical analyses were performed using the R statistical software (version 3.3.2). No missing values were present in the measures of interest. Mahalanobis distance was used for detection of univariate and multivariate outliers. No influential outlier was detected. Group differences in demographic variables were assessed by t-test for continuous variables and chi square tests for categorical variables. Univariate and bivariate models were used to investigate prediction of conversion from MCI to AD within five years as well as prediction of the time to conversion. Each bivariate model consisted of standardized values of hippocampal volume and one of four cognitive/functional measures as well as their interaction. The alpha level was set at < 0.05.

Prediction of AD conversion: Logistic regression analysis (package Stats; version 3.3.2 and package Caret; version 6.3-73) was used to quantify the magnitude of predictive values of the measures for predicting MCI conversion to AD. Univariate and bivariate models were applied. The odds ratios were used to quantify the magnitude of the main and interaction effects of the predictors. To graphically illustrate the effect of HCV, the probability of conversion for the cognitive/functional measures at different categories of HCV was investigated. Participants were categorized into three groups; small HCV for those with HCV

less than 5500 mm^3 (smaller than one SD), medium HCV for the volume between 5500 to 7500 mm^3 (within one SD), and large HCV for those with larger than 7500 mm^3 (larger than one SD).

Prediction of time to AD conversion: Cox proportional hazard analysis (package survival; version 2.40-1) was used to predict the time to AD conversion using the univariate and bivariate models. The Hazard ratio for a one SD change in the measures was used to quantify the magnitude of the main and interactive effects of the measures. In the case of the presence of interactive effect, to better interpret the effect the analyses were repeated with HCV as a categorical variable (small, medium and large) in the model. To graphically illustrate the contribution of cognitive/functional measures and HCV on probability of remaining MCI over time, separate Kaplan-Meier curves were plotted for different combinations of categorical levels of HCV (small, medium and large as defined above) and cognitive/functional measures (low and high). Cognitive/functional measures were categorized into low and high based on the median: 27 for MMSE, 13 for ADAS, 2 for FAQ, and 31 for RAVLT. Participants were categorized into six combinations for each cognitive/functional measure (Figure 1). For example, for ADAS, they were categorized into small HCV/low ADAS, small HCV/high ADAS, medium HCV/low ADAS, medium HCV/high ADAS, large HCV/low ADAS, and large HCV/high ADAS.

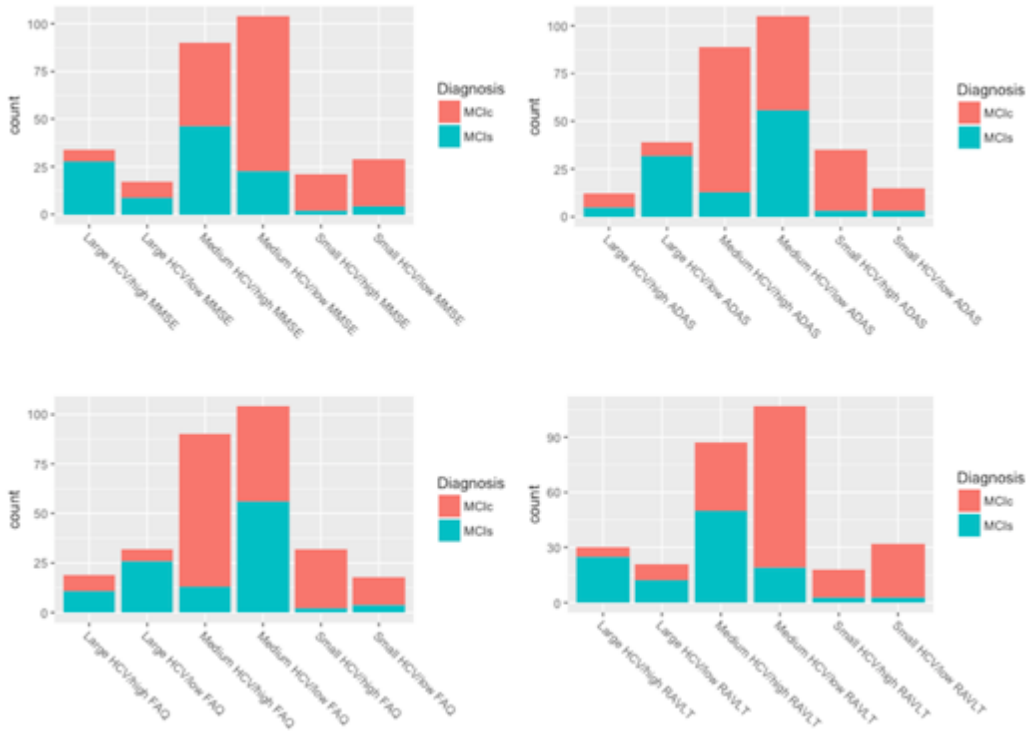


Figure 1: Number of participants at different hippocampal and cognitive/functional categories**:** In general the proportion of those who converted to AD are higher at small hippocampal volume regardless of the value of cognitive/functional measure. MCIc; mild cognitive impairment converted to Alzheimer's disease within five years, MCIs; mild cognitive impairment stable for five or more years, HCV; hippocampal volume adjusted by age, field strength and intracranial volume, MMSE; mini mental state examination, ADAS; Alzheimer disease assessment scale (cognitive subscale), RAVLT; Rey auditory verbal learning test (immediate memory subscale), FAQ; functional assessment questionnaire

Results

Participants' characteristics

Two hundred and ninety five MCI participants were categorized as MCI who subsequently converted to AD within five years (MCIc; n=183), and MCI who were stable for more than five years (MCIs; n=112). MCIs participants were about two years younger than MCIc but there were no significant differences in sex ratio or education between the

two groups. The proportion of APOE e4 carriers was significantly higher in MCIc than MCIs. All the measures of interest (HCV and cognitive/functional measures) were significantly different between the groups (Table 1).

Prediction of AD conversion

HCV, MMSE, ADAS, RAVLT, and FAQ were evaluated separately (univariate model) and all were significant predictors of AD conversion. Each cognitive/functional predictor remained a significant predictor of conversion from MCI to AD when HCV was added to the model, and HCV also remained a significant predictor. Additionally, HCV had additive effects with ADAS, RAVLT, and FAQ, whereas HCV and MMSE had interactive effects (Table 2). A graphical illustration (Figure 2) of the probability of conversion for the measures at three different categories of HCV (small, medium and large) suggests that having a medium to large HCV had a protective effect against conversion in MMSE from 24 to 30. However that protective effect was smaller at lower MMSE scores. The same pattern was demonstrated in the normal range of FAQ, i.e. having a medium to large HCV had a protective effect against conversion but the protection was lower when FAQ scores were closer to upper limit of the normal range. The pattern was relatively different for ADAS and RAVLT, where larger HCV was protective in medium ADAS or RAVLT.

Table 1: Participants characteristics and measurements

	MCIs	MCIc	Group difference
Sample size	112	183	
Age; year, Mean (SD)	71.95 (7.65)	74.31 (6.90)	YES
Age range, year	57 - 88	55 - 89	-
Male sex; N (%)	72 (64.29)	112 (61.20)	NO
Education; year, Mean (SD)	15.75 (3.03)	16.03 (2.73)	NO
APOE e4; N (%)	40 (35.71)	124 (67.76)	YES
One allele	32 (28.57)	93 (49.21)	YES
Two alleles	8 (7.14)	31 (17.32)	YES
Age at Diagnosis change; year, Mean (SD)	-	76.83 (7.05)	-
Time to Diagnosis change; year, Mean (SD)	-	2.40 (0.89)	-
MMSE, Mean (SD)	28.11 (1.49)	26.93(1.73)	YES
ADAS, Mean (SD)	13.45 (5.45)	20.19 (5.49)	YES
RAVLT immediate, Mean (SD)	38.40 (10.34)	28.85 (7.11)	YES
FAQ, Mean (SD)	1.75 (3.00)	4.96 (4.62)	YES
HCV[1], mm^3 ,Mean (SD)	7052.82 (909.03)	6223.92 (875.56)	YES

MCIc; mild cognitive impairment converted to Alzheimer's disease within five years, MCIs; mild cognitive impairment stable for five or more years, APOE e4; Apolipoprotein E allele 4, MMSE; mini mental state examination, ADAS; Alzheimer disease assessment scale (cognitive subscale), RAVLT; Rey auditory verbal learning test, FAQ; functional assessment questionnaire.

[1] Adjusted by age, field strength and intracranial volume

Table 2: Logistic Regression and Cox proportional hazard results: bivariate models

	Prediction of conversion				Prediction of time to conversion			
	Coef.	SE	OR (95% CI)	Z, P-value	Coef.	SE	HR (95% CI)	Z, P-value
HCV & MMSE								
HCV	-0.92	0.16	0.40 (0.29 -0.54)	-5.67, p<0.0001	-0.53	0.08	0.59 (0.51 – 0.68)	-6.89, p<00001
MMSE	-0.63	0.15	0.53 (0.39 – 0.71)	-4.17, p<0.0001	-0.40	0.08	0.66 (0.57 – 0.78)	-5.01, p<0.0001
HCV: MMSE	-0.35	0.17	0.71 (0.50 – 0.98)	-2.05, p<0.05	-0.25	0.08	0.78 (0.66 – 0.91)	-3.18, p=0.002
HCV & ADAS								
HCV	-0.66	0.17	0.52 (0.37 – 0.72)	-3.86, p=0.0001	-0.41	0.08	0.67 (0.57 – 0.78)	-5.05, p<0.0001
ADAS	1.18	0.19	3.26 (2.27 – 4.83)	6.16, p<0.0001	0.64	0.08	1.91 (1.62 – 2.25)	7.79, p<0.0001
HCV: ADAS	0.34	0.22	1.41 (0.92 – 2.14)	1.59, p=0.11	0.23	0.09	1.26 (1.07 – 1.49)	2.70, p<01
HCV & FAQ								
HCV	-0.95	0.17	0.39 (0.27 – 0.54)	-5.56, p<0.0001	-0.50	0.07	0.61 (0.53 – 0.70)	-6.99, p<0.0001
FAQ	1.05	0.22	2.84 (1.90 – 4.55)	4.72, p<0.0001	0.38	0.06	1.46 (1.30 – 1.65)	6.28, p<0.0001
HCV: FAQ	0.15	0.23	1.16 (0.73 – 1.77)	0.65, p=0.52	0.06	0.06	1.06 (0.95 – 1.19)	1.09, p=0.28
HCV & RAVLT								
HCV	-0.92	0.17	0.40 (0.28 – 0.55)	-5.44, p<0.0001	-0.49	0.08	0.61 (0.53 – 0.71)	-6.46, p<0.0001
RAVLT	-0.18	0.20	0.31 (0.21 – 0.44)	-6.04, p<0.0001	-0.75	0.10	0.47 (0.39 – 0.58)	-7.45, p<0.0001
HCV: RAVLT	-0.09	0.22	0.91 (0.59 – 1.40)	-0.41, p=0.68	-0.17	0.84	0.84 (0.70 – 1.01)	-1.84, p=0.07

MMSE; mini mental state examination (standardized), ADAS; Alzheimer disease assessment scale (standardized), RAVLT; Rey auditory verbal learning test (immediate; standardized), FAQ; functional assessment questionnaire (standardized), HCV; hippocampal volume adjusted by age, field strength and intracranial volume (standardized).

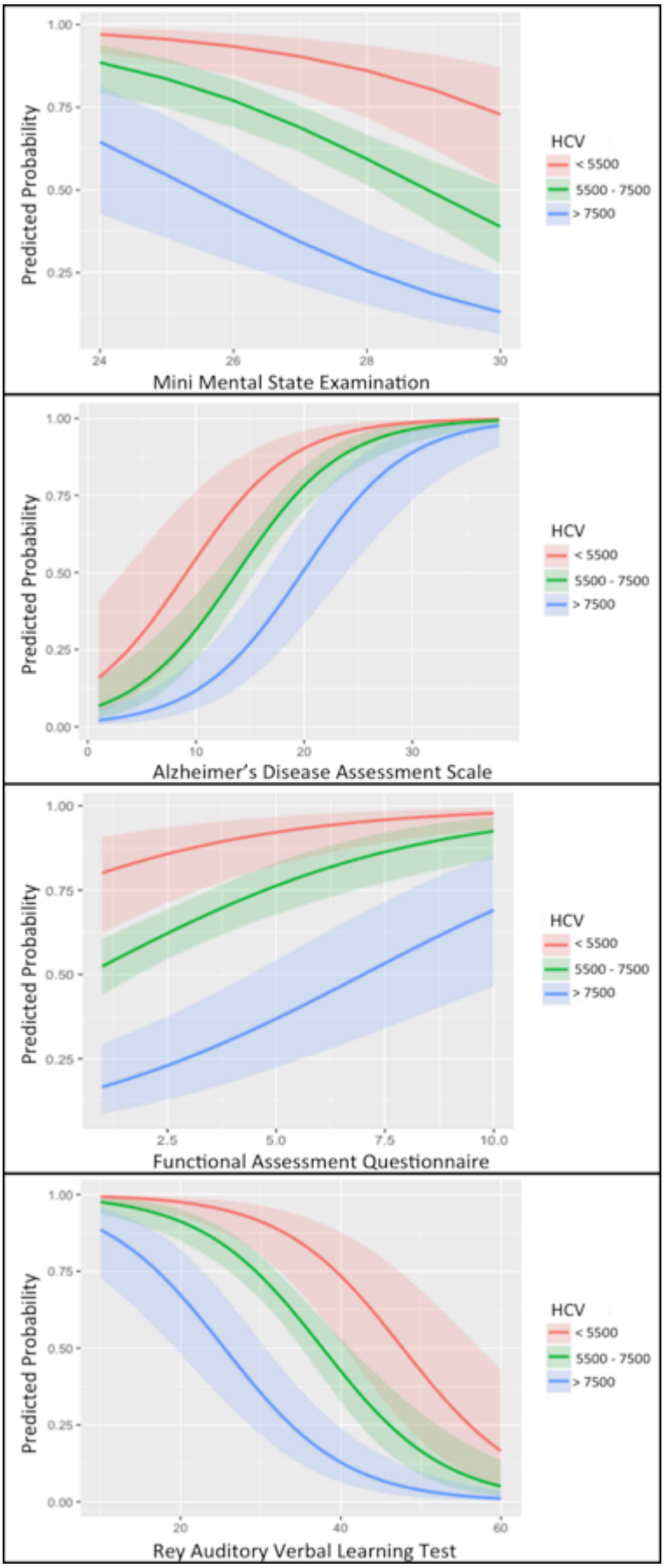


Figure 2: Predicted Probabilities of conversion to Alzheimer's: Predicted probabilities of cognitive measures at different hippocampal volumes. HCV has a reciprocal impact on predicted probability of the cognitive measures for conversion to Alzheimer's.

HCV; hippocampal volume adjusted by age, field strength and intracranial volume, MMSE; mini mental state examination, ADAS; Alzheimer disease assessment scale (cognitive subscale), RAVLT; Rey auditory verbal learning test (immediate memory subscale), FAQ; functional assessment questionnaire.

Table 3: Risk of conversion from MCI to AD over time (Cox Proportional Hazard) by hippocampal volume categories.

	Coef.	SE	HR (95%CI)	Z, P-value
MMSE & HCV				
Medium HCV category	**-0.59**	**0.18**	**0.55 (0.39-0.79)**	**-3.29, p=0.001**
Large HCV category	**-1.68**	**0.31**	**0.19 (0.10-0.35)**	**-5.35, p<0.0001**
MMSE	-0.07	0.14	0.93 (0.70-1.23)	-0.5, p=0.61
Medium HCV category: MMSE	**-0.38**	**0.17**	**0.68 (0.49-0.96)**	**-2.21, p=0.03**
Large HCV category: MMSE	-0.60	0.31	0.55 (0.30-1.00)	-1.95, 0.05
ADAS & HCV				
Medium HCV category	**-2.28**	**0.69**	**0.10 (0.03-0.39)**	**-3.33, p=0.0008**
Large HCV category	**-3.11**	**1.05**	**0.04 (0.01-0.35)**	**-2.96, p=0.003**
ADAS	0.03	0.03	1.03 (0.97-1.08)	0.94, p=0.35
Medium HCV category: ADAS	**0.10**	**0.03**	**1.10 (1.04-1.17)**	**3.14, p=0.003**
Large HCV category: ADAS	0.10	0.06	1.10 (0.99-1.23)	1.71, p=0.09
FAQ & HCV				
Medium HCV category	**-0.63**	**0.25**	**0.53 (0.33-0.87)**	**-2.53, p=0.01**
Large HCV category	**-2.08**	**0.44**	**0.13 (0.05-0.30)**	**-4.74, p<0.0001**
FAQ	0.06	0.03	1.06(1.00-1.13)	1.89, p=0.06
Medium HCV category: FAQ	0.03	0.04	1.03 (0.96-1.10)	0.81, p=0.42
Large HCV category: FAQ	0.08	0.06	1.09 (0.97-1.21)	1.47, p=0.14
RAVLT & HCV				
Medium HCV category	0.82	0.66	2.27 (0.62-8.28)	1.25, p=0.21
Large HCV category	1.11	1.42	3.02 (0.19-49.31)	0.78, p=0.44
RAVLT	**-0.04**	**0.02**	**0.96 (0.93-0.99)**	**-2.05, p=0.04**
Medium HCV category: RAVLT	**-0.04**	**0.02**	**0.96 (0.92-0.99)**	**-2.01, p=0.045**
Large HCV category: RAVLT	-0.09	0.05	0.92 (0.83-1.01)	-1.86, p=0.06

MMSE; mini mental state examination (standardized), ADAS; Alzheimer disease assessment scale (standardized), RAVLT; Rey auditory verbal learning test (immediate; standardized), FAQ; functional assessment questionnaire (standardized), HCV; hippocampal volume adjusted by age, field strength and intracranial volume (standardized).

Prediction of time to conversion

All the measures significantly predicted time to AD conversion in separate univariate analyses (likelihood ratio test between 33 to 90, df =1, $p<0.0001$). Each cognitive/functional predictor remained a significant predictor when HCV was added to the model, and HCV also remained a significant predictor. Additionally, HCV had additive effects with RAVLT, and FAQ, whereas HCV and MMSE, and HCV and ADAS had interactive effects (Table 2).

The analyses were repeated using categories of HCV (small, medium and large) in the models instead of HCV as a continuous variable (Table 3). The results revealed that MMSE was not a predictor of conversion in small HCV, and that having a medium to large HCV respectively associated with 45% and 81% lower risk of conversion from MCI to AD over time compared with small HCV. An additional 32% decrease in the risk of conversion was demonstrated for every one SD higher MMSE score in medium HCV but not in large HCV in comparison with small HCV. Similarly, ADAS was not predictive in small HCV and having a medium to large HCV was associated with 90% and 99.5% lower risk of conversion over time compared with small HCV. An additional 10% increase in the risk of conversion was demonstrated for every one SD higher ADAS score in medium HCV but not in large HCV in comparison with small HCV. In contrast, RAVLT was predictive in all HCV categories including small HCV, although an additional 4% decrease in the risk of conversion was detected for every one SD higher RAVLT in medium HCV.

Kaplan-Meier curves (Figure 3) revealed that the contribution of cognitive/functional measures in predicting the probability of remaining MCI over time was not constant at all HCV categories. For example, MMSE was not a determinant factor at small HCV, while it was at medium to large HCV.

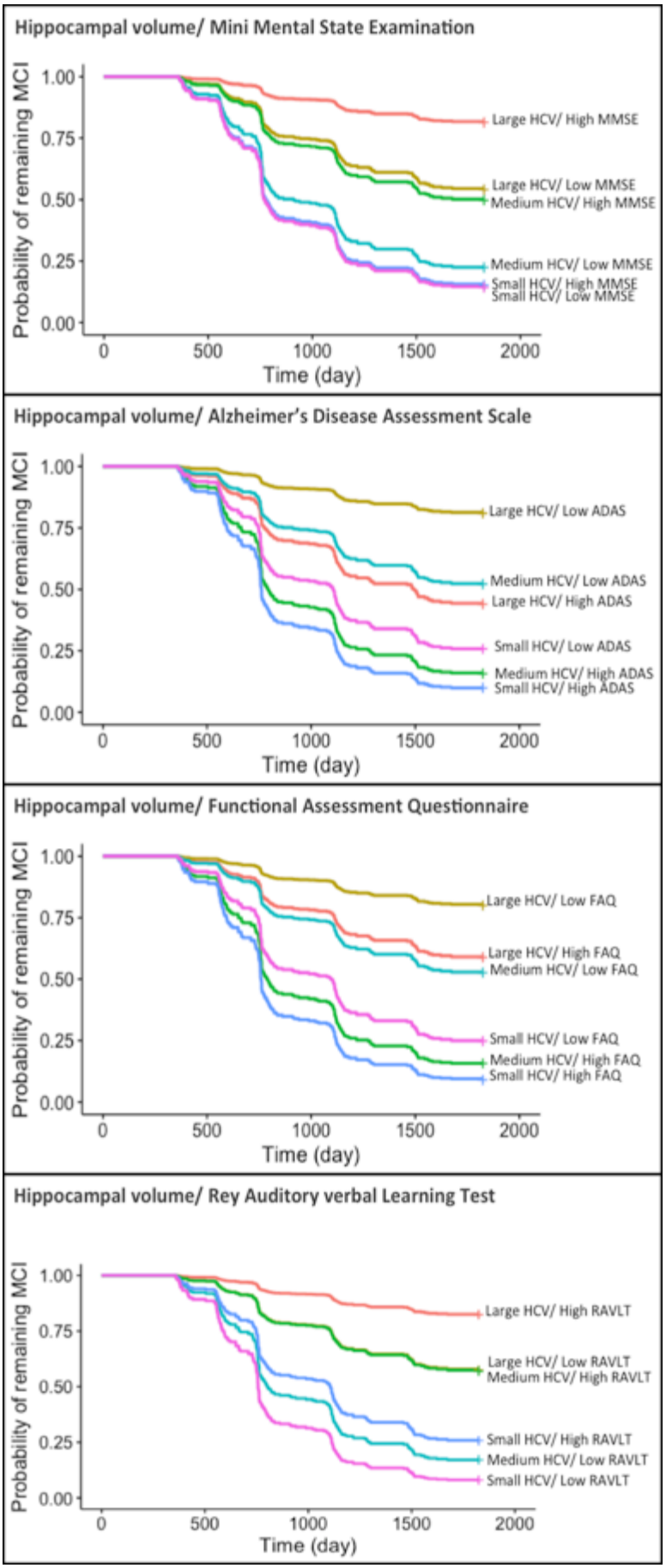


Figure 3: Kaplan-Meier plots for remaining stable over time: Illustrating the contribution of cognitive/functional measure and hippocampal volume on probability of remaining stable over time in MCI. Participants were categorized into six combinations based on three levels of HCV (small, medium and large) and two levels of cognitive/functional measures (low and high).

Discussion

This study aimed to investigate the predictive value of cognitive/functional measures in combination with hippocampal volume in order to predict conversion from MCI to AD within five years, as well as their capacity to predict time to conversion. The results demonstrated that the predictive value of cognitive/functional measures is dependent on HCV. The findings revealed that; (1) in predicting the conversion from MCI to AD, the predictive value of cognitive/functional measures was higher at lower HCV, while it was lower at higher HCV, and (2) in predicting the time to AD conversion, the cognitive/functional measures were somewhat more predictive when HCV was in the medium range (5500 mm^3 to 7500 mm^3) than at smaller or larger volumes, except for the immediate memory test that remained predictive across all HCV. The effect of HCV in predicting time to conversion was interactive with general cognitive measures (MMSE and ADAS) but additive with the functional assessment (FAQ) and immediate memory test (RAVLT).

These findings are important because they demonstrate that severity of cognitive impairment or subtle functional impairment and severity of neural pathology are both important in predicting probability of AD conversion. Although cognitive/functional performance is closely linked with neuropathology, the association is not straightforward. There is an imperfect overlap between cognitive deficit and pathology severity (Neuropathology Group of the Medical Research Council Cognitive Function and Aging Study, 2001). Individual variability in brain/cognitive reserve is the most likely explanation for this effect (Medaglia, Pasqualetti, Hamilton, Thompson-Schill, & Bassett, 2017; Steffener & Stern, 2012; Stern, 2009). Taking the severity of the pathology into account

when evaluating cognitive/functional performance is a practical way to take into account the moderating effect of brain/cognitive reserve.

There is accumulating evidence showing that individuals with larger brain/cognitive reserve may cope better with neural damage i.e. at a given level of observed pathology, cognitive impairment is lower in those with larger brain/cognitive reserve (Stern, 2009). Diversity in efficacy and capacity of neural networks as well as compensatory neural mechanisms such as using alternative neural networks may underlie this coping mechanism such that cognitive function may be maintained for some time in the context of increasing neurodegeneration. When brain/cognitive reserve is exhausted, further neurodegeneration cannot be compensated for and failure in cognitive processes clinically manifest as conversion from CN to MCI or MCI to AD (Steffener & Stern, 2012). Therefore, since individuals vary in their levels of brain/cognitive reserve, cognitive and functional performance alone is not a perfect predictor of decline. Cognitive reserve has been indirectly estimated in the literature by proxy variables including education, IQ, literacy, occupational complexity, participation in leisure activities and even personality variables (Steffener & Stern, 2012). However, the accurate measurement of brain/cognitive reserve is still the subject of ongoing research and much controversy (Steffener, Brickman, Rakitin, Gazes, & Stern, 2009; Steffener, Reuben, Rakitin, & Stern, 2011; Stern et al., 2008; Zarahn, Rakitin, Abela, Flynn, & Stern, 2007). Altogether, a practical way to deal with the concealing effect of cognitive reserve is to take into account the severity of neuropathology when evaluating cognitive/functional performance.

In addition to predicting the likelihood of converting from MCI to AD, the prediction of time to conversion is also of clinical significance but has proven difficult to achieve. Our results suggest that combining HCV and cognitive/functional measures is more effective in

predicting time to conversion. However, the effect of HCV differs for different cognitive/functional measures. It has an interactive effect with MMSE and ADAS but an additive effect with FAQ or RAVLT immediate. That is, the increase in the risk of AD conversion for each one-point decrease in the MMSE (or increase in ADAS) is not constant at different HCV values and is smaller at larger HCV. In contrast, the increase in the risk is constant for every one-point decrease on the RAVLT immediate (or higher FAQ) at any HCV values. This may be because HCV is more reflective of AD related pathology than MMSE and ADAS. As a consequence, at HCV less than 5500 mm^3, one unit difference in MMSE (or ADAS) is less influential than at larger HCV. This may explain the fact that MMSE and ADAS are not predictive of time to conversion at HCV less than 5500 mm^3 and at more than 7500 mm^3, but predictive in the mid-range of HCV (5500-7500 mm^3).

In contrast, the combined evaluation of performance in a specific domain (such as immediate memory) and the brain structure underpinning that performance (HCV) may provide a more precise evaluation of the degree of neurodegeneration and the level of brain/cognitive reserve exhaustion. This may explain our findings that RAVLT immediate and HCV are more sensitive predictors of time to conversion. It may also explain the lack of interactive effect between these two measures.

It is important to note that because MMSE, ADAS and FAQ evaluate performance across a larger number of neural networks, they may reflect the development of AD pathology across any of those networks and thus also predict the risk of AD conversion. However, because only part of their variability is related to hippocampal function, they do not appear to be as predictive of time to conversion than RAVLT immediate.

Many studies conducted to date have focused on combining MRI and cognitive/functional measures for improved diagnosis or prediction of AD conversion. Our

study, in contrast, investigated the nature of the interaction between MRI and cognitive/functional measures in predicting AD conversion and time to conversion. Understanding the relationship between structural and cognitive/functional measures not only emphasizes the benefit of combining these measures for diagnostic/ prognostic purpose, it may also help better conceptualize the impact of brain/cognitive reserve on clinical/MRI measures.

In conclusion, Alzheimer's disease is pathologically characterised by degenerative processes, the severity of which can be measured with neuroimaging techniques. The functional consequence of the degeneration can be concurrently assessed with cognitive/functional tools. A combination of both neuroimaging and cognitive/functional indexes are superior in predicting disease progression than either alone. However, the present findings indicate that the relative contribution of neuroimaging and cognitive/functional measures is not constant in predicting progression from MCI to AD. Cognitive/functional measures are predictors of conversion but their predictive values are not constant at all levels of hippocampal volumes. Additionally, the most effective combination of measures to predict time to conversion is likely to involve those that assess hippocampal volume in conjunction with one of its main functions, immediate memory.

Acknowledgement

Data collection and sharing for this work was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI; National Institutes of Health Grant U01 AG024904). 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 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.; Neu- roRx 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 co- ordinated 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 partly undertaken on the National Computational Infrastructure (NCI) facility in Canberra, Australia, which is supported by the Australian Commonwealth Government and was supported by an Australian Government Research Training Program (RTP) Scholarship.

References:

Arevalo-Rodriguez, I., Smailagic, N., Roque, I. F. M., Ciapponi, A., Sanchez-Perez, E., Giannakou, A., . . . Cullum, S. (2015). Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database Syst Rev*(3), CD010783. doi:10.1002/14651858.CD010783.pub2

Belleville, S., Fouquet, C., Hudon, C., Zomahoun, H. T. V., Croteau, J., & Consortium for the Early Identification of Alzheimer's, d.-Q. (2017). Neuropsychological Measures that Predict Progression from Mild Cognitive Impairment to Alzheimer's type dementia in Older Adults: a Systematic Review and Meta-Analysis. *Neuropsychol Rev, 27*(4), 328-353. doi:10.1007/s11065-017-9361-5

Devanand, D. P., Liu, X., Tabert, M. H., Pradhaban, G., Cuasay, K., Bell, K., . . . Pelton, G. H. (2008). Combining early markers strongly predicts conversion from mild cognitive impairment to Alzheimer's disease. *Biol Psychiatry, 64*(10), 871-879. doi:10.1016/j.biopsych.2008.06.020

Eckerstrom, C., Olsson, E., Borga, M., Ekholm, S., Ribbelin, S., Rolstad, S., . . . Malmgren, H. (2008). Small baseline volume of left hippocampus is associated with subsequent conversion of MCI into dementia: the Goteborg MCI study. *J Neurol Sci, 272*(1-2), 48-59. doi:10.1016/j.jns.2008.04.024

Estevez-Gonzalez, A., Kulisevsky, J., Boltes, A., Otermin, P., & Garcia-Sanchez, C. (2003). Rey verbal learning test is a useful tool for differential diagnosis in the preclinical phase of Alzheimer's disease: comparison with mild cognitive impairment and normal aging. *Int J Geriatr Psychiatry, 18*(11), 1021-1028. doi:10.1002/gps.1010

Falahati, F., Westman, E., & Simmons, A. (2014). Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging. *J Alzheimers Dis, 41*(3), 685-708. doi:10.3233/JAD-131928

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res, 12*(3), 189-198. doi:10.1016/0022-3956(75)90026-6

Gomar, J. J., Bobes-Bascaran, M. T., Conejero-Goldberg, C., Davies, P., Goldberg, T. E., & Neuroimaging, A. s. D. (2011). Utility of Combinations of Biomarkers, Cognitive Markers, and Risk Factors to Predict Conversion From Mild Cognitive Impairment to Alzheimer Disease in Patients in the Alzheimer's Disease Neuroimaging Initiative. *Archives of General Psychiatry, 68*(9), 961-969. doi:DOI 10.1001/archgenpsychiatry.2011.96

Ito, K., Hutmacher, M. M., & Corrigan, B. W. (2012). Modeling of Functional Assessment Questionnaire (FAQ) as continuous bounded data from the ADNI database. *Journal of Pharmacokinetics and Pharmacodynamics, 39*(6), 601-618. doi:10.1007/s10928-012-9271-3

Jack, C. R., Jr., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., . . . Weiner, M. W. (2008). The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging, 27*(4), 685-691. doi:10.1002/jmri.21049

Jack, C. R., Jr., Shiung, M. M., Weigand, S. D., O'Brien, P. C., Gunter, J. L., Boeve, B. F., . . . Petersen, R. C. (2005). Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic MCI. *Neurology, 65*(8), 1227-1231. doi:10.1212/01.wnl.0000180958.22678.91

Li, K., Chan, W., Doody, R. S., Quinn, J., Luo, S., & Alzheimer's Disease Neuroimaging, I. (2017). Prediction of Conversion to Alzheimer's Disease with Longitudinal Measures and Time-To-Event Data. *J Alzheimers Dis, 58*(2), 361-371. doi:10.3233/JAD-161201

Medaglia, J. D., Pasqualetti, F., Hamilton, R. H., Thompson-Schill, S. L., & Bassett, D. S. (2017). Brain and cognitive reserve: Translation via network control theory. *Neurosci Biobehav Rev, 75*, 53-64. doi:10.1016/j.neubiorev.2017.01.016

Moradi, E., Pepe, A., Gaser, C., Huttunen, H., Tohka, J., & Alzheimer's Disease Neuroimaging, I. (2015). Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *Neuroimage, 104*, 398-412. doi:10.1016/j.neuroimage.2014.10.002

Neuropathology Group of the Medical Research Council Cognitive Function and Aging Study. (2001). Pathological correlates of late-onset dementia in a multicentre, community-based population in England and Wales. Neuropathology Group of the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). *Lancet, 357*(9251), 169-175. doi:10.1016/s0140-6736(00)03589-3

Pandya, S. Y., Clem, M. A., Silva, L. M., & Woon, F. L. (2016). Does mild cognitive impairment always lead to dementia? A review. *J Neurol Sci, 369*, 57-62. doi:10.1016/j.jns.2016.07.055

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol, 56*(3), 303-308. doi:10.1001/archneur.56.3.303

Petersen, R. C., Thomas, R. G., Grundman, M., Bennett, D., Doody, R., Ferris, S., . . . Group, t. A. s. D. C. S. (2005). Vitamin E and Donepezil for the Treatment of Mild Cognitive Impairment. *the New England Journal of Medicine, 352*(23), 2379 -2388.

Pfeffer, R. I., Kurosaki, T. T., Harrah, C. H., Jr., Chance, J. M., & Filos, S. (1982). Measurement of functional activities in older adults in the community. *J Gerontol, 37*(3), 323-329. doi:10.1093/geronj/37.3.323

Pintzka, C. W., Hansen, T. I., Evensmoen, H. R., & Haberg, A. K. (2015). Marked effects of intracranial volume correction methods on sex differences in neuroanatomical structures: a HUNT MRI study. *Front Neurosci, 9*, 238. doi:10.3389/fnins.2015.00238

Rey, A. (1941). L'examen psychologique dans les cas d'encéphalopathie traumatique. (Les problems.) [The psychological examination in cases of traumatic encepholopathy. Problems]. *Archives de Psychologie, 28*, 215-285.

Rey, A. (1964). *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.

Steffener, J., Brickman, A. M., Rakitin, B. C., Gazes, Y., & Stern, Y. (2009). The impact of age-related changes on working memory functional activity. *Brain Imaging Behav, 3*(2), 142-153. doi:10.1007/s11682-008-9056-x

Steffener, J., Reuben, A., Rakitin, B. C., & Stern, Y. (2011). Supporting performance in the face of age-related neural changes: testing mechanistic roles of cognitive reserve. *Brain Imaging Behav, 5*(3), 212-221. doi:10.1007/s11682-011-9125-4

Steffener, J., & Stern, Y. (2012). Exploring the neural basis of cognitive reserve in aging. *Biochim Biophys Acta, 1822*(3), 467-473. doi:10.1016/j.bbadis.2011.09.012

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia, 47*(10), 2015-2028. doi:10.1016/j.neuropsychologia.2009.03.004

Stern, Y., Zarahn, E., Habeck, C., Holtzer, R., Rakitin, B. C., Kumar, A., . . . Brown, T. (2008). A Common Neural Network for Cognitive Reserve In Verbal and Object Working Memory In Young But Not Old. *Cereb Cortex, 18*(4), 959-967.

Tabatabaei-Jafari, H., Shaw, M. E., & Cherbuin, N. (2015). Cerebral atrophy in mild cognitive impairment: A systematic review with meta-analysis. *Alzheimer's &*

Dementia: Diagnosis, Assessment & Disease Monitoring, 1(4), 487-504. doi:10.1016/j.dadm.2015.11.002

Tabatabaei-Jafari, H., Walsh, E., Shaw, M. E., & Cherbuin, N. (2018). A simple and clinically relevant combination of neuroimaging and functional indexes for the identification of those at highest risk of Alzheimer's disease. *Neurobiol Aging, 69*, 102-110. doi:10.1016/j.neurobiolaging.2018.05.006

Zarahn, E., Rakitin, B., Abela, D., Flynn, J., & Stern, Y. (2007). Age-related changes in brain activation during a delayed item recognition task. *Neurobiol Aging, 28*(5), 784-798. doi:10.1016/j.neurobiolaging.2006.03.002

STUDY 5

REGIONAL BRAIN ATROPHY PREDICTS TIME TO CONVERSION TO ALZHEIMER'S DISEASE, DEPENDENT ON BASELINE VOLUME

Hossein Tabatabaei-Jafari[1], Marnie E. Shaw[2], Erin Walsh[1], Nicolas Cherbuin[1], For the Alzheimer's Disease Neuroimaging Initiative (ADNI)

[1] Centre for Research on Ageing, Health and Wellbeing, The Australian National University, Canberra, Australia

[2]College of Engineering and Computer Science, The Australian National University, Canberra, Australia

Published in Neurobiology of Aging, 83 (2019) 86 - 94

DOI: 10.1016/j.neurobiolaging.2019.08.033

Abstract

A key question for the design of clinical trials for Alzheimer's disease (AD) is whether the timing of conversion from mild cognitive impairment (MCI) to AD can be predicted. This is also an important question for the clinical management of MCI. This study aims to address this question by exploring the contribution of baseline brain volume and annual volume change, using Cox regression, in predicting the time to conversion. Individuals with MCI, who converted to AD (n=198), reverted to normal (n=38), or remained stable (n=96) for at least five years were included in this study. The results revealed that the volumes of all the brain regions considered were predictive of the time to conversion from MCI to AD. Annual change in volume was also predictive of the time to conversion but only when initial volumes were above a certain threshold. This is important because it suggests that reduction in atrophy rate, which is the outcome of some clinical trials, is not inevitably associated with delay in conversion from MCI to AD.

Introduction

Progressive neurodegeneration is a hallmark of Alzheimer's disease (AD). However, it is also prevalent in normal ageing (Fjell et al., 2014). One major difference is that the rate of degeneration in the pathological progression leading to AD is substantially higher than in normal ageing. A meta-analysis of longitudinal studies conducted in the last two decades revealed that the shrinkage rate in the prodromal stage of AD --mild cognitive impairment (MCI)-- is at least twice that observed in normal ageing (Hossein Tabatabaei-Jafari, Shaw, & Cherbuin, 2015). This is seen in the whole brain and even more so in brain regions typically more affected in the first stage of the disease, such as the hippocampus and entorhinal cortex. Moreover, degeneration begins decades before the disorder emerges clinically, sometimes even in early adulthood (Braak & Braak, 1997). These findings underpin the hope that early intervention aimed at decreasing brain shrinkage may stop, or at least slow down, further progression to clinical AD.

Several intervention trials, using nutrient supplements or medication, have been effective in reducing the atrophy rate in total or regional brain volumes in those with MCI (Douaud et al., 2013; Dubois et al., 2015; Kile et al., 2017; Kobe et al., 2016; Prins et al., 2014; Zhang, Miao, Li, Wu, & Ma, 2017). However, whether these changes can modify the course of AD progression and delay the time to conversion remains an unresolved question. To address such questions, it is necessary to better understand the contribution brain atrophy makes to the course of the disease, and particularly to the progression from MCI to AD.

In contrast to studies predicting conversion from MCI to AD, studies that have investigated the time to conversion are limited in number. They generally suggest an association between the pace of neurodegeneration and the time to AD conversion (Falahati et al., 2017; Jack et al., 2005; Liu, Chen, Yao, & Guo, 2017; Teipel, Kurth, Krause, Grothe,

& Alzheimer's Disease Neuroimaging, 2015). Most attempts have used spatial patterns of longitudinal volume loss (using machine learning) to successfully predict the time to conversion (Gavidia-Bovadilla, Kanaan-Izquierdo, Mataro-Serrat, Perera-Lluna, & Alzheimer's Disease Neuroimaging, 2017; Li et al., 2012; Risacher et al., 2010; Teipel et al., 2015; Thung et al., 2018). Falahati el al. developed a "severity index", based on degeneration in 34 measures of regional cortical thickness and 21 regional subcortical volumes and showed that it was predictive of the time to AD conversion for up to 3 years follow-up. The index showed 95% correct prediction of conversion within the first year and 80% over 3 years (Falahati et al., 2017). Global volume change such as whole brain atrophy and ventricular enlargement, but not regional brain atrophy rates (hippocampal and entorhinal cortex), has also been shown to be predictive of AD conversion but only for a short follow-up and in the context of a relatively small study (Jack et al., 2005). Although these limited numbers of studies are conceptually supportive of the idea that faster degeneration will lead to earlier conversion, the findings are based on a short-term follow-up and the approaches are complex and methodologically difficult to implement at individual level that is a requirement for clinical trials and clinical practice. Simple measures such as regional brain volume and regional atrophy rate investigated in a longer follow up may be more practical for individual evaluation, especially in a clinical setting or for clinical trials.

Therefore, strong evidence supporting the use of atrophy rate in the prediction of time to conversion from MCI to AD is still lacking. In addition, it is necessary to clarify the extent to which the predictive value of atrophy rate depends on baseline volume. This is needed because the clinical impact of any future degeneration is likely to be highly dependent on prior atrophy and/or brain reserve indexed by the current volume of a region of interest. To address these questions, the present study aimed to investigate the value of global as well as

regional baseline volume and atrophy rate and their interaction over long-term follow-ups in predicting conversion from MCI to AD.

Methodology

Study Participants

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 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 MCI and early AD.

All participants of ADNI 1/GO/2, who were diagnosed with MCI at baseline, remained stable for at least six months, and underwent MRI scanning more than twice, were considered for inclusion. Individuals with MCI who converted to AD (MCIc, n=198), reverted to cognitively normal (CN; MCIr, n=38), or remained stable for more than five years (MCIs, n=96) were included in this study.

Details of the diagnostic criteria can be found at the ADNI web site (http://www.adni-info.org/Scientists/AboutADNI.aspx). Briefly, participants were classified as MCI if they had an Mini-Mental State Examination (MMSE) score greater than 24, a CDR of 0.5, a report of subjective memory concern, an objective memory loss, preserved daily living activity and did not meet diagnostic criteria for dementia. Participants were classified as AD if they had a MMSE score less than 26, CDR 0.5 or above, and fulfilled criteria for clinically probable AD according to the Institute of Neurological and Communicative Diseases and

Stroke/ Alzheimer's Disease and Related Disorders Association. It is also important to note that a Geriatric Depression Scale score of less than 6 was a requirement for participation in the ADNI study (Petersen et al., 2010), so all participants had a GDS score of normal range.

Neuroimaging acquisition and processing

Participants underwent high-resolution MRI brain scans on 1.5 (N=889) or 3 T (N=872) scanners from General Electric, Siemens, or Philips (Milwaukee, WI, USA; Germany; the Netherlands respectively) using a standardized ADNI acquisition protocol for 3D MP-RAGE sequence (Jack et al., 2008). Images which had undergone specific ADNI preprocessing correction steps to standardize images from different sites and platforms, were obtained for this study: (1) Grad wrap; a specific correction of image geometry distortion due to non-linearity, (2) B1 non-uniformity; B1 calibration to correct the image intensity non-uniformity that results when RF transmission is performed with a more uniform body coil while reception is performed with a less uniform head coil, (3) N3 correction; a histogram peak sharpening algorithm applied after grad wrap and B1 correction. For MCI participants, only images acquired before conversion to AD or reversion to CN were included. The MRI scans of individual participants were acquired on the same scanner with the same parameters throughout the follow-up.

All scans were segmented with FreeSurfer version 5.3 (http://surfer.nmr.mgh.harvard.edu/), processed with the longitudinal pipeline. For each participant, all scans were initially processed by the default workflow. Then an unbiased template (an average template) was created from all time points. The unbiased template was used as a base for registering all the time point scans to reduce the random within-subject variation in the processing procedure of the longitudinal analysis. Finally, all time points were longitudinally processed.

The output-segmented images were visually checked. The criterion was a clear segmentation error as assessed by an experienced neuroscientist. Scans with segmentation errors were re-processed and would only be excluded if the error could not be corrected. Six scans with error were not correctable and excluded from the study.

Measurements

Four brain volumes were considered as regions of interest (ROI) in this study: (1) total whole brain volume (sum of the total gray and white matter), (2) total ventricular volume (sum of the lateral, third and fourth ventricular volumes), (3) total hippocampal volume (sum of the left and right hippocampus), and (4) total entorhinal cortex volume (sum of the left and right entorhinal cortex). Baseline volume and annual change rate (atrophy rate for the whole brain, hippocampus and entorhinal cortex and enlargement rate for the ventricles) of each ROI were investigated as the measures of interest for predicting time to conversion from MCI to AD.

The annual change rate for each ROI was computed by the least square linear regression method for each individual separately: brain volume (at each time point) was used as the dependent variable, with age at each time point (centred at 55, the minimum age at baseline) as the independent variable. The regression coefficient for age was considered as the volume change for each year increase in age in mm^3. The regression coefficient was used to compute the annual change rate in percent using the formula

$$100 \times (\text{the regression coefficient for age} / \text{baseline volume})$$

Because the results from our previous study revealed that baseline scores on the MMSE, the Alzheimer's Disease Assessment Scale (ADAS cognitive version), the Functional Assessment Questionnaire (FAQ), and the Rey Auditory Learning Test (RAVLT;

immediate memory subtype) were predictive of time to conversion from MCI to AD when also taking into account hippocampal volume (H. Tabatabaei-Jafari, Shaw, Walsh, Cherbuin, & Alzheimer's Disease Neuroimaging, 2019), the annual change rates of these measures were also evaluated to better characterise the participants.

While CSF level of amyloid β 1-42 and total and phosphorylated tau were only available for a sub-sample of participants (236 for amyloid β, 232 for total tau, and 236 for phosphorylated tau out of the 332 participants) they could not be included in analyses but are reported to better characterise the sample investigated.

Statistical analysis

Statistical analyses were performed using the R statistical software (version 3.3.2). Data were checked for missing values and for univariate and multivariate outliers using Mahalanobis distance. There were no missing values or outliers. Group differences in demographic variables were assessed by t-test for continuous variables and chi square tests for categorical variables. The alpha level was set at < 0.05.

Cox regression analysis (package survival; version 2.40-1) using time-to-event as time metric was used to investigate the predictive value of brain ROIs for time to conversion from MCI to AD. The event in the model was specified as happened if the individual converted to AD, thus MCIc were coded as 1 and MCIs and MCIr were coded as 0 in the model. For MCIs the time-to-event was the time from baseline to last scan while in MCIc and MCIr it was the time from baseline to diagnosis change (Change to AD for MCIc and change to CN for MCIr). One-sided Wald tests were used to test associations because only increase in the risk of conversion to AD was predicted. Baseline volume and annual change rate were considered as predictors of time to conversion and were standardised to reduce the variance inflation factor in the model. Baseline volumes were adjusted for age, sex, field strength and

intracranial volume using the residual method before adopted in the models (Pintzka, Hansen, Evensmoen, & Haberg, 2015).

Univariate models were used to investigate the association between brain measures and time to conversion. Four separate bivariate models, each consisting of standardized baseline volume, standardized annual change rate, and their interaction, were conducted for the whole brain, ventricles, hippocampus, and entorhinal cortex. Hazard ratios with 95% confidence intervals for a 1-SD difference in baseline volume and 1-SD change in annual change rate were used to quantify the magnitude of the effect. In the case of significant interaction between baseline volume and annual change rate (as continuous variables), to better conceptualise the interaction, participants were categorized into three groups based on their baseline volume (for each brain region separately) and the bivariate analyses were repeated with categorical baseline volume in the model. Categorization was based on the standard deviation (SD, round values): (1) small category: smaller than one SD below the mean, (2) medium category: one SD below and above the mean, and (3) large category: larger than one-SD above the mean. In addition, to better visualize the contribution of baseline volume and annual change rate in predicting conversion from MCI to AD, the density of those converted to AD over time was plotted across different stratified annual change rate for each baseline volume category separately.

Results

Participants' characteristics

Three hundred and thirty two participants (59% male), who were followed up for up to ten years (5.35 ±2.31 year), were categorized into MCIr, MCIs, and MCIc (Table 1). Individuals with MCIc were about three years older than other MCI. There was no

significant difference in education across the groups, but the proportion of males was somewhat lower in MCIr (47.37%) than in MCIs (60.42%) and MCIc (60.42 %). The proportion of individuals carrying the APOE e4 allele was significantly larger in MCIc than others, and more so for those with two e4 alleles (Table 1 & 2).

Baseline brain volumes and annual changes

For all ROIs baseline volumes and annual change rates were different between MCIc and other MCI types. Differences were most pronounced in the hippocampus, entorhinal cortex, and ventricles and followed the direction MCIr > MCIs > MCIc for volumes, and MCIr < MCIs < MCIc for change rates (Table 1 & 2).

Despite significant group differences, the distribution of the brain measures revealed a large overlap across the groups (Figure 1). When considered across the whole sample there was no significant correlation between baseline volume and annual change rate for the whole brain, and the ventricles. A moderate correlation was detected for the hippocampus (r=0.27), and a smaller correlation for the entorhinal cortex (r=0.12). However, when computed separately in each group, associations between baseline volume and annual change rate were only significant in MCIs for the hippocampus and entorhinal cortex as well as for the ventricles in MCIc (r=-0.19) (Table 3).

Table 1: Participants characteristic and measurements

	MCIr	MCIs	MCIc	Significant pair difference
Sample size	38	96	198	-
Age; year, Mean (SD)	69.30 (8.23)	71.65 (7.48)	74.25 (7.16)	MCIc vs. MCIr & MCIs
Age range, year	55 - 87	57 - 88	55 – 89	-
Male sex; N (%)	18 (47.37)	58 (60.42)	121 (61.42)	No difference
Education; year, Mean (SD)	16.68 (2.52)	15.88 (3.04)	16.01 (2.78)	No difference

APOE e4; N (%)				
One allele	17 (45)	22 (23)	102 (51.51)	MCIc vs. MCIr & MCIs MCIr vs. MCIs
Two alleles	1 (3)	6 (6)	32 (16.33)	
Number of scan points	11.4 (2.7)	7.7 (2.5)	5.9 (1.8)	MCIc vs. MCIr & MCIs MCIr vs. MCIs
Follow-up, Range; day	1082 - 3662	1850 - 3927	343 - 3690	-
Follow-up, Mean (SD)	1704 (676)	2381 (686)	1790 (869)	-
Time to Diagnosis change, Range; day	184 - 1583	-	357 - 3714	-
Time to DX change, Mean (SD)	762 (411)	-	1041 (603)	-
Brain Measures				
Whole Brain				
Baseline, mm^3	1081597 (35162)	1095415 (38165)	1070628 (42231)	MCIc vs. MCIs MCIr vs. MCIs
Annual change rate, %/y	-0.15 (1.20)	-0.55 (0.37)	-0.73 (1.26)	MCIc vs. MCIr MCIr vs. MCIs
Ventricles				
Baseline, mm^3	37782 (14261)	38808 (15115)	44744 (17982)	MCIc vs. MCIr & MCIs
Annual change rate, %/y	2.42 (3.94)	3.93 (2.66)	7.62 (5.74)	MCIc vs. MCIr & MCIs MCIr vs. MCIs
Hippocampus				
Baseline, mm^3	7229 (794)	7035 (953)	6127 (912)	MCIc vs. MCIr & MCIs
Annual change rate, %/y	0.13 (3.34)	-1.29 (1.10)	-3.12 (2.86)	MCIc vs. MCIr & MCIs MCIr vs. MCIs
Entorhinal cortex				
Baseline, mm^3	3787 (693)	3645 (644)	3224 (698)	MCIc vs. MCIr & MCIs
Annual change rate, %/y	-0.11 (5.05)	-1.75 (1.56)	-3.62 (5.95)	MCIc vs. MCIr & MCIs
Cognitive/functional measures				
MMSE				

Baseline	28.53 (1.50)	28.22 (1.42)	27.09 (1.78)	MCIc vs. MCIr & MCIs
Annual change, u/y	0.64 (1.94)	-0.15 (0.29)	-0.93 (1.94)	MCIc vs. MCIr & MCIs MCIr vs. MCIs
ADAS cog				
Baseline	10.66 (4.24)	12.08 (4.63)	19.94 (5.81)	MCIc vs. MCIr & MCIs
Annual change, u/y	-1.97 (4.31)	0.24 (0.57)	1.50 (3.98)	MCIc vs. MCIr & MCIs MCIr vs. MCIs
RAVLT immediate				
Baseline	43.55 (10.21)	39.70 (10.96)	29.64 (7.98)	MCIc vs. MCIr & MCIs
Annual change, u/y	-1.44 (7.88)	-0.46 (1.44)	-2.06 (6.14)	MCIc vs. MCIs
FAQ				
Baseline	0.87 (1.73)	1.60 (3.07)	4.61 (4.54)	MCIc vs. MCIr & MCIs
Annual change, u/y	-0.16 (2.35)	0.22 (0.69)	1.61 (3.55)	MCIc vs. MCIr & MCIs
CSF measures (baseline)				
Amyloid β level, pg/ml	211.54 (50.51)	198.23 (48.10)	143.19 (42.85)	MCIc vs. MCIr & MCIs
TAU	60.24 (29.31)	75.21 (49.97)	115.45 (55.86)	MCIc vs. MCIr & MCIs
P-TAU	27.69 (14.45)	32.21 (49.97)	49.61 (26.13)	MCIc vs. MCIr & MCIs

MCIc= mild cognitive impairment converted to Alzheimer's disease; MCIs= mild cognitive impairment remained stable for more than five years; MCIr= mild cognitive impairment reverted to cognitively. APOE e4; Apolipoprotein E allele 4;MMSE = mini-mental state examination; ADAS cog= Alzheimer Disease Assessment Scale (cognitive subscale); RAVLT = Rey Auditory Verbal Learning Test; FAQ = functional assessment questionnaire. CSF= cerebrospinal fluid; Aβ= amyloid-beta 1–42; TAU=total tau protein, P-TAU=phosphorylated tau protein. Baseline measures adjusted for age, sex, field strength and intracranial volume

Table 2: Statistics of participants' characteristics (pair comparison)

	MCIc vs. MCIs	MCIc vs. MCIr	MCIs vs. MCIr
Age	t(df=180.93)= 2.83, p<0.01	t(df=48.34)=3.46, p<0.01	t(df=62.56)=-1.53, p=0.13
Sex	X^2(df=1)=0, p=1	X^2(df=1)=1.95, p=0.16	X^2(df=1)=1.39, p=0.24
APOEe4	X^2(df=2)=39.05, p<0.0001	X^2(df=2)=8.29, p<0.05	X^2(df=2)=6.51, p<0.05
Education	t(df=173.95)=0.35, p=0.72	t(df=55.83)=-1.50, p=0.14	t(df=81.55)=1.58, p=0.12
Number of scan point	t(df=146.82)=-6.16, p<0.0001	t(df=43.65)=-11.94, p<0.0001	t(df=62.75)=7.35, p<0.0001
Brain Measures			
Whole Brain			
Baseline, mm^3	t(df=206.23)=-5.04, p<0.0001	t(df=59.47)=-1.70, p=0.09	t(df=73.34)=-2.00, p<0.05
Annual change rate, %/y	t(df=255.60)=-1.83, p=0.07	t(df=54.07)=-2.71, p<0.01	t(df=39.79)=-2.03, p<0.05
Ventricles			
Baseline, mm^3	t(df=220.14)=2.96, p<0.01	t(df=61.95)=2.63, p<0.05	t(df=71.70)=-0.37, p=0.71
Annual change rate, %/y	t(df=291.55)=7.54, p<0.0001	t(df=71.11)=6.87, p<0.0001	t(df=50.93)=-2.18, p<0.05
Hippocampus			
Baseline, mm^3	t(df=180.89)=-7.77, p<0.0001	t(df=57.39)=-7.64, p<0.0001	t(df=80.93)=1.20, p=0.23
Annual change rate, %/y	t(df=280.98)=-7.89, p<0.0001	t(df=47.99)=-5.62, p<0.0001	t(df=40.20)=2.56, p<0.05
Entorhinal cortex			
Baseline, mm^3	t(df=202.35)=-5.12, p<0.0001	t(df=52.45)=-4.58, p<0.0001	t(df=63.75)=1.09, p=0.28
Annual change rate, %/y	t(df=246.61)=-4.15, p<0.0001	t(df=54.98)=-3.56, p<0.001	t(df=39.38)=1.81, p=0.08

Cognitive/functional measures			
MMSE			
Baseline	t(df=230.58)=-5.91, p<0.0001	t(df=58.73)=-5.25, p<0.0001	t(df=64.49)=1.09, p=0.28
Annual change, u/y	t(df=211.03)=-5.40, p<0.0001	t(df=40.91)=-4.19, p<0.001	t(df=31.40)=2.35, p<0.05
ADAS cog			
Baseline	t(df=230.87)=12.50, p<0.0001	t(df=64.46)=11.44, p<0.0001	t(df=70.99)=-1.69, p=0.10
Annual change, u/y	t(df=223.32)=3.35, p<0.001	t(df=35.97)=3.75, p<0.001	t(df=30.48)=-2.84, p<0.01
RAVLT immediate			
Baseline	t(df=145.52)=-8.03, p<0.0001	t(df=46.06)=-7.95, p<0.0001	t(df=72.53)=1.93, p=0.06
Annual change, u/y	t(df=218.50)=-3.28, p<0.01	t(df=40.94)=-0.84, p=0.41	t(df=31.59)=-0.41, p=0.68
FAQ			
Baseline	t(df=261.49)=6.67, p<0.0001	t(df=150.96)=8.73, p<0.0001	t(df=116.37)=-1.75, p=0.08
Annual change, u/y	t(df=223.71)=5.28, p<0.0001	t(df=57.02)=3.65, p<0.001	t(df=32.80)=-0.92, p=0.37
CSF measures (baseline)			
Amyloid β level, pg/ml	t(df=129.55)=-8.08, p<0.0001	t(df=42.37)=-7.07, p<0.0001	t(df=57.29)=1.26, p=0.21
TAU	t(df=158.41)=5.22, p<0.0001	t(df=93.64)=7.73, p<0.0001	t(df=94)=-1.09, p=0.06
P-TAU	t(df=184.35)=5.47, p<0.0001	t(df=86.41)=6.42, p<0.0001	t(df=76.81)=-1.33, p=0.19

MCIc= mild cognitive impairment converted to Alzheimer's disease; MCIs= mild cognitive impairment remained stable for more than five years; MCIr= mild cognitive impairment reverted to cognitively. APOE e4; Apolipoprotein E allele 4;MMSE = mini-mental state examination; ADAS cog= Alzheimer Disease Assessment Scale (cognitive subscale); RAVLT = Rey Auditory Verbal Learning Test; FAQ = functional assessment questionnaire. CSF= cerebrospinal fluid; Aβ= amyloid-beta 1–42; TAU=total tau protein, P-TAU=phosphorylated tau protein. Baseline measures adjusted for age, sex, field strength and intracranial volume

Table 3: Correlation between baseline brain volume and annual change rate in the whole sample and across the groups

	All groups	MCIr	MCIs	MCIc
Whole brain	r=-0.03, p=0.54	r=0.23, p=0.17	r=-0.04, p=0.69	r=-0.11, p=0.13
Ventricles	r=-0.06, p=0.31	r=0.08, p=0.62	r=-0.01, p=0.99	r=-0.19, p<0.01
Hippocampus	r=0.27, p<0.0001	r=-0.09, p0.57	r=0.24, p<0.05	r=0.12, p=0.09
Entorhinal cortex	r=0.12, p<0.05	r=-0.11, p=0.50	r=0.32, p<0.01	r=0.05, p=0.53

Cognitive/functional Measures

Similar to brain measures, cognitive/functional measures were significantly different between MCIc and other MCI types. Differences were most pronounced in baseline volumes following the order MCIr <MCIs <MCIc. While, annual changes were significantly different between MCIc and other MCI types, differences between MCIr and MCIs did not followed a constant pattern (Table 1 & 2).

CSF measures

The pattern in CSF differences was consistent (in the sub-sample that data were available) across the groups. Amyloid β was significantly lower in MCIc than MCIs and MCIr, and total tau and phosphorylated tau were significantly greater in MCIc than MCIs and MCIr. These measures were not different between MCIr and MCIs (Table 1 & 2).

Prediction of time to AD conversion

Baseline volume and annual change rate for each brain region significantly predicted time to AD conversion ($Z > 5$, $p<0.01$) when they were evaluated separately (univariate

model). When baseline volume and annual change rate were tested in the same model (bivariate model) both measures remained significantly predictive in all ROIs. In addition, an interaction between annual change rate and baseline volume was detected. It means, in addition to a constant increase in the risk for each 1-SD decrease in ROIs' baseline volume (1-SD increase in ventricular volume) and 1-SD increase in annual volume loss there was an additional risk for each measure, which was dependent on the other measure (Table 4). To better conceptualise this interactive effect between the two measures, analyses were repeated with a categorical baseline volume (small, medium and large) and annual change rate in percent in the model (Table 5). Following are brief reports for each ROI separately.

Whole brain

Atrophy rate did not predict time to conversion in whole brain baseline volumes less than 1,040,000 mm^3, whereas it had significant predictive value at higher volumes. Medium to large whole brain volumes were associated with 61% and 72% lower risk of conversion from MCI to AD compared to small volumes. An additional 35% and 43% decrease in the risk of conversion were demonstrated for every one percent lower atrophy rate in medium and large volumes.

Ventricles

Enlargement rate did not predict time to conversion in ventricular baseline volumes larger than 55,000 mm^3, whereas it had significant predictive value at small volumes (lower than 28,000 mm^3). Medium to small volumes were respectively associated with 48% and 83% lower risk of conversion from MCI to AD compared to large volumes. An additional 14% increase in the risk of conversion was demonstrated for one percent greater enlargement rate in small volumes.

Hippocampus

Atrophy rate did not predict time to conversion in hippocampal baseline volumes less than 5500 mm^3, whereas it had significant predictive value at higher volumes. Medium to large volumes were associated with 69% and 95% lower risk of conversion from MCI to AD compared to small volumes. An additional 15% and 50% decrease in the risk of conversion were demonstrated for every one percent lower atrophy rate in medium and large volumes.

Entorhinal cortex

Atrophy rate did not predict time to conversion in entorhinal cortex baseline volumes less than 2800 mm^3, whereas it had significant predictive value at large volumes (larger than 4000 mm^3). Medium to large entorhinal cortex volumes were respectively associated with 47% and 86% lower risk of conversion from MCI to AD compared to small volumes. An additional 24% decrease in the risk of conversion was demonstrated for one percent lower atrophy rate in large entorhinal cortex baseline volumes.

Because APOEe4 carrier prevalence was significantly higher in MCIc than other MCI types, post hoc analyses were done to investigate the effect of APOEe4 on the predictive values of baseline volumes and annual change rates and their interactions. The result showed that APOEe4 genotype had no effect on the predictive values of these measures as well as their interactions.

Figure 2 demonstrates the distribution of individuals converted from MCI to AD within ten years using Cox analysis to estimate probability in each separate baseline category across stratified annual atrophy rates (enlargement rate for the ventricles). It reveals that at hippocampal baseline volumes less than 5500 mm^3, conversion within 3 years occurs

regardless of the atrophy rate. A similar but somewhat weaker pattern was observed for an entorhinal cortex volume smaller than 2800 mm^3, a whole brain volume smaller than 1,040,000 mm^3, and a ventricular volume larger than 55,000 mm^3. In contrast, atrophy rate (enlargement rate for the ventricles) is determinant of probability of conversion over time at medium to large baseline brain volumes (medium to small for the ventricles). It is especially noticeable for the hippocampus with atrophy rate more than the average.

Table 4: Cox proportional hazard

	Coef.	SE	HR (95% CI)	Z
Whole Brain				
Baseline volume	0.43	0.08	1.53 (1.31 – 1.79)	5.344, p<0.0001
Annual atrophy rate	0.32	0.09	1.38 (1.15 – 1.65)	4.024, p<0.0001
Interaction	0.21	0.07	1.24 (1.08 – 1.41)	3.110, p<0.01
Ventricles				
Baseline volume	0.25	0.07	1.29 (1.14 – 1.46)	3.919, p<0.0001
Annual enlargement rate	0.46	0.06	1.58 (1.41 – 1.77)	7.988, p<0.0001
Interaction	-0.16	0.08	0.85 (0.73– 0.99)	-2.133, p<0.05
Hippocampus				
Baseline volume	0.63	0.08	1.87 (1.60 – 2.19)	7.840, p<0.0001
Annual atrophy rate	0.66	0.10	1.94 (1.61 – 2.33)	7.001, p<0.0001
Interaction	0.45	0.09	1.56 (1.30 – 1.88)	4.755, p<0.0001
Entorhinal cortex				
Baseline volume	0.44	0.08	1.55 (1.33 – 1.80)	5.671, p<0.0001
Annual atrophy rate	0.56	0.10	1.75 (1.45 – 2.12)	5.768, p<0.0001
Interaction	0.30	0.10	1.35 (1.11 – 1.64)	3.061, p<0.01
HR; hazard ratio for 1-SD decrease in whole brain, hippocampal volume and entorhinal cortex volume and their annual rates as well as 1-SD increase in ventricular volume and its annual ventricular volume enlargement. All measures have been adjusted for age, field strength and intracranial volume.				

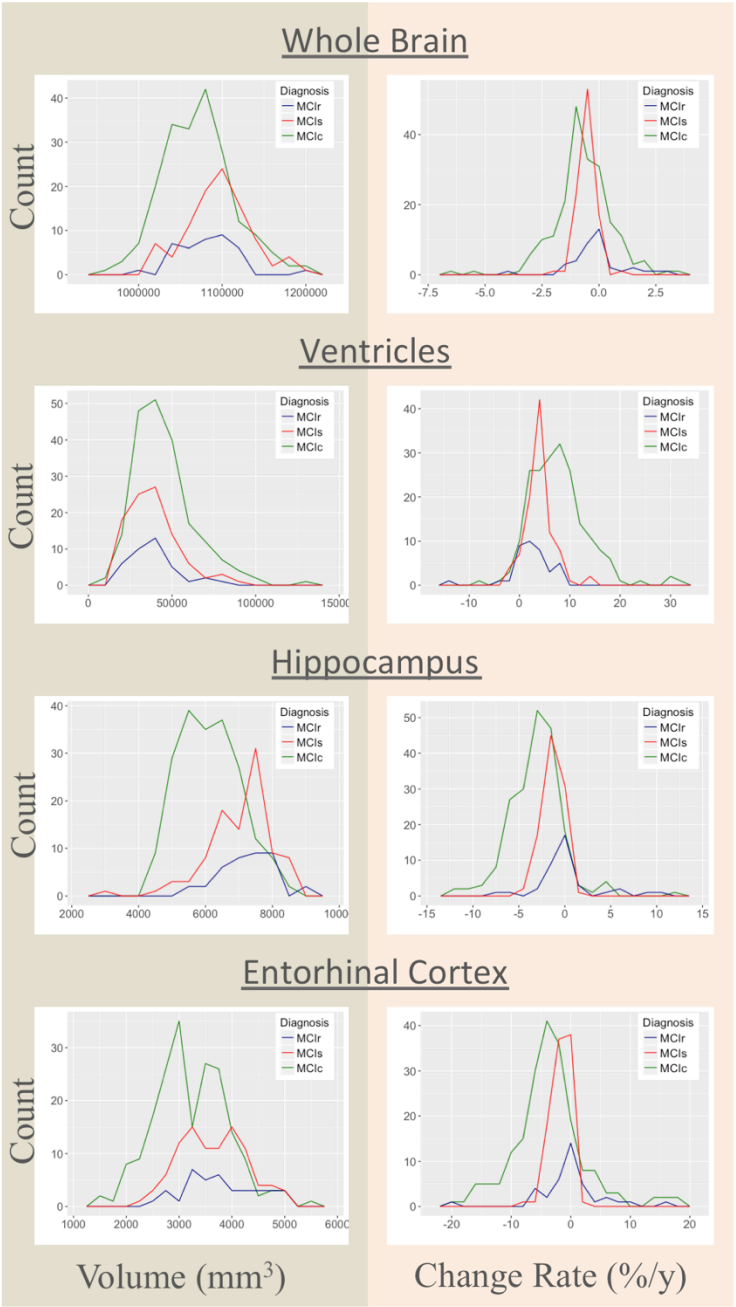


Figure 1: Frequency of baseline volumes and atrophy rates across the groups. Left column shows the overlap of baseline volumes and right column shows the overlap of atrophy rates (enlargement rate for the ventricles) across MCI groups.

MCIc= mild cognitive impairment who convert to Alzheimer's disease,

MCIs= mild cognitive impairment who remain stable for at least five years, and MCIr= mild cognitive impairment who revert to cognitively normal.

Table 5: Risk of conversion from MCI to AD over time (Cox Proportional Hazard ratios) in medium and large brain volume categories (small and medium categories in the ventricles) compared with the small brain volume category (large category in the ventricles).

	Coef.	SE	HR (95%CI)	Z, P-value
Whole Brain				
Medium whole brain	-0.96	0.21	0.39 (0.25 – 0.58)	-4.488, p<0.0001
Large whole brain	-1.28	0.32	0.28 (0.15 – 0.53)	-3.949, p<0.0001
Atrophy rate	0.08	0.18	1.08 (0.76 – 1.54)	0.438, p=0.66
Medium whole brain: atrophy rate	-0.44	0.21	0.65 (0.43 – 0.98)	-2.072, p<0.05
Large whole brain: atrophy rate	-0.56	0.34	0.57 (0.34 – 0.96)	-2.132, p<0.05
Ventricles				
Medium ventricles	-0.65	0.32	0.52 (0.28 – 0.97)	-2.068, p<0.05
Small ventricles	-1.76	0.48	0.17 (0.07 – 0.42)	-3.664, p<0.001
Enlargement rate	0.05	0.05	1.05 (0.96 – 1.15)	1.000, p=0.32
Medium ventricles: atrophy rate	0.04	0.05	1.04 (0.94 – 1.14)	0.757, p=0.45
Small ventricles: atrophy rate	0.13	0.06	1.14 (1.01 – 1.29)	2.065, p<0.05
Hippocampus				
Medium hippocampus	-1.16	0.28	0.31 (0.18 – 0.54)	-4.122, p<0.0001
Large hippocampus	-3.09	0.50	0.05 (0.02 – 0.12)	-6.252, p<0.0001
Atrophy rate	-0.04	0.05	0.96 (0.86 – 1.06)	-0.818, p=0.41
Medium hippocampus: atrophy rate	-0.16	0.06	0.85 (0.75 – 0.97)	-2.472, p<0.05
Large hippocampus: atrophy rate	-0.69	0.14	0.50 (0.38 – 0.66)	-5.022, p<0.0001
Entorhinal cortex				
Medium entorhinal	-0.64	0.22	0.53 (0.34 – 0.81)	-2.954, p<0.01
Large entorhinal	-1.97	0.35	0.14 (0.07 – 0.28)	-5.600, p<0.0001
Atrophy rate	-0.05	0.03	0.96 (0.91 – 1.01)	-1.623, p=0.11
Medium entorhinal: atrophy rate	-0.03	0.04	0.98 (0.90 – 1.04)	-0.900, p=0.37
Large entorhinal: atrophy rate	-0.27	0.07	0.76 (0.67 – 0.87)	-3.941, p<0.0001
Coef= coefficient, SE= standard error, HR= hazard ratio. All measures have been adjusted for age, sex, field strength and intracranial volume.				

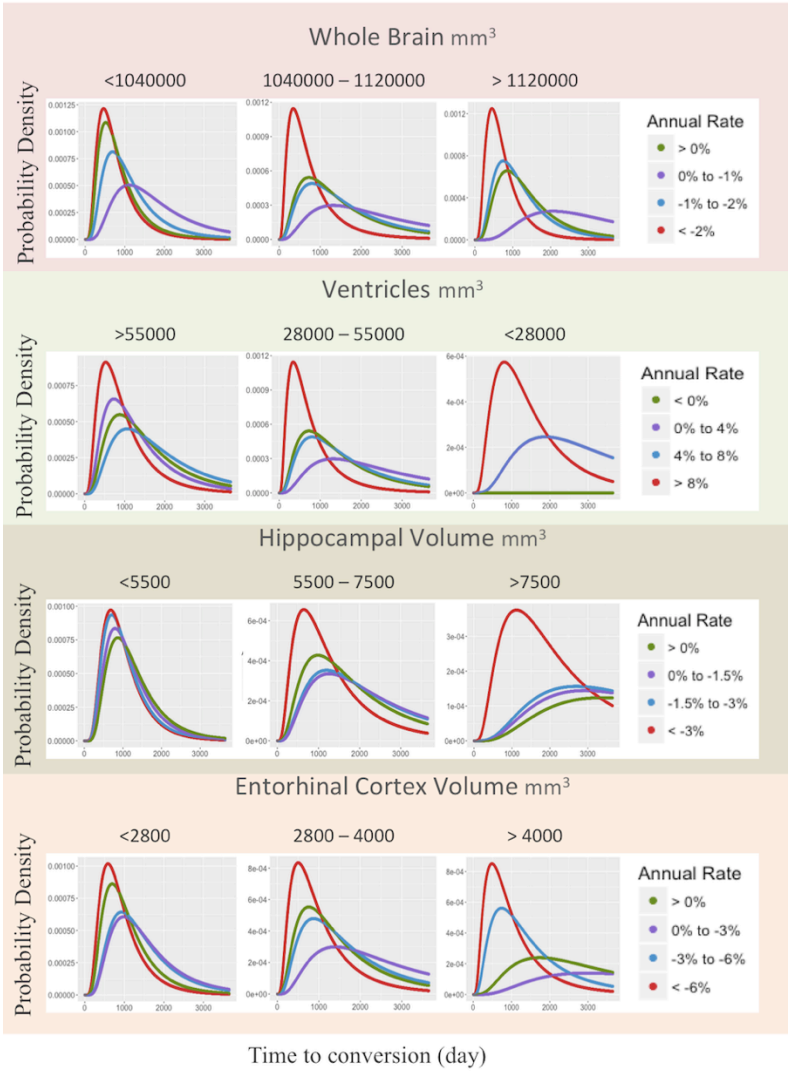


Figure 2: Distribution of probability of conversion over time: Separate illustration of probability density measured by Cox proportional models in four brain regions at three baseline categories across stratified respected annual atrophy rate (enlargement rate for the ventricles) within ten years. The figure shows that at hippocampal baseline volumes less than 5500 mm^3 conversions mostly happen within three years regardless of atrophy rate. Similar patterns but relatively less determinant are noticeable at entorhinal cortex volumes lower than 2800 mm^3, at whole brain baseline volumes lower than 1,040,000 mm^3, and at ventricle baseline volumes larger than 55,000 mm^3. In contrast, atrophy rate (enlargement rate for the ventricles) has an impact on probability of conversion over time at medium to large baseline brain volumes (medium to small for the ventricles), especially noticeable for the hippocampus with atrophy rate more than the average.

Discussion

This study aimed to investigate of the volume or change in volume over time of different brain regions could predict the time to conversion from MCI to AD. The main finding was that the baseline volumes of the whole brain, ventricles, hippocampus, and entorhinal cortex and their respective atrophy rates (enlargement rate for ventricles) were all significant predictors of earlier conversion. However, the predictive value of these ROIs' atrophy rates was highly dependent on their baseline volume.

Although volume and change in volume over time are predictive across all ROIs, the effect of baseline volume on the predictive value of volume change over time is more distinctive in the hippocampus than other ROIs (figure 1). Individuals with hippocampal volumes smaller than 5500 mm^3 mostly convert to AD within three years regardless of atrophy rate. This has an important implication for clinical trials aiming to delay AD conversion by reducing atrophy rate. In these trials, any treatment effects on brain atrophy rate should be interpreted in light of baseline volumes because at small hippocampal volumes, any reduction in atrophy rate is less likely to be associated with delay in disease progression. Indeed, it may be better for clinical trials to exclude individuals with small hippocampal volumes to identify interventions that can really delay the conversion by reducing volume loss. In addition, hippocampal volume can be used as a simple heuristics to identify those at risk of early conversion in clinical practice. However, it is important to note that the baseline brain volumes in this study were normalised for age, sex, field strength and

ICV, and therefore hippocampal threshold for small volume (i.e. 5500 mm^3) for any individual must be corrected with the provided formula[1].

Although we cannot shed light on specific reasons for this hippocampal threshold we speculate that volumes below this value are indicative of an accumulation of pathology, which makes conversion to AD all but inevitable. Regional accumulation of pathology is associated with concomitant spread of pathology to the adjacent brain regions. At early stages of the disease, neuropathology and brain atrophy is mainly limited to the medial and inferior temporal lobes (including hippocampus and entorhinal cortex) particularly in relation to tauopathy. As the disease progresses, degeneration spreads into more posterior regions of the temporal lobe and starts to spread to the parietal lobe. By the time of conversion to AD, atrophy has become more severe in the areas first affected and has spread further into the frontal lobes (Braak & Braak, 1991; Thal, Rub, Orantes, & Braak, 2002; J. L. Whitwell et al., 2007). Therefore, hippocampal volume below a certain threshold is not only indicative of pathology accumulation in this structure but also of spreading neurodegeneration in adjacent regions, which together indicate poorer prognosis.

By contrast, in those with larger ROI volumes, atrophy rate is a predictor of the time to conversion but is dependent on baseline volume. The pattern in larger volumes is also somewhat more distinctive in the hippocampus than other ROIs. Atrophy rate in those with medium to large hippocampal baseline volume (5500 mm^3 to 7500 mm^3) is determinant of the risk of AD conversion, whereas at volume larger than 7500 mm^3 atrophy rate more than

[1]Hippocampal threshold = Male => 3141+ 74.5*Age- 477.5* field strength - 0.0014* ICV

Female=>3438+ 74.5*Age-477.5*field strength - 0.0014* ICV

the average i.e. more than 3%/y is determinant. This can also be explained by the contribution of previous (reflected in baseline volume) and ensuing (reflecting in atrophy rate) neurodegeneration in prediction of progression. It is likely that at medium baseline volumes there is a balance between previous and ensuing neurodegeneration, thus both measures are determinant of the time to conversion. While at volume larger than 7500 mm^3, because of low level of previous degeneration only a large atrophy rate (more than the average of 3%/y) can be determinant of time to AD conversion.

The present results are particularly significant because they provide a guide on how structural imaging measures can assist in predicting conversion to AD as recommended by the National Institute on Aging and the Alzheimer's Association although to date they have been unable to advise on how this should or could be done (Jack et al., 2018). This approach also aligns with our understanding of AD's pathological progression, which recognizes MCI as a clinical stage of the disease continuum, rather than a distinct clinical entity with a higher risk of AD conversion (Albert et al., 2011; Dubois et al., 2016).

It is noteworthy that the selection of the brain ROIs in this study was based on the typical spread of the neurofibrillary tangles and neurodegeneration in the course of the disease. Typically, AD's neurofibrillary tangles aggregation and subsequent neurodegeneration originate in the transentorhinal cortex and spread through the hippocampus to subcortical structures and the lateral temporal, parietal and frontal association and primary cortices (Braak & Braak, 1991). However, there is some evidence demonstrating the presence of at least two atypical subtypes of AD that do not follow the typical pattern i.e. limbic-predominant AD and hippocampus sparing AD (Byun et al., 2015; Ferreira et al., 2017; Jennifer L. Whitwell et al., 2012). In the limbic-predominant AD fibrillary tangles and degeneration remain restrictively in medial temporal lobe and cortical

areas remain relatively preserved. The hippocampus and entorhinal cortex are severely involved and progression to the final stages of the disease is faster than the other subtypes (Ferreira et al., 2017; Murray et al., 2011). Thus, hippocampal atrophy would be expected to remain predictive of time to conversion, and to be consistent with the present findings. By contrast, in the hippocampus-sparing subtype the pathology originates in the lateral cortical areas and the medial temporal lobe including the hippocampus remains preserved and hippocampal atrophy is in line with that found in normal ageing (Ferreira et al., 2017). Thus hippocampal atrophy is not expected to be predictive of time to AD conversion. Of relevance to the present findings the possible presence of this subtype in the sample investigated – it affects approximately 10% of all AD cases in the population – may have negatively impacted the predictive value of the measures investigated, although probably only to a small extent.

To our knowledge, the present study is the first investigation of interaction between brain volume and annual change rate in predicting the time to conversion from MCI to AD. In addition, unlike previous studies, which investigated the prediction of conversion from MCI to AD within follow-ups of one to three years (Jack et al., 2005; Liu et al., 2017; McEvoy et al., 2011; Risacher et al., 2009), the follow-up time of the present study was up to ten years. However, these findings need replications in other population before their usefulness in clinical practice can ascertained.

Conclusion

These findings are among the first to demonstrate that simple structural imaging measures can make a useful contribution in predicting disease progression from MCI to AD. Importantly, they provide specific guidance on volumetric thresholds in specific brain structures, which can be used to inform clinical assessment. However, while this is an important first step, further investigation in different, more diverse and larger populations is

needed before recommendation for their routine use in clinical trials and clinical practice can be confidently made.

Funding:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

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). 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 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 co- ordinated 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 partly undertaken on the National Computational Infrastructure (NCI) facility in Canberra, Australia, which is supported by the Australian Commonwealth Government and was supported by an Australian Government Research Training Program (RTP) Scholarship.

References:

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., . . . Phelps, C. H. (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. *Alzheimers Dement, 7*(3), 270-279. doi:10.1016/j.jalz.2011.03.008

Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol, 82*, 239-259.

Braak, H., & Braak, E. (1997). Frequency of stages of Alzheimer-related lesions in different age categories. *Neurobiol Aging, 18*(4), 351-357.

Byun, M. S., Kim, S. E., Park, J., Yi, D., Choe, Y. M., Sohn, B. K., . . . Alzheimer's Disease Neuroimaging, I. (2015). Heterogeneity of Regional Brain Atrophy Patterns Associated with Distinct Progression Rates in Alzheimer's Disease. *PLoS One, 10*(11), e0142756. doi:10.1371/journal.pone.0142756

Douaud, G. l., Refsum, H., Jager, C. A. d., Jacoby, R., Nichols, T. E., Smith, S. M., & Smith, A. D. (2013). Preventing Alzheimer's disease-related gray matter atrophy by B-vitamin treatment. *PANS, 110*(23), 9523-9528.

Dubois, B., Chupin, M., Hampel, H., Lista, S., Cavedo, E., Croisile, B., . . . Hippocampus Study, G. (2015). Donepezil decreases annual rate of hippocampal atrophy in suspected prodromal Alzheimer's disease. *Alzheimers Dement, 11*(9), 1041-1049. doi:10.1016/j.jalz.2014.10.003

Dubois, B., Hampel, H., Feldman, H. H., Scheltens, P., Aisen, P., Andrieu, S., . . . Washington Dc, U. S. A. (2016). Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria. *Alzheimers Dement, 12*(3), 292-323. doi:10.1016/j.jalz.2016.02.002

Falahati, F., Ferreira, D., Muehlboeck, J. S., Eriksdotter, M., Simmons, A., Wahlund, L. O., & Westman, E. (2017). Monitoring disease progression in mild cognitive impairment: Associations between atrophy patterns, cognition, APOE and amyloid. *Neuroimage Clin, 16*, 418-428. doi:10.1016/j.nicl.2017.08.014

Ferreira, D., Verhagen, C., Hernandez-Cabrera, J. A., Cavallin, L., Guo, C. J., Ekman, U., . . . Westman, E. (2017). Distinct subtypes of Alzheimer's disease based on patterns of brain atrophy: longitudinal trajectories and clinical applications. *Sci Rep, 7*, 46263. doi:10.1038/srep46263

Fjell, A. M., Westlye, L. T., Grydeland, H., Amlien, I., Espeseth, T., Reinvang, I., . . . Walhovd, K. B. (2014). Accelerating Cortical Thinning: Unique to Dementia or Universal in Aging? *Cerebral Cortex, 24*(4), 919-934. doi:10.1093/cercor/bhs379

Gavidia-Bovadilla, G., Kanaan-Izquierdo, S., Mataro-Serrat, M., Perera-Lluna, A., & Alzheimer's Disease Neuroimaging, I. (2017). Early Prediction of Alzheimer's Disease Using Null Longitudinal Model-Based Classifiers. *PLoS One, 12*(1), e0168011. doi:10.1371/journal.pone.0168011

Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., . . . Contributors. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement, 14*(4), 535-562. doi:10.1016/j.jalz.2018.02.018

Jack, C. R., Jr., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., . . . Weiner, M. W. (2008). The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging, 27*(4), 685-691. doi:10.1002/jmri.21049

Jack, C. R., Jr., Shiung, M. M., Weigand, S. D., O'Brien, P. C., Gunter, J. L., Boeve, B. F., . . . Petersen, R. C. (2005). Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic MCI. *Neurology, 65*(8), 1227-1231. doi:10.1212/01.wnl.0000180958.22678.91

Kile, S., Au, W., Parise, C., Rose, K., Donnel, T., Hankins, A., . . . Ghassemi, A. (2017). IVIG treatment of mild cognitive impairment due to Alzheimer's disease: a randomised double-blinded exploratory study of the effect on brain atrophy, cognition and conversion to dementia. *J Neurol Neurosurg Psychiatry, 88*(2), 106-112. doi:10.1136/jnnp-2015-311486

Kobe, T., Witte, A. V., Schnelle, A., Lesemann, A., Fabian, S., Tesky, V. A., . . . Floel, A. (2016). Combined omega-3 fatty acids, aerobic exercise and cognitive stimulation prevents decline in gray matter volume of the frontal, parietal and cingulate cortex in patients with mild cognitive impairment. *Neuroimage, 131*, 226-238. doi:10.1016/j.neuroimage.2015.09.050

Li, Y., Wang, Y., Wu, G., Shi, F., Zhou, L., Lin, W., & Shen, D. (2012). Discriminant analysis of longitudinal cortical thickness changes in Alzheimer's disease using dynamic and network features. *Neurobiol Aging, 33*(2), 427 e415-430. doi:10.1016/j.neurobiolaging.2010.11.008

Liu, K., Chen, K., Yao, L., & Guo, X. (2017). Prediction of Mild Cognitive Impairment Conversion Using a Combination of Independent Component Analysis and the Cox Model. *Front Hum Neurosci, 11*, 33. doi:10.3389/fnhum.2017.00033

McEvoy, L. K., Holland, D., Donald J. Hagler, J., Fennema-Notestine, C., Brewer, J. B., & Dale, A. M. (2011). Mild Cognitive Impairment: Baseline and Longitudinal Structural MR Imaging Measures Improve Predictive Prognosis1. *radiology*. doi:10.1148/radiol.11101975/-/DC1

Murray, M. E., Graff-Radford, N. R., Ross, O. A., Petersen, R. C., Duara, R., & Dickson, D. W. (2011). Neuropathologically defined subtypes of Alzheimer's disease with distinct clinical characteristics: a retrospective study. *The Lancet Neurology, 10*(9), 785-796. doi:10.1016/s1474-4422(11)70156-9

Petersen, R. C., Aisen, P. S., Beckett, L. A., Donohue, M. C., Gamst, A. C., Harvey, D. J., . . . Weiner, M. W. (2010). Alzheimer's Disease Neuroimaging Initiative (ADNI) Clinical characterization. *Neurology, 74*(3), 201-209. doi:DOI 10.1212/WNL.0b013e3181cb3e25

Pintzka, C. W., Hansen, T. I., Evensmoen, H. R., & Haberg, A. K. (2015). Marked effects of intracranial volume correction methods on sex differences in neuroanatomical structures: a HUNT MRI study. *Front Neurosci, 9*, 238. doi:10.3389/fnins.2015.00238

Prins, N. D., van der Flier, W. A., Knol, D. L., Fox, N. C., Brashear, H., Nye, J. S., . . . Scheltens, P. (2014). The effect of galantamine on brain atrophy rate in subjects with mild cognitive impairment is modified by apolipoprotein E genotype: post-hoc analysis of data from a randomized controlled trial. *Alzheimers Res Ther, 6*(4), 47. doi:10.1186/alzrt275

Risacher, S. L., Saykin, A. J., West, J. D., Shen, L., Firpi, H. A., McDonald, B. C., & Alzheimer's Disease Neuroimaging, I. (2009). Baseline MRI predictors of conversion from MCI to probable AD in the ADNI cohort. *Curr Alzheimer Res, 6*(4), 347-361.

Risacher, S. L., Shen, L., West, J. D., Kim, S., McDonald, B. C., Beckett, L. A., . . . Alzheimer's Disease Neuroimaging, I. (2010). Longitudinal MRI atrophy biomarkers: relationship to conversion in the ADNI cohort. *Neurobiol Aging, 31*(8), 1401-1418. doi:10.1016/j.neurobiolaging.2010.04.029

Tabatabaei-Jafari, H., Shaw, M. E., & Cherbuin, N. (2015). Cerebral atrophy in mild cognitive impairment: A systematic review with meta-analysis. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 1*(4), 487-504. doi:10.1016/j.dadm.2015.11.002

Tabatabaei-Jafari, H., Shaw, M. E., Walsh, E., Cherbuin, N., & Alzheimer's Disease Neuroimaging, I. (2019). Cognitive/functional measures predict Alzheimer's disease, dependent on hippocampal volume. *J Gerontol B Psychol Sci Soc Sci*. doi:10.1093/geronb/gbz011

Teipel, S. J., Kurth, J., Krause, B., Grothe, M. J., & Alzheimer's Disease Neuroimaging, I. (2015). The relative importance of imaging markers for the prediction of Alzheimer's disease dementia in mild cognitive impairment - Beyond classical regression. *Neuroimage Clin, 8*, 583-593. doi:10.1016/j.nicl.2015.05.006

Thal, D. R., Rub, U., Orantes, M., & Braak, H. (2002). Phases of A beta-deposition in the human brain and its relevance for the development of AD. *Neurology, 58*(12), 1791-1800.

Thung, K. H., Yap, P. T., Adeli, E., Lee, S. W., Shen, D., & Alzheimer's Disease Neuroimaging, I. (2018). Conversion and time-to-conversion predictions of mild cognitive impairment using low-rank affinity pursuit denoising and matrix completion. *Med Image Anal, 45*, 68-82. doi:10.1016/j.media.2018.01.002

Whitwell, J. L., Dickson, D. W., Murray, M. E., Weigand, S. D., Tosakulwong, N., Senjem, M. L., . . . Josephs, K. A. (2012). Neuroimaging correlates of pathologically defined subtypes of Alzheimer's disease: a case-control study. *The Lancet Neurology, 11*(10), 868-877. doi:10.1016/s1474-4422(12)70200-4

Whitwell, J. L., Przybelski, S. A., Weigand, S. D., Knopman, D. S., Boeve, B. F., Petersen, R. C., & Jack, C. R., Jr. (2007). 3D maps from multiple MRI illustrate changing atrophy patterns as subjects progress from mild cognitive impairment to Alzheimer's disease. *Brain, 130*(Pt 7), 1777-1786. doi:10.1093/brain/awm112

Zhang, Y. P., Miao, R., Li, Q., Wu, T., & Ma, F. (2017). Effects of DHA Supplementation on Hippocampal Volume and Cognitive Function in Older Adults with Mild Cognitive Impairment: A 12-Month Randomized, Double-Blind, Placebo-Controlled Trial. *J Alzheimers Dis, 55*(2), 497-507. doi:10.3233/JAD-160439

DISCUSSION

> "Excessive reservations and paralyzing despondency have not helped the sciences to advance nor are they helping them to advance, but a healthy optimism that cheerfully searches for new ways to understand, as it is convinced that it will be possible to find them."
>
> Alois Alzheimer (1864-1915)

While the progressive nature of AD pathology is well documented, prediction of its clinical progression is still a key question. This thesis aimed to investigate the value of simple structural brain MRI markers whose use to predict AD progression is feasible in daily practice. The present findings suggest promising prognostic roles for the hippocampus in predicting the risk of conversion to AD and in predicting the time within which conversion occurs. The combination of hippocampal volume (normalised by cerebellar volume) and MMSE can help identify those at risk of conversion to AD within five years. Additionally, the combination of hippocampal volume and measures of short-term memory is predictive of time to AD conversion from MCI. Individuals with hippocampal volumes smaller than a certain threshold have poor prognosis and they are at high risk of conversion to AD in less than three years. Among a large number of studies that have investigated the predictive values of AD biomarkers, the present collection of studies are unique in providing simple guidelines, which can easily be used in clinical practice and clinical trials.

Review of the main findings

The aims of this thesis have been explored through a series of five studies, briefly reviewed as follows.

STUDY 1: This study was a systematic review with meta-analysis. It revealed that the hippocampus, entorhinal cortex, whole brain and ventricles are the brain regions most frequently investigated in longitudinal studies focused on MCI. Among these four brain regions the hippocampus was investigated the most (33 of 44 studies). Atrophy rates in these brain regions were found to be significantly higher in individuals with MCI than in normal ageing, particularly in the hippocampus where it was more than double. This suggests that volumetric measures of these brain regions, specifically hippocampal volume, may be useful biomarkers of disease progression.

Hippocampal volume comprises less than 1% of the whole brain volume (Standring, 2008), and meta-analysis revealed that hippocampal volume loss in MCI accounts for less than 2% of whole brain volume loss. However, despite its limited contribution to AD-related whole brain atrophy, the development of AD pathology in the hippocampus is associated with memory dysfunction, which is the core symptom in typical AD (Dubois et al., 2014; Livingston et al., 2017). Therefore, because the hippocampus is among the brain regions most affected by AD pathology from early stages (Braak & Braak, 1991, 1997), it is likely that change in this brain region closely reflects the progression in the course of the disease.

STUDY 2: The second study explored the extent to which the cerebellum is affected by AD-related pathology because previous histological studies have produced inconsistent evidence on the impact of AD pathology in this structure (Andersen, Korbo, & Pakkenberg, 1992; Wegiel et al., 1999). Longitudinal evaluation of cerebellar volume across CN, MCI,

and AD revealed that its atrophy rate in MCI was not different to that found in CN, while it was mildly higher in AD (0.56% in AD vs. 0.36% in CN).

The finding was consistent with previous PET studies that typically reported no difference in the cerebellar amyloid β and tau ligand uptake throughout the AD course (Jack et al., 2008; Jonasson et al., 2016; Rowe et al., 2007) which led to its adoption as a normalizing area for standardising uptake in other parts of the brain (Jonasson et al., 2016; Lopresti et al., 2005). Similarly, since cerebellar shrinkage in MCI is similar to CN, the cerebellum can be thought of as an area of reference with which the volume of other brain regions (i.e. those that are more strongly implicated in the disease) can be normalised.

STUDY 3: Based on two previous studies hippocampal volume normalized by cerebellar volume (HCCR) was considered as a biomarker of neurodegeneration in the third study. The study tested the predictive value of HCCR and MMSE in identifying individuals with MCI who were at high risk of conversion to AD within five years. It showed that there was a specific HCCR threshold at each MMSE score level which indicated a higher risk of progression to AD within five years. The HCCR threshold for an MMSE score of 25 was 7.5%, for 26 was 6.3%, for 27 was 5.3%, for 28 was 4.8%, and for MMSE scores of 29 and 30 was 4.3%.

The presence of different HCCR thresholds at different MMSE score levels indicate that the severity of neurodegeneration is not a reliable predictor of clinical progression by itself. In a similar way, the severity of cognitive dysfunction is not a reliable predictor of pathological progression. This is because there is an imperfect overlap between cognitive deficit and pathology severity in the course of AD (Neuropathology Group of the Medical Research Council Cognitive Function and Aging Study, 2001). This has been explained in terms individual variability in brain/cognitive reserve (Medaglia, Pasqualetti, Hamilton,

Thompson-Schill, & Bassett, 2017; Steffener & Stern, 2012; Stern, 2009). It postulates that individuals with larger brain/cognitive reserve may cope better with neural damage at a given level of observed pathology (Stern, 2009). The combination of neural measure (HCCR) and cognitive measure (MMSE) is a way to take into account the moderating effect of brain/cognitive reserve, and thus achieve a better prediction of clinical progression.

STUDY 4: Identifying MCI at risk of conversion to AD prompted questions about the feasibility of predicting the timing of conversion. Thus, the fourth study aimed to investigate the predictive values of hippocampal volume and cognitive measures i.e. MMSE, ADAS, FAQ, and RAVLT-immediate in the prediction of time to AD conversion. Result demonstrated that all these measures were predictive of time to conversion, however the predictive value of cognitive measures was dependent on hippocampal volume. At normalised hippocampal volumes lower than 5500 mm^3, MMSE, ADAS and FAQ were not predictive of time to conversion, while at higher volumes they made significant contributions. In contrast, RAVLT improved prediction at all hippocampal volumes including those below 5,500 mm^3.

The findings indicate that the relative contribution of hippocampal volume and cognitive measures is not constant in predicting time to AD conversion. As discussed above, individual variability in brain/cognitive reserve is the most likely explanation for this effect. However, when brain/cognitive reserve is exhausted, seemingly below a normalised hippocampal volume of 5500 mm^3, further neurodegeneration cannot be compensated for (Steffener & Stern, 2012) and most cognitive measures no longer contribute to predicting AD conversion. The practical implication of these findings is that at small-normalised hippocampal volumes (< 5,500 mm^3) conversion to AD in the near future appears to be inevitable. This pattern is particularly observed in the relationship between hippocampal

volume and MMSE, ADAS and FAQ, perhaps because these measures evaluate functions underpinned by a larger number of neural networks.

STUDY 5: Building on the findings from the fourth study, this study explored whether brain atrophy in MCI makes an additional contribution in predicting the timing of conversion to AD. It examined the four most investigated brain regions (based on the first study) which are known to be affected by AD pathology in the prodromal stages of the disease i.e. the hippocampus, entorhinal cortex, whole brain and ventricles. The atrophy rate (enlargement rate in the ventricles) in these ROIs was found to be predictive of time to AD conversion, but its contribution was dependent on ROIs baseline volumes. While the influence of baseline volumes on the predictive value of atrophy rate has been previously reported (Jack et al., 2005), the present finding is novel in showing the interactive effects between baseline volumes and atrophy rates. At volumes smaller than (in ventricles larger than) a certain threshold, the rate of further atrophy was no longer predictive of time to AD conversion. The threshold for the hippocampus was 5,500 mm^3, for the entorhinal cortex 2,800 mm^3, for the whole brain 1,040,000 mm^3, and for the ventricle 55,000 mm^3. Moreover, the interaction between volume and atrophy was more predictive in the hippocampus.

These findings suggest that the severity of neurodegeneration (i.e. the degeneration that has already occurred and is already reflected in baseline volumes) and the pace of further degeneration (ensuing atrophy rate) are both determinants of how fast individuals will progress from MCI to AD. However, when the present level of degeneration is severe enough as indexed by a brain volume that is less than a threshold specific to each region of interest, AD conversion is highly likely to occur in the near future and atrophy rate is no longer a useful determinant of timing of conversion. This is most likely because at small volumes brain/cognitive reserve may be exhausted and any further degeneration cannot be

compensated for and thus leads to severe loss of function, which impacts one's ability to conduct activities of daily living. This is consistent with the emergence of impairment in activities of daily living in the transition from MCI to AD.

Together, these studies suggest that the hippocampus, which has a major role in the core symptom of AD (memory dysfunction), is a useful and reliable biological marker of progression in the AD continuum. To reach to this final suggestion, the most appropriate statistical analysis has been used in each study with carefully considering their strengths and limitations in interpretation of the results. For example, a linear mixed effect model was used in study 2 to investigate cerebellar volume change across different cognitive status over time, linear discriminate analysis was used in study 3 to discriminate progressive MCI from CN and MCI with different prognosis, and Cox regression was used in study 4 & 5 to investigate the probability of conversion from MCI to AD over time.

Synthesis of the findings to address the aims of the thesis

The major aims and findings of this thesis address two important challenges in the field of ageing/AD i.e. identifying those at high risk of progression to AD and timing of progression. The contributions of the present results in clarifying these challenges are discussed in the following sections.

Prediction of conversion from MCI to AD

The use of a combination of cognitive and MRI measures, which was investigated in the third study, was not only found to be predictive of conversion from MCI to AD but was also found to have good psychometric properties. Indeed the combination of MMSE and HCCR predicted conversion to AD within five years with 83.5% accuracy (95%CI; 80.4% – 86.2%) with 65.2% sensitivity and 90.7% specificity. Moreover, 73.5% of those who were

identified as MCI converter and 86.8% of those who were identified as non-converter were correctly identified. The small number of false positive and false negative in these results is similar to findings reported in other studies in the literature (see study three) despite the fact they used much more elaborate and sophisticated approaches. Additionally, previous studies investigated follow-ups between 1.5 years to the maximum of 3 years (Chupin et al., 2009; Minhas, Khanum, Riaz, Alvi, & Khan, 2017; Suppa et al., 2016; Wei, Li, Fogelson, & Li, 2016), while the combination of HCCR and MMSE can predict conversion within 5 years. Nevertheless, this simple combination appears very effective in ruling out those who are not at high risk of progression to AD and ruling in those who will eventually progress to AD within 5 years. Given that the clinical diagnosis of MCI is associated with an overestimation of the risk of AD conversion, the combination of HCCR and MMSE is a practical approach to address this problem and provide some reassurance to the majority of those who are at lower risk of conversion to AD within 5 years. Although everybody may be at some risk of AD in old age, having some reassurance that it may not occur in the relatively near future is a significant achievement. However, it is important to note that there were a small number of MCI at risk of AD conversion who were not identified with the present combination. Therefore, prognosis should not entirely be based on the result of this combination and, as with other paraclinical assessments, the result of the combination of HCCR and MMSE should be interpreted in the context of the whole clinical picture of the patient.

Therefore, the present approach may have some important clinical indications. However, in order to confirm that a test is clinically useful it needs to meet some minimum criteria. This is somewhat dependent on tools already available and the intended use. Typically it is considered that to be clinically useful a test should have a specificity and sensitivity of approximately 80%. Power et al. suggest that to be useful a test specificity and sensitivity should add to more than 150% (Power et al., 2013). For some indications a higher

sensitivity and a somewhat lower specificity may be more useful or appropriate, while in others the reverse may be the case (Power, Fell, & Wright, 2013; Trevethan, 2017). In multiple sclerosis the accepted and widely used diagnostic test (i.e. conventional structural MRI test) has a sensitivity of 85% (35% to 100%) and a specificity of 80% (36% to 92%) (Schaffler et al., 2011; Whiting et al., 2006). In the present research a sensitivity of about 65% and a specificity of about 91% was achieved which is in line with MRI test for multiple sclerosis and meets Power et al. criterion (i.e. adds up to more than 150%). Given the proposed approach could be used in the clinic to supplement other diagnostic measure to provide additional guidance to patients as to the likelihood of progression to AD within 5 years, and given that no other specific and practical tests which perform better than the present approach are available, it is reasonable to conclude that it has good potential for use in the clinic provided external validation confirms these findings.

Although AD and MCI diagnoses were mainly established based on clinical assessment, current diagnostic guidelines, such as the National Institute on Aging and Alzheimer's Association's guideline (Albert et al., 2011; Jack et al., 2018) and the last version of the Diagnostic and Statistical Manual of Mental disorders (The American Psychiatric Association, 2013), highly recommend biological confirmation of MCI/AD diagnoses. Additionally, markers of neurodegeneration (mostly neuroimaging markers) were suggested for tracking MCI/AD progression (Jack et al., 2018). However, there is currently no evidence-based guidance outlining specific evidence-based approaches that can be feasibly implemented in daily practice. The practicality of using HCCR and MMSE to identify those at high risk of progression to AD within a clinically useful period makes this combination a valuable guide on how biological markers can assist in predicting AD progression.

Previous studies that used biological markers to predict conversion either used invasive approaches such as PET or CSF (Diniz, Pinto Junior, & Forlenza, 2008; Ma et al., 2014; Ritchie et al., 2017) or used complex image processing methods such as machine learning (Falahati, Westman, & Simmons, 2014; Moradi et al., 2015; Rathore, Habes, Iftikhar, Shacklett, & Davatzikos, 2017). In contrast, HCCR is a highly practical biological marker because MRI is more widely available and cheaper than PET imaging or spinal taps necessary to assay CSF markers which require the availability of inpatient care after the procedure to monitor possible complications. Volumetric processing of the MRI images is also widely available to clinicians or support staff with limited technical expertise as it can be easily set up on a desktop computer equipped with automatic segmentation software such as the FreeSurfer or even better, integrated in the scanner software. Additionally, the procedure is likely to be easily accepted by patients, especially if it can contribute to largely ruling out conversion to AD in the next 5 years, given that MRI is a non-invasive outpatient procedure and MMSE administration only takes 5 to 10 minutes (Arevalo-Rodriguez et al., 2015).

It is important to note that the approach, as it has been detailed above, is suggested as a complementary diagnostic test in a clinical context. Consequently, the composition of the ADNI study, which over-samples MCI participants, is appropriate as it is representative of patients consulting in a memory clinic. However, it may not be useful as a screening test for identifying those with MCI in the general population as over-sampling of MCI participants in the ADNI study may lead to higher positive and negative predictive value (Steinberg, Fine, & Chappell, 2009), as it is not reflective of the broader population.

Prediction of time to conversion from MCI to AD

The second main aim of the thesis was to shed light on the timing of AD progression in people at risk. The results of study 4 and 5 suggested that brain volume, brain atrophy, and

cognitive function can all individually predict with moderate accuracy how fast individuals with MCI may progress to AD. This is in line with the findings of previous studies investigating the predictive value of MRI brain measures (Evans et al., 2010; Ewers et al., 2012; Falahati et al., 2017; Jack et al., 2005; McEvoy et al., 2011), cognitive measures (Rabin et al., 2012; Silva et al., 2012; Silva et al., 2013), and their combination (Barnes, Cenzer, Yaffe, Ritchie, & Lee, 2014; Devanand et al., 2007; Ewers et al., 2012; Falahati et al., 2014; Moradi et al., 2015; D. Zhang, Shen, & Alzheimer's Disease Neuroimaging, 2012). However, the present findings provide a more detailed modelling of how these measures interact with each other and suggest simple guidelines on how to use them in combination, which can be used in day-to-day practice.

An implication of these results is that the combination of measures of brain structure and function can better index the disease progression than each individual measure. However, despite the overall better prediction, combining of brain volume and cognitive measure is particularly effective when hippocampal volume is close to or above the average for a certain age. In contrast, simple measures of broad cognitive function do not add substantially to the predictive value of hippocampal volume when the latter is significantly lower than the average. Although there is an association between hippocampal volume and severity of AD degeneration (Leung et al., 2010), this correlation is not perfect because of normal variation in hippocampal volume i.e. any given volume can only partially reflect the severity of hippocampal neurodegeneration. Therefore, taking cognitive function (that also has normal variation) into account can somewhat add to the prediction of progression. This is particularly true at a large hippocampal volume when individual differences are more likely to be due to normal variation, whereas a small hippocampal volume (standardized volume smaller than 5,500 mm^3) may be more likely the result of severe neurodegeneration than normal variation. Hence, cognitive function may not add to the prediction at small volumes.

It is important to note that neurodegeneration and decrease in brain volume (and hippocampal volume) is not exclusively linked to AD pathology. Therefore, the present findings are more likely to apply when AD pathology underlies differences in hippocampal volume. This thesis has the advantage of using data from Alzheimer's Disease Neuroimaging Initiative (ADNI) study, whose focus is exclusively on AD (Mueller et al., 2005). ADNI is a longitudinal project, to date over more than 15 years, with a large number of participants (more than 1000) who have been evaluated every 6 to 12 months (Weiner et al., 2015; Weiner et al., 2017). Therefore, the present findings are the product of analyses on a well-characterized large neuroimaging dataset, which can more reliably be generalized to individuals with MCI due to AD pathology than data from other sources which are more likely to include a wider range of underlying pathologies. The findings have some important implications for clinical practice and research, which are outlined in the following section.

Relevance for clinical practice and research

Identifying those with MCI at the highest risk of conversion to AD and estimating the time of conversion is likely to have important implications for clinical practice as well as clinical trials that aim to slow down disease progression. This is because a large proportion of individuals clinically diagnosed with MCI never convert to AD (Gao et al., 2014; Mitchell & Shiri-Feshki, 2009; Pandya, Clem, Silva, & Woon, 2016). A meta-analysis of 41 studies demonstrated that only 33.6% of those diagnosed with MCI at first assessment eventually progressed to AD with an annual progression rate of 5% to 10% (Mitchell & Shiri-Feshki, 2009), the remainder may revert to CN or remain stable. Some studies have reported that reversion from MCI to CN ranged from 30% to 50% within two to five years and up to 55% over 10 years (Ganguli, Dodge, Shen, & DeKosky, 2004; Han et al., 2012; Ravaglia et al., 2008; Roberts et al., 2014). Moreover, population studies have demonstrated that 37% to

67% of individuals with MCI remain stable over the course of 1.5 to 5 years (Manly et al., 2008; Pandya et al., 2016; Ravaglia et al., 2008). Despite the inconsistency about the proportion of each MCI outcome, these findings suggest that overall the proportion of MCI who remain stable or revert to CN is substantially higher than that of those progressing to AD. This indicates the importance of identifying (with a practical approach in day-to-day practice) those MCI who will convert to AD from those who will remain stable or revert to CN.

The findings of this thesis suggest practical approaches for identifying those at highest risk of AD conversion and timing of conversion. In ADNI, the MCI conversion rate is about 50% over 10 years with less than 10% of MCI reverting to CN and about 40% remaining stable even after 10 years (Girard et al., 2018). While the proportion of those MCI who revert to CN is lower in ADNI than what has been reported previously, the proportion of those who convert to AD is relatively consistent with previous reports in different populations (Gao et al., 2014; Mitchell & Shiri-Feshki, 2009; Pandya et al., 2016). Higher rate of reversion from MCI to CN in some datasets has been suggested to be due to false-positive (Park, Han, & Initiative, 2015), because the diagnosis is mainly based on clinical judgment and thus vulnerable to mis-classification (Park et al., 2015). Therefore, the lower rate of reversion in ADNI is likely to be due to lower rate of false positive in MCI diagnosis. This is likely because ADNI is a clinic-based study, whose focus is exclusively on AD pathology, and in which the diagnosis process implemented has been optimized through highly formalised clinical, cognitive and paraclinical evaluation (Weiner et al., 2015; Weiner et al., 2017).

Implementation in clinical practice

Since AD is a progressive disease for which no known cure is currently available, the focus of disease management should be individualized and based on the stage of the disease

the patient is in to achieve the greatest benefits possible. At the advanced stages of the disease i.e. mild-to-severe symptomatic AD, management is mostly symptomatic relief (Aaseth et al., 2016; Folch et al., 2016; Livingston et al., 2017; Yiannopoulou & Papageorgiou, 2013). In contrast, at the earlier stages of the disease the intention is to modify the progression and help the patient maintain maximum independence for longer and preferably at home (Livingston et al., 2017). Whereas, disease-modifying interventions are still under extensive research (Long & Holtzman, 2019), early diagnosis of those with MCI (who are relatively high functioning) provides an opportunity for the patients to prepare for what may come next, put their affairs in order, communicate their wishes to loved ones, organise support services, and thus enable them to stay longer in their home (Livingston et al., 2017). However, as discussed above, not all people diagnosed with MCI will progress to AD. Therefore, identifying those who are likely from those who are unlikely to be progressive can help clinicians adjust their management.

The combination of HCCR and MMSE may help clinicians to identify MCI who are likely to be progressive and hippocampal volume and RAVLT will assist them in predicting when conversion to AD may happen. These may contribute to better management and planning in the early stages of the disease. This is important because to date, there is no approved pharmacological or specific non-pharmacological treatment for MCI (Langa & Levine, 2014). Instead, the general recommendation is to minimize further neurological damage in the early stages of the disease by optimizing patients' general medical status combined with approaches that may maximize general function such as regular physical activity, cognitive/social stimulation, and cognitive training (Langa & Levine, 2014; Livingston et al., 2017).

In addition, identifying those who are unlikely to be progressive is also clinically valuable. These people may be particularly psychologically vulnerable because of fear of progression to one of the most devastating neurodegenerative disorders. This is even more likely if they also have to cope with the uncertainty of the speed of progression of their disease (Joosten-Weyn Banningh, Vernooij-Dassen, Rikkert, & Teunisse, 2008; Werner, 2012). Based on the present findings they can be reassured about the lower risk of AD conversion in the near future, and they can also be further motivated to eliminate or control any modifiable AD risk factors they may have to further reduce the risk of developing the disease (Livingston et al., 2017).

Implementation in clinical trials

The ability to identify those at higher risk of conversion to AD as demonstrated in this thesis has also major implication for the implementation of clinical trials because it can 1) help select the most suitable participants, 2) contribute to ensuring the greatest treatment effects can be detected, and 3) assist in decreasing the costs of drug development.

Despite more than 20 large phase 3 double-blind randomized clinical trials, no effective disease-modifying therapy for AD has been developed and the available approved medications (donepezil, rivastigmine, galantamine and memantine) are only for symptom relief (Long & Holtzman, 2019). As a consequence, the focus of clinical trials has moved to the earlier stages of the disease process i.e. the preclinical and early clinical stages (Long & Holtzman, 2019) including the MCI stage (Bateman et al., 2017; Reiman et al., 2011; Smith et al., 2010). However, the effectiveness of therapeutic interventions conducted on all those diagnosed with MCI may reduce because of the high proportions of individuals with MCI who normally and without any intervention remain stable or revert to CN (Kobe et al., 2016; Prins et al., 2014; Smith et al., 2010; Y. P. Zhang, Miao, Li, Wu, & Ma, 2017). Using the

combination of HCCR and MMSE allows targeting those who are progressive and most likely to benefit from interventions within the study period.

While progressive participants with MCI may be suitable for disease-modifying interventions, the treatment effect is likely to be different depending on how far along the disease they are (i.e. AD) before the intervention. Participants who are at high risk of conversion to AD in the near future may not gain a significant benefit from the intervention because of advanced pathology. Therefore, to achieve the greatest treatment effects it is better to include only those MCI who are less likely to convert to AD in the near future. It indicates that those with small hippocampal volumes may not be the optimal targets for such interventions because they are at high risk of early conversion regardless of the pace of further degeneration. Therefore, the present findings suggest that the combination of HCCR and MMSE can be used to identify individuals with progressive MCI (for example at a MMSE score of 26, only individuals with a HCCR less than 6% should be considered for intervention (see figure 2, study 3)). Furthermore, from these individuals only those who have hippocampal volumes larger than a standardized threshold (5,500 mm^3) might be investigated for the efficacy of intervention. Additionally, it may be better that the efficacy of intervention be evaluated with a combination of hippocampal volume and RAVLT immediate and not just by hippocampal volume or any other brain regional volume since the combination of structure and function will provide better estimation of the disease progression than any of them separately.

Narrowing the target population of disease-modifying trials to progressive MCI who will not convert to AD in the near future may assist to decrease the cost of drug development. This is because those who are not progressive are less likely to respond to treatment and thus having them in the sample will impact the outcome measure of the trial. In contrast excluding

them will lead to requiring a smaller sample size to achieve the same statistical power (Noordzij et al., 2010). Given that only half of those with MCI diagnosis are progressive, the sample size can be decreased to less than half for testing the actual therapeutic effect. This may not only decrease the cost of clinical trials, but also it can eliminate the risk of exposure to unknown harm of new interventions in those who are not the actual target of disease-modifying interventions.

Strengths and Limitations

The findings of this thesis are the results of the robust and systematic investigation of practical brain/cognitive measures using complex statistics that were conducted on a large, well-characterised longitudinal neuroimaging sample, i.e. ADNI dataset, and led to some clear-cut advice and biomarkers' thresholds that can easily be used by clinicians in clinical settings and by researchers in clinical trials. Despite using only one dataset, the special characteristics of ADNI study such as the accuracy of AD related diagnoses, neuroimaging protocols and comparability across all collecting centres, and a relatively long follow up period may justify generalization of the findings. However it is important to note that there is a limitation in the extent that the result can be generalized to a non-clinical population and also to those with atypical AD pathology.

Since ADNI is primarily a clinical study with an amnestic clinical population that uses stringent clinical assessment to establish the diagnosis (Weiner et al., 2017), the accuracy of MCI/AD diagnoses is higher compared with epidemiological studies (Mitchell & Shiri-Feshki, 2009). This is well reflected in the small ratio of MCI reversion to CN. Given that the findings of the present studies are founded on participants' diagnosis, the high level of accuracy of diagnosis in ADNI can allow drawing strong inferences from the findings.

Another characteristic of the ADNI dataset that can increase the reliability of the findings is careful neuroimaging protocols that ensure comparability across all centres that collecting data all over the United States and Canada (Weiner et al., 2015; Weiner et al., 2017). A standard structural MRI protocol with a high-resolution geometric MRI phantom for calibration has been used across the neuroimaging centres and longitudinal scans. The phantom consisted of polycarbonate spheres filled with water and copper sulphate in a precise geometrical pattern. It was scanned after each participant to detect linear and nonlinear spatial distortion, signal-to-noise ratio, and image contrast to identify any artefacts and problems for subsequent correction. This standard protocol allows comparability of images across different centres and over longitudinal scans. This increases the accuracy of predictions using the structural brain measures.

Additionally, ADNI data is unique in regard to its long follow up period. The project has now been running for 15 years. This leads to a high level of certainty in describing the participants in regard to their prognoses i.e. stable, reversible or progressive MCI. In the present studies a minimum of 5 years was specified to define MCI stability. Therefore, the results are highly reliable in differentiating progressive from non-progressive MCI compared with studies with up to 2 to 3 years follow up, which they could not role out the possibility of further progression in those who were stable for up to 2 to 3 years.

Despite the fact that ADNI project is a clinical-based study developed to investigate clinical-related challenges such as the main aims of this thesis, the extent to which the present results can be generalized to a non-clinical population is one of the limitations needs to be determined. For example the proportion of APOE e4 in the stable MCI (see the fifth study) was almost similar to APOEe4 proportion in the CN in ADNI (Petersen et al., 2010) i.e. 23% in stable MCI vs. 26.6% in CN. This may raise this speculation that those with stable MCI are

indeed CN who have a constant cognitive impairment that mistakenly diagnosed as MCI. However, the APOE e4 proportion in the ADNI's normal participants is more than the average in the USA normal population (6% to 12%) and also other part of the world (Heffernan, Chidgey, Peng, Masters, & Roberts, 2016; Singh, Singh, & Mastana, 2006). This indicates that normal participants in ADNI do not represent an epidemiologically selected real life population of the US. Therefore, in summary some of the present findings with epidemiological interpretation should not be generalized to the entire population and population-based studies should determine the preciseness of the information derived from ADNI.

Another important limitation, which is again related to the dataset used, is that ADNI participants represent a primarily amnestic clinical population with limited comorbidities, as those with cortical strokes, heart failure, substance abuse, cancer, and other pre-existing conditions are excluded from ADNI study. Therefore, the findings can only be generalized to those with typical AD and no serious comorbidities and should not be or at least should cautiously be generalized to all AD. This is because AD pathology typically originates in the medial temporal lobe and clinically emerges with memory dysfunction, whereas in an atypical subtype of AD (approximately 10% of all AD), called the hippocampus sparing subtype, the pathology originates in the lateral cortical areas and the medial temporal lobe including the hippocampus and thus memory function remains preserved (Ferreira et al., 2017). Therefore, since ADNI represents an amnestic clinical population, atypical AD, who do not present with memory dysfunction, may not be included in the ADNI participants. Indeed everybody with pre-existing conditions has been excluded from ADNI study. However, since comorbidities are common at old age, those who are diagnosed with AD may also have a secondary pathology, which may affect the course of the disease (Schneider, Arvanitakis, Bang, & Bennett, 2007; Schneider, Arvanitakis, Leurgans, & Bennett, 2009).

Since postmortem pathological evaluations is only available for a limited number of participants who passed away (64 people up to 2019), the nature of the prevalent pathology in ADNI participants is not yet completely clear. The early investigation on the limited number of autopsies revealed that all those who were diagnosed with AD or MCI at their last visits had AD pathology. However, coincident pathological evidence of dementia with Lewy bodies and medial temporal pathology (including hippocampal sclerosis) was another common finding (Toledo et al., 2013).

Taking all the strengths and limitations together and considering all the potential clinical benefits of the findings, it is important to acknowledge that they are based on a single dataset. Therefore, these findings will need to be validated in other datasets in future research in order for the clinical utility of the present approaches to be confirmed.

Future research directions

Concluding this thesis has also led to some important self-reflections, while I am relieved and proud to have reached the end of this particular journey, I also recognize that had I known when I started what I know now, I may have pursued a similar yet slightly different path. For example, I may have prioritized sourcing a different dataset to validate the present results in a different population as this would have strengthened results and provided needed evidence for implementation in clinical practice, I may have investigated more closely MCI reverters as this is a poorly understood population that merits more attention, or I may have attempted to contrast MRI/cognitive markers with other biological markers e.g. cerebrospinal fluid markers.

The present findings raise a number of key questions regarding the nature of the pathology affecting those with MCI who revert to CN, the contribution of second pathology

in AD progression and the predictive value of spread of pathology instead of regional severity, which should be addressed by future research.

It is not yet clear whether the pathology of those who remain stable or revert to CN is different from those who convert to AD. The progressive nature of AD pathology suggests that individuals with MCI should eventually progress to AD (Jack et al., 2018; Livingston et al., 2017). However, there is no specified time frame for the conversion even in the presence of confirmatory evidence of AD pathology. In this research, conversion within five years was considered for practical reasons (e.g. possibility of earlier intervention compared with previous studies) it does not mean that all those who remained stable MCI after five years did not have substantial AD pathology. Therefore, future studies should aim to compare pathological difference between MCI who remain stable or revert to normal with MCI who convert to AD within five years as well as to cognitively normal people.

Secondly, while the pace of progression from MCI to AD is assumed to be due to the pace of progression in AD pathology, the possibility that clinical progression is driven by the worsening of an additional pathology cannot be ruled out. Indeed, previous studies have revealed that about a large number of individuals diagnosed with AD present with substantial non-AD pathology, most commonly vascular (Schneider et al., 2007; Schneider et al., 2009). Therefore, the contribution of a secondary pathology to progression and timing of progression from MCI to AD adds further uncertainty that needs to be clarified. This is important because management of those MCI with an additional pathology may be different from management of those with pure AD pathology (Biessels, 2016; O'Brien & Thomas, 2015; Staszewski et al., 2017; Sun, 2018).

Another area requiring further investigation is whether the local accumulation of pathology (as indexed by hippocampal atrophy) or the spread of pathology to new brain

regions is most predictive of progression to clinical AD. It is important to address this question because previous studies using machine-learning approaches suggest that the pattern of progression in neurodegeneration is a significant determinant of clinical progression (Falahati et al., 2014; Moradi et al., 2015). Since change in the pattern of neurodegeneration is indeed the result of change in the local accumulation of pathology and the spread of pathology to new brain regions, measures of regional pathology and spread of pathology may be each separately predictive of AD progression. The present research has explored the predictive value of local degeneration in brain regions mostly affected in AD pathology including the hippocampus. Therefore, further study needs to investigate the predictive value of the spread of pathology and its relationship with regional progression.

Conclusion

This thesis has demonstrated that it is possible to identify those in the early stages of the Alzheimer continuum, who can still live independently but who are at higher risk of further substantial decline in neurocognitive functioning in the near future. Importantly, it showed that combinations of simple measures are precise enough for evaluating the risk as well as the pace of progression in this people. The main advantage of using this combination of biomarkers is its feasibility and practical implementation in day-to-day practice and clinical trials.

These findings have important implications for clinical practice as clinicians can easily use these measures to target those with a poor prognosis and avoid an unnecessary more intensive treatment for those likely to be following a relatively benign course of the disease. This evidence is also likely to help policy makers and managers of health systems direct the available resources to those who need them the most. The benefits of these measures are not limited to their use in clinical practice, as they can also be used in clinical trials, for example,

to target only those individuals who are most likely to benefit from interventions i.e. those with substantial risk of progression to dementia.

The fact that the majority of individuals with cognitive impairment may be at lower risk of progression to Alzheimer's than originally thought based on their cognitive assessment emphasizes the necessity of concomitant neurobiological evaluation to confirm that impairment is due to AD-related neurodegeneration. MRI constitutes a non-invasive tool providing robust and valid evaluation of neurobiological changes in Alzheimer's pathology. Thus, the present findings provide useful guidance on how to use the interplay between function and structure to identify those at higher risk and to predict their progression over time.

References:

Aaseth, J., Alexander, J., Bjorklund, G., Hestad, K., Dusek, P., Roos, P. M., & Alehagen, U. (2016). Treatment strategies in Alzheimer's disease: a review with focus on selenium supplementation. *Biometals, 29*(5), 827-839. doi:10.1007/s10534-016-9959-8

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., . . . Phelps, C. H. (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. *Alzheimers Dement, 7*(3), 270-279. doi:10.1016/j.jalz.2011.03.008

Andersen, B. B., Korbo, L., & Pakkenberg, B. (1992). A Quantitative Study of the Human Cerebellum With Unbiased Stereological Techniques. *J Comp Neurol, 326*, 549-560.

Arevalo-Rodriguez, I., Smailagic, N., Roque, I. F. M., Ciapponi, A., Sanchez-Perez, E., Giannakou, A., . . . Cullum, S. (2015). Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database Syst Rev*(3), CD010783. doi:10.1002/14651858.CD010783.pub2

Barnes, D. E., Cenzer, I. S., Yaffe, K., Ritchie, C. S., & Lee, S. J. (2014). A point-based tool to predict conversion from mild cognitive impairment to probable Alzheimer's disease. *Alzheimer's & Dementia, 10*(6), 646-655. doi:10.1016/j.jalz.2013.12.014

Bateman, R. J., Benzinger, T. L., Berry, S., Clifford, D. B., Duggan, C., Fagan, A. M., . . . Network, D.-T. P. C. f. t. D. I. A. (2017). The DIAN-TU Next Generation Alzheimer's prevention trial: Adaptive design and disease progression model. *Alzheimers Dement, 13*(1), 8-19. doi:10.1016/j.jalz.2016.07.005

Biessels, G. J. (2016). Diagnosis and treatment of vascular damage in dementia. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 1862*(5), 869-877. doi:10.1016/j.bbadis.2015.11.009

Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol, 82*, 239-259.

Braak, H., & Braak, E. (1997). Frequency of stages of Alzheimer-related lesions in different age categories. *Neurobiol Aging, 18*(4), 351-357.

Chupin, M., Gerardin, E., Cuingnet, R., Boutet, C., Lemieux, L., Lehericy, S., . . . Alzheimer's Disease Neuroimaging, I. (2009). Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. *Hippocampus, 19*(6), 579-587. doi:10.1002/hipo.20626

Devanand, D. P., Pradhaban, G., Liu, X., Khandji, A., Santi, S. D., Segal, S., . . . Leon, M. J. d. (2007). Hippocampal and entorhinal atrophy in mild cognitive impairmentPrediction of Alzheimer disease. *Neurology, 68*, 828-836.

Diniz, B. S., Pinto Junior, J. A., & Forlenza, O. V. (2008). Do CSF total tau, phosphorylated tau, and beta-amyloid 42 help to predict progression of mild cognitive impairment to Alzheimer's disease? A systematic review and meta-analysis of the literature. *World J Biol Psychiatry, 9*(3), 172-182. doi:10.1080/15622970701535502

Dubois, B., Feldman, H. H., Jacova, C., Hampel, H., Molinuevo, J. L., Blennow, K., . . . Cummings, J. L. (2014). Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *The Lancet Neurology, 13*(6), 614-629. doi:10.1016/s1474-4422(14)70090-0

Evans, M. C., Barnes, J., Nielsen, C., Kim, L. G., Clegg, S. L., Blair, M., . . . Fox, N. C. (2010). Volume changes in Alzheimer's disease and mild cognitive impairment: cognitive associations. *Eur Radiol, 20*(3), 674-682. doi:10.1007/s00330-009-1581-5

Ewers, M., Walsh, C., Trojanowski, J. Q., Shaw, L. M., Petersen, R. C., Jack, C. R., . . . Neuroimagin, N. A. A. s. D. (2012). Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. *Neurobiol Aging, 33*(7), 1203-+. doi:10.1016/j.neurobiolaging.2010.10.019

Falahati, F., Ferreira, D., Muehlboeck, J. S., Eriksdotter, M., Simmons, A., Wahlund, L. O., & Westman, E. (2017). Monitoring disease progression in mild cognitive impairment: Associations between atrophy patterns, cognition, APOE and amyloid. *Neuroimage Clin, 16*, 418-428. doi:10.1016/j.nicl.2017.08.014

Falahati, F., Westman, E., & Simmons, A. (2014). Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging. *J Alzheimers Dis, 41*(3), 685-708. doi:10.3233/JAD-131928

Folch, J., Petrov, D., Ettcheto, M., Abad, S., Sanchez-Lopez, E., Garcia, M. L., . . . Camins, A. (2016). Current Research Therapeutic Strategies for Alzheimer's Disease Treatment. *Neural Plast, 2016*, 8501693. doi:10.1155/2016/8501693

Ganguli, M., Dodge, H. H., Shen, C., & DeKosky, S. T. (2004). Mild cognitive impairment, amnestic type: an epidemiologic study. *Neurology, 63*(1), 115-121. doi:10.1212/01.wnl.0000132523.27540.81

Gao, S., Unverzagt, F. W., Hall, K. S., Lane, K. A., Murrell, J. R., Hake, A. M., . . . Hendrie, H. C. (2014). Mild cognitive impairment, incidence, progression, and reversion: findings from a community-based cohort of elderly African Americans. *Am J Geriatr Psychiatry, 22*(7), 670-681. doi:10.1016/j.jagp.2013.02.015

Girard, H., Potvin, O., Nugent, S., Dallaire-Theroux, C., Cunnane, S., Duchesne, S., & Alzheimer's Disease Neuroimaging, I. (2018). Faster progression from MCI to probable AD for carriers of a single-nucleotide polymorphism associated with type 2 diabetes. *Neurobiol Aging, 64*, 157 e111-157 e117. doi:10.1016/j.neurobiolaging.2017.11.013

Han, J. W., Kim, T. H., Lee, S. B., Park, J. H., Lee, J. J., Huh, Y., . . . Kim, K. W. (2012). Predictive validity and diagnostic stability of mild cognitive impairment subtypes. *Alzheimers Dement, 8*(6), 553-559. doi:10.1016/j.jalz.2011.08.007

Heffernan, A. L., Chidgey, C., Peng, P., Masters, C. L., & Roberts, B. R. (2016). The Neurobiology and Age-Related Prevalence of the epsilon4 Allele of Apolipoprotein E in Alzheimer's Disease Cohorts. *J Mol Neurosci, 60*(3), 316-324. doi:10.1007/s12031-016-0804-x

Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., . . . Contributors. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement, 14*(4), 535-562. doi:10.1016/j.jalz.2018.02.018

Jack, C. R., Jr., Lowe, V. J., Senjem, M. L., Weigand, S. D., Kemp, B. J., Shiung, M. M., . . . Petersen, R. C. (2008). 11C PiB and structural MRI provide complementary information in imaging of Alzheimer's disease and amnestic mild cognitive impairment. *Brain, 131*(Pt 3), 665-680. doi:10.1093/brain/awm336

Jack, C. R., Jr., Shiung, M. M., Weigand, S. D., O'Brien, P. C., Gunter, J. L., Boeve, B. F., . . . Petersen, R. C. (2005). Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic MCI. *Neurology, 65*(8), 1227-1231. doi:10.1212/01.wnl.0000180958.22678.91

Jonasson, M., Wall, A., Chiotis, K., Saint-Aubert, L., Wilking, H., Sprycha, M., . . . Lubberink, M. (2016). Tracer kinetic analysis of (S)-18F-THK5117 as a PET tracer for assessing tau pathology. *J Nucl Med*. doi:10.2967/jnumed.115.158519

Joosten-Weyn Banningh, L., Vernooij-Dassen, M., Rikkert, M. O., & Teunisse, J. P. (2008). Mild cognitive impairment: coping with an uncertain label. *Int J Geriatr Psychiatry, 23*(2), 148-154. doi:10.1002/gps.1855

Kobe, T., Witte, A. V., Schnelle, A., Lesemann, A., Fabian, S., Tesky, V. A., . . . Floel, A. (2016). Combined omega-3 fatty acids, aerobic exercise and cognitive stimulation prevents decline in gray matter volume of the frontal, parietal and cingulate cortex in patients with mild cognitive impairment. *Neuroimage, 131*, 226-238. doi:10.1016/j.neuroimage.2015.09.050

Langa, K. M., & Levine, D. A. (2014). The Diagnosis and Management of Mild Cognitive Impairment. *Jama, 312*(23), 2551. doi:10.1001/jama.2014.13806

Leung, K. K., Barnes, J., Ridgway, G. R., Bartlett, J. W., Clarkson, M. J., Macdonald, K., . . . Ourselin, S. (2010). Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease. *Neuroimage, 51*(4), 1345-1359. doi:10.1016/j.neuroimage.2010.03.018

Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., . . . Mukadam, N. (2017). Dementia prevention, intervention, and care. *The Lancet, 390*(10113), 2673-2734. doi:10.1016/s0140-6736(17)31363-6

Long, J. M., & Holtzman, D. M. (2019). Alzheimer Disease: An Update on Pathobiology and Treatment Strategies. *Cell, 179*(2), 312-339. doi:10.1016/j.cell.2019.09.001

Lopresti, B. J., Klunk, W. E., Mathis, C. A., Hoge, J. A., Ziolko, S. K., Lu, X., . . . Price, J. C. (2005). Simplified quantification of Pittsburgh Compound B amyloid imaging PET studies: a comparative analysis. *J Nucl Med, 46*(12), 1959-1972.

Ma, Y., Zhang, S., Li, J., Zheng, D. M., Guo, Y., Feng, J., & Ren, W. D. (2014). Predictive accuracy of amyloid imaging for progression from mild cognitive impairment to Alzheimer disease with different lengths of follow-up: a meta-analysis. [Corrected]. *Medicine (Baltimore), 93*(27), e150. doi:10.1097/MD.0000000000000150

Manly, J. J., Tang, M. X., Schupf, N., Stern, Y., Vonsattel, J. P., & Mayeux, R. (2008). Frequency and course of mild cognitive impairment in a multiethnic community. *Ann Neurol, 63*(4), 494-506. doi:10.1002/ana.21326

McEvoy, L. K., Holland, D., Donald J. Hagler, J., Fennema-Notestine, C., Brewer, J. B., & Dale, A. M. (2011). Mild Cognitive Impairment: Baseline and Longitudinal Structural MR Imaging Measures Improve Predictive Prognosis1. *radiology*. doi:10.1148/radiol.11101975/-/DC1

Medaglia, J. D., Pasqualetti, F., Hamilton, R. H., Thompson-Schill, S. L., & Bassett, D. S. (2017). Brain and cognitive reserve: Translation via network control theory. *Neurosci Biobehav Rev, 75*, 53-64. doi:10.1016/j.neubiorev.2017.01.016

Minhas, S., Khanum, A., Riaz, F., Alvi, A., & Khan, S. A. (2017). A Nonparametric Approach for Mild Cognitive Impairment to AD Conversion Prediction: Results on Longitudinal Data. *IEEE J Biomed Health Inform, 21*(5), 1403-1410. doi:10.1109/JBHI.2016.2608998

Mitchell, A. J., & Shiri-Feshki, M. (2009). Rate of progression of mild cognitive impairment to dementia - meta-analysis of 41 robust inception cohort studies. *Acta Psychiatrica Scandinavica, 119*(4), 252-265. doi:10.1111/j.1600-0447.2008.01326.x

Moradi, E., Pepe, A., Gaser, C., Huttunen, H., Tohka, J., & Alzheimer's Disease Neuroimaging, I. (2015). Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *Neuroimage, 104*, 398-412. doi:10.1016/j.neuroimage.2014.10.002

Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C., Jagust, W., . . . Beckett, L. (2005). The Alzheimer's Disease Neuroimaging Initiative. *Neuroimaging Clin N Am, 15*(4), 869-xii.

Neuropathology Group of the Medical Research Council Cognitive Function and Aging Study. (2001). Pathological correlates of late-onset dementia in a multicentre, community-based population in England and Wales. Neuropathology Group of the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). *Lancet, 357*(9251), 169-175. doi:10.1016/s0140-6736(00)03589-3

Noordzij, M., Tripepi, G., Dekker, F. W., Zoccali, C., Tanck, M. W., & Jager, K. J. (2010). Sample size calculations: basic principles and common pitfalls. *Nephrol Dial Transplant, 25*(5), 1388-1393. doi:10.1093/ndt/gfp732

O'Brien, J. T., & Thomas, A. (2015). Vascular dementia. *The Lancet, 386*(10004), 1698-1706. doi:10.1016/s0140-6736(15)00463-8

Pandya, S. Y., Clem, M. A., Silva, L. M., & Woon, F. L. (2016). Does mild cognitive impairment always lead to dementia? A review. *J Neurol Sci, 369*, 57-62. doi:10.1016/j.jns.2016.07.055

Park, M. H., Han, C., & Initiative, t. A. s. D. N. (2015). Is there an MCI reversion to cognitively normal? Analysis of Alzheimer's disease biomarkers profiles. *international Psychiatrics*

, 27(3), 429-437. doi:10.1017/S1041610214002129

Petersen, R. C., Aisen, P. S., Beckett, L. A., Donohue, M. C., Gamst, A. C., Harvey, D. J., . . . Weiner, M. W. (2010). Alzheimer's Disease Neuroimaging Initiative (ADNI) Clinical characterization. *Neurology, 74*(3), 201-209. doi:DOI 10.1212/WNL.0b013e3181cb3e25

Power, M., Fell, G., & Wright, M. (2013). Principles for high-quality, high-value testing. *Evid Based Med, 18*(1), 5-10. doi:10.1136/eb-2012-100645

Prins, N. D., van der Flier, W. A., Knol, D. L., Fox, N. C., Brashear, H., Nye, J. S., . . . Scheltens, P. (2014). The effect of galantamine on brain atrophy rate in subjects with mild cognitive impairment is modified by apolipoprotein E genotype: post-hoc analysis of data from a randomized controlled trial. *Alzheimers Res Ther, 6*(4), 47. doi:10.1186/alzrt275

Rabin, L. A., Wang, C., Katz, M. J., Derby, C. A., Buschke, H., & Lipton, R. B. (2012). Predicting Alzheimer's disease: neuropsychological tests, self-reports, and informant reports of cognitive difficulties. *J Am Geriatr Soc, 60*(6), 1128-1134. doi:10.1111/j.1532-5415.2012.03956.x

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, 155*, 530-548. doi:10.1016/j.neuroimage.2017.03.057

Ravaglia, G., Forti, P., Montesi, F., Lucicesare, A., Pisacane, N., Rietti, E., . . . Mecocci, P. (2008). Mild cognitive impairment: epidemiology and dementia risk in an elderly Italian population. *J Am Geriatr Soc, 56*(1), 51-58. doi:10.1111/j.1532-5415.2007.01503.x

Reiman, E. M., Langbaum, J. B., Fleisher, A. S., Caselli, R. J., Chen, K., Ayutyanont, N., . . . Tariot, P. N. (2011). Alzheimer's Prevention Initiative: a plan to accelerate the evaluation of presymptomatic treatments. *J Alzheimers Dis, 26 Suppl 3*, 321-329. doi:10.3233/JAD-2011-0059

Ritchie, C., Smailagic, N., Noel-Storr, A. H., Ukoumunne, O., Ladds, E. C., & Martin, S. (2017). CSF tau and the CSF tau/ABeta ratio for the diagnosis of Alzheimer's disease

dementia and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database Syst Rev, 3*, CD010803. doi:10.1002/14651858.CD010803.pub2

Roberts, R. O., Knopman, D. S., Mielke, M. M., Cha, R. H., Pankratz, V. S., Christianson, T. J., . . . Petersen, R. C. (2014). Higher risk of progression to dementia in mild cognitive impairment cases who revert to normal. *Neurology, 82*(4), 317-325. doi:10.1212/WNL.0000000000000055

Rowe, C. C., Ng, S., Ackermann, U., Gong, S. J., Pike, K., Savage, G., . . . Villemagne, V. L. (2007). Imaging beta-amyloid burden in aging and dementia. *Neurology, 68*(20), 1718-1725. doi:10.1212/01.wnl.0000261919.22630.ea

Schaffler, N., Kopke, S., Winkler, L., Schippling, S., Inglese, M., Fischer, K., & Heesen, C. (2011). Accuracy of diagnostic tests in multiple sclerosis--a systematic review. *Acta Neurol Scand, 124*(3), 151-164. doi:10.1111/j.1600-0404.2010.01454.x

Schneider, J. A., Arvanitakis, Z., Bang, W., & Bennett, D. A. (2007). Mixed brain pathologies account for most dementia cases in community-dwelling older persons. *Neurology, 69*(24), 2197-2204. doi:10.1212/01.wnl.0000271090.28148.24

Schneider, J. A., Arvanitakis, Z., Leurgans, S. E., & Bennett, D. A. (2009). The neuropathology of probable Alzheimer disease and mild cognitive impairment. *Ann Neurol, 66*(2), 200-208. doi:10.1002/ana.21706

Silva, D., Guerreiro, M., Maroco, J., Santana, I., Rodrigues, A., Bravo Marques, J., & de Mendonca, A. (2012). Comparison of four verbal memory tests for the diagnosis and predictive value of mild cognitive impairment. *Dement Geriatr Cogn Dis Extra, 2*, 120-131. doi:10.1159/000336224

Silva, D., Guerreiro, M., Santana, I., Rodrigues, A., Cardoso, S., Maroco, J., & de Mendonca, A. (2013). Prediction of long-term (5 years) conversion to dementia using neuropsychological tests in a memory clinic setting. *J Alzheimers Dis, 34*(3), 681-689. doi:10.3233/JAD-122098

Singh, P. P., Singh, M., & Mastana, S. S. (2006). APOE distribution in world populations with new data from India and the UK. *Ann Hum Biol, 33*(3), 279-308. doi:10.1080/03014460600594513

Smith, A. D., Smith, S. M., Jager, C. A. d., Whitbread, P., Johnston, C., Agacinski, G., . . . Refsum, H. (2010). Homocysteine-Lowering by B Vitamins Slows the Rate of Accelerated Brain Atrophy in Mild Cognitive Impairment: A Randomized Controlled Trial. *PLoS One, 5*(9). doi:10.1371/journal.pone.0012244

Standring, S. (2008). Gray's Anatomy The Anatomical Basis of Clinical Practice, 40th Ed *40th*.

Staszewski, J., Piusinska-Macoch, R., Brodacki, B., Skrobowska, E., Macek, K., & Stepien, A. (2017). Vascular parkinsonism and vascular dementia are associated with an increased risk of vascular events or death. *Arch Med Sci Atheroscler Dis, 2*(1), e16-e23. doi:10.5114/amsad.2017.68549

Steffener, J., & Stern, Y. (2012). Exploring the neural basis of cognitive reserve in aging. *Biochim Biophys Acta, 1822*(3), 467-473. doi:10.1016/j.bbadis.2011.09.012

Steinberg, D. M., Fine, J., & Chappell, R. (2009). Sample size for positive and negative predictive value in diagnostic research using case-control designs. *Biostatistics, 10*(1), 94-105. doi:10.1093/biostatistics/kxn018

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia, 47*(10), 2015-2028. doi:10.1016/j.neuropsychologia.2009.03.004

Sun, M. K. (2018). Potential Therapeutics for Vascular Cognitive Impairment and Dementia. *Curr Neuropharmacol, 16*(7), 1036-1044. doi:10.2174/1570159X15666171016164734

Suppa, P., Hampel, H., Kepp, T., Lange, C., Spies, L., Fiebach, J. B., . . . Alzheimer's Disease Neuroimaging, I. (2016). Performance of Hippocampus Volumetry with FSL-FIRST for Prediction of Alzheimer's Disease Dementia in at Risk Subjects with Amnestic Mild Cognitive Impairment. *J Alzheimers Dis, 51*(3), 867-873. doi:10.3233/JAD-150804

The American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC.

Toledo, J. B., Cairns, N. J., Da, X., Chen, K., Carter, D., Fleisher, A., . . . Alzheimer's Disease Neuroimaging, I. (2013). Clinical and multimodal biomarker correlates of ADNI neuropathological findings. *Acta Neuropathol Commun, 1*, 65. doi:10.1186/2051-5960-1-65

Trevethan, R. (2017). Sensitivity, Specificity, and Predictive Values: Foundations, Pliabilities, and Pitfalls in Research and Practice. *Front Public Health, 5*, 307. doi:10.3389/fpubh.2017.00307

Wegiel, J., Wisniewski, H. M., Dziewiatkowski, J., Badmajew, E., Tarnawski, M., Reisberg, B., . . . Miller, D. C. (1999). Cerebellar atrophy in Alzheimer's disease-clinicopathological correlations. *Brain Res, 818*(1), 41-50.

Wei, R., Li, C., Fogelson, N., & Li, L. (2016). Prediction of Conversion from Mild Cognitive Impairment to Alzheimer's Disease Using MRI and Structural Network Features. *Front Aging Neurosci, 8*, 76. doi:10.3389/fnagi.2016.00076

Weiner, M. W., Veitch, D. P., Aisen, P. S., Beckett, L. A., Cairns, N. J., Cedarbaum, J., . . . Alzheimer's Disease Neuroimaging, I. (2015). Impact of the Alzheimer's Disease Neuroimaging Initiative, 2004 to 2014. *Alzheimers Dement, 11*(7), 865-884. doi:10.1016/j.jalz.2015.04.005

Weiner, M. W., Veitch, D. P., Aisen, P. S., Beckett, L. A., Cairns, N. J., Green, R. C., . . . Alzheimer's Disease Neuroimaging, I. (2017). The Alzheimer's Disease Neuroimaging Initiative 3: Continued innovation for clinical trial improvement. *Alzheimers Dement, 13*(5), 561-571. doi:10.1016/j.jalz.2016.10.006

Werner, P. (2012). Mild Cognitive Impairment and Caregiver Burden: A Critical Review and Research Agenda. *Public Health Reviews*.

Whiting, P., Harbord, R., Main, C., Deeks, J. J., Filippini, G., Egger, M., & Sterne, J. A. (2006). Accuracy of magnetic resonance imaging for the diagnosis of multiple sclerosis: systematic review. *BMJ, 332*(7546), 875-884. doi:10.1136/bmj.38771.583796.7C

Yiannopoulou, K. G., & Papageorgiou, S. G. (2013). Current and future treatments for Alzheimer's disease. *Therapeutic Advances in Neurological Disorders, 6*(1), 19-33. doi:10.1177/1756285612461679

Zhang, D., Shen, D., & Alzheimer's Disease Neuroimaging, I. (2012). Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers. *PLoS One, 7*(3), e33182. doi:10.1371/journal.pone.0033182

Zhang, Y. P., Miao, R., Li, Q., Wu, T., & Ma, F. (2017). Effects of DHA Supplementation on Hippocampal Volume and Cognitive Function in Older Adults with Mild Cognitive Impairment: A 12-Month Randomized, Double-Blind, Placebo-Controlled Trial. *J Alzheimers Dis, 55*(2), 497-507. doi:10.3233/JAD-160439

APPENDICES

Printed copies of published manuscripts

ELSEVIER

Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 1 (2015) 487-504

Alzheimer's & Dementia

Neuroimaging

Cerebral atrophy in mild cognitive impairment: A systematic review with meta-analysis

Hossein Tabatabaei-Jafari*, Marnie E. Shaw, Nicolas Cherbuin

Centre for Research on Ageing, Health and Wellbeing, Australian National University, Canberra, ACT, Australia

Abstract

Introduction: Although mild cognitive impairment (MCI) diagnosis is mainly based on cognitive assessment, reliable estimates of structural changes in specific brain regions, that could be contrasted against normal brain aging and inform diagnosis, are lacking. This study aimed to systematically review the literature reporting on MCI-related brain changes.

Methods: The MEDLINE database was searched for studies investigating longitudinal structural changes in MCI. Studies with compatible data were included in the meta-analyses. A qualitative review was conducted for studies excluded from meta-analyses.

Results: The analyses revealed a 2.2-fold higher volume loss in the hippocampus, 1.8-fold in the whole brain, and 1.5-fold in the entorhinal cortex in MCI participants.

Discussion: Although the medial temporal lobe is likely to be more vulnerable to MCI pathology, atrophy in this brain area represents a relatively small proportion of whole brain loss, suggesting that future investigations are needed to identify the source of unaccounted volume loss in MCI.

Keywords: Mild cognitive impairment; Brain atrophy; Hippocampus; Entorhinal cortex; MRI

1. Introduction

Although Alzheimer's disease (AD) was first characterized more than 100 years ago, little concrete progress has been made toward an effective cure of this progressive disorder. Identification of mild cognitive impairment (MCI) as a prodromal phase of AD has raised hopes of the possibility of preventing or modifying progressive neurodegeneration in AD. Indeed, initial attempts at early therapeutic interventions have reported some successes in the early phase of MCI [1,2].

Clinically, MCI is defined based on the detection of cognitive decline greater than that expected at any given age and less than that observed in dementia in the context of preserved activities of daily living and the absence of other neurological disorders. However, clinical evaluation is complicated by heterogeneity in cognitive reserve and diversity in daily function. Considering that each cognitive measure is designed to target a particular brain function, selecting which cognitive measures are appropriate to assess functional decline in the MCI trajectory is a matter of concern not only for diagnostic purposes but also in the evaluation of clinical trials [3]. Besides higher uncertainty in characterizing MCI based on functional impairment [4], cognitive evaluation is not currently informative enough for demonstrating patterns of deterioration that will accurately discriminate those who will remain stable from those who will convert to AD or other dementias. Therefore, without a better understanding of the neurologic basis of the disorder, as well as the identification of structural biomarkers, reliable detection of MCI and estimation of future risk of dementia remain elusive.

Assuming that impairment in cognitive function is the result of neurodegeneration, monitoring structural brain changes may be beneficial in understanding the pathophysiology of MCI. Recent development in neuroimaging

The authors have reported no conflicts of interest. This study is not industry sponsored.

*Corresponding author. Tel.: +61-2-6125-9032; Fax: +61-2-6125-1558.

E-mail address: hossein.tabatabaei@anu.edu.au

http://dx.doi.org/10.1016/j.dadm.2015.11.002

technologies has provided an opportunity to investigate structural biomarkers in living subjects. In the past two decades, the use of magnetic resonance imaging (MRI) to assess cerebral structure has become widespread. Most early studies have used a cross-sectional design and have suggested that, although the presence of structural differences in any particular brain area is not specific to MCI or AD (i.e. it can also be observed in "normal" aging), the pattern of regional atrophy rates and the topological progression of atrophy are quite characteristic, particularly in AD [5]. Moreover, these studies also revealed that regional atrophy rates are different in MCI and AD [6]. Consequently, identification of regionally specific atrophy rates in MCI may be beneficial for detecting the early stage of AD development, as well as evaluating the magnitude of expected structural changes in clinical trials.

Available longitudinal studies have identified a subset of brain areas that may be involved in MCI pathology. An important next step is to combine, contrast, and integrate the findings from different studies to produce normative information on regional atrophy rates, and to identify the most sensitive anatomic biomarkers characteristic for MCI. As far as we are aware, no study has systematically summarized these findings to date. Therefore, the aim of this study was to systematically review the literature concerning MCI-related structural brain changes.

2. Methodology

This systematic review was conducted based on an established methodology [7], using prespecified search terms and inclusion and exclusion criteria, and was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines [8].

To retrieve all references relating to longitudinal brain structural changes in MCI published in the MEDLINE database, a literature search was conducted through the PubMed portal in two stages, (1) at the beginning of the study (2) and at the end of February 2015 to update pooled data with the most recent published studies. The following search string was used for both searches; *(Brain or Cerebral or Cortical) And (Mild Cognitive Impairment Or MCI Or Cognitive disorder Or Neurocognitive disorder Or Cognitive decline Or Cognition) And (Structur* Or Volum* Or Thickness Or MRI Or Neuroimaging) And (Atrophy Or Change Or Longitudinal Or shrinkage)*. Both literal and Medical Subject Heading searches were performed. Searches were limited to studies published in English and focusing on human subjects.

2.1. Selection criteria and selection process

To be selected, studies were required to use a longitudinal methodology with two or more structural MRI scans conducted over a follow-up of 12 months or more. As MCI status defined the group being compared with healthy controls (HC), cognitive status of HC and MCI was required to be stable between all time points. Studies were required to use Peterson or Winblad criteria for MCI diagnosis. Cross-sectional, experimental, and review articles were excluded. Studies were also excluded if they had a combined total of less than 30 HC and MCI participants. All retrieved articles were first screened by title and abstract and irrelevant studies were excluded. The full text of all remaining articles was double screened by two reviewers (H.T.-J. and M.E.S.) against selection criteria.

2.2. Data extraction and structural measures

Two reviewers (H.T.-J. and M.E.S.) extracted data from all included articles and any disagreement was resolved by consensus. Data extracted consisted of (1) study design including sample source, number of participants in each group, type of structural measurement, and follow-up period; (2) participants' demographics including age, gender ratio, *APOE* ε4 ratio, years of education, dropout rate, MCI subtype for MCI groups, subjective memory complaint for HC, and handedness; (3) measurement details including number of scans, scan intervals, follow-up period, MRI parameters, segmentation method, and method of analysis; and (4) study results including areas of interest (left and right) and effect sizes (left, right, and total).

All structural measures were evaluated, and studies were categorized according to the following structural measurements; voxel-based morphometry (VBM), volumetry, tensor-based morphometry (TBM), cortical thickness, sulcal morphometry, diffusion tensor imaging (DTI), white matter hyperintensities (WMH), susceptibility weighted imaging (SWI), and other structural measures.

Studies meeting the selection criteria were assessed for quality using the Newcastle-Ottawa scale [9]. The Newcastle-Ottawa scale is an instrument for assessing the quality of studies included in a systematic review. Each study was evaluated on eight items classified into three categories including the selection of the study groups, the comparability of the groups, and the ascertainment of outcome of interest. Each quality item was awarded by a star (except two for comparability) and for each study up to nine stars in total.

2.3. Multiple reports

In the case of multiple reports for the same cohort, or any overlap of participants, an annual change rate estimate from only one publication was used in any single analysis. The most appropriate reports were selected based on recency, availability of effect size and moderators, sample size, and methodology. Studies that reported effect sizes (or provided them after contact) were the first priority and from those the most recent study with the largest sample size was selected. If there was more than one study similar in sample size and recency, the one with the highest quality rating was selected.

When different studies on the same cohort reported on different brain areas, estimates from the same cohort but from different studies might be used in different analyses.

2.4. Statistical analysis

The R statistical software (version 3.1.1) was used for the statistical analysis, and the metafor package (version 1.9-4) was used for meta-analysis. The annual percentage mean atrophy rate was considered as the effect size, and calculation of required standard error (SE) for meta-analysis was based on the standard deviation and number of participants in each group for each individual study. Availability of mean annual atrophy rate (%/year), either reported or computed based on other reported results, was the essential requirement for the meta-analysis. Where insufficient data were available for inclusion in the meta-analysis, authors were contacted directly to seek additional information.

2.4.1. Meta-analysis

It was assumed that the heterogeneity in the atrophy rates across reviewed studies was the impact of the between-study and within-study heterogeneities, and the random effects for between- and within-studies were normally distributed. A random-effects model using the restricted maximum likelihood estimator was applied for all analyses. Random-effects model was chosen based on the assumption that cerebral atrophy rates (effect size) are not similar in population with different characteristics and there is no single effect size representative of all population but an array of effect sizes. Therefore, each included study was assumed to represent a random sample of a particular effect size and a random-effects model estimates a mean of the distribution of these effect sizes [10]. Separate meta-analyses were performed for healthy and MCI atrophy rates and also for the mean difference in atrophy rate between MCI and healthy controls (MCI-HC) across each cerebral region.

Heterogeneity across studies was assessed with the Q and I^2 statistics. P value $<.01$ considered as significant heterogeneity in the Q test and in the I^2 statistic values of 25%, 50%, and 75% were suggestive of low, moderate, and high heterogeneity, respectively. Heterogeneity in the atrophy rates was also assumed to be in part the result of disparities in age, sex ratio, *APOE* ε4 ratio, and education levels in the studies' participants as well as scan intervals and different segmentation approaches. Therefore, these variables were investigated as possible moderators for subgroup and meta-regression analyses. Subgroup analyses were conducted to investigate the impact of manual versus automated segmentations. Meta-regression analyses using a mixed-effect model were conducted to determine the influence of moderators.

To identify studies contributing excessively to heterogeneity, sensitivity analyses were conducted using the leave-one-out method. Visual evaluation of asymmetry of the funnel plots was used to assess the bias in the meta-analyses results toward publication of studies with significant outcomes. The trim-and-fill method was used to estimate the number of missing studies (representative of unreported effect sizes) in the meta-analysis to estimate adjusted effect sizes [11].

3. Results

3.1. Literature search and studies included in the review

The search strategy identified 5220 unique citations. After exclusion of irrelevant studies based on title and abstracts, 219 publications remained for full-text assessment. A further 151 studies did not meet the inclusion criteria and were excluded leaving 68 studies for further analysis (Fig. 1).

Of the studies included, 45 assessed brain structure with volumetry, nine with cortical thickness, and 18 with a wide variety of structural measurements including sulcal morphometry, VBM, TBM, DTI, WMH, SWI, and quantitative scaling methods such as the medial temporal atrophy scale (MTAS) [12] and the brain atrophy and legion index (BALI) [13] (Table 1).

3.2. Study quality

All studies except one, which was rated 6 [14], were rated as high quality (eight or nine stars) based on the Newcastle-Ottawa scale (Table 1). Fifty-four of 68 studies fulfilled the maximum of nine stars, two studies were rated as not representative of the population due to a higher rate of medical diseases in the participants, and one study did not describe the derivation of the HC. Twelve studies only controlled for age to establish comparability between controls and MCI participants.

3.3. Multiple reports

A number of multiple reports were identified. Forty-six studies reported on participants taking part in the Alzheimer's Disease Neuroimaging Initiative (ADNI; to date up to 229 HC and 395 MCI), four studies used Mayo AD research center and AD patient registry data (up to 91 HC and 72 MCI), and one study used a mixture of ADNI and Mayo data. There was also an overlap of participants in two studies reported by Henneman et al. [15,16]. A total of 15 publications reported on separate independent cohorts including in total 629 HC and 571 MCI participants from 10 countries across four continents (eight in Europe, five in North America, one in Asia, and one in Australia).

3.4. Compatible studies for meta-analysis

A sufficient number of compatible studies was only available for meta-analysis of volumetric measurements. Quantitative report of structural measures in VBM and TBM studies were not comparable. Brain areas investigated by

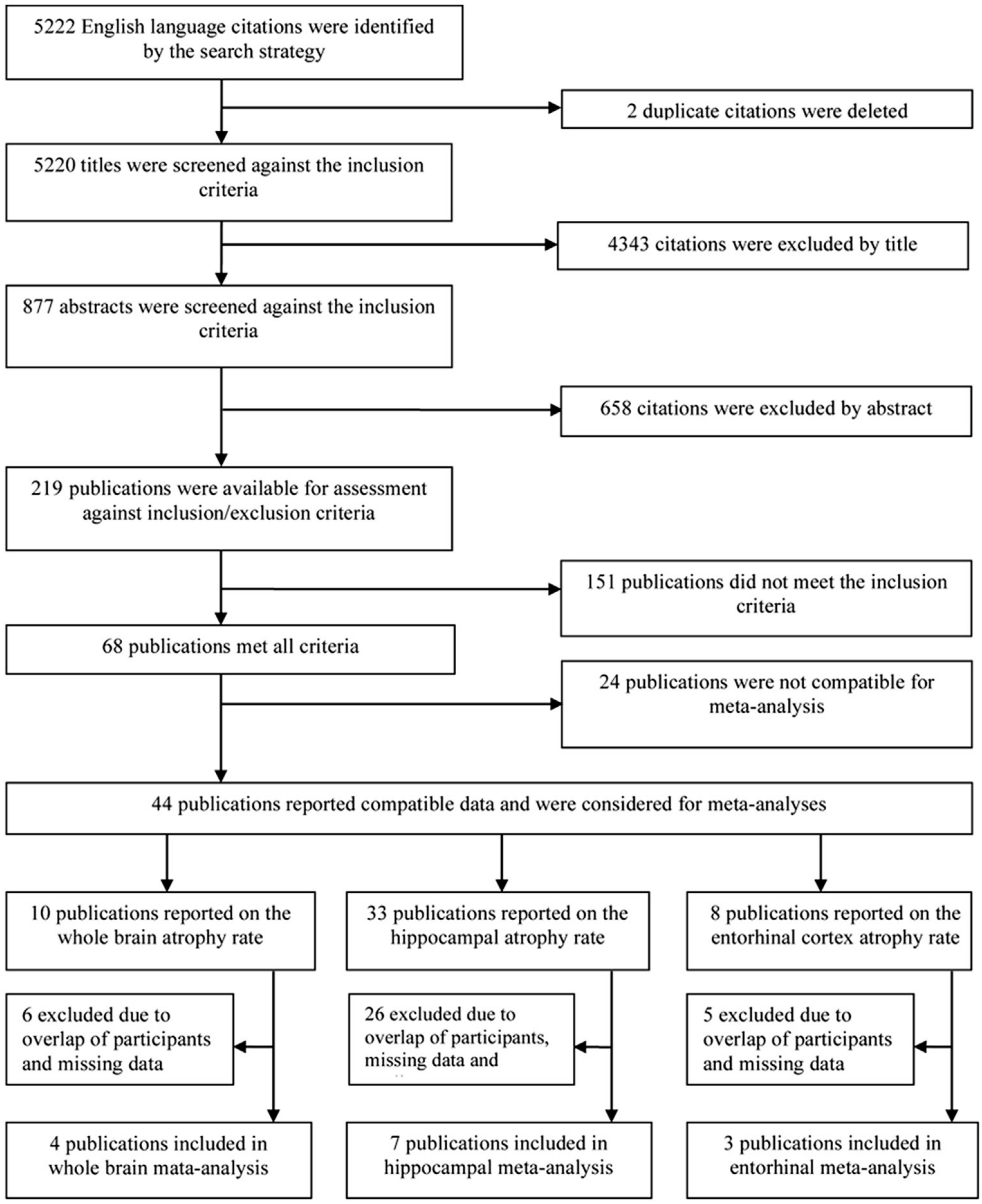


Fig. 1. Screening and selection process for studies included in the systematic review and the meta-analyses.

cortical thickness or DTI studies were not anatomically compatible. There was only one study in each given category of sulcal morphometry, WMH, and SWI. Finally, studies using MTAS and BALI scales were all based on the same cohort except for one study (Table 1). Therefore, of the 68 studies that met the selection criteria, 24 studies could not be included in the meta-analyses, leaving 44 volumetric studies for inclusion. Because too few sporadic reports of laterality were available, this factor could not be investigated. There was also no report of handedness.

Volumetric studies evaluated a wide variety of brain regions including the whole brain, hippocampus, entorhinal cortex, ventricles, parahippocampal gyrus, amygdala, fusiform gyrus, superior temporal, medial lateral and inferior temporal lobes, medial and lateral orbitofrontal cortex, superior frontal cortex, cingulate cortex, and parietal and occipital lobes. Besides the first four measures, other brain areas were investigated sporadically. Three of 44 studies evaluated brain areas incompatible with other studies and were not considered for meta-analysis. Forty-one studies were identified as potentially compatible and were included in meta-analyses. These studies evaluated annual atrophy rate of the whole brain (n = 10), the hippocampus (n = 33), and the entorhinal cortex (n = 10), as well as annual expansion rate of the ventricles (n = 14).

Of 41 studies, 29 were excluded because of overlap in participants and one because of missing data which could not be obtained from authors (Table 1). Although three

Table 1
Studies included in the review

#	Study	Address	Measurement	Cohort	Newcastle-Ottawa quality assessment scale†								Compatible for meta-analysis		
					Selection				Comparability	Outcome					
					Q-1	Q-2	Q-3	Q-4	Q-5	Q-6	Q-7	Q-8	Yes/no	In meta-analysis	Details
1	Madsen et al.	Neurobiology of aging 36(2015)532–541	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Ventricle meta-analysis due to overlap of participants
2	Lorenzi et al.	Neurobiology of Aging 36(2015)542–552	SVF	ADNI	*	*	*	*	**	*	*	*	No	No	No quantitative structural measure
3	Toledo et al.	Acta Neuropathol 127(2014)621–632	???	ADNI	*	*	*	*	**	—	*	*	No	No	Incomplete report of structural data
4	Teiple et al.	Neurobiology of aging 35(2014)482–491	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
5	Mulder et al.	Neurology 92(2014) 169–181	Volumetry	ADNI	*	*	*	*	*	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
6	Marshal et al.	Journal of Alzheimer's disease 41(2014) 719–728	Cortical thickness	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible brain area with other studies
7	Manning et al.	PLOS ONE May (2014) Vol 9/issue5e97608	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
8	Lilemark et al.	BMC Medical imaging (2014)14–21	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	WB and Hip meta-analyses due to overlap of participants
9	Kljajevic et al.	Neurobiology of Aging 35(2014)1973–1981	Volumetry, cortical thickness	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible brain area with other studies
10	Insel et al.	Alzheimer's & Dementia (2014)1–9	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip and ERC meta-analyses due to overlap of participants
11	Guo et al.	Journal of Alzheimer's Disease 42(2014) 691–703	BALI	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible with independent study
12	Aguilar et al.	Frontiers in Aging Neuroscience July 2014/Vol6/Article 145	Volumetry, cortical thickness	AddNeuroMed	*	*	*	*	**	*	*	*	Yes	No	Missing and mismatch data
13	Nowrangi et al.	Alzheimer's & Dementia 9(2013)519–528	DTI	Community-dwelling volunteers	*	*	*	*	*	*	*	*	No	No	Incompatible brain area with other studies
14	Guo et al	Alzheimer's & Dementia 9(2013)580–586	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	WB meta-analysis due to overlap of participants
15	Franko et al.	PLOS ONE Aug. (2014) Vol 8/issue 8/e71354	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
16	Adaszewski et al.	Neurobiology of Aging 34(2013)2815–2826	VBM, SVM	ADNI	*	*	*	*	*	*	*	*	No	No	No quantitative structural measures
17	Villemagne et al.	Lansent Neural 12(2013) 357–367	Volumetry	AIBL	*	*	*	*	**	*	*	*	Yes	Yes	Hip meta-analysis

(*Continued*)

Table 1
Studies included in the review (*Continued*)

					Newcastle-Ottawa quality assessment scale[†]								Compatible for meta-analysis		
					Selection				Comparability	Outcome					
#	Study	Address	Measurement	Cohort	Q-1	Q-2	Q-3	Q-4	Q-5	Q-6	Q-7	Q-8	Yes/no	In meta-analysis	Details
18	Song et al.	J Neurosurg Psychiatry 84(2013)71–78	MTAS, BALI	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible with independent study
19	Selnes et al.	Journal of Alzheimer's Disease 33(2013) 723–739	DTI	Memory clinics	*	*	*	*	**	*	*	*	No	No	Incompatible brain area with other studies
20	Liu et al.	NeuroImage 74(2013) 337–342	Sulcal morphology, cortical thickness	MAS	*	*	*	*	*	*	*	*	No	No	Incompatible brain area with other studies
21	Gutman et al.	NeuroImage 70(2013) 386–401	Volumetry	ADNI	*	*	*	*	*	*	*	*	Yes	No	Ventricles meta-analysis due to overlap of participants
22	Zhang et al.	Dement Geriat Cogn Disord 33(2012) 318–326	MTAS, BALI	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible with independent study
23	Yao et al.	PLOS ONE (2012) Vol 7/Issue 11/e48973	Cortical thickness	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible brain area with other studies
24	Schuff et al.	Neurobiology of Aging 33(2012)845–855	Volumetry	ADNI	*	*	*	*	*	*	*	*	Yes	Yes	ERC meta-analysis
25	McDonald et al.	Neurobiology of Aging 33(2012)242–253	Volumetry	ADNI	*	*	*	*	**	*	*	*	No	No	Due to mismatch of brain areas
26	Li et al.	Neurobiology of Aging 33(2012) 427 e15–30	Cortical thickness	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible brain area with other studies
27	Leung et al.	NeuroImage 59(2012) 3995–4005	Volumetry	ADNI	*	*	*	*	*	*	*	*	No	No	Mismatch data
28	Andrawis et al.	Neurobiology of Aging 33(2012)856–866	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
29	Zhang et al.	Journal of Alzheimer's Disease 26(2011) 359–367	BALI	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible with independent study
30	Tosun et al.	Journal of Alzheimer's Disease 26(2011) 77–90	Cortical thickness	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible brain area with other studies
31	Skup et al.	NeuroImage 56(2011) 890–906	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip and ERC meta-analyses due to overlap of participants
32	Mouiha et al.	Neuroscience Letters 495(2011)6–10	Volumetry	ADNI	*	*	*	*	*	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
33	Lo et al.	Arch Neurol. Oct. (2011) Vol 68, No. 10	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
34	Desikan et al.	Ann Neurol 70(2011) 657–661	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	ERC meta-analysis due to overlap of participants

(*Continued*)

Table 1
Studies included in the review (*Continued*)

#	Study	Address	Measurement	Cohort	Newcastle-Ottawa quality assessment scale[†]								Compatible for meta-analysis		
					Selection				Comparability	Outcome					
					Q-1	Q-2	Q-3	Q-4	Q-5	Q-6	Q-7	Q-8	Yes/no	In meta-analysis	Details
35	Chiang et al.	Alzheimer's & Dementia 7(2011)514–520	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
36	Villain et al.	Brain 133(2010) 3301–3314	VBM	Memory Clinic	*	*	*	*	**	*	*	*	No	No	No quantitative structural measure
37	Vemuri et al.	Neurology 75(2010) 143–151	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	WB meta-analysis due to overlap of participants
38	Tosun et al.	Neurobiology of Aging 31(2010)1340–1354	Volumetry, cortical thickness	ADNI	*	*	*	*	**	*	*	*	Yes	No	Ventricle, hip and ERC meta-analyses due to overlap of participants
39	Stoub et al.	Neurobiology of Aging 31(2010)1089–1098	Volumetry	RADC & ROS and MAP	*	*	*	*	**	*	*	*	Yes	No	Hip and ERC meta-analyses due to overlap of participants
40	Schott et al.	Neurobiology of Aging 31(2010)1452–1462	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	Yes	WB and hip meta-analyses
41	Prestia et al.	Journal of Alzheimer's Disease 22(2010) 1339–1349	VBM	TOMC	*	*	*	*	**	*	*	*	No	No	No quantitative structural measure
42	Leung et al.	NeuroImage 51(2010) 1345–1359	Volumetry	ADNI	*	*	*	*	*	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
43	Hua et al.	Neurobiology of Aging 31(2010)1463–1480	TBM	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible brain area with other studies
44	Ho et al.	Human Brain Mapping 31(2010)499–514	TBM	ADNI	*	*	*	*	*	*	*	*	No	No	Incompatible brain area with other studies
45	Evans et al.	Eur Radiol. 20(2010) 674–682	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	WB and ventricle meta-analyses due to overlap of participants
46	Desikan et al.	PLOS ONE (2010) Vol 5/Issue 9/e12853	Volumetry, Cortical Thickness	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
47	Carmichael et al.	Arch Neurol. 67(2010) 1370–1378	WMH	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible with independent study
48	Beckett et al.	Alzheimer's & Dementia 6(2010)257–264	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip and ventricle meta-analyses due to overlap of participants
49	Ayaz et al.	Journal of Magnetic Resonance Imaging 31(2010)142–148	SWI	???	—	—	*	*	*	*	*	*	No	No	Incompatible with independent study
50	Archer et al.	Int J Geriatr Psychiatry 25(2010)1119–1126	Volumetry	Hospital & memory clinic	*	*	*	*	**	*	*	*	Yes	Yes	WB and hip meta-analyses
51	Apostolova et al.	NeuroImage 51(2010) 488–499	Volumetry	ADNI	*	*	*	*	**	*	*	*	No	No	Hip meta-analysis due to overlap of participants

(*Continued*)

Table 1
Studies included in the review (*Continued*)

					Newcastle-Ottawa quality assessment scale[†]								Compatible for meta-analysis		
					Selection				Comparability	Outcome					
#	Study	Address	Measurement	Cohort	Q-1	Q-2	Q-3	Q-4	Q-5	Q-6	Q-7	Q-8	Yes/no	In meta-analysis	Details
52	Wang et al.	Psychiatric research neuroimaging 171(2009)221–231	Volumetry	Neurological clinic	*	*	*	*	**	*	*	*	Yes	Yes	Hip meta-analysis
53	Sluimer et al.	Eur Radio. 19(2009) 2826–2833	Volumetry	Memory clinic	*	*	*	*	**	*	*	*	No	No	Incompatible brain area with other studies
54	Schuff et al.	Brain 132(2009) 1067–1077	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
55	Morra et al.	NeuroImage 45(2009) s3–15	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants
56	Leow et al.	NeuroImage 45(2009) 645–655	TBM	ADNI	*	*	*	*	**	*	*	*	No	No	Incompatible brain area with other studies
57	Jack Jr. et al.	Brain 132(2009) 1355–1365	Volumetry	Mayo, ADNI	*	*	*	*	**	*	*	*	Yes	No	Ventricle meta-analysis due to overlap of participants
58	Hua et al.	NeuroImage 48(2009) 668–681	TBM	ADNI	*	*	*	*	*	*	*	*	No	No	Incompatible brain area with other studies
59	Holland et al.	PNAS (2009) Vol 106/No. 49/20,954–20,959	Volumetry	ADNI	*	*	*	*	**	*	*	*	Yes	No	WB, ventricle, hip and ERC meta-analyses due to overlap of participants
60	Henneman et al.	Neurology 73(2009) 935–940	Volumetry	Memory clinic	*	*	*	*	**	*	*	*	Yes	Yes	Hip meta-analysis
61	Henneman et al.	Neurology 72(2009) 999–1007	Volumetry	Memory clinic	*	*	*	*	**	*	*	*	Yes	Yes	WB meta-analysis
62	Brys et al.	Journal of Alzheimer's Disease 16(2009) 351–362	VBM, MTL-rBS	AD research center	*	*	*	*	**	*	*	*	No	No	Incompatible with independent study
63	Jack Jr. et al.	Neurology 70(2008) 1740–1752	Volumetry	Mayo	*	*	*	*	**	*	*	*	Yes	No	WB and ventricle meta-analyses due to overlap of participants
64	Eckerstrom et al.	Journal of the Neurological sciences 272(2008)48–59	Volumetry	Goteborg MCI study	—	*	*	*	**	*	*	*	Yes	Yes	Hip meta-analysis
65	Desikan et al.	Neurology 71(2008) 819–825	Volumetry	Community-dwelling volunteers	*	*	*	*	**	*	*	*	Yes	Yes	Hip and ERC meta-analyses
66	Jack Jr. et al.	Neurology 65(2005) 1227–1231	Volumetry	Mayo	*	*	*	*	**	*	*	*	Yes	Yes	WB, hip and ERC meta-analyses
67	Jack Jr. et al.	Neurology 62(2004) 591–600	Volumetry	Mayo	*	*	*	*	**	*	*	*	Yes	No	WB, ventricle, hip and ERC meta-analyses due to overlap of participants

(*Continued*)

Table 1
Studies included in the review (*Continued*)

#	Study	Address	Measurement	Cohort	Newcastle-Ottawa quality assessment scale[†]										
					Selection				Comparability	Outcome			Compatible for meta-analysis		
					Q-1	Q-2	Q-3	Q-4	Q-5	Q-6	Q-7	Q-8	Yes/no	In meta-analysis	Details
68	Jack Jr. et al.	Neurology 55(2000) 484–489	Volumetry	Mayo	*	*	*	*	**	*	*	*	Yes	No	Hip meta-analysis due to overlap of participants

Abbreviations: ADNI, Alzheimer's disease Neuroimaging Initiative; SVF, stationary velocity field; Hip, hippocampus; WB, whole brain; ERC, entorhinal cortex; BALI, brain atrophy and lesion index; AddNeuroMed, six European sites compatible with the US ADNI study; DTI, diffusion tensor imaging; VBM, voxel-based morphometry; SVM, support vector machine; AIBL, Australian imaging, biomarker, and lifestyle; MTAS, the medial temporal atrophy scale; MAS, Sydney memory aging study; RADC, Rush Alzheimer's Disease Center, ROS and MAP, Religious Order Study and Rush Memory and Aging Project; TOMC, The Transitional Outpatient Memory Clinic; TBM, tensor-based morphometry; WMH, white matter hyperintensities; SWI, susceptibility weighted imaging; MTL-rBS, medial temporal lobe atrophy using regional boundary shift; AD, Alzheimer's disease; MCI, mild cognitive impairment.

*A 'star system' for a quick visual assessment. Stars awarded for each quality item.

[†]Q-1; Representativeness of the exposed cohort, Q-2; Selection of the non-exposed cohort, Q-3; Ascertainment of exposure, Q-4; Demonstration of interest was not present at the start of the study, Q-5; Comparability of cohorts on the basis of the design or analysis, Q-6; Assessment of outcome, Q-7; Was follow-up long enough for outcomes to occur, Q-8; Adequacy of follow-up of cohorts.

studies were available for ventricle expansion analysis, reported expansion rates did not use the same units (mL/year vs. %/year) and requests for more information from authors was not successful. Therefore, meta-analysis could not be conducted for this region. Final numbers of studies included in the meta-analyses were four for whole brain, eight for hippocampal, and three for entorhinal cortex atrophy (Table 2).

3.4.1. Whole brain atrophy

Four studies [16–19], which were included for whole brain analysis (Fig. 2), surveyed 351 control and 466 MCI participants over an average follow-up of 1.30 years (range 1.00–1.80). Estimated mean atrophy rates were 1.02%/year (SE = 0.13) for MCI and 0.57%/year (SE = 0.03) for controls. Thus, the additional annual total brain atrophy attributable to MCI above the effect of "normal" aging was 0.46%/year (SE = 0.10). There was no significant heterogeneity (based on the Q test) for whole brain atrophy rates in HC and MCI after removing the effect attributable to normal aging (MCI-HC). The proportion of real observed variance (not related to random error) between studies (I^2) was moderate in MCI-HC and high in MCI (Table 3).

3.4.2. Hippocampal atrophy

Of eight studies [15,17–23], which were included for hippocampal meta-analysis, one study [22] reported an increase in hippocampal volume in MCI and a decrease in volume in HC as well as standard deviations larger than twice the mean atrophy rates. These characteristics were interpreted as being potentially methodologically problematic and after further investigation, the study was excluded from the meta-analysis because it was remarkably different in quality and design compared with other studies in the group, including gender proportion misbalance and high level of medical illness in the participants.

The remaining seven studies estimated hippocampal atrophy rates for 487 HC and 540 MCI participants with an average follow up of 1.97 years (range 1–3.8) (Fig. 2). The estimated mean atrophy rates were 2.53%/year (SE = 0.33) for MCI, 1.12%/year (SE = 0.16) for controls, and 1.35%/year (SE = 0.19) for MCI after removing the effect attributable to normal aging. Significant heterogeneity was found for hippocampal atrophy rates in MCI and MCI-HC but not in HC. The proportion of real observed variance (not related to random error) between studies (I^2) was moderate to high in all groups (Table 3).

3.4.3. Entorhinal cortex annual atrophy

Three studies [19,23,24], which were included for entorhinal cortex meta-analysis (Fig. 2), surveyed 257 controls and 258 MCI participants, followed up for 2.28 years (range 1.25–3.00). Estimated mean atrophy rates were 3.75%/year (SE = 1.60) for MCI and 2.41%/year (SE = 1.30) for HC. After removing the effect attributable to normal aging, the mean atrophy rate exclusively associated with MCI was 1.13%/year (SE = 0.33). Significant

Table 2
Studies included in meta-analyses

	Measurement					Participants		Age		Female %		*APOE* ε4 %		Change rate	
First author, year	WB	Hip	ERC	Vent*	Recruit	HC	MCI	HC	MCI	HC	MCI	HC	MCI	HC	MCI
Villemagne, 2013		✓			AIBL	112	32	71.2 (7.2)	74.2 (6.6)	48.21	43.75	46	65	−0.911 (1.15) %/y	−2.15 (1.33) %/y
Schuff, 2012			✓		ADNI	147	164	76 (5)	75 (7)	49.66	37.8	22	45	−1.6 (0.4) %/y	−2.4 (0.4)
Schott, 2010	✓				ADNI	199	334	76 (5.1)	74.9 (7.2)	46.73	36.53	28.64	53.3	−0.592 (0.581) %/y	−1.08 (0.84) %/y
		✓												−1.01 (1.72) %/y	−2.63 (2.35) %/y
				✓										−1.43 (1.63) mL/y	−285 (2.75) mL/y
Archer, 2010	✓				Clinic	27	16	62.3 (8.3)	67.1 (6.9)	51.85	31.25	18.5	75	−0.47 (0.67) %/y	−1 (0.81) %/y
		✓												−0.78 (0.91) %/y	−2.8 (1.68) %/y
				✓										−1.14 (1.73) mL/y	−3.62 (2.33) mL/y
Wang, 2009		✓			Clinic	20	39	75.1 (3.7)	75.6 (3.6)	45	20.51	20	26.5	−1 (0.7) %/y	−2.1 (1.5) %/y
Henneman, 2009a		✓			Clinic	19	25	66 (9)	71 (6)	42.11	56	47	71	−2 (1.5) %/y	−3.7 (1.2) %/y
Henneman, 2009b	✓				Clinic	34	44	67 (9)	71 (6)	47.06	47.72	_	_	−0.6 (0.6) %/y	−1.3 (0.9) %/y
Eckerstrom[†], 2008		✓			GMS	19	15	?	?	?	?	?	?	−0.168 (0.464) mL/y	+0.082 (0.329) mL/y
Desikan, 2008		✓			Media	19	22	69.7 (3.7)	70.1 (4.4)	63.16	59.1	31.6	31.8	−0.71 (0.88) %/y	−1.13 (1.01) %/y
			✓											−0.68 (1.4) %/y	−1.92 (2.12) %/y
Jack Jr, 2005	✓				MAYO	91	72	80.5 (?)	78.7 (?)	60.44	43.06	_	_	−0.5 (0.7) %/y	−0.7 (1) %/y
		✓												−1.7 (1.4) %/y	−3.3 (2.7) %/y
			✓											−5 (3.6) %/y	−7 (4.3) %/y
				✓										−2.4 (2) %/y	−3.3 (2.3) %/y

Abbreviations: WB, whole brain; hi, hippocampus; ERC, entorhinal cortex; vent, ventricles; HC, healthy controls; MCI, mild cognitive impairment; AIBL, Australian Imaging, Biomarker, and Lifestyle; ADNI, Alzheimer's Disease Neuroimaging Initiative; GMS, Goteborg MCI study; MAYO, Mayo AD research center and AD patient registry.

Measures provided as mean (standard deviation).

*Ventricular studies were not matched in atrophy rate unit and excluded from the meta-analyses.

[†]This study was an outlier and excluded from final hippocampal meta-analysis.

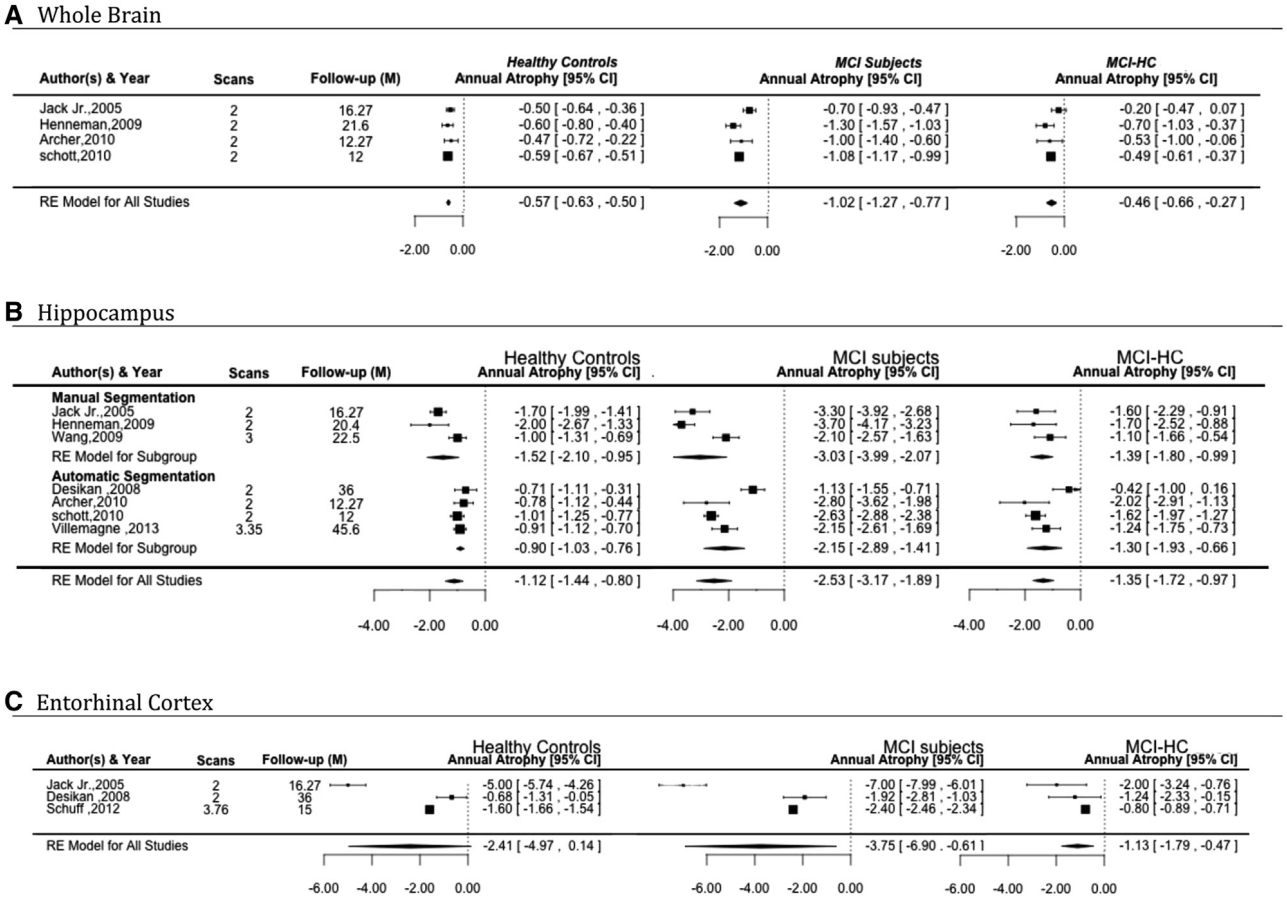


Fig. 2. Forest plots of atrophy rates for (A) whole brain, (B) hippocampus, and (C) entorhinal cortex in healthy controls, MCI, and the difference in atrophy rate between MCI and healthy controls (MCI-HC). Studies are ordered by year of publication. Abbreviations: MCI, mild cognitive impairment; CI, confidence interval.

heterogeneity was identified in entorhinal cortex atrophy rates in MCI and HC but not MCI-HC. The proportion of real observed variance (not related to random error) between studies (I^2) was moderate to high in all groups (Table 3).

3.4.4. Sensitivity analyses

The influence of single studies was investigated with leave-one-out analyses. Globally, the analysis revealed no particularly influential study and showed consistency in reported estimates.

3.4.5. Publication bias

Some evidence of publication bias was detected based on the funnel plot asymmetry diagnostic and the trim-and-fill test. The funnel plots revealed some degree of asymmetry for all three groups of analyses (the whole brain, hippocampus, and entorhinal), and the trim-and-fill method identified one or two missing studies in each analysis group. One missing study was identified in the whole brain and hippocampal analyses and two studies in entorhinal analysis, representing 20%, 12.5%, and 40% of included studies, respectively. Although asymmetry and presence of missing studies suggest some publication bias toward studies reporting higher atrophy rates, the differences between actual and reported atrophy rates were generally small, particularly for the hippocampus (Fig. 3).

3.4.6. Subgroup and meta-regression analyses

The influence of segmentation methods (automatic vs. manual), MCI subtype (amnestic MCI vs. MCI), female proportion, *APOE* ε4 genotype, and sample mean age on pooled estimates was investigated by subgroup meta-analyses and meta-regression on hippocampal volumetry only, as too few studies were available for other regions of interest (Table 3). Subgroup analyses showed that the estimated mean hippocampal atrophy rates in studies [15,19,21] using manual segmentation were significantly higher than studies [17,18,20,23] using automatic segmentation (Fig. 2 and Table 3) by 68% in HC, 40% in MCI, and 7% in MCI-HC. Additionally, subgroup analysis of MCI subtypes (amnestic MCI vs. MCI) showed significantly higher hippocampal atrophy rate in amnestic MCI [19,21] compared with MCI (all subtypes) [15,17,18,20,23] (2.68%/year [SE = 0.66] vs. 2.47%/year [SE = 0.42]) in MCI

Table 3
Random-effect models of whole brain, hippocampus, and entorhinal cortex atrophy rates in healthy controls, MCI, and in MCI after removing the effect attributed to normal aging and subgroup and meta-regression analyses of hippocampal atrophy rate in MCI after removing the effect attributed to normal aging

Random-effects model

Brain areas	K	Age	Estimate %/year	SE	95% CI		Z-value	*P* value	T^2	T	I^2 %	H^2	Test for heterogeneity df	Q	*P* value
Whole brain (K = 4)															
HC	351	71.45	−0.5665	0.0328	−0.6308	−0.5023	−17.2757	<.0001	0	0	0	1.0	3	1.8707	.5997
MCI	466	72.92	−1.0203	0.1263	−1.2679	−0.7727	−8.0772	<.0001	0.0477	0.2185	79.98	4.99	3	12.6691	.0053
MCI-HC	—	—	−0.4634	0.0987	−0.6569	−0.2699	−4.6944	<.0001	0.0194	0.1393	51.86	2.08	3	5.7540	.1242
Entorhinal cortex (K = 3)															
HC	257	75.40	−2.4146	1.3036	−4.9696	0.1505	−1.8522	.0640	5.0168	2.2398	98.81	83.72	2	89.1356	<.0001
MCI	258	74.60	−3.754	1.6065	−6.9028	−0.6052	−2.3367	.0195	7.5905	2.7551	98.51	67.29	2	83.2905	<.0001
MCI-HC	—	—	−1.1301	0.3373	−1.7911	−0.4691	−3.3509	.0008	0.1936	0.4400	52.49	2.10	2	4.1965	.1227
Hippocampus (K = 7)															
HC	487	71.54	−1.1197	0.1622	−1.4376	−0.8019	−6.9048	<.0001	0.1513	0.3890	86.22	7.26	6	34.2283	<.0001
MCI	540	73.09	−2.5303	0.3261	−3.1694	−1.8912	−7.7598	<.0001	0.6741	0.8211	92.87	14.02	6	78.1854	<.0001
MCI-HC	—	—	−1.3450	0.1906	−1.7186	−0.9715	−7.0571	<.0001	0.1556	0.3945	64.69	2.83	6	16.5628	.0110

Subgroup and meta-regression analyses

Hippocampus; MCI-HC	K	Age	Coef	SE	95% CI		Z-value	*P* value	T^2	T	I^2 %	H^2	R^2	Residual hetrogeneity df	QE	*P* value
Model 1																
Automatic segmentation	4	71.57	−1.2900	0.2682	−1.8156	−0.7644	−4.8106	<.0001	0.2019	0.4494	69.90	3.32	—	5	16.5244	.0055
Manual segmentation	3	75.1	−1.4383	0.3289	−2.0829	−0.7936	−4.3730	<.0001								
Model 2																
aMCI	2	77.15	−1.3337	0.3939	−2.1057	−0.5618	−3.3863	.0007	0.2091	0.4572	70.73	3.42	—	5	16.4832	.0056
MCI	5	71.46	−1.3562	0.2488	−1.8438	−0.8686	−5.4510	<.0001								
Model 3																
Intercept	—	—	−0.9973	5.0703	−10.9349	8.9403	−0.1967	.8441	0.0384	0.1960	36.24	1.57	79.71	2	2.9841	.2249
Age	—	—	0.0006	0.0640	−0.1249	0.1261	0.0093	.9926								
Female rate	—	—	0.0209	0.0132	−0.0050	0.0467	1.5821	.1136								
APOE ε4 rate	—	—	−0.0233	0.0088	−0.0406	−0.0061	−2.6477	.0081								

Abbreviations: SE, standard error; CI, confidence interval; T, standard deviation of true effects; df, degrees of freedom; HC, healthy control; MCI, mild cognitive impairment; Coef, coefficient; aMCI = amnestic MCI; r^2, proportion of observed dispersion accounted for by the model; H^2, total variability/sampling variability; R^2, heterogeneity accounted for the moderator(s); Q, heterogeneity; QE, residual heterogeneity.

participants. After removing the effect attributable to normal aging, the hippocampal atrophy rate was significantly higher in analyses including all generic/unspecified MCI (1.35%/year, SE = 0.25) compared with those including amnestic MCI only (1.33%/year, SE = 0.39). However, the atrophy rate difference was relatively small especially in MCI-HC analyses, and also numbers of studies in each subgroup were limited. In addition, (as it is notified in the discussion) studies, which were not specific in detecting MCI subtype, generally used cognitive measures that commonly used for detecting amnestic MCI in other studies.

The influence of age, female gender, and *APOE* ε4 rate on hippocampal atrophy was separately investigated in HC, MCI, and MCI-HC. Except for *APOE* ε4, which significantly predicted the unexplained variance (55.38%) in annual atrophy rate, age and female gender did not contribute substantially to the heterogeneity detected between studies. A mixed-effects model using age, female gender, and *APOE* ε4 rate as moderators accounted for 79.7% of heterogeneity in hippocampal atrophy rate in MCI-HC; however, only *APOE* ε4 rate was a significant moderator of atrophy rate (Table 3).

3.5. Incompatible studies

3.5.1. Ventricular expansion

Although it was not possible to produce a pooled estimate of ventricular expansion rate because of insufficient reports of separate cohorts, the remaining studies reported very similar estimates [17,18,25] of, on average, twofold (3.30%/year vs. 2.40%/year in one report and 2.85 mL/year vs. 1.43 mL/year and 3.62 mL/year vs. 1.14 mL/year in two other reports) increase in expansion rate in MCI compared with HC. When considering that whole brain volume is about 1200–1500 mL, reported ventricular expansion rate is approximately 0.1%/year of the whole

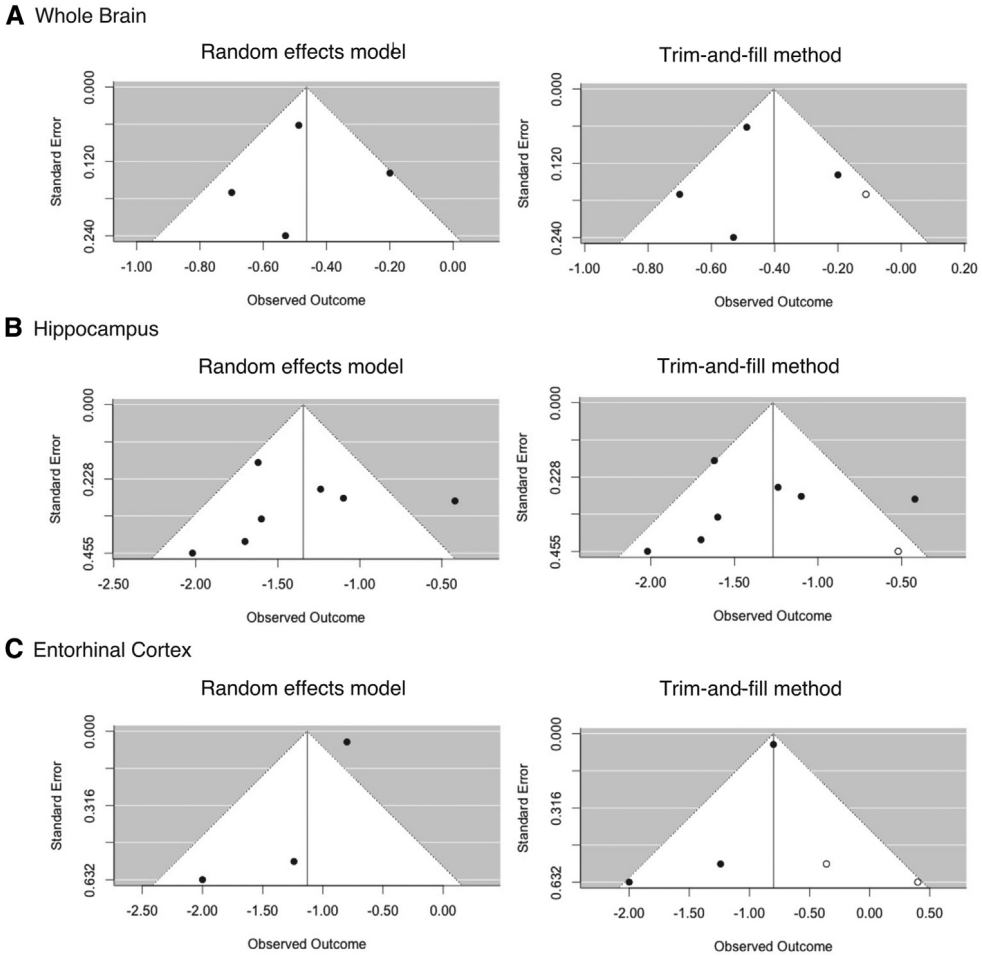


Fig. 3. Funnel plots of (A) whole brain, (B) hippocampus, and (C) entorhinal cortex using random-effects model (left column) and trim-and-fill method (right column). Filled circles represent included studies in the meta-analyses, and open circles represent possible missing studies.

brain volume in HC and 0.2%/year of the whole brain volume in MCI.

3.5.2. *Gray matter atrophy*

Besides the hippocampus and entorhinal cortex, which were the focus of most volumetric studies, there were also sporadic reports of volume loss for other parts of the brain including the parahippocampus, amygdala, and fusiform gyrus [23], lateral temporal lobe [26], cingulate [23,26], insula [6], parietal lobe [6,23,26], frontal and occipital lobes [6,26]. Atrophy rates in these regions were less than the average hippocampal atrophy rate and also differed based on the clinical outcome. Volume loss in the temporal and parietal lobes was higher for MCI subjects who had converted to AD within 4–5 years compared with stable MCI (lowest Cohen d for the inferior parietal lobe = 0.53 and largest for the hippocampus = 1.39) [23]. However, in clinically diagnosed AD, the atrophy rate in the medial temporal lobe was less than in MCI, whereas volume loss in frontal, parietal, and occipital regions was greater in MCI than AD [6].

3.5.3. *Cortical thickness and sulcal morphometry*

Cortical thickness was the second most commonly reported structural measure. Reports covered almost all parts of the brain but without quantitative estimates amenable to meta-analysis. Overall, studies revealed that controls and MCI participants demonstrated a similar spatial distribution of cortical loss, specifically in the parahippocampal cortex, middle/inferior temporal gyrus, supramarginal gyrus, angular gyrus, and superior frontal gyrus [27]. However, these studies suggested that atrophy rates were higher (no report of effect size) in MCI than controls, mainly in the temporal, superolateral parietal, and frontal lobes [28,29]. The only available longitudinal sulcal morphometry study showed an almost twofold higher rate of superior frontal and superior temporal sulcal widening in MCI compared with controls [30].

3.5.4. *White matter*

A minority of studies evaluated longitudinal changes in white matter. Recent DTI studies demonstrated a loss of integrity (increase in mean diffusivity) in the white matter

fiber tracts [31] particularly in the fornix (fitted mean changes in mean diffusivity over 12 months of 0.003 in controls vs. 0.051 in MCI), inferior and anterior cingulum (fitted mean changes in mean diffusivity over 6 months of −0.003 in controls vs. 0.013 in MCI) [32], in MCI compared with controls. DTI studies were limited in number and restricted to regions of interest evaluation.

4. Discussion

This study aimed to systematically review the literature on longitudinal structural brain changes specific to stable MCI. The main findings of this review were that (1) atrophy rates were 1.5–2.2 times larger in MCI participants than HC; (2) atrophy rate estimates were greater when assessed with manual than automatic segmentation; and (3) age, sex, and *APOE* ε4 were the most important moderators and together explained almost 80% of the between-study heterogeneity.

4.1. Global and local atrophy

Whole brain annual atrophy rate in MCI was twice that observed in controls. After removing the effect of normal aging, MCI-related shrinkage was estimated at 0.46%/year or almost 5 mL per year. This finding was consistent with studies reporting approximately 0.1%/year ventricular expansion in MCI in addition to that observed in normal aging [17,18,25], when considering that 20%–25% of the whole brain shrinkage is accounted for ventricular expansion [33].

Shrinkage in the whole brain is not necessarily the result of homogenous atrophy in all parts of the brain. Studies using measurement of cortical thickness and gray/white matter density in different parts of the brain demonstrated that atrophy rates in different brain regions were different and that some areas were more susceptible to neurodegeneration in normal aging as well as MCI-related degeneration [6,21,29,30]. Studies suggested that in MCI, noticeable atrophy was restricted to the medial temporal lobe, whereas frontal lobe and sensory motor cortices remained less atrophic until late in AD [34,35]. Additionally, previous evidence suggested that medial temporal lobe atrophy was higher in MCI participants who converted to AD compared with those with stable MCI [34,36].

It is important to consider that most reviewed studies used general diagnostic criteria to recruit MCI participants and did not investigate MCI subtypes. However, study design and cognitive tests, which were used in these studies, suggested that there was probably a higher prevalence of amnestic MCI in MCI participants. Therefore, reported findings are likely to be more representative of amnestic MCI than other MCI subtypes.

The hippocampus and entorhinal cortex were two of the most commonly investigated subregions of the medial temporal lobe, and direct evaluation of the medial temporal lobe volume change was not an issue in volumetric studies. Therefore, there is no estimation of the whole medial temporal lobe atrophy rate in the literature. However, overall atrophy rates in these medial temporal lobe subregions were similar to the whole brain atrophy rate, i.e., approximately twice in MCI compared with HC. Although, to our knowledge, there is no other systematic review of brain areas atrophy rates in MCI, a systematic review estimating annual hippocampal atrophy rate in healthy aging across the life span revealed hippocampal annual atrophy rate of 1.12%/year in healthy aging over the age of 70 years [7], which is consistent with the present findings. The roles of the hippocampus and entorhinal cortex in memory function have been known for a long time and the association between atrophy rates in these regions and cognitive decline has been well documented in MCI. However, the mean estimates of annual atrophy rates in these regions do not explain a 5-mL annual reduction in the whole brain volume. The cerebral atrophy observed in MCI above that detected in normal aging was 1.35%/year in the hippocampus and 1.35%/year in the entorhinal cortex. This indicates a total annual volume loss of about 0.07 mL in these areas [33], which covers less than 1.5% (of 5 mL) of the whole brain annual volume loss. This suggests that volume loss in areas well known for memory and cognition may only be the tip of the iceberg. In summary, although most available evidence has suggested that high rates of atrophy are mostly restricted to the medial temporal lobe in stable MCI, this conclusion might be due to underinvestigation of other cerebral regions.

4.2. Gray matter and white matter

Apart from medial temporal lobe atrophy, decrease in gray matter volume was reported in the lateral temporal, parietal, and frontal lobes [37]. These findings are consistent with reports demonstrating cortical thinning in the superolateral parietal lobe and some regions of the frontal cortex [29] as well as sulcal widening in the superior temporal and superior frontal sulci [30]. There are also sporadic reports suggesting decrease in the volume of the parahippocampal gyrus, amygdala, fusiform gyrus, superior temporal lobe [23], lateral temporal lobe [26], inferior temporal lobe [23], frontal lobe [6,26], cingulate [26], parietal and occipital lobes [6,26], and insula [6]. Therefore, although higher atrophy rates have been prominently reported in the medial temporal lobe and the atrophy rate in this region was positively associated with cognitive decline, brain atrophy is also widely distributed to other parts of the temporal, parietal, and frontal lobes. Nonetheless, in spite of the widespread gray matter atrophy, estimated atrophy rates in these areas alone cannot explain the whole brain atrophy rate. Indeed, the gray matter forms less than half of the brain tissue and atrophy rates as high as the atrophy rate in the hippocampus are needed in all parts of the gray matter to explain the total brain volume loss.

Therefore, atrophy of white matter is likely to significantly contribute to whole brain atrophy, especially

because axonal integrity depends on cell body viability in the gray matter and theoretically cell loss in gray matter atrophy should have an impact on white matter integrity. Loss of integrity in the white matter fiber tracts, particularly in the fornix and anterior and inferior cingulum, has been detected by DTI studies [31,32]. These studies are limited in number and restricted in the selection of regions of interest. A relationship between hippocampal gray matter atrophy and subsequent disruption in the uncinated fasciculus and the cingulum bundle has also been reported [37].

Although too few studies investigating white matter atrophy were available for review and for reliable assessment of their magnitude, they suggest that white matter is not spared from MCI pathology. However, the rate of atrophy in white matter and its association with gray matter and whole brain volume loss are some important unanswered questions. White matter forms the dominant proportion of brain structure, which reflects the importance of connection and networks in neural structure and consequently brain function. Therefore, it is essential that more investigations focus on these questions.

Furthermore, although neuroimaging studies largely interpret their results in relation to neural tissue, the brain also consists of connective tissue forming the brain's structural frame, supporting neural content and providing nutrients to neural tissue. This structural frame has an important role in preserving neural integrity and brain function. Therefore, any change in brain connective tissue may affect the structure and function of the neural system. The effect of aging on connective tissues in other parts of the body, including the skin, has been well documented, but the involvement of brain connective tissue in aging and age-related disorders needs to be evaluated in more detail. In summary, further longitudinal investigation of non–gray matter (e.g., white matter and connective tissue) atrophy might be informative and may help explain gaps in our understanding of pathologic processes associated with MCI and dementia.

4.3. Segmentation method

We investigated the impact of segmentation methodologies (manual vs. automated) through meta-regression analyses and found that manual segmentation of the hippocampus resulted in larger atrophy rate estimates compared with automatic segmentation using FreeSurfer. Although previous studies suggested that automatic segmentation with FreeSurfer resulted in a larger estimation of hippocampal volume in comparison with manual segmentation of the same images [38,39], atrophy rates have been reported to be lower in investigations using automatic segmentation [40]. As detailed in Fig. 2, differences between manual and automatic estimations of hippocampal atrophy are bigger in HC than MCI participants (68% compared with 40%), and in MCI (after removing the effect of normal aging), the difference is remarkably less than HC (7% compared with 68%).

As suggested by Wenger et al. [39], automatic segmentation may classify some nonhippocampal tissue—with lower atrophy rate—as hippocampal tissue. This would explain how the automatic approach could result in higher volume estimates but lower atrophy rate. A systematic review by Fraser et al. [7], estimating annual hippocampal atrophy rate in healthy aging across the life span, also detected a similar difference between manual tracing and automatic FreeSurfer segmentation and suggested that most studies using manual tracing excluded the tail of hippocampus and estimate the atrophy rate based on the atrophy of the head of the hippocampus. They concluded that hippocampal atrophy in HC was mostly restricted to the head of the hippocampus, rather than the tail; therefore, manual approaches, which excluded the tail, were likely to estimate a lower atrophy rate compared with automatic FreeSurfer approaches, which included the tail. In summary, although manual tracing is traditionally considered as the gold standard method of hippocampal volume estimation, the difference between manual tracing and automatic approaches appears to be largely related to the subregions included in each method, rather than the accuracy of estimation.

4.4. Moderators

An important question is whether study-specific factors such as age, female gender, and *APOE* ε4 influenced the reported estimates of brain atrophy in MCI. To investigate this question, we performed a mixed-effects model analysis for hippocampal atrophy rates (the largest analysis group). The results showed that these moderators accounted for almost 80% of the observed heterogeneity between studies, with *APOE* ε4 showing the largest moderating effect.

Moderating effects of age on brain atrophy have been well documented, although the pattern of association needs more investigation. It seems that this association is nonlinear and that the atrophy rate in stable MCI is larger at younger than older ages [41] although this was not confirmed in our meta-regression, possibly be due to a narrow age range as well as small number of studies in the meta-regression. Indeed, research consistent with this finding suggests that a higher whole brain atrophy rate is present in female compared with male individuals with MCI as well as in HC [41]. However, although this appears to be the case across the brain, it may not apply at regional levels. This is the likely reason we did not find a gender effect in our hippocampal meta-regression. Previous evidence revealed that in different brain regions are different in male and female not only in MCI but also in HC. For example, atrophy rates for the thalamus, caudate nucleus, and right middle temporal gyrus are higher in male MCI, compared with female, and atrophy

rates in the left middle temporal gyrus and precuneus are higher in female MCI than male [42]. Our finding that *APOE* ε4 genotype is a significant moderator and is associated with a higher rate of hippocampal atrophy in MCI is consistent with reviewed longitudinal studies that were not included in the meta-analysis. Moreover, the effect appears to become more salient across the disease process with MCI and AD showing that the *APOE* ε4 genotype is associated with faster atrophy rates [43,44], particularly in the hippocampus [45,46]. Association between *APOE* ε4 genotype and greater atrophy rate has been reported previously in HC [47]. Thus, all parts of the brain do not seem to have a similar vulnerability to the effect of *APOE* ε4 genotype and brain areas primarily involved in AD pathology, i.e., medial temporal lobe and particularly the hippocampus, are more affected, although the pattern of vulnerability is disease-stage specific [48,49]. *APOE* ε4 genotype is also associated with lower level of β-amyloid [50] and higher level of total and phosphorylated tau proteins [49] in cerebrospinal fluid. All these biomarkers are shown to be associated with faster regional brain atrophy (particularly the hippocampus) together and separately [48–52].

4.5. Strength and limitations of the study

A broad search of the literature (e.g., using a wide range of search terms) and inclusion of all available studies (using all sorts of structural measurements) were major strengths of this review. Special care was taken to combine studies with compatible measurements—to investigate pooled estimates of atrophy rates—and an attempt was made to comprehensively integrate incompatible findings and to summarize available knowledge about structural changes in MCI pathology. However, the review was limited by a relatively small number of available studies that could be included in the meta-analyses, particularly where whole brain and entorhinal cortex analyses are concerned. Additionally, many brain regions (such as the cerebellum) could not be analyzed because of lack of evidence and should be the focus of future studies. In addition, owing to the small number of studies in the meta-analysis, in relation to the number of moderators, it was recognized that estimates of moderator effects might be imprecise. The review was limited to comparing stable HC and prevalent MCI and data related to healthy participants converting to MCI and MCI participants converting to AD were insufficient to consider them in the present investigation.

5. Conclusion

To our knowledge, this is the first systematic review of longitudinal studies investigating MCI-related brain structural changes. The analyses revealed that the whole brain shrinks approximately two times faster in MCI participants compared with matched healthy people of the same age. Additionally, the medial temporal lobe regions—particularly the entorhinal cortex and hippocampus—are remarkably affected in AD pathology and associated with risk factors including *APOE* ε4 genotype and female gender. These regions demonstrate an atrophy rate of 1.5–2.2%/year times for MCI compared with HC. Although the medial temporal lobe was reported as the region highly involved in AD-related neurodegeneration, estimated atrophy rates in this region do not convincingly explain the amount of annual whole brain volume loss observed in MCI. Further investigation of other components of neural tissue, including white matter and non-neural brain tissue (e.g., connective tissue), is needed.

Acknowledgments

This study was funded by Australian Research Council project grant number 120101705. The funding sources were not involved in the design, collection, analysis or interpretation of data; or in the writing of the report or the decision to submit.

H.T.-J. and N.C. contributed to the design of the study. H.T.-J. and M.E.S. contributed to the data screening and extraction. H.T.-J. conducted all statistical analyses and managed all aspects of article preparation and submission. M.E.S. contributed to the article preparation. N.C. provided methodological input and theoretical expertise and contributed to writing and editing of the article.

RESEARCH IN CONTEXT

1. Systematic review: The authors reviewed literature investigating longitudinal structural brain changes in mild cognitive impairment (MCI). Studies largely investigated the preclinical phase of Alzheimer's disease and mostly focused on the medial temporal lobe.
2. Interpretation: In MCI, our analyses revealed a mean shrinkage of 5 mL/year in the whole brain above normal aging. Hippocampus and entorhinal cortex contributed less than 1.5% to the whole brain volume loss. Gray matter atrophy reported for other parts of the brain cannot explain a 5-mL annual whole brain volume loss. Atrophy in posterior parts of the brain (including the cerebellum) have been largely unstudied and may be important for explaining total annual volume loss in MCI.
3. Future directions: This review proposes a framework for generation of new studies regarding (1) atrophy rates specific to the cerebellum and white matter in MCI (2) and the role of non-neuronal brain tissue (i.e. connective tissue) changes in MCI pathology.

References

[1] Douaud Gl, Refsum H, de Jager CA, Jacoby R, Nichols TE, Smith SM, et al. Preventing Alzheimer's disease-related gray matter atrophy by B-vitamin treatment. Proc Natl Acad Sci U S A 2013;110:9523–8.

[2] Smith AD, Smith SM, de Jager CA, Whitbread P, Johnston C, Agacinski G, et al. Homocysteine-lowering by B vitamins slows the rate of accelerated brain atrophy in mild cognitive impairment: A randomized controlled trial. PLoS One 2010;5:e12244.

[3] Snyder PJ, Kahle-Wrobleski K, Brannan S, Miller DS, Schindler RJ, DeSanti S, et al. Assessing cognition and function in Alzheimer's disease clinical trials: Do we have the right tools? Alzheimers Dement 2014;10:853–60.

[4] Park MH, Han C. Is there an MCI reversion to cognitively normal? Analysis of Alzheimer's disease biomarkers profiles. Int Psychogeriatr 2015;27:429–37.

[5] Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol 1991;82:239–59.

[6] Sluimer JD, van der Flier WM, Karas GB, van Schijndel R, Barnes J, Boyes RG, et al. Accelerating regional atrophy rates in the progression from normal aging to Alzheimer's disease. Eur Radiol 2009; 19:2826–33.

[7] Fraser MA, Shaw ME, Cherbuin N. A systematic review and meta-analysis of longitudinal hippocampal atrophy in healthy human ageing. Neuroimage 2015;112:364–74.

[8] Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. Ann Intern Med 2009;151:264–70.

[9] Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomized studies in metaanalysis. Available at: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed December 17, 2015.

[10] Borenstein M, Hedges LV, Higgins JP, Rothstein HR. Introduction to meta. Hoboken, NJ: John Wiley & Sons; 2009.

[11] Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000;56:455–63.

[12] Song X, Mitnitski A, Zhang N, Chen W, Rockwood K. Dynamics of brain structure and cognitive function in the Alzheimer's disease neuroimaging initiative. J Neurol Neurosurg Psychiatry 2013;84:71–8.

[13] Zhang N, Song X, Zhang Y. Combining structural brain changes improves the prediction of Alzheimer's disease and mild cognitive impairment. Dement Geriatr Cogn Disord 2012;33:318–26.

[14] Ayaz M, Boikov AS, Haacke EM, Kido DK, Kirsch WM. Imaging cerebral microbleeds using susceptibility weighted imaging: One step toward detecting vascular dementia. J Magn Reson Imaging 2010; 31:142–8.

[15] Henneman WJ, Vrenken H, Barnes J, Sluimer IC, Verwey NA, Blankenstein MA, et al. Baseline CSF p-tau levels independently predict progression of hippocampal atrophy in Alzheimer disease. Neurology 2009;73:935–40.

[16] Henneman WJ, Sluimer JD, Barnes J, van der Flier WM, Sluimer IC, Fox NC, et al. Hippocampal atrophy rates in Alzheimer disease: Added value over whole brain volume measures. Neurology 2009; 72:999–1007.

[17] Schott JM, Bartlett JW, Barnes J, Leung KK, Ourselin S, Fox NC. Reduced sample sizes for atrophy outcomes in Alzheimer's disease trials: Baseline adjustment. Neurobiol Aging 2010;31:1452–62. 1462.e1–2.

[18] Archer HA, Kennedy J, Barnes J, Pepple T, Boyes R, Randlesome K, et al. Memory complaints and increased rates of brain atrophy: Risk factors for mild cognitive impairment and Alzheimer's disease. Int J Geriatr Psychiatry 2010;25:1119–26.

[19] Jack CR Jr, Shiung MM, Weigand SD, O'Brein PC, Gunter JL, Boeve BF, et al. Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic MCI. Neurology 2005; 65:1227–31.

[20] Villemagne VL, Burnham S, Bourgeat P, Brown B, Ellis KA, Szoeke C, et al. Amyloid beta deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: A prospective cohort study. Lancet Neurol 2013;12:357–67.

[21] Wang PN, Liu HC, Lirng JF, Lin KN, Wu ZA. Accelerated hippocampal atrophy rates in stable and progressive amnestic mild cognitive impairment. Psychiatry Res 2009;171:221–31.

[22] Eckerstrom C, Olsson E, Borga M, Ekholm S, Ribbelin S, Rolstad S, et al. Small baseline volume of left hippocampus is associated with subsequent conversion of MCI into dementia: The Goteborg MCI study. J Neurol Sci 2008;272:48–59.

[23] Desikan RS, Fischl B, Cabral HJ, Kemper TL, Guttmann CRG, Blacker D, et al. MRI measures of temporoparietal regions show differential rates of atrophy during prodromal AD. Neurology 2008; 71:819–25.

[24] Schuff N, Tosun D, Insel PS, Chiang GC, Truran D, Aisen PS, et al. Nonlinear time course of brain volume loss in cognitively normal and impaired elders. Neurobiol Aging 2012;33:845–55.

[25] Jack CR Jr, Lowe VJ, Weigand SD, Wiste HJ, Senjem ML, Knopman DS, et al. Serial PIB and MRI in normal, mild cognitive impairment and Alzheimer's disease: Implications for sequence of pathological events in Alzheimer's disease. Brain 2009;132:1355–65.

[26] McDonald CR, Gharapetian L, McEvoy LK, Fennema-Notestine C, Hagler DJ Jr, Holland D, et al. Relationship between regional atrophy rates and cognitive decline in mild cognitive impairment. Neurobiol Aging 2012;33:242–53.

[27] Li Y, Wang Y, Wu G, Shi F, Zhou L, Lin W, et al. Discriminant analysis of longitudinal cortical thickness changes in Alzheimer's disease using dynamic and network features. Neurobiol Aging 2012;33:427.e15–30.

[28] Marshall GA, Lorius N, Locascio JJ, Hyman BT, Rentz DM, Johnson KA, et al. Regional cortical thinning and cerebrospinal biomarkers predict worsening daily functioning across the Alzheimer's disease spectrum. J Alzheimers Dis 2014;41:719–28.

[29] Yao Z, Hu B, Liang C, Zhao L, Jackson M. A longitudinal study of atrophy in amnestic mild cognitive impairment and normal aging revealed by cortical thickness. PLoS One 2012;7:e48973.

[30] Liu T, Sachdev PS, Lipnicki DM, Jiang J, Cui Y, Kochan NA, et al. Longitudinal changes in sulcal morphology associated with late-life aging and MCI. Neuroimage 2013;74:337–42.

[31] Selnes P, Aarsland D, Bjornerud A, Gjerstad L, Wallin A, Hessen E, et al. Diffusion tensor imaging surpasses cerebrospinal fluid as predictor of cognitive decline and medial temporal lobe atrophy in subjective cognitive impairment and mild cognitive impairment. J Alzheimers Dis 2013;33:723–36.

[32] Nowrangi MA, Lyketsos CG, Leoutsakos JM, Leoutsakos JMS, Oishi K, Albert M, et al. Longitudinal, region-specific course of diffusion tensor imaging measures in mild cognitive impairment and Alzheimer's disease. Alzheimers Dement 2013;9:519–28.

[33] Standring S. Gray's anatomy: The anatomical basis of clinical practice. Churchill Livingstone, London, 40th ed. 2008.

[34] Leow AD, Yanovsky I, Parikshak N, Hua X, Lee S, Toga AW, et al. Alzheimer's disease neuroimaging initiative: A one-year follow up study using tensor-based morphometry correlating degenerative rates, biomarkers and cognition. Neuroimage 2009;45:645–55.

[35] Hua X, Lee S, Yanovsky I, Leow AD, Chou YY, Ho AJ, et al. Optimizing power to track brain degeneration in Alzheimer's disease and mild cognitive impairment with tensor-based morphometry: An ADNI study of 515 subjects. Neuroimage 2009;48:668–81.

[36] Evans MC, Barnes J, Nielsen C, Kim LG, Clegg SL, Blair M, et al. Volume changes in Alzheimer's disease and mild cognitive impairment: Cognitive associations. Eur Radiol 2010;20:674–82.

[37] Villain N, Fouquet M, Baron JC, Me źenge, Landeau B, de La Sayette V, et al. Sequential relationships between grey matter and white matter atrophy and brain metabolic abnormalities in early Alzheimer's disease. Brain 2010;133:3301–14.

[38] Cherbuin N, Anstey KJ, glade-Meslin CR, Sachdev PS. In vivo hippocampal measurement and memory: A comparison of manual tracing and automated segmentation in a large community-based sample. PLoS One 2009;4:e5265.
[39] Wenger E, Martensson J, Noack H, Bodammer NC, Kuhn S, Schaefer S, et al. Comparing manual and automatic segmentation of hippocampal volumes: Reliability and validity issues in younger and older brains. Hum Brain Mapp 2014;35:4236–48.
[40] Mulder ER, de Jong RA, Knol DL, van Schijndel RA, Cover KS, Visser PJ, et al. Hippocampal volume change measurement: Quantitative assessment of the reproducibility of expert manual outlining and the automated methods FreeSurfer and FIRST. Neuroimage 2014;92:169–81.
[41] Hua X, Hibar DP, Lee S, Toga AW, Jack CR Jr, Weiner MW, et al. Sex and age differences in atrophic rates: An ADNI study with n=1368 MRI scans. Neurobiol Aging 2010;31:1463–80.
[42] Skup M, Zhu H, Wang Y, Giovanello KS, Lin J, Shen D, et al. Sex differences in grey matter atrophy patterns among AD and aMCI patients: Results from ADNI. Neuroimage 2011;56:890–906.
[43] Aguilar C, Muehlboeck JS, Mecocci P, Vellas B, Tsolaki M, Kloszewska I, et al. Application of a MRI based index to longitudinal atrophy change in Alzheimer disease, mild cognitive impairment and healthy older individuals in the AddNeuroMed cohort. Front Aging Neurosci 2014;6:145.
[44] Vemuri P, Wiste HJ, Weigand SD, Knopman DS, Trojanowski JQ, Shaw LM, et al. Serial MRI and CSF biomarkers in normal aging, MCI, and AD. Neurology 2010;75:143–51.
[45] Manning EN, Barnes J, Cash DM, Bartlett JW, Leung KK, Ourselin S, et al. APOE epsilon4 is associated with disproportionate progressive hippocampal atrophy in AD. PLoS One 2014;9:e97608.
[46] Andrawis JP, Hwang KS, Green AE, Kotlerman J, Elashoff D, Morra JH, et al. Effects of ApoE4 and maternal history of dementia on hippocampal atrophy. Neurobiol Aging 2012;33:856–66.
[47] Morra JH, Tu Z, Apostolova LG, Green AE, Avedissian C, Madsen SK, et al. Automated mapping of hippocampal atrophy in 1-year repeat MRI data from 490 subjects with Alzheimer's disease, mild cognitive impairment, and elderly controls. Neuroimage 2009;45:S3–15.
[48] Tosun D, Schuff N, Shaw LM, Trojanowski JQ, Weiner MW. Relationship between CSF biomarkers of Alzheimer's disease and rates of regional cortical thinning in ADNI data. J Alzheimers Dis 2011;26(Suppl 3):77–90.
[49] Tosun D, Schuff N, Truran-Sacrey D, Shaw LM, Trojanowski JQ, Aisen P, et al. Relations between brain tissue loss, CSF biomarkers, and the ApoE genetic profile: A longitudinal MRI study. Neurobiol Aging 2010;31:1340–54.
[50] Schuff N, Woerner N, Boreta L, Kornfield T, Shaw LM, Trojanowski JQ, et al. MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. Brain 2009;132:1067–77.
[51] Chiang GC, Insel PS, Tosun D, Schuff N, Truran-Sacrey D, Raptentsetsang ST, et al. Impact of apolipoprotein E4-cerebrospinal fluid beta-amyloid interaction on hippocampal volume loss over 1 year in mild cognitive impairment. Alzheimers Dement 2011;7:514–20.
[52] Toledo JB, Da X, Weiner MW, Wolk DA, Xie SX, Arnold SE, et al. CSF Apo-E levels associate with cognitive decline and MRI changes. Acta Neuropathol 2014;127:621–32.

◆ Human Brain Mapping 38:3141–3150 (2017) ◆

The Cerebellum Shrinks Faster Than Normal Ageing in Alzheimer's Disease but not in Mild Cognitive Impairment

Hossein Tabatabaei-Jafari,* Erin Walsh, Marnie E. Shaw, and Nicolas Cherbuin, for the Alzheimer's Disease Neuroimaging Initiative (ADNI)

Centre for Research on Ageing, Health and Wellbeing, The Australian National University, Canberra, Australia

Abstract: *Background:* While acceleration in age-related cerebral atrophy has been well documented in Alzheimer's disease, the cerebellar contributions to this effect have not been thoroughly investigated. Objective: This study investigated cerebellar volume and atrophy rate using magnetic resonance imaging in individuals with normal cognition (CN), mild cognitive impairment (MCI), and Alzheimer's disease (AD). Methods: Two hundred twenty-nine CN, 398 MCI and 191 AD participants of stage I ADNI database with screening scans were evaluated for cerebellar volume. Of those, 758 individuals with two or more follow-up scans were categorized into stable, converted, and reverted CN, MCI and AD and evaluated for cerebellar atrophy rate. Results: Cerebellar volume was 2.5% larger in CN than in those with AD but there were no differences between CN and MCI and MCI and AD in cross-sectional analysis. Similarly, the atrophy rate was 49% larger in AD and 64% larger in MCI who converted to AD but no difference was detected between CN and MCI. There were no association between education and APOEe4 and cerebellar volume or cerebellar atrophy across the diagnostic groups. Conclusion: Cerebellar atrophy contributes to Alzheimer's clinical progression but mostly at the late stage of the disease. However, even in the late stage shrinkage rate is less than the average of the shrinkage in the cerebrum and is not associated with AD moderators. This suggests that cerebellar involvement is secondary to cerebral involvement and can be due to network connection spread regardless of the primary pathology. *Hum Brain Mapp 38:3141–3150, 2017.* © 2017 Wiley Periodicals, Inc.

Key words: Alzheimer's disease; mild cognitive impairment; cerebellar atrophy; cerebellum; magnetic resonance imaging

Contract grant sponsor: Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant); Contract grant number: U01 AG024904; Contract grant sponsor: Department of Defense (DOD ADNI); Contract grant number: W81XWH-12-2-0012; Contract grant sponsor: National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering; Contract grant sponsor: Australian Research Council (to N.C.); Contract grant number: 120101705; Contract grant sponsor: Canadian Institutes of Health Research.

*Correspondence to: Hossein Tabatabaei-Jafari, Centre for Research on Ageing, Health and Wellbeing, Australian National University, Florey Building 54, Mills Road, Acton, ACT 2601, Australia.
E-mail: hossein.tabatabaei@anu.edu.au

Received for publication 19 November 2016; Revised 27 February 2017; Accepted 11 March 2017.

DOI: 10.1002/hbm.23580

Published online 21 March 2017 in Wiley Online Library (wileyonlinelibrary.com).

INTRODUCTION

The human cerebellum is a brain structure well known for its role in motor function and recently has drawn attention for its implication in cognitive functions [Schmahmann and Sherman, 1998; Stoodley, 2012; Weier et al., 2014; Wolf et al., 2009]. It is connected to almost all parts of the nervous system, comprises more than 50% of the total brain neurons, but surprisingly contributes to only 10% of the whole brain volume [Andersen et al., 1992]. This mismatch is a reflection of the difference in neural architecture. Gray matter makes up 80% of the cerebellar volume (compared with less than half for the cerebrum) [Hoogendam et al., 2012] and consists of densely packed small granular neurons tightly folded which are less diverse compared to those of the cerebral cortex. In contrast to the variety of cytoarchitectonic organisation observed in different regions of the cerebral cortex, all regions of the cerebellar cortex appear similar in histological sections [Standring, 2008]. Specific histological architecture in addition to rich connections to the other parts of the brain makes the cerebellum an important region to investigate in the context of neurodegenerative disorders.

Pathologically, Alzheimer's disease (AD) is characterized by abnormal intra and extra cellular protein aggregations, i.e., intracellular tau phosphorylation and extracellular β-amyloid deposition. Studies using positron emission tomography (PET) revealed significant correlations between postmortem and in vivo presence and density of amyloid plaques and phosphorylated tau: ^{11}C-labeled Pittsburgh compound B (^{11}C-PiB) [Driscoll et al., 2012] and Florbetapir-PET imaging [Clark et al., 2011] for β-amyloid deposition and labelled THK5117-PET [Lemoine et al., 2015] for aggregated hyperphosphorylated tau. PET studies suggested no difference in the cerebellar uptake in AD and cognitively normal (CN) participants [Jack et al., 2008b; Jonasson et al., 2016; Rowe et al., 2007] and therefore it has been adopted as a normalizing area for standardized uptake values (SUVs) [Jonasson et al., 2016; Lopresti et al., 2005].

Although AD related shrinkage and neuronal death are thought to be associated with and possibly due to β-amyloid deposition and tau aggregation [Wang et al., 2002], their topological patterns and progression are different [Braak and Braak, 1991; Thal et al., 2002]. Moreover, the pattern of regional brain atrophy in AD does not follow precisely either β-amyloid or tau topological patterns [Sluimer et al., 2009]. Therefore, normal level of β-amyloid deposition and tau aggregation may not rule out the presence of neuronal loss or shrinkage in the cerebellum. A recent postmortem stereological study suggested no significant differences in the cerebellar total Purkinje and granular cell number nor in the volume of the granular layer between severely demented Alzheimer's disease (AD) and normal individuals [Andersen et al., 2012]. However, this finding is inconsistent with a previous study that showed a significant reduction in the granular layer in AD [Wegiel et al., 1999] although both studies reported significant reduction in whole cerebellar volume. These somewhat inconsistent findings may be due to the fact that these studies were postmortem (cross-sectional) with low sample sizes (20 and 16 subjects, respectively) in qualitatively different cohorts and thus afforded low statistical power.

To bypass the inevitable limitations of post mortem studies (single measurement occasion and small sample size), structural neuroimaging techniques including magnetic imaging are the best available option for longitudinal examination of brain volume change over time. Our recent published systematic review [Tabatabaei-Jafari et al., 2015] revealed that there is no morphological longitudinal study aimed at comparing cerebellar structural change in normal ageing and cognitively impaired populations including mild cognitive impairment (MCI) and Alzheimer's disease. Therefore, the main aim of this study is to evaluate cross-sectional and longitudinal structural differences in the cerebellum across cognitively different populations including CN, MCI, and AD.

METHODOLOGY

Study Participants

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 MCI and early AD.

All individuals participating in ADNI1 study who underwent MRI screening and diagnostic evaluations were included in the cross-sectional analysis and categorized into three diagnostic groups: CN, MCI, and AD. Participants with additional scans in follow-up assessments were included in the longitudinal analysis and categorised into more specific diagnostic groups according to the diagnosis at the first and last scanning time points. Details of the diagnostic criteria can be found on the ADNI web site (http://www.adni-info.org/Scientists/AboutADNI.aspx). Briefly, participants were categorized as CN if they had a Mini Mental State Examination (MMSE) score higher than 24, a Clinical Dementia Rating (CDR) of 0 and were not diagnosed with MCI, dementia or depression. Participants were categorized as MCI if they had a MMSE score higher than 24, a subjective report of memory concern, a measured objective memory loss, a CDR of 0.5, absence of dementia and preserved daily living activities. Participants were categorized as AD if they had a MMSE score lower than 26, a CDR of 0.5 or 1.0, and fulfilled criteria for clinically probable AD according to the Institute of Neurological and Communicative Diseases and Stroke/Alzheimer's Disease and Related Disorders Association. Participants with follow-up evaluation were categorized

into stable, converted or reverted CN, MCI, and AD according to the first and last time points diagnoses: stable if the first and last evaluation were similar, converted if the last evaluation progressed to declined cognitive diagnosis and reverted if the last evaluation was improved.

Image Acquisition

Participants underwent a high-resolution MRI scans of the brain on 1.5 T scanners from General Electric, Siemens, or Philips (Milwaukee, WI; Germany; The Netherlands, respectively) across multiple scanners using a standardized MRI protocol for 3D MP-RAGE sequences [Jack et al., 2008a] and following parameters: TR = 2,400 ms, minimum full TE, TI = 1,000 ms, flip angle = 8°, 24 cm field of view, acquisition matrix of 192 × 192 × 166 and yielding 1.25 × 1.25 × 1.2 mm^3.

Segmentation and Image Analysis

Volumetric segmentation were conducted by the ADNI team at the University of California, San Francisco using FreeSurfer version 5.1 for longitudinal analyses [Reutera et al., 2012]. The cerebellum was automatically segmented into gray matter and white matter. Sum values of the gray and white matter were considered as hemisphere volume and total of left and right were considered as cerebellar volumes.

Statistical Analysis

The R statistical software (version 3.1.1) was used for the cross-sectional and longitudinal analyses. The intra-class correlation coefficient (ICC) for the repeated longitudinal cerebellar volumes measurements was 0.98 (95%CI 0.9803–0.9843), which indicates that most of the variance (~96%) occurs between participants while only 4% occurs within participants.

Nonparametric locally weighted scatterplot smoothing (LOWESS) was used to visually inspect the data to determine whether linear models were appropriate. The LOWESS approach uses weighted least squares (giving more weight to points near the point whose response is being estimated) to estimate the mean response value at each time point and provide a smooth line representing the relationship between dependent and explanatory variables, when there are no assumptions about the relationship. The LOWESS plots for cerebellar volume versus age suggested that linear modeling of the relationship between cerebellar volume and age was appropriate for cross-sectional and longitudinal analyses since little departure from linearity was observed across groups except for CNc, which assumed to be due to low sample size i.e. 19 participants (Fig. 1).

The lme4 package (version 1.1-7) was used to conduct linear regressions analyses. In cross-sectional analyses, multiple linear regressions were conducted to investigate the cross-sectional relationship between cerebellar volume and clinical diagnosis status. Cerebellar volume was applied as dependent variable and age (centred on 55, the youngest participants at baseline), gender, education, APOE e4, diagnosis and intracranial volume (ICV) were considered as explanatory variables. In longitudinal analyses, mixed effects models were applied with the same explanatory variables for linear regressions in addition to a random effect by scanner and two random effects by subjects: a random intercept and a random slope for age at each time point. The random slope of time (centred age at each time point) was tested in a minimally controlled model and if statistically significant was included in the model as random effect [Bernal-Rusiel et al., 2013]. A time by clinical diagnosis group interaction effect was tested to determine whether the rate of change in cerebellar volume differed between groups. Fixed effect of age on cerebellar volume for each diagnostic group was considered as cerebellar atrophy rate.

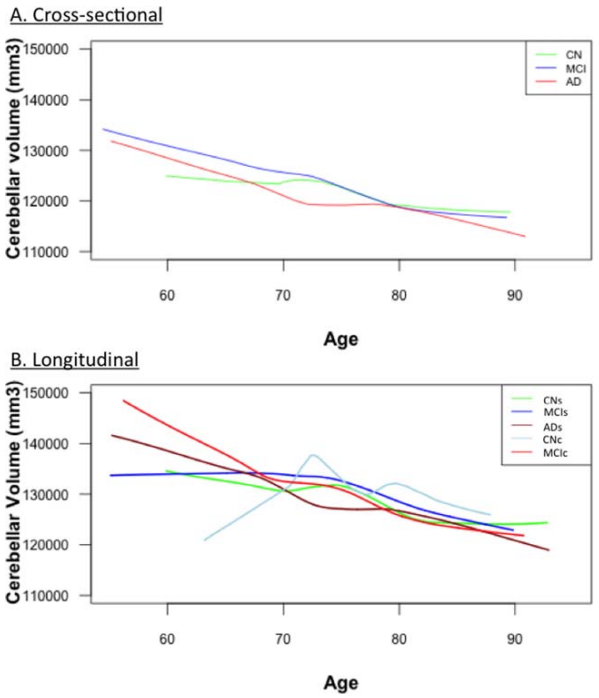


Figure 1.

Locally weighted smoothed mean measurement trajectory (LOWESS plot) of cerebellar volumes vs. age. (**A**) Three clinical groups including cognitively normal (CN), mild cognitive impairment (MCI), and Alzheimer's disease (AD) in cross-sectional level. (**B**) Five clinical groups including stable cognitively normal (CNs), cognitively normal converted to mild cognitive impairment (CNc), stable mild cognitive impairment (MCIs), mild cognitive impairment converted to AD (MCIc), and stable Alzheimer's disease (ADs) in serial scans. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE I. Demographic information

	Cross-sectional ($N = 818$)			Longitudinal ($N = 758$)						
	CN	MCI	AD	CNs	CNc	MCIs	MCIc	CN to AD	ADs	MCIr	ADr
No. participants	229	398	191	196	19	193	161	2	172	13	2
Age at baseline (yr) (SD)	75.87 (5.02)	74.74 (7.39)	75.27 (7.46)	75.76 (5.03)	77.45 (5.22)	75.00 (7.42)	74.73 (6.71)	80.55 (3.61)	75.12 (7.61)	73.43 (9.96)	79.50 (4.38)
Male sex, *n* (%)	119 (52)	257 (65)	100 (52)	103 (53)	11 (58)	124 (64)	101 (63)	0 (0)	91 (53)	9 (69)	2 (100)
Education (yr) (SD)	16.07 (2.86)	15.64 (3.03)	14.70 (3.15)	16.13 (2.88)	15.95 (2.39)	15.45 (3.15)	15.84 (2.81)	15.00 (0.00)	14.77 (3.14)	16.00 (2.42)	16.00 (0.00)
APOEe4 (*n*) (%)	61 (27)	212 (53)	127 (67)	49 (25)	8 (42)	92 (48)	104 (65)	1 (50)	117 (68)	4 (31)	1 (50)
MMSE at baseline (SD)	29.11 (0.99)	27.03 (1.78)	23.31 (2.04)	29.08 (1.06)	29.32 (0.75)	27.22 (1.77)	26.71 (1.72)	29.5 (0.71)	23.40 (1.97)	27.85 (1.77)	26.00 (0.00)

CN, cognitively normal; MCI, mild cognitive impairment; AD, Alzheimer's disease; CNs, stable cognitively normal, cognitively normal converted to mild cognitive impairment; MCIs, stable mild cognitive impairment; MCIc, mild cognitive impairment converted to Alzheimer's disease; Ads, stable Alzheimer's disease; MCIr, mild cognitive impairment reverted to cognitively normal; ADr, Alzheimer's disease reverted to mild cognitive impairment; APOEe4, Apolipoprotein alleles e4 genotype; MMSE, mini-mental state examination.

The final models were visually checked for any obvious deviations from homoscedasticity, normality of residuals, and linearity. Likelihood ratio test of the model with the effect in question against the model without was used to determine statistical significance.

RESULTS

Demography

Cross-sectional

Eight hundred eighteen participants were categorized into CN, MCI, and AD. There were no significant differences in age across the groups, but significant differences in gender and APOE e4 distributions among the diagnostic groups. The male ratio was higher in MCI and, as expected, APOEe4 frequencies were significantly higher in MCI and AD. AD participants were significantly less educated than CN (Table I).

Longitudinal

Of 818 participants with screening scans 758, who had one or more follow-up scans and cognitive tests, were included in the longitudinal part. They were categorized into different diagnostic groups according to the first and last time points diagnoses: stable CN (CNs), CN converted to MCI (CNc), stable MCI (MCIs), MCI converted to AD (MCIc), stable AD (ADs), CN converted to AD, MCI reverted to CN (MCIr), and AD reverted to MCI (ADr). There were no significant differences in age and education across the diagnostic groups except for education between CNs and ADs. Pearson χ^2 test revealed no significant difference in gender distribution but a significant difference in APOEe4 distributions between diagnostic groups. APOEe4 distributions were higher in MCIs than CNs and in ADs than CNs. The mean follow-up period across the groups was 2.54 (1.20) years, which was shorter in MCIs and ADs compared with CNs.

Cross-Sectional Results

A significant association between cognitive diagnosis and cerebellar volume ($F_{(2,811)} = 3.95$, $P < 0.01$) was detected.

Pairwise comparisons demonstrated (3,400 mm^3; ~2.5%) larger cerebellar volume in CN compared to AD ($F_{(1,413)} = 9.82$, $P < 0.001$), but no differences between CN and MCI ($F_{(1,620)} = 3.40$ $P > 0.1$), and MCI and AD ($F_{(1,582)} = 1.62$, $P > 0.1$). Table II presents the mean ICV-adjusted cerebellar volumes and the fixed effect of age for the three diagnostic groups. Although, the average cerebellar volume was significantly smaller in AD compared to CN and MCI, the slope of decrease in cerebellar volume for each year increase in age was only 0.41% (CN; 0.34%, MCI; 0.42%, AD; 0.38%) and was not significantly different across groups ($F_{(2,809)} = 0.28$, $P > 0.5$) and in pair-wise comparisons ($F < 0.5$, $P > 0.1$). When all explanatory variables were included, the linear regression

TABLE II. Cerebellar measures

	Cross-sectional ($N = 818$)			Longitudinal ($N = 758$)							
	CN	MCI	AD	CNs	CNc	MCIs	MCIc	CN to AD	ADs	MCIr	ADr
No. participants	229	398	191	196	19	193	161	2	172	13	2
Baseline volume (mm^3) (SD)[a]											
Total	123,249.80	122,394.90	120,706.90	131,276.60	132,820.00	130,319.50	130,154.90	131,354.80	129,524.10	131,603.50	114,157.80
	(10,018.69)	(10,632.40)	(10,591.60)	(11,163.98)	(9,450.25)	(11,317.70)	(11,680.98)	(778.99)	(10,838.42)	(10,703.89)	(17,439.08)
Left	61,739.12	60,420.98	60,193.35	65,382.32	66,270.67	64,842.59	64,934.41	63,636.93	64,600.50	65,556.00	57,525.16
	(6,539.80)	(6,431.27)	(7,254.14)	(5,724.23)	(4,846.23)	(5,688.15)	(5,939.78)	(154.53)	(5,454.41)	(5,127.40)	(9,387.68)
Right	62,498.44	61,213.30	60,910.64	65,894.26	66,549.30	65,476.95	65,220.46	65,717.89	64,923.62	66,047.48	56,632.64
	(6,404.02)	(6,584.98)	(7,329.76)	(5,570.00)	(4,843.58)	(5,784.85)	(5,943.95)	(624.47)	(5,525.67)	(5,681.92)	(8,051.40)
Follow-up period (yr) (SD)	—	—	—	3.12	3.52	2.41	2.84	2.99	1.61	2.68	3.45
				(1.18)	(0.73)	(1.24)	(1.08)	(0.00)	(0.62)	(1.09)	(0.76)
No. scan	1	1	1	4.81	5.21	4.76	5.38	5.00	3.48	4.31	4.00
				(1.15)	(0.86)	(1.57)	(1.37)	(0.00)	(0.38)	(1.44)	(0.00)
Last scan volume (mm^3) (SD)[b]											
Total	—	—	—	129,686.50	130,245.70	129,016.00	127,633.00	130,699.20	128,100.80	131,491.30	118,023.80
				(11,392.32)	(9,597.18)	(11,286.36)	(11,669.41)	(3,139.66)	(10,866.31)	(10,366.15)	(7,945.15)
Left	—	—	—	64,607.51	64,839.90	64,181.13	63,673.82	65,943.72	63,894.04	65,287.25	59,466.88
				(5,804.00)	(5,017.74)	(5,647.45)	(5,896.68)	(1,230.14)	(5,467.48)	(5,149.04)	(4,355.83)
Right	—	—	—	65,078.97	65,405.81	64,834.86	63,959.14	64,755.50	64,206.73	66,204.10	58,556.96
				(5,704.68)	(4,766.31)	(5,813.87)	(5,960.80)	(1,909.52)	(5,560.26)	(5,307.67)	(3,589.32)
Coef. of age (CS)/atrophy rate (longi)[c] (mm^3/yr) (SE)											
Total	−417.90	−531.60	−463.50	511.98	615.84	498.71	833.04	—[d]	747.84	373.60	—§
	(128.80)	(70.33)	(100.70)	(53.27)	(120.36)	(68.52)	(79.52)		(80.79)	(201.40)	
Left	−192.90	−256.30	−238.30	249.79	347.77	249.66	412.00	—	362.69	153.72	
	(66.25)	(35.39)	(50.17)	(28.39)	(62.36)	(37.03)	(41.59)		(41.92)	(83.20)	
Right	−225.00	−275.30	−225.30	262.54	271.53	248.74	419.13	—	374.76	200.00	
	(64.35)	(36.09)	(51.90)	(26.66)	(66.57)	(33.28)	(39.83)		(40.90)	(95.16)	

[a]The mean of cerebellar volume adjusted by the intra cranial volume for the cross-sectional part and baseline volume adjusted by the intra cranial volume for the longitudinal part.
[b]Adjusted by the intra cranial volume.
[c]Fixed effects of age for the cross-sectional data; extracted from the linear regression adjusted by intra cranial volume, gender, education and APOEe4, and atrophy rate for the longitudinal data; extracted from the linear mixed effects model adjusted by intra cranial volume, gender, education and APOEe4.
[d]Insufficient data for calculation.
CN, cognitively normal; MCI, mild cognitive impairment; AD, Alzheimer's disease; CNs, stable cognitively normal; cognitively normal converted to mild cognitive impairment; MCIs, stable mild cognitive impairment; MCIc, mild cognitive impairment converted to Alzheimer's disease; Ads, stable Alzheimer's disease; MCIr, mild cognitive impairment reverted to cognitively normal; ADr, Alzheimer's disease reverted to mild cognitive impairment; APOEe4, Apo lipoprotein alleles e4 genotype; MMSE; mini-mental state examination.

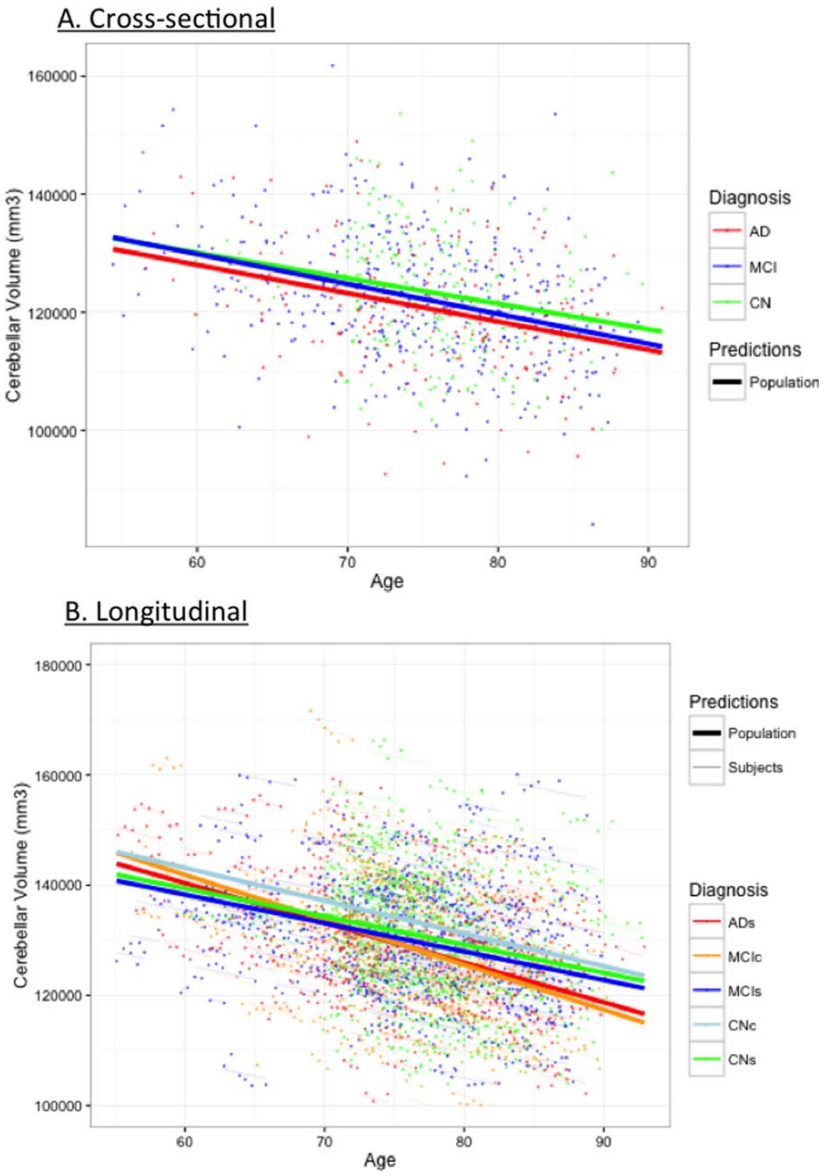


Figure 2.

Linear prediction of the cerebellar volumes for age at time points. (**A**) Prediction of the cerebellar volumes in three clinical groups including cognitively normal (CN), mild cognitive impairment (MCI) and Alzheimer's disease (AD) in cross-sectional level. (**B**) Prediction in subject and group (population) levels in five diagnostic groups including stable cognitively normal (CNs), cognitively normal converted to mild cognitive impairment (CNc), stable mild cognitive impairment (MCIs), and mild cognitive impairment converted to AD (MCIc) illustrating different slops for the diagnostic groups. [Color figure can be viewed at wileyonlinelibrary.com]

model explained 44.7% of the variance in cerebellar volume ($F_{(8,809)} = 83.61$, $P < 0.0001$) mostly explained by ICV (37.9%) with 7.7% explained by age alone, and 0.7% by clinical group.

The scatters plot presenting the association between age and cerebellar volume for each group also revealed an initial overlap of CN and MCI regression lines followed by deviation of MCI regression line to AD line suggesting that cerebellar volumes are highly similar in CN and MCI at younger ages but lower in MCI in older individuals (Fig. 2A). In contrast the AD regression line while following a similar

TABLE III. Pair-wise comparison of group diagnosis

	CNs vs. CNc		CNs vs. MCIs		CNs vs. MCIc		CNs vs. ADs	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Intercept	142,212.26	146,127.2	139,769.58	139,091.58	137,510.50	139,282.56	138,821.74	141,244.50
	(4,460.56)	(6,078.7)	(3,247.02)	(3,445.00)	(3,521.38)	(3,755.78)	(3,172.06)	(3,523.62)
Volume slope in	—	−4,286.9	—	1,745.53	—	−3,236.69	—	−1,504.24
CNs (mm^3/yr) (SE)		(4,450.6)		(2,164.33)		(2,257.59)		(2,371.78)
Pr(>|t|)	—	0.3366[a]	—	0.421[a]	—	0.1525[a]	—	0.5263[a]
Shrinkage slope in	—	100.9	—	−19.72	—	312.57	—	214.37
CNs (mm^3/yr) (SE)		(163.6)		(84.54)		(91.80)		(94.66)
Pr(>|t|)	—	0.5382[a]	—	0.816[a]	—	**0.0007**[d]	—	**0.0240**[b]
Loglik	−9,793	−9,793	−1,7438	−1,7438	−1,6960	−1,6951	−14,515	−14,509
Chisq		1.0195		1.3075		17.126		10.884
Chi *df*		2		2		2		2
Pr (>Chisq)		0.6006[a]		0.5201[a]		**0.0001911**[d]		**0.004332**[c]

[a]Significant code: >0.1.
[b]Significant code: <0.01.
[c]Significant code: <0.001.
[d]Significant code: <0.0001.
CN, cognitively normal; CNs, stable cognitively normal; CNc, cognitively normal converted to mild cognitive impairment; MCIs, stable mild cognitive impairment; MCIc, mild cognitive impairment converted to Alzheimer's disease; ADs, stable Alzheimer's disease.

slope had a clearly different intercept suggesting a constant smaller cerebellar volume in AD across the age span investigated. Similar patterns were demonstrated for the left and right cerebellar volumes (Table II).

Longitudinal Results

The linear mixed model achieved a good fit and fixed factors in the model explained 43% (marginal R^2) while fixed and random factors together explained 99% (conditional R^2) of variance in cerebellar atrophy. A significant negative fixed effect of age was detected ($\chi^2_{(1,9)} = 586.99$, $P < 0.0001$); each year beyond age 55 was associated with a 0.47% lower cerebellar volume compared to baseline. Additionally, a significant random effect of age on cerebellar volume ($\chi^2_{(2,18)} = 227.92$, $P < 0.0001$) and interaction between age and diagnosis ($\chi^2_{(7,25)} = 22.72$, $P < 0.01$) were detected. The model revealed no differences in cerebellar volume across the diagnostic groups ($\chi^2_{(7,18)} = 11.31$, $P > 0.1$), i.e., the average of cerebellar volumes in CNs, CNc, MCIs, MCIc, and ADs were not significantly different. However, a significant effect of cognitive diagnosis on cerebellar atrophy rates was detected ($\chi^2_{(7,25)} = 22.71$, $P < 0.001$). There was also a significant effect of gender on cerebellar volume (1,18) = 14.12, $P < 0.001$) with less shrinkage in male.

An annual shrinkage of 0.36% (SE = 0.04) was detected in CNs individuals. A pairwise comparison revealed that it was not significantly different in MCIs (0.36%/year, SE = 0.05) and CNc (0.42%/year, SE = 0.08); however, it was about 49% larger in ADs (0.53%/year, SE = 0.06). Similarly, the atrophy rate was about 64% larger in MCIc (0.62%/year, SE = 0.06) compared to CNs (Tables II and III). The annual atrophy was also about 53% larger in ADs than MCIs ($\chi^2_{(2,13)} = 8.67$ $P < 0.01$) and 68% larger in MCIc than MCIs ($\chi^2_{(2,13)} = 12.57$, $P < 0.001$; Table II). CN who converted to AD, MCI who reverted to CN and AD who reverted to MCI were excluded from pairwise comparison due to small samples sizes. Atrophy trajectories across groups are presented in Figure 2B.

Similar patterns of findings were observed for the left and right cerebellar volumes (Table II), as well as left and right cerebellar gray matter and white matter volumes.

DISCUSSION

This study aimed to investigate cerebellar shrinkage in normal ageing and preclinical (MCI) and clinical phases of AD. It revealed that cerebellar shrinkage occurs mostly in the late stages of the disease. The main findings were that (1) in cross-sectional analyses cerebellar volume was larger in CN compared to AD but not compared to MCI, (2) in longitudinal analyses cerebellar atrophy was higher in ADs and MCIc compared to CNs but not in CNc and MCIs, and (3) APOEe4 was not a significant predictor of baseline cerebellar volume nor of cerebellar atrophy across clinical groups.

Cross-Sectional

The smaller cerebellar volume observed in AD compared to CN and no difference between MCI and CN are in agreement with available cross-sectional studies reporting smaller cerebellar volume in AD [Kusbeci et al., 2009; Moller et al., 2013] but normal volume in MCI [Thomann et al., 2008; Yoon et al., 2013]. This discrepancy is consistent with the documented progression of AD pathology. However, the cerebellum can be parsed functionally and morphologically into different subdivisions and it is likely that AD pathology targets each subdivision differently. Previous voxel-based morphometric studies showed bilateral lower gray matter density in lobule VI [Colloby et al., 2014] and Crus I/II [Guo et al., 2016] in AD compared with CN, suggesting that network-selective vulnerability underlies the cerebellar neurodegeneration [Guo et al., 2016]. Regardless of selective or nonselective volume loss in the cerebellum and its subregions, cross-sectional approach needs to be affirmed by tracking atrophy in a longitudinal approach.

Longitudinal

The negative association between age and cerebellar volume is consistent with that demonstrated in the cross-sectional analysis (0.41%/year in cross-sectional and 0.47% in longitudinal). Pairwise analyses demonstrated significantly larger cerebellar atrophy rates in ADs and MCIc but not in CNc and MCIs compared to CNs. This pattern of results is suggestive of an increasing rate of cerebellar atrophy with progression of AD pathology. It is also consistent with the chronological development of AD pathology with progressive spreading of tau fibrillatory tangles (Braak stages), amyloid deposition, and subsequently gradual decline in cognitive function [Murray et al., 2015]. As Thal et al. demonstrated, clinically diagnosed AD occurs in the amyloid phase 3 to 5 while the cerebellar involvement mostly occurs in the fifth phase [Thal et al., 2002]. Thus, the available evidence suggests that the cerebellum is relatively spared of neurodegeneration in the preclinical stages of the disease and gradually becomes affected as the clinical presentation fully develops. However, it remains unclear whether association of the cerebellum with AD clinical progression is due to spreading of fibrillary tangle and/or amyloid deposition, or secondary to cerebral neurodegeneration.

Although the findings suggest shrinkage in the cerebellum with ageing and larger cerebellar atrophy in ADs compared with CNs and MCIs, it is worthy to consider that cerebellar atrophy in the diagnostic groups were less than that reported for whole brain atrophy (CNs: 0.36%/year versus 0.57%/year; MCIs: 0.36%/year versus 1.02%/year; ADs: 0.53%/year versus 1.90%/year) [Henneman et al., 2009; Tabatabaei-Jafari et al., 2015]. This is in contrast to brain areas characteristics for AD pathology, including hippocampus and entorhinal cortex, for which atrophy rates are roughly 200% higher for MCI and 300% higher for AD compared to normal ageing [Desikan et al., 2008; Tabatabaei-Jafari et al., 2015], further emphasising the relative resistance of the cerebellum to AD related degeneration. However, despite the small effect size and partial resistance, the cerebellum is not intact in AD pathology and future investigation is needed to elucidate the impact of cerebellar atrophy on uptake measurement when using the cerebellum to standardise FDG uptake in PET studies.

Covariates and Correlates

Age is a common predictor for CN and AD-related brain atrophy and all cognitive groups in the current study were matched for age. However, they were differences in gender distribution, education and APOEe4 alleles—the most well-known risk factors of AD pathology—as were expected. An effect of sex on cerebellar volume was detected such that males showed less cerebellar atrophy than females. However, no significant association between education or APOEe4 alleles and cerebellar volume were detected. APOEe4 is a known moderator of hippocampal atrophy in AD pathology [Tabatabaei-Jafari et al., 2015], therefore it might have been expected that carrying the APOEe4 allele would be associated with increased cerebellar atrophy. However, this was not the case in our findings. It may indicate that while neurodegeneration in the cerebrum is directly related to the development of neurofibrillary tangles and β-amyloid deposition which occurs at higher rates in APOEe4 carriers, cerebellar atrophy is the product of secondary processes associated with cerebral neuronal loss, Wallerian degeneration, and widespread disconnection. To clarify this question future investigations need to further elucidate the impact of risk factors in different AD clinical stages.

Strengths and Limitation

This study is unique in using in vivo evaluation of the cerebellum with a reasonable follow up period in a relatively large sample while computing both cross-sectional and longitudinal estimates and using advanced and well-controlled mixed-effects models. Most AD related cerebellar studies conducted to date have been postmortem or if in vivo, cross-sectional in design, thus raising questions as to the precision and generalizability of their estimates. Consequently, the present study fills an important gap. However, it should be noted that this investigation was restricted to the gray and white matter volumes in the left and right cerebellum and therefore do not provide information on the cerebellar subregions.

CONCLUSION

The cerebellum is often thought to be spared from neurodegenerative processes but the present findings indicate that this is not the case. The present findings demonstrate that although the cerebellum is not significantly affected in

the preclinical phase of AD (i.e. MCI), it is affected in the clinical phase. However, acceleration in atrophy rate is less than the average of the atrophy in the cerebrum and it is not associated with AD moderators (education and APOEe4 status). These findings in addition to previous evidence of network-selective vulnerability of the cerebellum suggest that AD-related cerebellar atrophy might be secondary to the development of AD pathology in the cerebrum rather than the cerebellum itself. Therefore, modifying interventions targeting the non-specific network progression is a potential therapeutic option additional to interventions targeting the specific pathological process.

ACKNOWLEDGMENTS

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. H.T.J.: study concept, design, data management, statistical analyses, and managed all aspects of manuscript preparation and submission. E.W.: statistical analysis, and manuscript preparation. M.E.S.: image analysis, and manuscript preparation. N.C.: study design, methodological input and theoretical expertise, and manuscript preparation. 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.

REFERENCES

Andersen BB, Korbo L, Pakkenberg B (1992): A quantitative study of the human cerebellum with unbiased stereological techniques. J Comp Neurol 326:549–560.

Andersen K, Andersen BB, Pakkenberg B (2012): Stereological quantification of the cerebellum in patients with Alzheimer's disease. Neurobiol Aging 33:197.e11–197.e120.

Bernal-Rusiel JL, Greve DN, Reuter M, Fischl B, Sabuncu MR, Initiative ADNI. (2013): Statistical analysis of longitudinal neuroimage data with linear mixed effects models. Neuroimage 1: 249–260.

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

Clark CM, Schneider JA, Bedell BJ, Beach TG, Bilker WB, Mintun MA, Pontecorvo MJ, Hefti F, Carpenter AP, Flitter ML, Krautkramer MJ, Kung HF, Coleman RE, Doraiswamy PM, Fleisher AS, Sabbagh MN, Sadowsky CH, Reiman EP, Zehntner SP, Skovronsky DM (2011): Use of florbetapir-PET for imaging beta-amyloid pathology. JAMA 305:275–283.

Colloby SJ, O'Brien JT, Taylor JP (2014): Patterns of cerebellar volume loss in dementia with Lewy bodies and Alzheimers disease: A VBM-DARTEL study. Psychiatry Res 223:187–191.

Desikan RS, Fischl B, Cabral HJ, Kemper TL, Guttmann CR, Blacker D, Hyman BT, Albert MS, Killiany RJ (2008): MRI measures of temporoparietal regions show differential rates of atrophy during prodromal AD. Neurology 71:819–825.

Driscoll I, Troncoso JC, Rudow G, Sojkova J, Pletnikova O, Zhou Y, Kraut MA, Ferrucci L, Mathis CA, Klunk WE, O'Brien RJ, Davatzikos C, Wong DF, Resnick SM (2012): Correspondence between in vivo (11)C-PiB-PET amyloid imaging and postmortem, region-matched assessment of plaques. Acta Neuropathol 124:823–831.

Guo CC, Tan R, Hodges JR, Hu X, Sami S, Hornberger M (2016): Network-selective vulnerability of the human cerebellum to Alzheimer's disease and frontotemporal dementia. Brain 139: 1527–1538.

Henneman WJ, Sluimer JD, Barnes J, van der Flier WM, Sluimer IC, Fox NC, Scheltens P, Vrenken H, Barkhof F (2009): Hippocampal atrophy rates in Alzheimer disease: Added value over whole brain volume measures. Neurology 72:999–1007.

Hoogendam YY, Geest JNvd, Lijn Fvd, Lugt Avd, Niessen WJ, Krestin GP, Hofman A, Vernooij MW, Breteler MMB, Ikram MA (2012): Determinants of cerebellar and cerebral volume in the general elderly population. Neurobiol Aging 33:2774–2781.

Jack CR Jr, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, Borowski B, Britson PJ, J LW, Ward C, Dale AM, Felmlee JP, Gunter JL, Hill DL, Killiany R, Schuff N, Fox-Bosetti S, Lin C, Studholme C, DeCarli CS, Krueger G, Ward HA, Metzger GJ, Scott KT, Mallozzi R, Blezek D, Levy J, Debbins JP, Fleisher AS, Albert M, Green R, Bartzokis G, Glover G, Mugler J, Weiner MW (2008a): The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J Magn Reson Imaging 27:685–691.

Jack CR Jr, Lowe VJ, Senjem ML, Weigand SD, Kemp BJ, Shiung MM, Knopman DS, Boeve BF, Klunk WE, Mathis CA, Petersen RC (2008b): 11C PiB and structural MRI provide complementary information in imaging of Alzheimer's disease and amnestic mild cognitive impairment. Brain 131:665–680.

Jonasson M, Wall A, Chiotis K, Saint-Aubert L, Wilking H, Sprycha M, Borg B, Thibblin A, Eriksson JP, Sorensen J, Antoni G, Nordberg A, Lubberink M (2016): Tracer kinetic

analysis of (S)-18F-THK5117 as a PET tracer for assessing tau pathology. J Nucl Med 57:574–581.

Kusbeci OY, Bas O, Gocmen-Mas N, Karabekir HS, Yucel A, Ertekin T, Yazici AC (2009): Evaluation of cerebellar asymmetry in Alzheimer's disease: A stereological study. Dementia Geriatr Cognit Disord 28:1–5.

Lemoine L, Saint-Aubert L, Marutle A, Antoni G, Eriksson JP, Ghetti B, Okamura N, Nennesmo I, Gillberg PG, Nordberg A (2015): Visualization of regional tau deposits using (3)H-THK5117 in Alzheimer brain tissue. Acta Neuropathol Commun 3:40.

Lopresti BJ, Klunk WE, Mathis CA, Hoge JA, Ziolko SK, Lu X, Meltzer CC, Schimmel K, Tsopelas ND, DeKosky ST, Price JC (2005): Simplified quantification of Pittsburgh compound B amyloid imaging PET studies: A comparative analysis. J Nucl Med 46:1959–1972.

Moller C, Vrenken H, Jiskoot L, Versteeg A, Barkhof F, Scheltens P, van der Flier WM (2013): Different patterns of gray matter atrophy in early- and late-onset Alzheimer's disease. Neurobiol Aging 34:2014–2022.

Murray ME, Lowe VJ, Graff-Radford NR, Liesinger AM, Cannon A, Przybelski SA, Rawal B, Parisi JE, Petersen RC, Kantarci K, Ross OA, Duara R, Knopman DS, Jack CR Jr, Dickson DW (2015): Clinicopathologic and 11C-Pittsburgh compound B implications of Thal amyloid phase across the Alzheimer's disease spectrum. Brain 138:1370–1381.

Reutera, M., Schmansky, N.J., Rosasa, D., Fischl, B. (2012) Within-subject template estimation for unbiased longitudinal image analysis. Neuroimage, 61:1402–1418.

Rowe CC, Ng S, Ackermann U, Gong SJ, Pike K, Savage G, Cowie TF, Dickinson KL, Maruff P, Darby D, Smith C, Woodward M, Merory J, Tochon-Danguy H, O'Keefe G, Klunk WE, Mathis CA, Price JC, Masters CL, Villemagne VL (2007): Imaging beta-amyloid burden in aging and dementia. Neurology 68:1718–1725.

Schmahmann JD, Sherman JC (1998): The cerebellar cognitive affective syndrome. Brain 121:561–579.

Sluimer JD, van der Flier WM, Karas GB, van Schijndel R, Barnes J, Boyes RG, Cover KS, Olabarriaga SD, Fox NC, Scheltens P, Vrenken H, Barkhof F (2009): Accelerating regional atrophy rates in the progression from normal aging to Alzheimer's disease. Eur Radiol 19:2826–2833.

Standring S. (2008): Gray's Anatomy. The Anatomical Basis of Clinical Practice, 40th ed. Churchill Livingstone, London.

Stoodley CJ (2012): The cerebellum and cognition: Evidence from functional imaging studies. Cerebellum (London, England) 11: 352–365.

Tabatabaei-Jafari H, Shaw ME, Cherbuin N (2015): Cerebral atrophy in mild cognitive impairment: A systematic review with meta-analysis. Alzheimer's Dementia 1:487–504.

Thal DR, Rub U, Orantes M, Braak H (2002): Phases of A beta-deposition in the human brain and its relevance for the development of AD. Neurology 58:1791–1800.

Thomann PA, Schlafer C, Seidl U, Santos VD, Essig M, Schroder J (2008): The cerebellum in mild cognitive impairment and Alzheimer's disease - A structural MRI study. J Psychiatr Res 42: 1198–1202.

Wang H-Y, D'Andrea MR, Nagele RG (2002): Cerebellar diffuse amyloid plaques are derived from dendritic A-beta 42 accumulations in Purkinje cells. Neurobiol Aging 23:213–223.

Wegiel J, Wisniewski HM, Dziewiatkowski J, Badmajew E, Tarnawski M, Reisberg B, Mlodzik B, De Leon MJ, Miller DC (1999): Cerebellar atrophy in Alzheimer's disease-clinicopathological correlations. Brain Res 818:41–50.

Weier K, Penner IK, Magon S, Amann M, Naegelin Y, Andelova M, Derfuss T, Stippich C, Radue E-W, Kappos L, Sprenger T (2014): Cerebellar abnormalities contribute to disability including cognitive impairment in multiple sclerosis. PLoS One 9: e86916.

Wolf U, Rapoport MJ, Schweizer TA (2009): Evaluating the affective component of the cerebellar cognitive affective syndrome. J Neuropsychiatry Clin Neurosci 21:245–253.

Yoon CW, Seo SW, Park JS, Kwak KC, Yoon U, Suh MK, Kim GH, Shin JS, Kim CH, Noh Y, Cho H, Kim MJ, Kim JH, Roh JH, Lee JM, Na DL (2013): Cerebellar atrophy in patients with subcortical-type vascular cognitive impairment. Cerebellum (London, England) 12:35–42.

Neurobiology of Aging 69 (2018) 102–110

Contents lists available at ScienceDirect

Neurobiology of Aging

journal homepage: www.elsevier.com/locate/neuaging

A simple and clinically relevant combination of neuroimaging and functional indexes for the identification of those at highest risk of Alzheimer's disease

Hossein Tabatabaei-Jafari*, Erin Walsh, Marnie E. Shaw, Nicolas Cherbuin, For the Alzheimer's Disease Neuroimaging Initiative (ADNI)[1]

Centre for Research on Ageing, Health and Wellbeing, The Australian National University, Canberra, Australia

ARTICLE INFO

Article history:
Received 19 January 2018
Received in revised form 4 May 2018
Accepted 4 May 2018
Available online xxx

Keywords:
Hippocampus to cerebellum volume ratio
Mild cognitive impairment
MRI
Classification
Alzheimer's disease

ABSTRACT

The current challenge in clinical practice is to identify those with mild cognitive impairment (MCI), who are at greater risk of Alzheimer's disease (AD) conversion in the near future. The aim of this study was to assess a clinically practical new hippocampal index—hippocampal volume normalized by cerebellar volume (hippocampus to cerebellum volume ratio) used alone or in combination with scores on the Mini—Mental State Examination, as a predictor of conversion from MCI to AD. The predictive value of the HCCR was also contrasted to that of the hippocampal volume to intracranial volume ratio. The findings revealed that the performance of the combination of measures was significantly better than that of each measure used individually. The combination of Mini—Mental State Examination and hippocampal volume, normalized by the cerebellum or by intracranial volume, accurately discriminated individuals with MCI who progress to AD within 5 years from other MCI types (stable, reverters) and those with intact cognition (area under receiver operating curve of 0.88 and 0.89, respectively). Normalization by cerebellar volume was as accurate as normalization by intracranial volume with the advantage of being more practical, particularly for serial assessments.

1. Introduction

Mild cognitive impairment (MCI) refers to modest cognitive decline along with preserved daily activities (Association, 2013). Although many people with MCI live largely normal lives, they are at higher risk of developing Alzheimer's disease (AD) than those without MCI (Forlenza et al., 2013). The available evidence suggests that less than half of patients diagnosed with MCI may progress to AD in a 5-year period while the rest remain stable or reverse to cognitively normal (CN) status (Falahati et al., 2014; Pandya et al., 2016). Generally, there is an expectation of eventual conversion from MCI to AD due to the progressive nature of the neurodegenerative processes involved, and MCI stability can depend on the duration of follow-up (Ganguli, 2013). Reversion to CN status is still an unresolved question but may relate to the relatively unspecific nature of diagnostic criteria, interaction with comorbid conditions, and/or variability in the pathological process (Park et al., 2015). Thus, the current clinical challenge is to discriminate individuals with MCI who are more likely to convert to AD.

In their revised position, the National Institute on Aging and the Alzheimer's Association (NIA-AA) considered MCI and AD as different stages of the AD continuum rather than 2 distinct clinical entities (Albert et al., 2011; Jack et al., 2018). In 2011, NIA-AA reviewed diagnostic guidelines and suggested that, owing to greater diagnostic uncertainty earlier in the AD continuum, MCI diagnosis should be supported by biological markers reflecting AD pathology (Albert et al., 2011). In 2018, the NIA-AA work group further qualified this position and recommended that biological markers should reflect neuropathological processes that define the disease instead of simply supporting the diagnosis (Jack et al., 2018). Based on this expert consensus, the work group recommended that AD biomarkers should be incorporated into MCI/AD diagnostic criteria. The NIA-AA work group identified 3 types of AD biomarkers directly related to the underlying pathological processes. The biomarkers include (1) amyloid-β deposition including cortical amyloid positron emission tomography (PET) ligand

* Corresponding author at: Centre for Research on Ageing, Health and Wellbeing, Australian National University, Florey Building 54, Mills Road, Acton, ACT 2601 Australia. Tel.: (612) 61259032; fax: (612) 61251558.
E-mail address: hossein.tabatabaei@anu.edu.au (H. Tabatabaei-Jafari).

[1] 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.

0197-4580/$ – see front matter © 2018 Elsevier Inc. All rights reserved.
https://doi.org/10.1016/j.neurobiolaging.2018.05.006

bonding (F^{18}–flutemetamol PET) and low cerebrospinal fluid (CSF) $A\beta_{42}$; (2) aggregated tau including cortical tau PET ligand bonding (flortaucipir-PET) and elevated CSF phosphorylated tau (P-tau); and (3) neurodegeneration or neural injury including PET-detected hypometabolism (fluorodeoxyglucose-PET), CSF total tau (T-tau), and cortical/volume atrophy on magnetic resonance imaging (MRI) scan (Jack et al., 2018).

Much research has been conducted to evaluate amyloid-β deposition, tau aggregation, and hypometabolism using PET scans and CSF biomarkers—separately or in combination—to classify MCI at risk of AD conversion, with some promising performance (Mitchell, 2009; Ritchie et al., 2017; Vandenberghe et al., 2013; Yuan et al., 2009). However, these methods are invasive and, especially for PET imaging, have limited availability in clinical practice. Ideally, a practical biomarker should be widely available, accurate, cost-effective, relatively simple to interpret, easy to use, and be acceptable to patients while not imposing an excessive burden. It is important that—before assessing a new biomarker—clear criteria for selection be established, and the likelihood of meeting them be considered. As a minimum, the proposed new biomarker should perform at least similar to simple, noninvasive, and currently available biomarkers.

A type of noninvasive and more widely available biomarker is provided by structural brain measurement obtained using MRI. Cerebral cortical thickness and hippocampal measures are the most predictive and practical MRI methods to date (Falahati et al., 2014; Rathore et al., 2017). Although cerebral cortical thickness has been shown to be more predictive compared to volumetric measures based on single brain regions, it requires agreement on a standard pattern of cerebral cortical thickness in AD to be adoptable in clinical practice. Hippocampal volume, which has been shown to be a moderate predictor of AD conversion with a sensitivity of 67% and specificity of 72%, has the advantage of being less invasive compared to a CSF biomarker, less costly than a PET scan, and more widely available and clinically easier to use compared to cortical atrophy measures (Chupin et al., 2009). However, using hippocampal volume in the clinical setting is less straightforward compared to the use of this measure in a research setting.

Hippocampal volume needs to be normalized by or adjusted for intracranial volume (ICV) (Whitwell et al., 2001) to control for intersubject (Barnes et al., 2010) and gender (Pintzka et al., 2015) variations in head size, as well as variation in head size estimations in serial scans (Whitwell et al., 2001). The most widely used method in neuroimaging research is adjustment for ICV using its inclusion as a covariate in regression analyses. A less commonly used normalization approach is dividing the hippocampal volume by another volume that can be accurately measured and is not significantly impacted by neurodegenerative processes, typically ICV. In this study, we investigate normalization by cerebellar volume (hippocampus to cerebellar volume ratio) as an alternative approach, to correct for head size/premorbid brain volume as the cerebellum has been shown to be little affected by age-related atrophy in the absence of clinical dementia. Neurodegeneration in AD gradually progresses from the medial temporal lobe to the parietal and frontal lobes and then to the posterior parts of the brain. The cerebellum is among the last brain regions affected by AD pathology (Thal et al., 2002). We have recently shown that cerebellar atrophy is not different in MCI compared to normal aging (Tabatabaei-Jafari et al., 2017). Furthermore, while cerebellar atrophy increases in AD, it remains lower in other regions and particularly in the medial temporal lobe (Tabatabaei-Jafari et al., 2017). Thus, using the cerebellum as a reference area should be both methodologically robust and practical in a clinical context. Importantly, regional brain volume is more accurately measured than ICV using semi-automated methods, such as FreeSurfer (Heinen et al., 2016), and unlike ICV also less affected by field strength (Heinen et al., 2016; Nordenskjold et al., 2013) and segmentation method (Hansen et al., 2015; Keihaninejad et al., 2010; Malone et al., 2015).

Although hippocampal volume is not sufficiently accurate to be clinically useful as a single predictor of MCI who progress to AD, it is a useful benchmark. If other measures sufficiently improve the predictive value of hippocampal volume, they may be worth for further consideration. The Mini—Mental State Examination (MMSE) may be a good candidate. A recent Cochrane review indicated that the weighted sensitivity and specificity of the MMSE for conversion from MCI to AD are 54% and 80% in a limited number of available studies (Arevalo-Rodriguez et al., 2015). Moreover, evidence suggests that a combination of cognitive measures and hippocampal volume can improve the predictive value of hippocampal volume for predicting AD conversion in MCI (Devanand et al., 2008). Therefore, such a combination is also likely to improve on the classification performance of hippocampal volume for identifying MCI who convert to AD in short term from all those who do not convert.

In the present study, we investigated the classification performance of MMSE and hippocampal volume normalized by cerebellar volume or ICV both individually and in combination, to identify individuals with MCI who will convert to AD within 5 years. We expected that these combinations of measures would have classification accuracies high enough to be useful in clinical practice.

2. Methodology

2.1. Study participants

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 a principal investigator, Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD.

A total of 1289 participants with MCI ($n = 872$) or CN ($n = 417$) at baseline were considered for inclusion. All MCI participants who were stable for at least 6 months after baseline and converted to AD or reverted to CN within 5 years (confirmed with 2 consecutive stable diagnoses) or were stable for at least 5 years were included. Participants who were CN at baseline and were stable throughout the study were also included.

Based on diagnosis and diagnostic change, participants were categorized into 4 groups: (1) MCIc ($N = 187$), MCI patients who converted to AD in less than 5 years; (2) MCIs ($N = 112$), MCI patients who were stable for 5 years or more; (3) MCIr ($N = 39$), MCI patients who reverted to CN in less than 5 years; and (4) CN ($N = 322$), patients who remained cognitively healthy for the whole follow-up period.

Details of the diagnostic criteria can be found at the ADNI web site (http://www.adni-info.org/Scientists/AboutADNI.aspx). Briefly, participants were classified as CN if they had an MMSE greater than 24, had a clinical dementia rating (CDR) of 0, and did not meet diagnostic criteria for MCI, dementia, or depression. Participants were classified as MCI if they had an MMSE greater than 24, had a CDR of 0.5, had a subjective report of memory concern, had an objective memory loss, had preserved daily living activity, and did not meet diagnostic criteria for dementia. AD participants have MMSE scores less than 26, have a CDR of 0.5 or 1.0, and fulfill criteria for clinically probable AD according to the Institute of Neurological and Communicative Diseases and Stroke/Alzheimer's Disease and Related Disorders Association.

2.2. Neuroimaging acquisition and processing

Participants underwent high-resolution MRI brain scans on 1.5 (N = 335) or 3T (N = 325) scanners from General Electric, Siemens, or Philips (Milwaukee, WI; Germany; the Netherlands, respectively) using a standardized ADNI acquisition protocol for 3D MP-RAGE sequence (Jack et al., 2008). Baseline images that had undergone specific ADNI preprocessing correction steps to standardize images from different sites and platforms were obtained for this study: (1) grad wrap, a specific correction of image geometry distortion due to nonlinearity; (2) B1 nonuniformity, B1 calibration to correct the image intensity nonuniformity that results when RF transmission is performed with a more uniform body coil while reception is performed with a less uniform head coil; and (3) N3 correction, a histogram peak–sharpening algorithm applied after grad wrap and B1 correction. We conducted automatic volumetric segmentation using FreeSurfer (version 5.3, http://surfer.nmr.mgh.harvard.edu/), and the output images were visually checked for the hippocampal and cerebellar segmentations. The criterion was a clear segmentation error assessed by an experienced neuroscientist. Scans with segmentation errors were reprocessed and would only be excluded if the error could not be corrected. In this sample, no image was excluded.

2.3. Measurements

ICV was computed by the sum of the whole brain gray and white matter and CSF volumes. Total cerebellar volume was computed by summing the left and right cerebellar gray and white matter. Total hippocampal volume was the sum of the volumes of the left hippocampus and right hippocampus. Hippocampus to intracranial volume ratio (HCICV) was the ratio of total hippocampal volume to intracranial volume adjusted for age and field strength. Hippocampus to cerebellar volume ratio (HCCR) was the ratio of total hippocampal volume to total cerebellar volume adjusted for age and field strength. No significant correlation was detected between HCICV (correlation = −0.09) or HCCR (correlation = −0.09) and ICV. There was a moderate correlation between hippocampal volume and MMSE (r = 0.35, Supplementary Fig. 1). The residual method was used for all adjustments implemented by running a regression line between raw ratios and the variables using the whole data (Pintzka et al., 2015).

2.4. Statistical analysis

Statistical analyses were performed using the R statistical software (version 3.3.2). Data were checked for missing values and univariate and multivariate outliers using Mahalanobis distance. Discriminant analysis was used to estimate the predictive value of HCICV, HCCR, MMSE, and their combination for clinical status. The DiscriMiner package (version 0.01-29, https://CRAN.R-project.org/package=DiscriMiner) was used for descriptive discrimination and the MASS (version 7.3-45, http://www.stats.ox.ac.uk/pub/MASS4) and Caret package (version 6.3-73, https://CRAN.R-project.org/package=caret) for predictive discrimination (classification). Data were evaluated for normality of all measures (Q-Q plot), linearity, and multicollinearity and singularity (variation inflation factor) assumptions of discriminant analysis, which were all satisfied. Statistically significant heterogeneity of variance-covariance matrices was observed (Box's M-test; $\chi^2 > 51.19$, $p < 0.001$), and therefore, a quadratic classification procedure was used because linear discriminant analysis is known to perform poorly in the presence of heterogeneous covariance matrices (Tabachnick and Fidell, 2013).

For binary classification analyses using quadratic classification procedure, MCIc was contrasted with (1) CN, MCIs, and MCIr pooled together; (2) CN alone; and (3) MCIs and MCIr pooled together and CN was contrasted with MCIs and MCIr pooled together. The stability of the classification procedure was checked by a 10-fold cross-validation. The sample randomly partitioned into 10 equal-size subsamples. Nine subsamples (combined) were used as training data, and the remaining single subsample was retained as the validation data to evaluate predictive model. The process was repeated 10 times, with each of the 10 subsamples used only once as the validation data. The average of the results was provided with confidence interval. We measured reliability using the Kappa coefficient, a chance-corrected measure of agreement between the reference classification (categorized by long-term clinical follow-up) and predictive classification (classifications based on study measures) (Fritz and Wainner, 2001). The receiver operating characteristic (ROC) curve (package pROC version 1.9.1, http://www.biomedcentral.com/1471-2105/12/77/) and the area under the curve (AUC) were used to estimate the discriminant capacity of each model and DeLong's test was used to compare different models (Tabachnick and Fidell, 2013).

3. Results

3.1. Demography and brain measures

The average age of all participants together was 73.76 (SD = 6.80). Participants within the 4 diagnostic groups were similar in age, except for MCIr who were 3 to 5 years younger. *APOE e4* genotype was significantly higher, and MMSE scores were lower in the MCI subgroups than those in the CN group. The average time for MCIc to convert to AD and MCIr to revert to CN was similar at about 2 years. Baseline imaging measures showed that there was a trend of ascending hippocampal volume (adjusted for age, field strength, and ICV), HCICV, and HCCR values in MCIc, MCIs, MCIr, and CN. No such trend was detected for cerebellar volume (Table 1).

3.2. Discriminant analyses; descriptive statistics

Discriminant analyses were conducted to evaluate discriminative performance of the HCICV-MMSE and HCCR-MMSE models. Two discriminant functions were calculated for each model separately. The first function significantly distinguished among the diagnostic groups (HCICV-MMSE: F[6, 1310] = 74.556, HCCR-MMSE: F[6, 1310] = 70.096) and accounted for 99.6% of prediction of MCIc from CN, MCIs, and MCIr (first function's eigenvalue/sum of all eigenvalues × 100) in both models, whereas the second function was not effective in distinguishing CN, MCIs, and MCIr. Predictive values of the combination of HCICV and MMSE or HCCR and MMSE were almost equal (equal standardized coefficient correlation of predictors and discriminant functions) in the first discriminant functions for distinguishing among the groups (Supplementary Table 1).

The binary classification analyses revealed that HCICV, HCCR, and MMSE were equally predictive of MCIc with loadings of more than 0.5 on the discriminant functions (standardized coefficient correlation) with large effect sizes (canonical R^2 and eigenvalue) in all contrasts. In comparison, the standardized coefficients in CN group contrasted with MCIs and MCIr groups were more than 0.5, but because the effect sizes were very low, the discriminant functions were not effective in separating the groups (Supplementary Table 1).

Table 1
Characteristics: demographic information, MMSE, and brain measures. Trends of decrease in the average of MMSE and hippocampal measures are noticeable across the groups.

Diagnostic group	CN	MCIr	MCIs	MCIc	Test of significance ($p < 0.05$)	
Sample size	322	39	112	187	Across groups	Significant pairs
Age; y, mean (SD)	74.55 (5.80)	69.33 (8.32)	72.08 (7.65)	74.31 (7.02)	$F(3) = 10.09$[a]	CN vs. MCIr CN vs. MCIs MCIc vs. MCIr
Age range, y	59–90	55–87	57–88	55–89		
Male sex; N (%)	158 (49)	17 (44)	72 (64)	113 (60)	$\chi^2(3) = 12.68$	All pairs are different
Education, y; mean (SD)	16.38 (2.74)	16.87 (2.38)	15.75 (3.03)	16.09 (2.73)	$F(3) = 2.285$	No difference in pairs
APOE e4; N (%)	82 (25)	19 (49)	40 (36)	127 (68)	$\chi^2(3) = 90.63$[a]	All pairs are different
One allele	75 (23)	18 (46)	32 (29)	96 (51)		
Two alleles	7 (2)	1 (3)	8 (7)	31 (17)		
Age at DX change, y; mean (SD)	-	71.38 (8.31)	-	76.74 (7.15)	-	MCIc vs. MCIr
Time to DX change, y; mean (SD)	-	2.06 (1.14)	-	2.43 (0.91)	-	-
Measures						
MMSE; mean (SD)	29.08 (1.14)	28.85 (1.23)	28.11 (1.48)	26.95 (1.72)	$F(3) = 95.22$[a]	MCIc vs. CN MCIs vs. CN MCIr vs. MCIc MCIs vs. MCIc
Hippocampus, mm^3, mean (SD)[b]	7510.06 (807.29)	7210.85 (756.46)	7052.82 (909.03)	6240.78 (888.32)	$F(3) = 89.32$[a]	MCIc vs. CN MCIc vs. MCIr MCIc vs. MCIs MCIs vs. CN
Cerebellum, mm^3; mean (SD)[b]	121937.60 (9539.73)	120522.40 (9840.47)	121318.00 (10,337.83)	122673.50 (10,510.29)	$F(3) = 0.458$	No difference in pairs
HCICV; mean (SD)	0.50 (0.06)	0.47 (0.05)	0.46 (0.07)	0.41 (0.06)	$F(3) = 87.86$[a]	MCIc vs. CN MCIc vs. MCIr MCIc vs. MCIs MCIs vs. CN
HCCR; mean (SD)	6.21(0.73)	5.99 (0.68)	5.85 (0.94)	5.09 (0.79)	$F(3) = 79.83$[a]	MCIc vs. CN MCIc vs. MCIr MCIc vs. MCIs MCIs vs. CN

Key: APOE e4, apolipoprotein E allele 4; CN, cognitively normal; DX, diagnosis; HCCR, hippocampus to cerebellum volume ratio × 100 adjusted by age and field strength; HCICV, hippocampus to intracranial volume ratio × 100 adjusted by age and field strength; MCIc, mild cognitive impairment converted to Alzheimer's disease in 5 years; MCIr, mild cognitive impairment reverted to normal; MCIs, mild cognitive impairment stable for 5 years or more; MMSE, Mini–Mental State Examination.
[a] Indicates significance at $p < 0.0001$.
[b] Adjusted by age, field strength, and intracranial volume.

3.3. Discriminant analysis; classification

3.3.1. Individual predictor classification

HCICV, HCCR, and MMSE performed similarly in identifying diagnostic groups when tested individually and classified participants of the 4 diagnostic groups into 2 groups: CN and MCIc. A high proportion of CN and MCIc were correctly classified, whereas the majority of MCIs and MCIr were classified as CN and the remainder as MCIc (Table 2).

In binary classifications (Table 3), classification performance of MMSE, HCICV, and HCCR was generally comparable and more specific than sensitive for detecting MCIc from the other 3 groups: classification accuracy from 77.6% to 78.9%, specificity from 90.9% to 92%, and sensitivity from 41.2% to 47.1%. Similar trends were demonstrated in all other contrasts. ROC analyses demonstrated no statistically significant difference between AUC for MMSE, HCICV, and HCCR based on Delong's test in all contrasts (Table 3 and Fig. 1).

Table 2
Group classification performance: predictors separate MCIc from CN but cannot separate MCIs and MCIr from others and majority of them were classified as CN and minority as MCIc

References	MMSE				HCICV				HCCR				HCICV + MMSE				HCCR + MMSE			
	CN	MCIc	MCIs	MCIr	CN	MCIc	MCIs	MCIr	CN	MCIc	MCIs	MCIr	CN	MCIc	MCIs	MCIr	CN	MCIc	MCIs	MCIr
Prediction																				
CN	293	71	76	33	272	70	78	29	283	69	83	34	290	42	75	34	293	43	73	34
MCIc	29	116	36	6	50	117	34	10	39	118	29	5	27	144	37	5	25	142	36	5
MCIs	0	0	0	0	0	0	0	0	0	0	0	0	5	1	0	0	4	2	3	0
MCIr	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sensitivity, %	91.0	62.0	-	-	84.5	62.57	-	-	87.9	63.1	-	-	90.1	77.0	-	-	91.0	75.9	-	-
Specificity, %	46.8	85.0	-	-	47.6	80.13	-	-	45.0	84.6	-	-	55.3	85.4	-	-	55.6	86.1	-	-
Pos Pred Value %	62.0	62.0	-	-	60.6	55.45	-	-	60.3	61.8	-	-	65.8	67.6	-	-	66.1	68.3	-	-
Neg Pred Value %	84.5	85.0	-	-	76.3	84.41	-	-	79.6	85.3	-	-	85.4	90.4	-	-	86.6	90.1	-	-
Prevalence, %	48.8	28.3	17.0	5.9	48.8	28.33	17.0	5.9	48.8	28.3	17.0	5.9	48.8	28.3	17.0	5.9	48.8	28.3	17.0	5.9
Accuracy (95% CI)	62.0 (58.1–65.7)				58.94 (55.1–62.7)				60.8 (56.9–64.5)				65.8 (62.0–69.4)				66.4 (62.6–67.0)			
Kappa, %	33.3				28.90				31.3				41.1				42.1			

Key: 95% CI, 95% confidence interval; CN, cognitively normal; HCCR, hippocampus to cerebellum volume ratio adjusted for age and field strength; HCICV, hippocampus to intracranial volume ratio adjusted for age and field strength; MCIc, mild cognitive impairment converted to Alzheimer's disease in 5 years; MCIr, mild cognitive impairment reverted to normal; MCIs, mild cognitive impairment stable for 5 years or more; MMSE, Mini–Mental State Examination; Neg Pred value, negative predictive value; Pos Pred Value, positive predictive value.

Table 3
Contrast classification performance: MCIc contrasted separately with all 3 groups together, other 2 MCI groups, and CN alone. CN also contrasted with MCIs and MCIr together. In MCIc contrasts (with all groups or CN alone), predictors were mostly specific than being sensitive when they were not in combinations while combinations improved all classification performances.

Measurements	Classification accuracy % (95% CI)	Kappa, %	McNemar test, *p*-value	Sensitivity, %	Specificity, %	Positive predictive value, %	Negative predictive value, %	LR^+	LR^-	AUROC (95% CI)
MCIc vs. [CN + MCIs + MCIr]										
MMSE	77.6 (74.2–80.7)	37.5	<0.0001	41.2	92.0	67.0	69.8	5.2	0.6	0.80 (0.76–0.84)
HCICV	78.9(75.6–82.0)	44.0	<0.0001	50.3	90.3	67.1	82.1	5.2	0.6	0.82 (0.79–0.86)
HCCR	78.5 (75.2–81.6)	41.8	<0.0001	47.1	90.9	67.2	81.3	5.2	0.6	0.82 (0.78–0.85)
HCICV + MMSE	83.2 (80.1–86.0)	56.6	0.008	62.6	91.3	74.1	86.1	7.2	0.4	0.89 (0.86–0.91)
HCCR + MMSE	83.5 (80.4–86.2)	57.9	0.0554	65.2	90.7	73.5	86.8	7.0	0.4	0.88 (0.85–0.91)
MCIc vs. CN										
MMSE	80.4 (76.6–83.7)	55.7	<0.0001	62.1	91.0	80.0	80.5	6.9	0.4	0.84 (0.81–0.88)
HCICV	76.4 (72.5–80.1)	48.12	0.0828	62.6	84.5	70.1	79.5	4.0	0.4	0.86 (0.82–0.89)
HCCR	78.8 (75.0–82.3)	52.8	0.0053	63.1	87.9	75.2	80.4	5.2	0.4	0.85 (0.81–0.88)
HCICV + MMSE	85.5 (82.1–88.4)	68.3	0.2010	77.0	90.4	82.3	87.1	8.0	0.3	0.93 (0.90–0.95)
HCCR + MMSE	86.1 (82.7–88.9)	69.4	0.0576	76.5	91.6	84.1	87.0	9.1	0.3	0.92 (0.89–0.94)
MCIc vs. [MCIs + MCIr]										
MMSE	66.6 (61.3–71.6)	33.6	0.0084	62.0	72.2	73.4	60.6	2.2	0.5	0.72 (0.67–0.77)
HCICV	69.2 (64.0–74.1)	36.8	0.0241	78.6	57.6	69.7	68.5	1.9	0.4	0.75 (0.69–0.80)
HCCR	69.8 (64.6–74.7)	38.1	0.0376	78.6	58.9	70.3	69.0	1.9	0.4	0.75 (0.70–0.81)
HCICV + MMSE	74.6 (69.6–79.1)	48.3	0.5898	78.6	69.5	76.2	72.4	2.6	0.3	0.81 (0.76–0.85)
HCCR + MMSE	72.8 (67.7–77.5)	44.9	0.9170	75.9	68.9	75.1	69.8	2.4	0.4	0.80 (0.75–0.85)
CN vs. [MCIs + MCIr]										
MMSE	70.8 (66.5–74.9)	22.2	<0.0001	73.0	58.9	90.7	28.5	1.8	0.5	0.66 (0.61–0.72)
HCICV	69.1 (64.8–73.3)	12.7	<0.0001	93.8	16.6	70.6	55.6	1.1	0.4	0.65 (0.60–0.70)
HCCR	69.3 (65.0–73.5)	11.2	<0.0001	95.7	13.3	70.2	58.8	1.1	0.3	0.61 (0.55–0.66)
HCICV + MMSE	70.4 (66.01–74.5)	20.4	<0.0001	91.0	26.5	72.5	62.0	1.2	0.3	0.70 (0.65–0.75)
HCCR + MMSE	71.7 (67.4–75.7)	23.8	<0.0001	91.9	28.5	73.3	62.3	1.3	0.3	0.68 (0.63–0.73)

Key: 95% CI, 95% confidence interval; AUROC, area under receiver operating characteristic curve; CN, cognitively normal; HCCR, hippocampus to cerebellum volume ratio adjusted for age and field strength; HCICV, hippocampus to intracranial volume ratio adjusted for age and field strength; LR^+, positive likelihood ratio; LR^-, negative likelihood ratio; MCIc, mild cognitive impairment converted to Alzheimer's disease in 5 years; MCIr, mild cognitive impairment reverted to normal; MCIs, mild cognitive impairment stable for 5 years or more; MMSE, Mini–Mental State Examination.

Importantly, using ICV ratio to normalize the hippocampus or using regression to adjust for ICV was separately assessed, which was found to have little impact on the classification results (Supplementary Fig. 2).

3.3.2. Combined predictors classification

The combination of predictors (hippocampal and MMSE) improved almost all aspects of classification performance, but as for individual predictor models, classification was optimal in classifying participants into 2 groups: CN and MCIc. A high proportion of CN and MCIc were correctly classified, whereas a majority of MCIs and MCIr were misclassified as CN and a minority as MCIc (Table 2).

Almost all aspects of classification performance in all binary classifications that identified MCIc from other groups (i.e., MCIc vs. pooled of others, MCIc vs. CN, and MCIc vs. pooled of MCIs and MCIr) were improved with the combination of HCICV or HCCR and MMSE, when compared with the individual predictor. By contrast, combination models did not show improvement in discriminating CN group from pooled MCIs and MCIr groups (Table 3).

The discrimination ability (AUC of ROC analyses) of combinations of HCICV or HCCR and MMSE was significantly better than each predictor individually (Delong's test; $z < -4$, $p < 0.001$), while there was no significant difference between the HCICV-MMSE and HCCR-MMSE models. In addition, analyses suggested that there was no difference in discrimination ability between the combination models and MMSE alone in separating CN group from MCIs and MCIr groups. By contrast, the combination of hippocampal ratios (HCICV or HCCR) and MMSE was significantly better in discriminating MCIc from pooled MCIs and MCIr (Table 3 and Fig. 1). Additional analyses investigating the ability to discriminate MCI who convert within specified time periods (1–5 years) revealed that performance was better in the first 3 years of follow-up compared to the final 2 years (Supplementary Table 2).

Classification performance of the predictors in combination (HCCR-MMSE and HCICV-MMSE), for discriminating MCIc from other groups in all contrasts was generally substantial: classification accuracy for MCIc versus all other groups was more than 83% with sensitivity between 65.2% and 62.6%, with a specificity of 90.7%–91.3% and an AUC of 0.88–0.89. The performance was even better when discriminating MCIc from CN (Table 3).

Based on the partition plots in Fig. 2, individuals with MMSE scores of less than 25 were mostly classified as MCIc regardless of the HCICV and HCCR values. For individuals with higher MMSE values, lower hippocampal ratios were observed in those who were classified as MCIc. For example, for an MMSE score equal to 25, HCICV needed to be less than 0.6% or HCCR less than 7.5%, to be classified as MCIc. The thresholds for HCICV or HCCR were 0.5% and 6.3% for an MMSE of 26, 0.42% and 5.3% for 27, 0.38% and 4.8% for MMSE for 28. HCICV or HCCR needed to be less than 0.35% and 4.3%, respectively, for MCIc diagnosis, when MMSE scores were 29–30. These thresholds were slightly smaller for discriminating MCIc from CN.

4. Discussion

This study aimed to investigate the performance of hippocampal volume normalized to cerebellar volume as a new measure for the clinical discrimination of MCI individuals at risk of AD conversion within 5 years. A combination of HCCR and MMSE was most effective in identifying MCI at risk of conversion. The main findings were that (1) the combination of HCCR or HCICV and MMSE and MMSE performed better in classifying MCI at risk of AD conversion than each measure individually; (2) the classification performance of HCCR and MMSE was similar to that of HCICV and MMSE; and (3) CN and

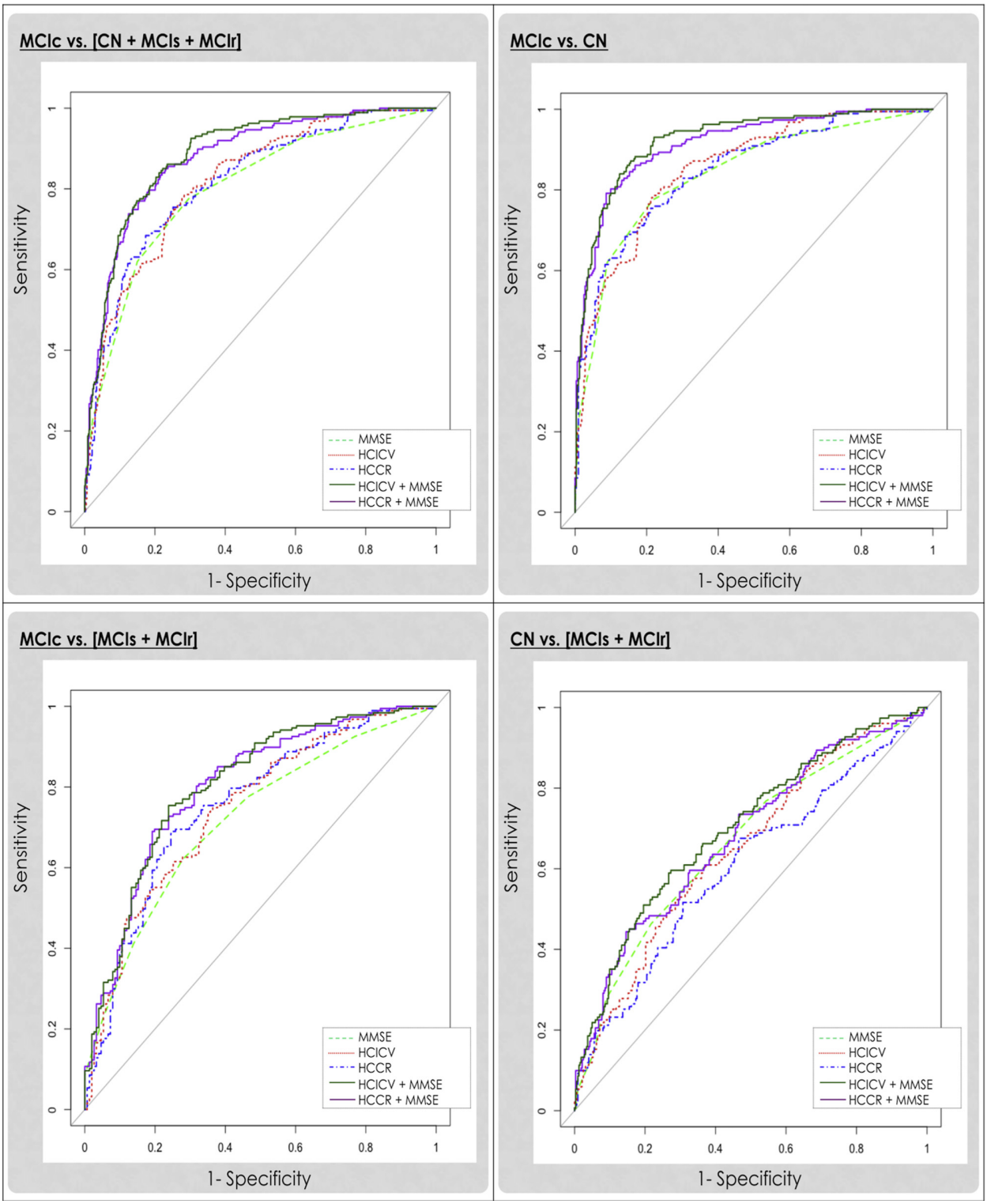


Fig. 1. Receiver operating characteristic (ROC) curve for group membership: Area under the curve (AUC) revealed that in mild cognitive impairment converted to Alzheimer (MCIc) contrasted with pooled of other groups (upper left) or cognitively normal (CN) alone (upper right), combination of Mini–Mental State Examination (MMSE) and hippocampus to intracranial volume ratio (HCICV) or hippocampus to cerebellum volume ratio (HCCR) was better than each predictor separately. This was partially true for MCIc contrasted with pooled of other mild cognitive impairment (MCI) groups (lower left), while not true for CN contrasted with other MCI groups (lower right).

MCI who did not convert to AD within 5 years did not differ statistically in their normalized hippocampal measures at a particular MMSE score.

Among the brain areas implicated in AD neuropathology, hippocampal shrinkage is most predictive of AD-related cognitive dysfunction (Jack et al., 2000), and MMSE is the most widely used

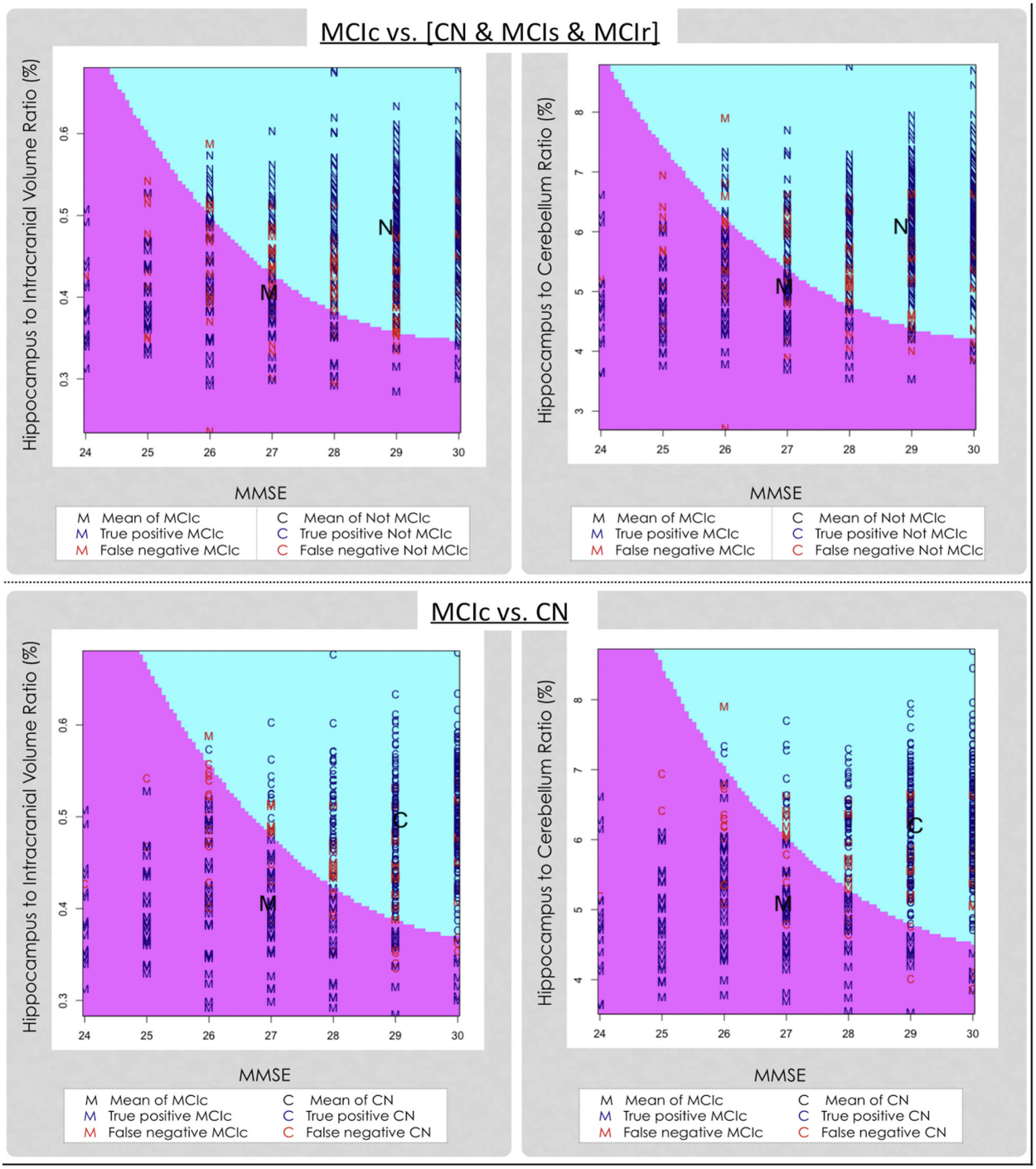


Fig. 2. Partition plots: Thresholds of different hippocampus to intracranial volume (HCICV, right) or hippocampus to cerebellum ratios (HCCR, left) based on different Mini–Mental State Examination (MMSE) scores, which separate mild cognitive impairment converted to Alzheimer (MCIc) from the pooled of cognitively normal (CN) and other mild cognitive impairment (MCI) groups (upper) and from CN alone (lower).

screening instrument for AD/dementia. We found that HCCR, a new normalized hippocampal measure, performed similar to HCICV in classification performance. Although none of HCICV, HCCR, or MMSE reliably identified MCI individuals who progressed to AD alone, we confirmed that HCICV or HCCR in combination with MMSE were effective in differentiating MCI patients who progressed to AD from CN and MCI patients who did not progress.

Both combinations were similar in performance and revealed a high level of classification accuracy, particularly for discriminating between MCIc and CN. However, classification accuracy only reflects the proportion of true results (positive or negative) in the sample. To be practical and useful, a test needs to be sensitive and specific. Our results revealed that of those with MCIc, 65.2%–62.6% were correctly identified (satisfactory sensitivity) by the combination models (HCCR + MMSE or HCICV + MMSE), while 91.3%–90.7%

of nonconverters (CN, MCIs, and MCIr) were correctly identified (high specificity). Furthermore, in those who were positively identified as MCIc, the likelihood of being truly MCIc was about nine-fold that of those who were falsely identified as MCIc (high positive likelihood ratio). For those who were positively identified as MCIc, the likelihood of being MCIc was close to a third that of those who were falsely identified as not having MCIc (low negative likelihood ratio). Therefore, not only was the overall accuracy of the combinations high, but the probabilities of false positive/negative results were also acceptable. Altogether, the combinations of hippocampal measures and MMSE are likely to be better than any single measure in identifying individuals with MCI at risk of AD conversion but also effective in ruling out those individuals unlikely to convert within 5 years.

Interestingly, using either a combination of HCICV and MMSE or HCCR and MMSE resulted in similar performance. This is important because it indicates that normalization of hippocampal volume by ICV or cerebellar volume is equally effective and thus validates our approach. ICV estimation is more sensitive to scanning parameters and segmentation methods than cerebellar volume. This is probably because ICV segmentation relies on the correct identification of the boundary between the subarachnoid space and CSF fluid whose contrast is more variable to that between cerebellar gray matter and CSF. Thus, cross-sectional comparison between patients (or longitudinal within patients) assessed with different scanning parameters may be more difficult when using the ICV ratio. Consequently, in these contexts, normalization by cerebellar volume may be more reliable and preferable.

The classification performance of HCICV and MMSE was in agreement with previous studies (in spite of different study parameters) that revealed a sensitivity of 67% and specificity of 72% for ICV-adjusted hippocampal volume and a sensitivity of 54% and specificity of 80% for MMSE in identifying MCIc from CN (Arevalo-Rodriguez et al., 2015; Chupin et al., 2009). Better performance for the combination models was consistent and comparable with a previous study that showed better prediction of a combination of hippocampal volume, entorhinal cortex volume, MMSE, informant report of functioning questionnaire, the University of Pennsylvania Smell Identification Test, and Selective Reminding Test immediate recall score with a sensitivity of 70% and a specificity of 90% (Devanand et al., 2008). In addition, the models' performances were comparable with other studies with combination of multiple modalities (including MRI and cognitive measures), which mostly had many predictors in each modality (Costafreda et al., 2011; Ferrarini et al., 2009; Moradi et al., 2015; Zhang et al., 2011). This suggests that adding more predictors into a model may not necessarily improve classification performance when the predictors are from a single domain. Therefore, similar to the comparability of the current findings with previous studies that used complex combinations of predictors, the combination of HCCR and MMSE has the advantage of being easily implementable and interpretable and thus may facilitate clinical adoption.

It is interesting to note that MCIs and MCIr did not differ from CN based on the combination of HCICV or HCCR and MMSE while they differed from MCIc. This suggests that those who are not at actual risk of short-term AD conversion are not substantially different from CN. A measure of concurrent decline in function and structure is likely to be a better predictor of AD conversion in short term.

Most classification studies conducted to date were predominantly based on multidomain/multivariate predictors and thus too complex to be easily adoptable in clinical practice. This study stands out in its use of a combination of simple structural (HCCR) and functional (MMSE) measures with a potential diagnostic value for identifying MCI subjects at risk of converting to AD in 5 years easily applicable in clinical practice.

5. Conclusion

The need to evaluate AD-related biological markers for identifying those at risk of AD conversion and to include them in MCI diagnosis has been well documented. However, there is no agreement on a biomarker that can be effectively applied in clinical practice. In the present study, we show that a combination of one brain biomarker, either HCCR or HCICV, and MMSE can accurately identify individuals at risk of AD conversion within 5 years. Moreover, normalization by cerebellar volume is as precise as normalization by intracranial volume with the advantage of being more practical in a clinical setting.

Disclosure statement

The authors have no actual or potential conflicts of interest.

Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI; National Institutes of Health grant U01 AG024904). 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 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 partly undertaken on the National Computational Infrastructure (NCI) facility in Canberra, Australia, which is supported by the Australian Commonwealth Government and by an Australian Government Research Training Program (RTP) Scholarship.

Authors' contributions: Statistical analysis was carried out by HT-J. TJ contributed to the design of the study, contributed to data management, conducted statistical analyses, and managed all aspects of article preparation and submission. EW contributed to the statistical analysis and article preparation. MES contributed to the image analysis and article preparation. NC contributed to the design of the study, provided methodological input and theoretical expertise, and contributed to writing and editing of the article.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.neurobiolaging.2018.05.006.

References

Albert, M.S., DeKosky, S.T., Dickson, D., Dubois, B., Feldman, H.H., Fox, N.C., Gamst, A., Holtzman, D.M., Jagust, W.J., Petersen, R.C., Snyder, P.J., Carrillo, M.C., Thies, B., Phelps, C.H., 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. Alzheimer's Dement. 7, 270–279.

Arevalo-Rodriguez, I., Smailagic, N., Roque, I.F.M., Ciapponi, A., Sanchez-Perez, E., Giannakou, A., Pedraza, O.L., Bonfill Cosp, X., Cullum, S., 2015. Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). Cochrane Database Syst. Rev. CD010783.

Association, T.A.P., 2013. Diagnostic and Statistical Manual of Mental Disorders, fifth ed. Washington, DC.

Barnes, J., Ridgway, G.R., Bartlett, J., Henley, S.M., Lehmann, M., Hobbs, N., Clarkson, M.J., MacManus, D.G., Ourselin, S., Fox, N.C., 2010. Head size, age and gender adjustment in MRI studies: a necessary nuisance? NeuroImage 53, 1244–1255.

Chupin, M., Gerardin, E., Cuingnet, R., Boutet, C., Lemieux, L., Lehericy, S., Benali, H., Garnero, L., Colliot, O. Alzheimer's Disease Neuroimaging, I, 2009. Fully automatic hippocampus segmentation and classification in Alzheimer's disease and mild cognitive impairment applied on data from ADNI. Hippocampus 19, 579–587.

Costafreda, S.G., Dinov, I.D., Tu, Z., Shi, Y., Liu, C.Y., Kloszewska, I., Mecocci, P., Soininen, H., Tsolaki, M., Vellas, B., Wahlund, L.O., Spenger, C., Toga, A.W., Lovestone, S., Simmons, A., 2011. Automated hippocampal shape analysis predicts the onset of dementia in mild cognitive impairment. NeuroImage 56, 212–219.

Devanand, D.P., Liu, X., Tabert, M.H., Pradhaban, G., Cuasay, K., Bell, K., Leon, M.J.D., Doty, R.L., Stern, Y., Pelton, G.H., 2008. Combining early markers strongly predicts conversion from mild cognitive impairment to Alzheimer's disease. Biol. Psychiatry 64, 871–879.

Falahati, F., Westman, E., Simmons, A., 2014. Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging. J. Alzheimers Dis. 41, 685–708.

Ferrarini, L., Frisoni, G.B., Pievani, M., Reiber, J.H., Ganzola, R., Milles, J., 2009. Morphological hippocampal markers for automated detection of Alzheimer's disease and mild cognitive impairment converters in magnetic resonance images. J. Alzheimers Dis. 17, 643–659.

Forlenza, O.V., Diniz, B.S., Stella, F., Teixeira, A.L., Gattaz, W.F., 2013. Mild cognitive impairment (part 1): clinical characteristics and predictors of dementia. Rev. Bras. Psiquiatr. 35, 178–185.

Fritz, J.M., Wainner, R.S., 2001. Examining diagnostic tests: an evidence-based perspective. Phys. Ther. 81, 1546–1564.

Ganguli, M., 2013. Can the DSM-5 framework enhance the diagnosis of MCI? Neurology 81, 2045–2050.

Hansen, T.I., Brezova, V., Eikenes, L., Haberg, A., Vangberg, T.R., 2015. How does the accuracy of intracranial volume measurements affect normalized brain volumes? Sample size estimates based on 966 subjects from the HUNT MRI cohort. Am. J. Neuroradiol. 36, 1450–1456.

Heinen, R., Bouvy, W.H., Mendrik, A.M., Viergever, M.A., Biessels, G.J., de Bresser, J., 2016. Robustness of automated methods for brain volume measurements across different MRI field strengths. PLoS One 11, e0165719.

Jack Jr., C.R., Bennett, D.A., Blennow, K., Carrillo, M.C., Dunn, B., Haeberlein, S.B., Holtzman, D.M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J.L., Montine, T., Phelps, C., Rankin, K.P., Rowe, C.C., Scheltens, P., Siemers, E., Snyder, H.M., Sperling, R. Contributors, 2018. NIA-AA research framework: toward a biological definition of Alzheimer's disease. Alzheimer's & dementia. J. Alzheimers Assoc. 14, 535–562.

Jack Jr., C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., L Whitwell, J., Ward, C., Dale, A.M., Felmlee, J.P., Gunter, J.L., Hill, D.L., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C.S., Krueger, G., Ward, H.A., Metzger, G.J., Scott, K.T., Mallozzi, R., Blezek, D., Levy, J., Debbins, J.P., Fleisher, A.S., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., Weiner, M.W., 2008. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J. Magn. Reson. Imaging 27, 685–691.

Jack Jr., C.R., Petersen, R.C., Xu, Y., O'Brien, P.C., Smith, G.E., Ivnik, R.J., Boeve, B.F., Tangalos, E.G., Kokmen, E., 2000. Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. Neurology 55, 484–489.

Keihaninejad, S., Heckemann, R.A., Fagiolo, G., Symms, M.R., Hajnal, J.V., Hammers, A. Alzheimer's Disease Neuroimaging, I, 2010. A robust method to estimate the intracranial volume across MRI field strengths (1.5T and 3T). NeuroImage 50, 1427–1437.

Malone, I.B., Leung, K.K., Clegg, S., Barnes, J., Whitwell, J.L., Ashburner, J., Fox, N.C., Ridgway, G.R., 2015. Accurate automatic estimation of total intracranial volume: a nuisance variable with less nuisance. NeuroImage 104, 366–372.

Mitchell, A.J., 2009. CSF phosphorylated tau in the diagnosis and prognosis of mild cognitive impairment and Alzheimer's disease: a meta-analysis of 51 studies. J. Neurol. Neurosurg. Psychiatry 80, 966–975.

Moradi, E., Pepe, A., Gaser, C., Huttunen, H., Tohka, J. Initiative, A.s.D.N, 2015. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. NeuroImage 104, 398–412.

Nordenskjold, R., Malmberg, F., Larsson, E.M., Simmons, A., Brooks, S.J., Lind, L., Ahlstrom, H., Johansson, L., Kullberg, J., 2013. Intracranial volume estimated with commonly used methods could introduce bias in studies including brain volume measurements. NeuroImage 83, 355–360.

Pandya, S.Y., Clem, M.A., Silva, L.M., Woon, F.L., 2016. Does mild cognitive impairment always lead to dementia? A review. J. Neurol. Sci. 369, 57–62.

Park, M.H., Han, C. Initiative, t.A.s.D.N, 2015. Is there an MCI reversion to cognitively normal? Analysis of Alzheimer's disease biomarkers profiles. Int. Psychiatry 27, 429–437.

Pintzka, C.W., Hansen, T.I., Evensmoen, H.R., Haberg, A.K., 2015. Marked effects of intracranial volume correction methods on sex differences in neuroanatomical structures: a HUNT MRI study. Front. Neurosci. 9, 238.

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 155, 530–548.

Ritchie, C., Smailagic, N., Noel-Storr, A.H., Ukoumunne, O., Ladds, E.C., Martin, S., 2017. CSF tau and the CSF tau/ABeta ratio for the diagnosis of Alzheimer's disease dementia and other dementias in people with mild cognitive impairment (MCI). Cochrane Database Syst. Rev. CD010803.

Tabachnick, B.G., Fidell, L.S., 2013. Using Multivariate Statistics, sixth ed. Pearson, London, UK.

Tabatabaei-Jafari, H., Walsh, E., Shaw, M.E., Cherbuin, N. Alzheimer's Disease Neuroimaging, I, 2017. The cerebellum shrinks faster than normal ageing in Alzheimer's disease but not in mild cognitive impairment. Hum. Brain Mapp. 38, 3141–3150.

Thal, D.R., Rub, U., Orantes, M., Braak, H., 2002. Phases of A beta-deposition in the human brain and its relevance for the development of AD. Neurology 58, 1791–1800.

Vandenberghe, R., Nelissen, N., Salmon, E., Ivanoiu, A., Hasselbalch, S., Andersen, A., Korner, A., Minthon, L., Brooks, D.J., Van Laere, K., Dupont, P., 2013. Binary classification of (1)(8)F-flutemetamol PET using machine learning: comparison with visual reads and structural MRI. NeuroImage 64, 517–525.

Whitwell, J.L., Crum, W.R., Watt, H.C., Fox, N.C., 2001. Normalization of cerebral volumes by use of intracranial volume: Implications for longitudinal quantitative MR imaging. Am. J. Neuroradiol. 22, 1483–1489.

Yuan, Y., Gu, Z.X., Wei, W.S., 2009. Fluorodeoxyglucose-positron-emission tomography, single-photon emission tomography, and structural MR imaging for prediction of rapid conversion to Alzheimer disease in patients with mild cognitive impairment: a meta-analysis. Am. J. Neuroradiol. 30, 404–410.

Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D. Alzheimer's Disease Neuroimaging, I., 2011. Multimodal classification of Alzheimer's disease and mild cognitive impairment. NeuroImage 55 (3), 856–867.

Supplementary table 1: Descriptive discriminants analysis of predictors: Predictors are generally significant with loadings more than 0.5 on the discriminant functions (standardised coefficient correlation) with large effect sizes (canonical R^2 and eigenvalue) across all groups as well as in all contrasts except for CN contrasting pooled of MCIs and MCIr, which the effect sizes are small and not effective in separating the groups.

		Correlation of predictors with discriminant functions (std. coef.)		Pooled within-group correlation among predictors	Univariate significance
		1	2	MMSE	F [DF]
All groups					
HCICV + MMSE					
	HCICV	-0.675	-0.740	0.024	87.86[3, 656]
	MMSE	-0.705	-0.711		95.22[3, 656]
	Canonical R^2	0.443	0.002		
	Eigenvalue	0.793	0.002		
HCCR + MMSE					
	HCCR	-0.644	0.768	0.035	79.83 [3, 656]
	MMSE	-0.715	-0.703		95.22 [3, 656]
	Canonical R^2	0.425	0.003		
	Eigenvalue	0.738	0.003		
MCIc vs. [CN, MCIs and MCIr]					
HCICV + MMSE					
	HCICV	0.666		0.108	221.73 [1,658]
	MMSE, std. coef	0.683			231.08 [1,658]
	Canonical R^2	0.385			
	Eigenvalue	0.627			
HCCR + MMSE					
	HCCR	-0.655		0.119	215.79 [1,658]
	MMSE	-0.683			231.08 [1,658]
	Canonical R^2	0.379			
	Eigenvalue	0.609			
MCIc vs. CN					
HCICV + MMSE					
	HCICV	-0.700		0.013	275.73 [1,507]
	MMSE	-0.707			281.25 [1,507]
	Canonical R^2	0.521			
	Eigenvalue	1.088			
HCCR + MMSE					
	HCCR	-0.676		0.032	256.09 [1,507]
	MMSE	-0.710			281.25 [1,507]
	Canonical R^2	0.505			
	Eigenvalue	1.019			
MCIc vs [MCIs and MCIr]					
HCICV + MMSE					

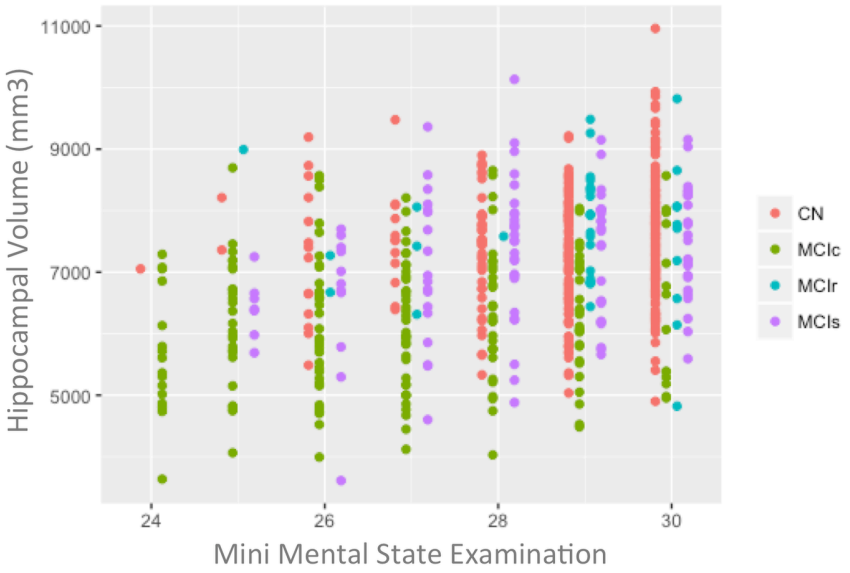
	HCICV	-0.717		0.082	69.94[1, 336]
	MMSE	-0.647			58.46 [1, 336]
	Canonical R^2	0.263			
	Eigenvalue	0.356			
HCCR + MMSE					
	HCCR	-0.717		0.142	75.00 [1, 336]
	MMSE	-0.610			58.46 [1, 336]
	Canonical R^2	0.260			
	Eigenvalue	0.352			
CN vs. [MCIs and MCIr]					
HCICV + MMSE					
	HCICV	-0.617		0.080	30.45 [1, 471]
	MMSE	-0.734			40.72 [1, 471]
	Canonical R^2	0.122			
	Eigenvalue	0.139			
HCCR + MMSE					
	HCCR	-0.498		0.069	17.47 [1, 471]
	MMSE	-0.826			40.72[1, 471]
	Canonical R^2	0.103			
	Eigenvalue	0.115			

CN; cognitively normal, MCIr; mild cognitive impairment reverted to normal, MCIc; mild cognitive impairment converted to Alzheimer's disease in five years, MCIs; mild cognitive impairment stable for five years or more, APOE e4; Apolipoprotein E allele 4, MMSE; mini mental status examination, HCICV; Hippocampus to intracranial volume ratio × 100 adjusted for age and field strength, HCCR; Hippocampus to Cerebellum volume ratio ×100 adjusted for age and field strength, std.coef; standardized coefficient

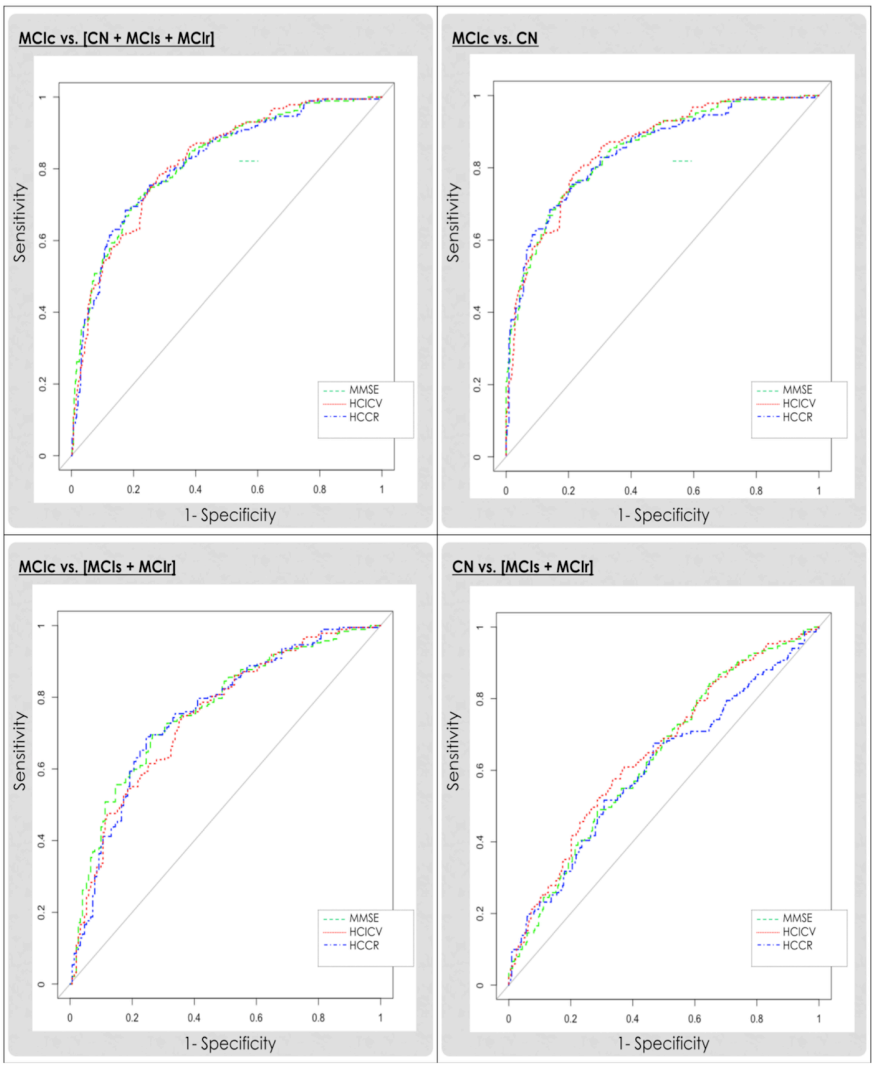
Supplementary Table-2. Area under receiver operating characteristic curve of mild cognitive impairment convert to Alzheimer's disease in one up to five years vs. pooled of mild cognitive impairment remain stable for five years or more and those who revert to cognitively normal

MCIc vs. [MCIs & MCIr]	AUROC (95% CI)	
	HCICV-MMSE	HCCR-MMSE
MCI convert in year 1	0.89 (0.74 – 0.99)	0.93 (0.85 – 0.99)
MCI convert in year 2	0.75 (0.67 – 0.82)	0.75 (0.67 - 0.82)
MCI convert in year 3	0.78 (0.72 – 0.85)	0.79 (0.72 – 0.85)
MCI convert in year 4	0.72 (0.63 – 0.80)	0.75 (0.67 – 0.83)
MCI convert in year 5	0.66 (0.54 – 0.79)	0.68 (0.56 – 0.81)

MCIc; mild cognitive impairment convert to Alzheimer's disease, MCIs; mild cognitive impairment remain stable for five years or more, MCIr; mild cognitive impairment revert to cognitively normal, AUROC; area under receiver operating characteristic curve, HCICV; hippocampus to intracranial volume ratio adjusted for age and field strength, HCCR; hippocampus to cerebellar volume ratio adjusted for age and field strength, MMSE; mini mental state examination.



Supplementary figure 1: Total hippocampal volume (mm^3) at different mini mental state examination scores in four diagnostic groups



Supplementary Figure 2: Receiver Operating Characteristic (ROC) curve for the three predictors: Area under curves reveal similar discrimination for the predictors in all group contrasts.

Journals of Gerontology: Psychological Sciences
cite as: *J Gerontol B Psychol Sci Soc Sci*, 2019, Vol. XX, No. XX, 1–10
doi:10.1093/geronb/gbz011
Advance Access publication January 21, 2019

Original Research Report

Cognitive/Functional Measures Predict Alzheimer's Disease, Dependent on Hippocampal Volume

Hossein Tabatabaei-Jafari, MD,[1,*] Marnie E. Shaw, PhD,[2] Erin Walsh, PhD,[1] and Nicolas Cherbuin PhD[1] for the Alzheimer's Disease Neuroimaging Initiative (ADNI)

[1]Centre for Research on Ageing, Health and Wellbeing and [2]College of Engineering and Computer Science, The Australian National University, Canberra, Australia.

*Address correspondence to: Hossein Tabatabaei-Jafari, MD, Centre for Research on Ageing, Health and Wellbeing, Australian National University, Florey Building 54, Mills Road, Acton, ACT 2601 Australia. E-mail: hossein.tabatabaei@anu.edu.au

Received: April 18, 2018; Editorial Decision Date: January 12, 2019

Decision Editor: Nicole Anderson, PhD, CPsych

Abstract

Objectives: This study aimed to investigate the predictive value of cognitive/functional measures in combination with hippocampal volume (HCV) on the probability of conversion from mild cognitive impairment (MCI) to Alzheimer's disease (AD).
Methods: The Rey Auditory Verbal Learning Test for immediate memory, Mini-Mental State Examination, a functional assessment for independent daily activities and Alzheimer's Disease Assessment Scale were used as cognitive/functional measures and HCV as neuroimaging measure. Logistic regression and Cox proportional hazard analyses were used to explore the measures' predictive values for AD conversion and time to conversion.
Results: The probability of conversion from MCI to AD was associated with cognitive function, but this was moderated by HCV: higher at lower HCV and lower at higher HCV. General cognitive/functional measures were less predictive than immediate memory in predicting time to conversion to AD at small HCVs.
Conclusion: Effectiveness of cognitive measures and subtle functional abnormality in predicting conversion from MCI to AD is dependent on HCV, thus combined evaluation should be considered. A combination of HCV and immediate memory appear to perform best in predicting time to conversion.

Key words: Brain/cognitive reserve, Hippocampus, Mild cognitive impairment, MRI, Neuropsychological tests

Alzheimer's disease (AD) is a progressive degenerative disorder that involves cognitive decline severe enough to substantially impair daily activities. Cognitive decline accompanied by preserved daily activities has been specified as mild cognitive impairment (MCI), and is commonly known to be the prodromal phase of AD (Petersen et al., 1999). Approximately half of those with MCI progress to AD within 5 years (Pandya, Clem, Silva, & Woon, 2016). Identifying those who will progress to AD and predicting time to conversion remains an important clinical challenge.

Cognitive and functional performance is the central component of AD/MCI diagnostic. Thus, it is to be expected that cognitive performance is a sensitive predictor of conversion from MCI to AD (Belleville et al., 2017). A combination of measures from a range of domains typically provides a better predictor of disease progression (Belleville et al., 2017). Additionally, although intact daily function is the main clinical differentiator of MCI and AD diagnosis, subtle decline in daily function, while it remains in the normal range, is still predictive of conversion from MCI to AD

(Gomar et al., 2011; Li et al., 2017). Furthermore, a combination of cognitive/functional measures with neuroimaging measures has been reported to produce significantly higher predictive accuracy (Devanand et al., 2008; Falahati, Westman, & Simmons, 2014; Moradi et al., 2015). Our recent study showed that the combination of a new hippocampal index—hippocampus to cerebellum volume ratio, HCCR—and Mini-Mental State Examination (MMSE) could reliably identify those who progress from MCI to AD within 5 years with an area under receiver operating characteristic curve of 0.9 (Tabatabaei-Jafari, Walsh, Shaw, & Cherbuin, 2018).

Cognitive/functional impairment is positively associated with neurodegeneration, but this association is not straightforward and there is a mismatch between the extent of neural pathology and the severity of cognitive/functional impairment (Steffener & Stern, 2012). Although a combination of cognitive performance and neuroimaging measures has been previously shown to have a higher predictive value compared with either measure alone, the relative contribution of these measures to each other across their range is not well understood. To answer these important questions, this study aimed to investigate the predictive value of cognitive/functional measures across the range of hippocampal volumes (HCVs), in those who have a diagnosis of MCI and convert to AD within 5 years. HCV and cognitive/functional measures were selected on the basis of established associations with MCI and AD (Jack et al., 2005; Li et al., 2017; Tabatabaei-Jafari, Shaw, & Cherbuin, 2015; Tabatabaei-Jafari et al., 2018). We hypothesized that HCV would moderate the predictive value of cognitive/functional performance. Additionally, we aimed to investigate how well a combination of these measures would predict time to conversion from MCI to AD.

Method

Study Participants

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 MCI and early AD.

All participants of ADNI 1/GO/2 who were diagnosed with MCI at baseline were considered for inclusion. Those who were stable for at least 6 months after baseline diagnosis were included if they converted to AD within 5 years (MCIc; n = 183) or remained stable for more than 5 years (MCIs; n = 112).

Details of the diagnostic criteria can be found at the ADNI web site (http://www.adni-info.org/Scientists/AboutADNI.aspx). Briefly, participants were classified as MCI if they had an MMSE greater than 24, a CDR of 0.5, a subjective report of memory concern, an objective memory loss, preserved daily living activity and did not meet diagnostic criteria for dementia. Participants were classified as having AD if they had MMSE scores less than 26, CDR 0.5 or 1.0 and fulfill criteria for clinically probable AD according to the Institute of Neurological and Communicative Diseases and Stroke/Alzheimer's Disease and Related Disorders Association.

Neuroimaging Acquisition and Processing

Participants underwent high-resolution MRI brain scans on 1.5 (N = 165) or 3 T (N = 130) scanners from General Electric, Siemens, or Philips (Milwaukee, WI; Germany; the Netherlands, respectively) using a standardized ADNI acquisition protocol for 3D MP-RAGE sequence (Jack et al., 2008). Images which had undergone specific ADNI preprocessing correction steps to standardize images from different sites and platforms, were obtained for this study: (a) Grad wrap; a specific correction of image geometry distortion due to nonlinearity, (b) B1 nonuniformity; B1 calibration to correct the image intensity nonuniformity that results when RF transmission is performed with a more uniform body coil while reception is performed with a less uniform head coil, (c) N3 correction; a histogram peak sharpening algorithm applied after grad wrap and B1 correction.

FreeSurfer version 5.3 (http://surfer.nmr.mgh.harvard.edu/) was used for automatic volumetric segmentation. The output images were visually checked for accurate segmentation.

Measures

One neuroimaging, three cognitive and one functional measure that have been extensively used for diagnostic purposes and cognitive and functional evaluation in clinical trials (Estévez-Gonzalez, Kulisevsky, Boltes, Otermin, & Garcia-Sanchez, 2003; Ito, Hutmacher, & Corrigan, 2012; Petersen et al., 2005) and with established associations with AD and predictive of MCI conversion (Ito et al., 2012; Li et al., 2017) were considered.

Hippocampal Volume

The hippocampus is one of the first brain areas to be impacted by AD pathology, and one of the areas with greatest shrinkage over the course of the disease (Tabatabaei-Jafari et al., 2015). It is also the most sensitive structural predictor of AD conversion in MCI individuals (Eckerstrom et al., 2008). Therefore, HCV, the total volume of the left and right hippocampi adjusted for age, field strength, and ICV using the residual regression method described elsewhere (Pintzka, Hansen, Evensmoen, & Haberg, 2015) was investigated as neuroimaging predictor.

Mini-Mental State Examination

The MMSE (Folstein, Folstein, & McHugh, 1975) is the most widely used screening instrument for AD/dementia (Arevalo-Rodriguez et al., 2015). It consists of 11 items with total scores ranging between 0 and 30, which lower scores reflecting more severe cognitive impairment. The items evaluate orientation in time and space (10 points), immediate recall (3 points), attention and calculation (5 points), delayed recall (3 points), language naming (2 points), following command (3 points), repetition (1 point), reading (1 point), writing (1 point), and visuospatial (1 point).

The Alzheimer's Disease Assessment Scale

The modified 13-items Alzheimer's Disease Assessment Scale (ADAS)-cog version (Petersen et al., 2005) was used here to assess general cognitive function. The modified ADAS consists of word recall (10 items), commends (5 items), construction (5 items), naming (5 items), ideational praxis (5 items), orientation (8 items), word recognition (12 items), recall instruction (5 items), spoken language (5 items), word finding (5 items), comprehension (5 items), delayed word recall (10 items), and number cancelation (5 items) in total 85 scores, which the higher score the severest the cognitive impairment.

Rey Auditory Verbal Learning Test

The Rey Auditory Verbal Learning Test (RAVLT) was used to evaluate episodic memory (Rey, 1941, 1964). It involves free recall of a list of 15 words in any order over five sequential trials. It is followed by recall of a second list of 15 words. Finally, the participant is asked to remember as many words as possible from the first list immediately following the second list recall and after 30 min. The scoring system of the RAVLT based on the correct number of words in each trial (5 in total) and evaluates a wide diversity of learning and memory functions including immediate memory, learning, and forgetting. The immediate recall score, RAVLT immediate, was considered for this study based on our introductory analyses that showed better predictive value for immediate memory compared with RAVLT learning and percentage of forgetting. The RAVLT immediate was computed as the total scores of trials 1–5.

The Functional Assessment Questionnaire

The Functional Assessment Questionnaire (FAQ) assesses abilities of daily living with total scores ranging from 0 to 30. A score of 0 indicates "no impairment" and 30 "severely impaired" (Ito et al., 2012; Pfeffer, Kurosaki, Harrah, Chance, & Filos, 1982). The total FAQ score is the sum of 10 daily activities, with each activity being rated from 0 to 3 (0 = *normal*, 1 = *has difficulty but does by self*, 2 = *requires assistance*, 3 = *dependent*). Evaluated activities are (a) writing checks, paying bills, or balancing a checkbook, (b) assembling tax records, business affairs, or other papers, (c) shopping alone for clothes, household necessities, or groceries, (d) playing a game of skill such as bridge or chess or working on a hobby, (e) heating water, making a cup of coffee, turning off the stove, (6) preparing a balanced meal, (f) keeping track of current events, (g) paying attention to and understanding a TV program, book, or magazine, (h) remembering appointments, family occasions, holidays, medications, and (i) traveling out of the neighborhood, driving, or arranging to take public transportation.

Statistical Analysis

Statistical analyses were performed using the R statistical software (version 3.3.2). No missing values were present in the measures of interest. Mahalanobis distance was used for detection of univariate and multivariate outliers. No influential outlier was detected. Group differences in demographic variables were assessed by *t*-test for continuous variables and chi square tests for categorical variables. Univariate and bivariate models were used to investigate prediction of conversion from MCI to AD within 5 years as well as prediction of the time to conversion. Each bivariate model consisted of standardized values of HCV and one of four cognitive/functional measures as well as their interaction. The alpha level was set at <0.05.

Prediction of AD Conversion

Logistic regression analysis (package Stats; version 3.3.2 and package Caret; version 6.3–73) was used to quantify the magnitude of predictive values of the measures for predicting MCI conversion to AD. Univariate and bivariate models were applied. The odds ratios were used to quantify the magnitude of the main and interaction effects of the predictors. To graphically illustrate the effect of HCV, the probability of conversion for the cognitive/functional measures at different categories of HCV was investigated. Participants were categorized into three groups; small HCV for those with HCV less than 5,500 mm^3 (smaller than 1 *SD*), medium HCV for the volume between 5500 and 7500 mm^3 (within 1 *SD*), and large HCV for those with larger than 7,500 mm^3 (larger than 1 *SD*).

Prediction of Time to AD Conversion

Cox proportional hazard analysis (package survival; version 2.40-1) was used to predict the time to AD conversion using the univariate and bivariate models. The hazard ratio for 1 *SD* change in the measures was used to quantify the magnitude of the main and interactive effects of the measures. In the case of the presence of interactive effect, to better interpret the effect the analyses were

repeated with HCV as a categorical variable (small, medium, and large) in the model. To graphically illustrate the contribution of cognitive/functional measures and HCV on probability of remaining MCI over time, separate Kaplan–Meier curves were plotted for different combinations of categorical levels of HCV (small, medium, and large as defined above) and cognitive/functional measures (low and high). Cognitive/functional measures were categorized into low and high based on the median: 27 for MMSE, 13 for ADAS, 2 for FAQ, and 31 for RAVLT. Participants were categorized into six combinations for each cognitive/functional measure (Supplementary Figure 1). For example, for ADAS, they were categorized into small HCV/low ADAS, small HCV/high ADAS, medium HCV/low ADAS, medium HCV/high ADAS, large HCV/low ADAS, and large HCV/high ADAS.

Table 1. Participants Characteristics and Measurements

Characteristics/Measures	MCIs	MCIc	Group difference
Sample size	112	183	
Age; year, mean (*SD*)	71.95 (7.65)	74.31 (6.90)	Yes
Age range, year	57–88	55–89	—
Male sex; *N* (%)	72 (64.29)	112 (61.20)	No
Education; year, mean (*SD*)	15.75 (3.03)	16.03 (2.73)	No
APOE e4; *N* (%)	40 (35.71)	124 (67.76)	Yes
One allele	32 (28.57)	93 (49.21)	Yes
Two alleles	8 (7.14)	31 (17.32)	Yes
Age at diagnosis change; year, mean (*SD*)	—	76.83 (7.05)	—
Time to diagnosis change; year, mean (*SD*)	—	2.40 (0.89)	—
MMSE, mean (*SD*)	28.11 (1.49)	26.93(1.73)	Yes
ADAS, mean (*SD*)	13.45 (5.45)	20.19 (5.49)	Yes
RAVLT immediate, mean (*SD*)	38.40 (10.34)	28.85 (7.11)	Yes
FAQ, mean (*SD*)	1.75 (3.00)	4.96 (4.62)	Yes
HCV,[a] mm^3, mean (*SD*)	7052.82 (909.03)	6223.92 (875.56)	Yes

Note. MCIc = mild cognitive impairment converted to Alzheimer's disease within 5 years; MCIs = mild cognitive impairment stable for 5 or more years; APOE e4 = apolipoprotein E allele 4; MMSE = mini-mental state examination; ADAS = Alzheimer Disease Assessment Scale (cognitive subscale); RAVLT = Rey Auditory Verbal Learning Test; FAQ = functional assessment questionnaire.
[a]Adjusted by age, field strength, and intracranial volume.

Table 2. Logistic Regression and Cox Proportional Hazard Results: Bivariate Models

	Prediction of conversion				Prediction of time to conversion			
Variables	Coef.	*SE*	OR (95% CI)	*Z*, *p* value	Coef.	*SE*	HR (95% CI)	*Z*, *p* value
HCV and MMSE								
HCV	**−0.92**	**0.16**	**0.40 (0.29–0.54)**	**−5.67, $p < .0001$**	**−0.53**	**0.08**	**0.59 (0.51–0.68)**	**−6.89, $p < .00001$**
MMSE	**−0.63**	**0.15**	**0.53 (0.39–0.71)**	**−4.17, $p < .0001$**	**−0.40**	**0.08**	**0.66 (0.57–0.78)**	**−5.01, $p < .0001$**
HCV: MMSE	**−0.35**	**0.17**	**0.71 (0.50–0.98)**	**−2.05, $p < .05$**	**−0.25**	**0.08**	**0.78 (0.66–0.91)**	**−3.18, $p = .002$**
HCV and ADAS								
HCV	**−0.66**	**0.17**	**0.52 (0.37–0.72)**	**−3.86, $p = .0001$**	**−0.41**	**0.08**	**0.67 (0.57–0.78)**	**−5.05, $p < .0001$**
ADAS	**1.18**	**0.19**	**3.26 (2.27–4.83)**	**6.16, $p < .0001$**	**0.64**	**0.08**	**1.91 (1.62–2.25)**	**7.79, $p < .0001$**
HCV: ADAS	0.34	0.22	1.41 (0.92–2.14)	1.59, $p = .11$	**0.23**	**0.09**	**1.26 (1.07–1.49)**	**2.70, $p < .01$**
HCV and FAQ								
HCV	**−0.95**	**0.17**	**0.39 (0.27–0.54)**	**−5.56, $p < .0001$**	**−0.50**	**0.07**	**0.61 (0.53–0.70)**	**−6.99, $p < .0001$**
FAQ	**1.05**	**0.22**	**2.84 (1.90–4.55)**	**4.72, $p < .0001$**	**0.38**	**0.06**	**1.46 (1.30–1.65)**	**6.28, $p < .0001$**
HCV: FAQ	0.15	0.23	1.16 (0.73–1.77)	0.65, $p = 0.52$	0.06	0.06	1.06 (0.95–1.19)	1.09, $p = 0.28$
HCV and RAVLT								
HCV	**−0.92**	**0.17**	**0.40 (0.28–0.55)**	**−5.44, $p < .0001$**	**−0.49**	**0.08**	**0.61 (0.53–0.71)**	**−6.46, $p < .0001$**
RAVLT	**−0.18**	**0.20**	**0.31 (0.21–0.44)**	**−6.04, $p < .0001$**	**−0.75**	**0.10**	**0.47 (0.39–0.58)**	**−7.45, $p < .0001$**
HCV: RAVLT	−0.09	0.22	0.91 (0.59–1.40)	−0.41, $p = .68$	−0.17	0.84	0.84 (0.70–1.01)	−1.84, $p = .07$

Note. MMSE = mini-mental state examination (standardized); ADAS = Alzheimer Disease Assessment Scale (standardized); RAVLT = Rey Auditory Verbal Learning Test (immediate; standardized); FAQ = Functional Assessment Questionnaire (standardized); HCV = hippocampal volume adjusted by age, field strength, and intracranial volume (standardized).

Results

Participants' Characteristics

Two hundred ninety-five MCI participants were categorized as MCI who subsequently converted to AD within 5 years (MCIc; n = 183), and MCI who were stable for more than 5 years (MCIs; n = 112). MCIs participants were about 2 years younger than MCIc but there were no significant differences in sex ratio or education between the two groups. The proportion of APOE e4 carriers was significantly higher in MCIc than MCIs. All the measures of interest (HCV and cognitive/functional measures) were significantly different between the groups (Table 1).

Prediction of AD Conversion

HCV, MMSE, ADAS, RAVLT, and FAQ were evaluated separately (univariate model) and all were significant predictors of AD conversion. Each cognitive/functional predictor remained a significant predictor of conversion from MCI to AD when HCV was added to the model, and HCV also remained a significant predictor. Additionally, HCV had additive effects with ADAS, RAVLT, and FAQ, whereas HCV and MMSE had interactive effects (Table 2). A graphical illustration (Figure 1) of the probability of conversion for the measures at three different categories of HCV (small, medium, and large) suggests that having a medium to large HCV had a protective effect against conversion in MMSE from 24 to 30. However that protective effect was smaller at lower MMSE scores. The same pattern was demonstrated in the normal range of FAQ, that is, having a medium to large HCV had a protective effect against conversion but the protection was lower when FAQ scores were closer to upper limit of the normal range. The pattern was relatively different for ADAS and RAVLT, where larger HCV was protective in medium ADAS or RAVLT.

Prediction of Time to Conversion

All the measures significantly predicted time to AD conversion in separate univariate analyses (likelihood ratio test between 33 and 90, df = 1, p < .0001). Each cognitive/functional predictor remained a significant predictor when HCV was added to the model, and HCV also remained a significant predictor. Additionally, HCV had additive effects with RAVLT, and FAQ, whereas HCV and MMSE, and HCV and ADAS had interactive effects (Table 2).

The analyses were repeated using categories of HCV (small, medium, and large) in the models instead of HCV as a continuous variable (Table 3). The results revealed that MMSE was not a predictor of conversion in small HCV, and that having a medium to large HCV, respectively, associated with 45% and 81% lower risk of conversion from MCI to AD over time compared with small HCV. An additional 32% decrease in the risk of conversion was

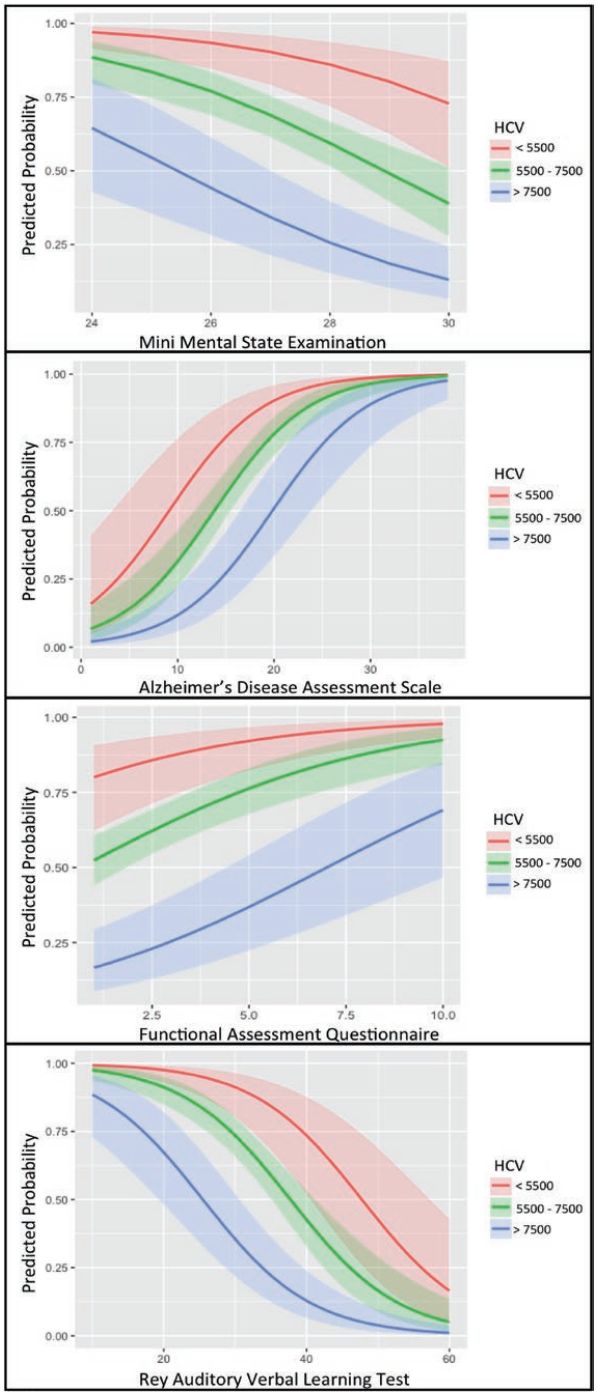


Figure 1. Predicted probabilities of conversion to Alzheimer's: Predicted probabilities of cognitive measures at different hippocampal volumes. HCV has a reciprocal impact on predicted probability of the cognitive measures for conversion to Alzheimer's. HCV = hippocampal volume adjusted by age, field strength, and intracranial volume; MMSE = mini-mental state examination; ADAS = Alzheimer Disease Assessment Scale (cognitive subscale); RAVLT = Rey Auditory Verbal Learning Test (immediate memory subscale); FAQ = Functional Assessment Questionnaire.

demonstrated for every 1 *SD* higher MMSE score in medium HCV but not in large HCV in comparison with small HCV. Similarly, ADAS was not predictive in small HCV and having a medium to large HCV was associated with 90% and 99.5% lower risk of conversion over time compared with small HCV. An additional 10% increase in the risk of conversion was demonstrated for every 1 *SD* higher ADAS score in medium HCV but not in large HCV in comparison with small HCV. In contrast, RAVLT was predictive in all HCV categories including small HCV, although an additional 4% decrease in the risk of conversion was detected for every 1 *SD* higher RAVLT in medium HCV.

Kaplan–Meier curves (Figure 2) revealed that the contribution of cognitive/functional measures in predicting the probability of remaining MCI over time was not constant at all HCV categories. For example, MMSE was not a determinant factor at small HCV, while it was at medium to large HCV.

Discussion

This study aimed to investigate the predictive value of cognitive/functional measures in combination with HCV to predict conversion from MCI to AD within 5 years, as well as their capacity to predict time to conversion. The results demonstrated that the predictive value of cognitive/functional measures is dependent on HCV. The findings revealed that (a) in predicting the conversion from MCI to AD, the predictive value of cognitive/functional measures was higher at lower HCV, while it was lower at higher HCV, and (b) in predicting the time to AD conversion, the cognitive/functional measures were somewhat more predictive when HCV was in the medium range (5,500 to 7,500 mm^3) than at smaller or larger volumes, except for the immediate memory test that remained predictive across all HCV. The effect of HCV in predicting time to conversion was interactive with general cognitive measures (MMSE and ADAS) but additive with the functional assessment (FAQ) and immediate memory test (RAVLT).

These findings are important because they demonstrate that severity of cognitive impairment or subtle functional impairment and severity of neural pathology are both important in predicting probability of AD conversion. Although cognitive/functional performance is closely linked with neuropathology, the association is not straightforward. There is an imperfect overlap between cognitive deficit and pathology severity (Neuropathology Group, Medical Research Council Cognitive & Aging, 2001). Individual variability in brain/cognitive reserve is the most likely explanation for this effect (Medaglia,

Table 3. Risk of Conversion from MCI to AD OverTime (Cox Proportional Hazard) by Hippocampal Volume Categories

Variables	Coef.	SE	HR (95% CI)	*Z*, *p* Value
MMSE and HCV				
Medium HCV category	**−0.59**	**0.18**	**0.55 (0.39–0.79)**	**−3.29, *p* = .001**
Large HCV category	**−1.68**	**0.31**	**0.19 (0.10–0.35)**	**−5.35, *p* < .0001**
MMSE	−0.07	0.14	0.93 (0.70–1.23)	−0.5, *p* = .61
Medium HCV category: MMSE	**−0.38**	**0.17**	**0.68 (0.49–0.96)**	**−2.21, *p* = .03**
Large HCV category: MMSE	−0.60	0.31	0.55 (0.30–1.00)	−1.95, *p* = .05
ADAS and HCV				
Medium HCV category	**−2.28**	**0.69**	**0.10 (0.03–0.39)**	**−3.33, *p* = .0008**
Large HCV category	**−3.11**	**1.05**	**0.04 (0.01–0.35)**	**−2.96, *p* = .003**
ADAS	0.03	0.03	1.03 (0.97–1.08)	0.94, *p* = .35
Medium HCV category: ADAS	**0.10**	**0.03**	**1.10 (1.04–1.17)**	**3.14, *p* = .003**
Large HCV category: ADAS	0.10	0.06	1.10 (0.99–1.23)	1.71, *p* = .09
FAQ and HCV				
Medium HCV category	**−0.63**	**0.25**	**0.53 (0.33–0.87)**	**−2.53, *p* = .01**
Large HCV category	**−2.08**	**0.44**	**0.13 (0.05–0.30)**	**−4.74, *p* < .0001**
FAQ	0.06	0.03	1.06(1.00–1.13)	1.89, *p* = .06
Medium HCV category: FAQ	0.03	0.04	1.03 (0.96–1.10)	0.81, *p* = .42
Large HCV category: FAQ	0.08	0.06	1.09 (0.97–1.21)	1.47, *p* = .14
RAVLT and HCV				
Medium HCV category	0.82	0.66	2.27 (0.62–8.28)	1.25, *p* = .21
Large HCV category	1.11	1.42	3.02 (0.19–49.31)	0.78, *p* = .44
RAVLT	**−0.04**	**0.02**	**0.96 (0.93–0.99)**	**−2.05, *p* = .04**
Medium HCV category: RAVLT	**−0.04**	**0.02**	**0.96 (0.92–0.99)**	**−2.01, *p* = .045**
Large HCV category: RAVLT	−0.09	0.05	0.92 (0.83–1.01)	−1.86, *p* = .06

Note. MMSE = mini-mental state examination (standardized); ADAS = Alzheimer Disease Assessment Scale (standardized); RAVLT = Rey Auditory Verbal Learning Test (immediate; standardized); FAQ = Functional Assessment Questionnaire (standardized); HCV = hippocampal volume adjusted by age, field strength, and intracranial volume (standardized).

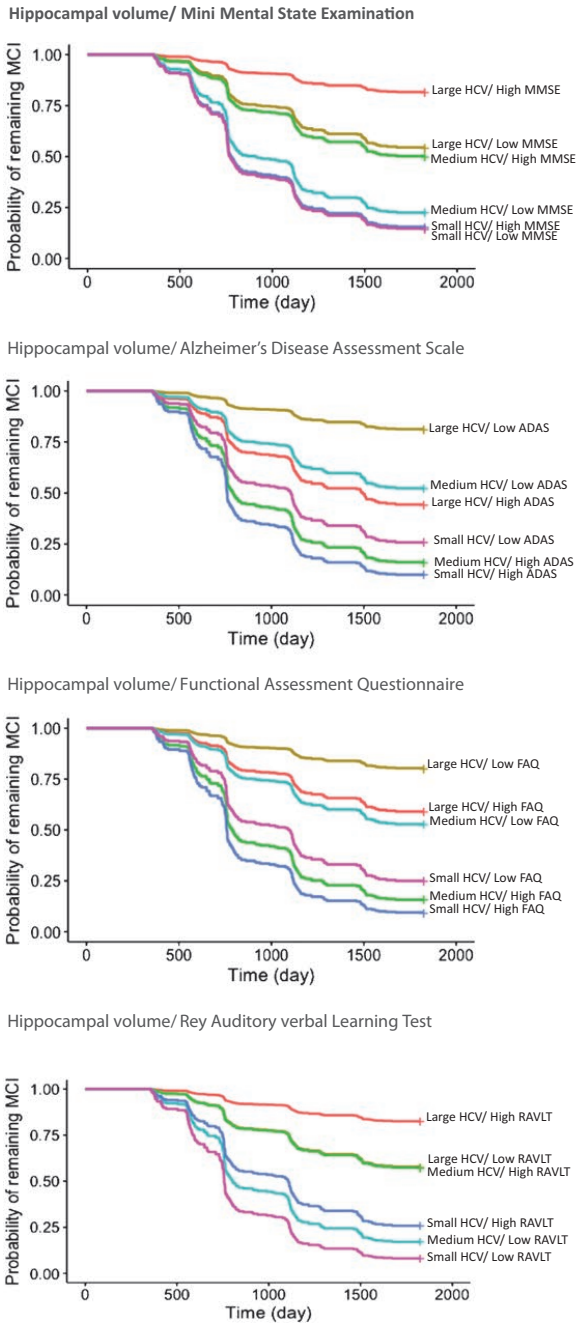


Figure 2. Kaplan–Meier plots for remaining stable over time: Illustrating the contribution of cognitive/functional measure and hippocampal volume on probability of remaining stable over time in MCI. Participants were categorized into six combinations based on three levels of HCV (small, medium, and large) and two levels of cognitive/functional measures (low and high). HCV = hippocampal volume adjusted by age, field strength, and intracranial volume; MMSE = mini-mental state examination; ADAS = Alzheimer Disease Assessment Scale (cognitive subscale); RAVLT = Rey Auditory Verbal Learning Test (immediate memory subscale); FAQ = Functional Assessment Questionnaire.

Pasqualetti, Hamilton, Thompson-Schill, & Bassett, 2017; Steffener & Stern, 2012; Stern, 2009). Taking the severity of the pathology into account when evaluating cognitive/functional performance is a practical way to take into account the moderating effect of brain/cognitive reserve.

There is accumulating evidence showing that individuals with larger brain/cognitive reserve may cope better with neural damage, that is, at a given level of observed pathology, cognitive impairment is lower in those with larger brain/cognitive reserve (Stern, 2009). Diversity in efficacy and capacity of neural networks as well as compensatory neural mechanisms such as using alternative neural networks may underlie this coping mechanism such that cognitive function may be maintained for some time in the context of increasing neurodegeneration. When brain/cognitive reserve is exhausted, further neurodegeneration cannot be compensated for and failure in cognitive processes clinically manifest as conversion from CN to MCI or MCI to AD (Steffener & Stern, 2012). Therefore, since individuals vary in their levels of brain/cognitive reserve, cognitive and functional performance alone is not a perfect predictor of decline. Cognitive reserve has been indirectly estimated in the literature by proxy variables including education, IQ, literacy, occupational complexity, participation in leisure activities and even personality variables (Steffener & Stern, 2012). However, the accurate measurement of brain/cognitive reserve is still the subject of ongoing research and much controversy (Steffener, Brickman, Rakitin, Gazes, & Stern, 2009; Steffener, Reuben, Rakitin, & Stern, 2011; Stern et al., 2008; Zarahn, Rakitin, Abela, Flynn, & Stern, 2007). Altogether, a practical way to deal with the concealing effect of cognitive reserve is to take into account the severity of neuropathology when evaluating cognitive/functional performance.

In addition to predicting the likelihood of converting from MCI to AD, the prediction of time to conversion is also of clinical significance but has proven difficult to achieve. Our results suggest that combining HCV and cognitive/functional measures is more effective in predicting time to conversion. However, the effect of HCV differs for different cognitive/functional measures. It has an interactive effect with MMSE and ADAS but an additive effect with FAQ or RAVLT immediate. That is, the increase in the risk of AD conversion for each one-point decrease in the MMSE (or increase in ADAS) is not constant at different HCV values and is smaller at larger HCV. In contrast, the increase in the risk is constant for every one-point decrease on the RAVLT immediate (or higher FAQ) at any HCV values. This may be because HCV is more reflective of AD related pathology than MMSE and ADAS. As a consequence, at HCV less than 5,500 mm^3, one unit difference in MMSE (or ADAS) is less influential than at larger HCV. This may explain the fact that MMSE and ADAS are not predictive of time to conversion at HCV less than 5,500 mm^3 and at more than 7,500 mm^3, but predictive in the mid-range of HCV (5,500–7,500 mm^3).

In contrast, the combined evaluation of performance in a specific domain (such as immediate memory) and the brain

structure underpinning that performance (HCV) may provide a more precise evaluation of the degree of neurodegeneration and the level of brain/cognitive reserve exhaustion. This may explain our findings that RAVLT immediate and HCV are more sensitive predictors of time to conversion. It may also explain the lack of interactive effect between these two measures.

It is important to note that because MMSE, ADAS, and FAQ evaluate performance across a larger number of neural networks, they may reflect the development of AD pathology across any of those networks and thus also predict the risk of AD conversion. However, because only part of their variability is related to hippocampal function, they do not appear to be as predictive of time to conversion than RAVLT immediate.

Many studies conducted to date have focused on combining MRI and cognitive/functional measures for improved diagnosis or prediction of AD conversion. Our study, in contrast, investigated the nature of the interaction between MRI and cognitive/functional measures in predicting AD conversion and time to conversion. Understanding the relationship between structural and cognitive/functional measures not only emphasizes the benefit of combining these measures for diagnostic/ prognostic purpose, it may also help better conceptualize the impact of brain/cognitive reserve on clinical/MRI measures.

In conclusion, AD is pathologically characterized by degenerative processes, the severity of which can be measured with neuroimaging techniques. The functional consequence of the degeneration can be concurrently assessed with cognitive/functional tools. A combination of both neuroimaging and cognitive/functional indexes are superior in predicting disease progression than either alone. However, the present findings indicate that the relative contribution of neuroimaging and cognitive/functional measures is not constant in predicting progression from MCI to AD. Cognitive/functional measures are predictors of conversion but their predictive values are not constant at all levels of HCVs. Additionally, the most effective combination of measures to predict time to conversion is likely to involve those that assess hippocampal volume in conjunction with one of its main functions, immediate memory.

Supplementary Material

Supplementary data are available at *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences* online.

Funding

Data collection and sharing for this work was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI; National Institutes of Health Grant U01 AG024904). 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 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.; Neu-roRx 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 partly undertaken on the National Computational Infrastructure (NCI) facility in Canberra, Australia, which is supported by the Australian Commonwealth Government and was supported by an Australian Government Research Training Program (RTP) Scholarship.

Acknowledgments

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.

Conflict of Interest

None reported.

References:

Arevalo-Rodriguez, I., Smailagic, N., Roque, I. F. M., Ciapponi, A., Sanchez-Perez, E., Giannakou, A.,…Cullum, S. (2015). Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database Systematic Reviews*, (3), CD010783. doi:10.1002/14651858.CD010783.pub2

Belleville, S., Fouquet, C., Hudon, C., Zomahoun, H. T. V., & Croteau, J.; Consortium for the Early Identification of Alzheimer's Disease-Quebec. (2017). Neuropsychological measures that predict progression from mild cognitive impairment to Alzheimer's type dementia in older adults: A systematic review and meta-analysis. *Neuropsychology Review,* **27**, 328–353. doi:10.1007/s11065-017-9361-5

Devanand, D. P., Liu, X., Tabert, M. H., Pradhaban, G., Cuasay, K., Bell, K.,…Pelton, G. H. (2008). Combining early markers strongly predicts conversion from mild cognitive impairment

to Alzheimer's disease. *Biological Psychiatry*, **64**, 871–879. doi:10.1016/j.biopsych.2008.06.020

Eckerstrom, C., Olsson, E., Borga, M., Ekholm, S., Ribbelin, S., Rolstad, S.,…Malmgren, H. (2008). Small baseline volume of left hippocampus is associated with subsequent conversion of MCI into dementia: The Goteborg MCI study. *Journal of the Neurological Sciences*, **272**(1–2), 48–59. doi:10.1016/j.jns.2008.04.024

Estévez-González, A., Kulisevsky, J., Boltes, A., Otermín, P., & García-Sánchez, C. (2003). Rey verbal learning test is a useful tool for differential diagnosis in the preclinical phase of Alzheimer's disease: Comparison with mild cognitive impairment and normal aging. *International Journal of Geriatric Psychiatry,* **18**, 1021–1028. doi:10.1002/gps.1010

Falahati, F., Westman, E., & Simmons, A. (2014). Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging. *Journal of Alzheimer's Disease,* **41**, 685–708. doi:10.3233/JAD-131928

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research,* **12**, 189–198. doi:10.1016/0022-3956(75)90026-6

Gomar, J. J., Bobes-Bascaran, M. T., Conejero-Goldberg, C., Davies, P., & Goldberg, T. E.; Alzheimer's Disease Neuroimaging Initiative. (2011). Utility of combinations of biomarkers, cognitive markers, and risk factors to predict conversion from mild cognitive impairment to Alzheimer disease in patients in the Alzheimer's disease neuroimaging initiative. *Archives of General Psychiatry,* **68**, 961–969. doi:10.1001/archgenpsychiatry.2011.96

Ito, K., Hutmacher, M. M., & Corrigan, B. W. (2012). Modeling of Functional Assessment Questionnaire (FAQ) as continuous bounded data from the ADNI database. *Journal of Pharmacokinetics and Pharmacodynamics,* **39**, 601–618. doi:10.1007/s10928-012-9271-3

Jack, C. R., Jr., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D.,…Weiner, M. W. (2008). The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging,* **27**, 685–691. doi:10.1002/jmri.21049

Jack, C. R., Jr., Shiung, M. M., Weigand, S. D., O'Brien, P. C., Gunter, J. L., Boeve, B. F.,…Petersen, R. C. (2005). Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic MCI. *Neurology,* **65**, 1227–1231. doi:10.1212/01.wnl.0000180958.22678.91

Li, K., Chan, W., Doody, R. S., Quinn, J., & Luo, S.; Alzheimer's Disease Neuroimaging Initiative. (2017). Prediction of conversion to Alzheimer's disease with longitudinal measures and time-to-event data. *Journal of Alzheimer's Disease,* **58**, 361–371. doi:10.3233/JAD-161201

Medaglia, J. D., Pasqualetti, F., Hamilton, R. H., Thompson-Schill, S. L., & Bassett, D. S. (2017). Brain and cognitive reserve: Translation via network control theory. *Neuroscience and Biobehavioral Reviews,* **75**, 53–64. doi:10.1016/j.neubiorev.2017.01.016

Moradi, E., Pepe, A., Gaser, C., Huttunen, H., & Tohka, J.; Alzheimer's Disease Neuroimaging Initiative. (2015). Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *Neuroimage,* **104**, 398–412. doi:10.1016/j.neuroimage.2014.10.002

Neuropathology Group, Medical Research Council Cognitive Function and Ageing Study. (2001). Pathological correlates of late-onset dementia in a multicentre, community-based population in England and Wales. Neuropathology Group of the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). *Lancet,* **357**(9251), 169–175. doi:10.1016/s0140-6736(00)03589-3

Pandya, S. Y., Clem, M. A., Silva, L. M., & Woon, F. L. (2016). Does mild cognitive impairment always lead to dementia? A review. *Journal of the Neurological Sciences,* **369**, 57–62. doi:10.1016/j.jns.2016.07.055

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology,* **56**, 303–308. doi:10.1001/archneur.56.3.303

Petersen, R. C., Thomas, R. G., Grundman, M., Bennett, D., Doody, R., Ferris, S.,…Thal, L. J.; Alzheimer's Disease Cooperative Study Group. (2005). Vitamin E and donepezil for the treatment of mild cognitive impairment. *The New England Journal of Medicine,* **352**, 2379–2388. doi:10.1056/NEJMoa050151

Pfeffer, R. I., Kurosaki, T. T., Harrah, C. H., Jr., Chance, J. M., & Filos, S. (1982). Measurement of functional activities in older adults in the community. *Journal of Gerontology,* **37**, 323–329. doi:10.1093/geronj/37.3.323

Pintzka, C. W., Hansen, T. I., Evensmoen, H. R., & Haberg, A. K. (2015). Marked effects of intracranial volume correction methods on sex differences in neuroanatomical structures: A HUNT MRI study. *Frontiers in Neuroscience,* **9**, 238. doi:10.3389/fnins.2015.00238

Rey, A. (1941). L'examen psychologique dans les cas d'encéphalopathie traumatique. (Les problems.) [The psychological examination in cases of traumatic encepholopathy. Problems]. *Archives de Psychologie,* **28**, 215–285.

Rey, A. (1964). *L'examen Clinique en Psychologie.* Paris: Presses Universitaires de France.

Steffener, J., Brickman, A. M., Rakitin, B. C., Gazes, Y., & Stern, Y. (2009). The impact of age-related changes on working memory functional activity. *Brain Imaging and Behavior,* **3**, 142–153. doi:10.1007/s11682-008-9056-x

Steffener, J., Reuben, A., Rakitin, B. C., & Stern, Y. (2011). Supporting performance in the face of age-related neural changes: Testing mechanistic roles of cognitive reserve. *Brain Imaging and Behavior,* **5**, 212–221. doi:10.1007/s11682-011-9125-4

Steffener, J., & Stern, Y. (2012). Exploring the neural basis of cognitive reserve in aging. *Biochimica et Biophysica Acta,* **1822**, 467–473. doi:10.1016/j.bbadis.2011.09.012

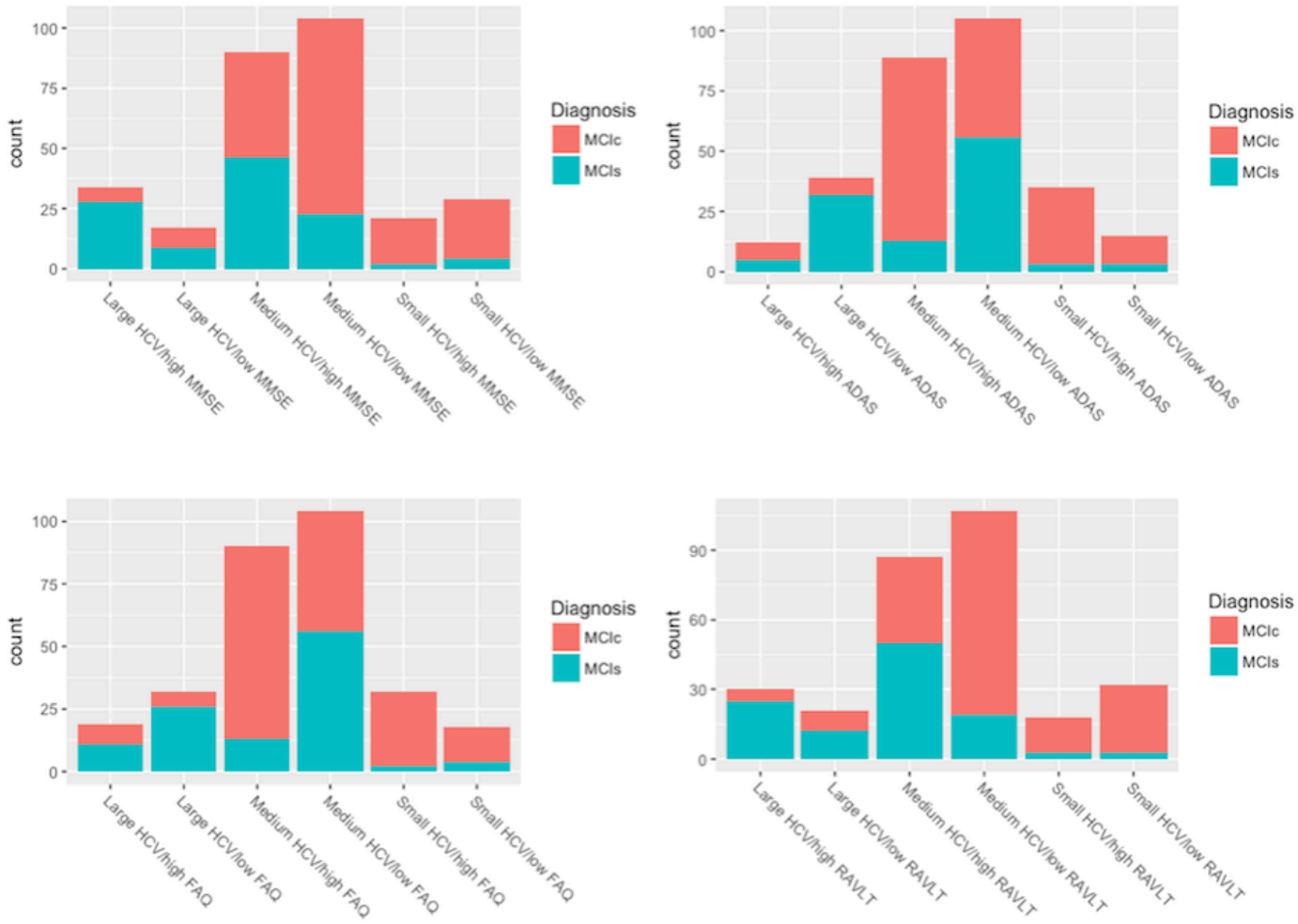
Stern, Y. (2009). Cognitive reserve. *Neuropsychologia,* **47**, 2015–2028. doi:10.1016/j.neuropsychologia.2009.03.004

Stern, Y., Zarahn, E., Habeck, C., Holtzer, R., Rakitin, B. C., Kumar, A.,…Brown, T. (2008). A common neural network for cognitive reserve in verbal and object working memory in young but not old. *Cerebral Cortex (New York, N.Y.: 1991),* **18**, 959–967. doi:10.1093/cercor/bhm134

Tabatabaei-Jafari, H., Shaw, M. E., & Cherbuin, N. (2015). Cerebral atrophy in mild cognitive impairment: A systematic review with meta-analysis. *Alzheimer's & Dementia (Amsterdam, Netherlands)*, **1**, 487–504. doi:10.1016/j.dadm.2015.11.002

Tabatabaei-Jafari, H., Walsh, E., Shaw, M. E., & Cherbuin, N.; Alzheimer's Disease Neuroimaging Initiative (ADNI). (2018). A simple and clinically relevant combination of neuroimaging and functional indexes for the identification of those at highest risk of Alzheimer's disease. *Neurobiology of Aging*, **69**, 102–110. doi:10.1016/j.neurobiolaging.2018.05.006

Zarahn, E., Rakitin, B., Abela, D., Flynn, J., & Stern, Y. (2007). Age-related changes in brain activation during a delayed item recognition task. *Neurobiology of Aging*, **28**, 784–798. doi:10.1016/j.neurobiolaging.2006.03.002



Supplementary Figure 1: Number of participants at different hippocampal and cognitive/functional categories: In general the proportion of those who converted to AD are higher at small hippocampal volume regardless of the value of cognitive/functional measure.

MCIc; mild cognitive impairment converted to Alzheimer's disease within five years, MCIs; mild cognitive impairment stable for five or more years, HCV; hippocampal volume adjusted by age, field strength and intracranial volume, MMSE; mini mental state examination, ADAS; Alzheimer disease assessment scale (cognitive subscale), RAVLT; Rey auditory verbal learning test (immediate memory subscale), FAQ; functional assessment questionnaire.

Neurobiology of Aging 83 (2019) 86–94

Contents lists available at ScienceDirect

Neurobiology of Aging

journal homepage: www.elsevier.com/locate/neuaging

Regional brain atrophy predicts time to conversion to Alzheimer's disease, dependent on baseline volume

Hossein Tabatabaei-Jafari [a,*], Marnie E. Shaw [b], Erin Walsh [a], Nicolas Cherbuin [a], For the Alzheimer's Disease Neuroimaging Initiative (ADNI)[1]

[a] Centre for Research on Ageing, Health and Wellbeing, The Australian National University, Canberra, Australia
[b] Research School of Electrical, Energy and Materials Engineering, College of Engineering and Computer Science, The Australian National University, Canberra, Australia

ARTICLE INFO

Article history:
Received 15 January 2019
Received in revised form 29 August 2019
Accepted 31 August 2019
Available online 10 September 2019

Keywords:
Hippocampus
Mild cognitive impairment
Alzheimer's disease
Time to conversion
Interaction
MRI

ABSTRACT

A key question for the design of clinical trials for Alzheimer's disease (AD) is whether the timing of conversion from mild cognitive impairment (MCI) to AD can be predicted. This is also an important question for the clinical management of MCI. This study aims to address this question by exploring the contribution of baseline brain volume and annual volume change, using Cox regression, in predicting the time to conversion. Individuals with MCI, who converted to AD (n = 198), reverted to normal (n = 38), or remained stable (n = 96) for at least five years, were included in this study. The results revealed that the volumes of all the brain areas considered were predictive of the time to conversion from MCI to AD. Annual change in volume was also predictive of the time to conversion but only when initial volumes were above a certain threshold. This is important because it suggests that reduction in atrophy rate, which is the outcome of some clinical trials, is not inevitably associated with delay in conversion from MCI to AD.

1. Introduction

Progressive neurodegeneration is a hallmark of Alzheimer's disease (AD). However, it is also prevalent in normal aging (Fjell et al., 2014). One major difference is that the rate of degeneration in the pathological progression leading to AD is substantially higher than in normal aging. A meta-analysis of longitudinal studies conducted in the last two decades revealed that the shrinkage rate in the prodromal stage of AD—mild cognitive impairment (MCI)—is at least twice that observed in normal aging (Tabatabaei-Jafari et al., 2015). This is seen in the whole brain and even more so in brain areas typically more affected in the first stage of the disease, such as the hippocampus and entorhinal cortex. Moreover, degeneration begins decades before the disorder emerges clinically, sometimes even in early adulthood (Braak and Braak, 1997). These findings underpin the hope that early intervention aimed at decreasing brain shrinkage may stop, or at least slow down, further progression to clinical AD.

Several intervention trials, using nutrient supplements or medication, have been effective in reducing the atrophy rate in total or regional brain volumes in those with MCI (Douaud et al., 2013; Dubois et al., 2015; Kile et al., 2017; Kobe et al., 2016; Prins et al., 2014; Zhang et al., 2017). However, whether these changes can modify the course of AD progression and delay the time to conversion remains an unresolved question. To address such questions, it is necessary to better understand the contribution brain atrophy makes to the course of the disease and particularly to the progression from MCI to AD.

In contrast to studies predicting conversion from MCI to AD, studies that have investigated the time to conversion are limited in number. They generally suggest an association between the pace of neurodegeneration and the time to AD conversion (Falahati et al., 2017; Jack et al., 2005; Liu et al., 2017; Teipel et al., 2015). Most attempts have used spatial patterns of longitudinal volume loss (using machine learning) to successfully predict the time to conversion (Gavidia-Bovadilla et al., 2017; Li et al., 2012; Risacher et al., 2010; Teipel et al., 2015; Thung et al., 2018). Falahati el al. developed a "severity index", based on degeneration in 34 measures of regional cortical thickness and 21 regional subcortical volumes and showed

* Corresponding author at: Centre for Research on Ageing, Health and Wellbeing, Florey Building 54, Mills Road, Australian National University, Acton, Australian Capital Territory, 2601 Australia. Tel: (612) 61259032; fax: (612) 61251558.
E-mail address: hossein.tabatabaei@anu.edu.au (H. Tabatabaei-Jafari).

[1] 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 in the Supplemental file.

0197-4580/$ – see front matter © 2019 Elsevier Inc. All rights reserved.
https://doi.org/10.1016/j.neurobiolaging.2019.08.033

that it was predictive of the time to AD conversion for up to 3 years follow-up. The index showed 95% correct prediction of conversion within the first year and 80% over 3 years (Falahati et al., 2017). Global volume change such as whole brain atrophy and ventricular enlargement, but not regional brain atrophy rates (hippocampal and entorhinal cortex), has also been shown to be predictive of AD conversion but only for a short follow-up and in the context of a relatively small study (Jack et al., 2005). Although these limited numbers of studies are conceptually supportive of the idea that faster degeneration will lead to earlier conversion, the findings are based on a short-term follow-up and the approaches are complex and methodologically difficult to implement at individual level that is a requirement for clinical trials and clinical practice. Simple measures such as regional brain volume and regional atrophy rate investigated in a longer follow-up may be more practical for individual evaluation, especially in a clinical setting or for clinical trials.

Therefore, strong evidence supporting the use of atrophy rate in the prediction of time to conversion from MCI to AD is still lacking. In addition, it is necessary to clarify the extent to which the predictive value of atrophy rate depends on baseline volume. This is needed because the clinical impact of any future degeneration is likely to be highly dependent on prior atrophy and/or brain reserve indexed by the current volume of a region of interest. To address these questions, the present study aimed to investigate the value of global as well as regional baseline volume and atrophy rate and their interaction over long-term follow-ups in predicting conversion from MCI to AD.

2. Methodology

2.1. Study participants

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 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, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD.

All participants of ADNI 1/GO/2, who were diagnosed with MCI at the baseline, remained stable for at least six months, and underwent MRI scanning more than twice, were considered for inclusion. Individuals with MCI who converted to AD (MCIc, n = 198), reverted to cognitively normal (CN; MCIr, n = 38), or remained stable for more than five years (MCIs, n = 96) were included in this study.

Details of the diagnostic criteria can be found at the ADNI web site (http://www.adni-info.org/Scientists/AboutADNI.aspx). Briefly, participants were classified as MCI if they had a Mini–Mental State Examination (MMSE) score greater than 24, a CDR of 0.5, a report of subjective memory concern, an objective memory loss, preserved daily living activity, and did not meet diagnostic criteria for dementia. Participants were classified as AD if they had an MMSE score less than 26, CDR of 0.5 or above, and fulfilled criteria for clinically probable AD according to the Institute of Neurological and Communicative Diseases and Stroke/Alzheimer's Disease and Related Disorders Association. It is also important to note that a Geriatric Depression Scale score of less than 6 was a requirement for participation in the ADNI study (Petersen et al., 2010), so all participants had a Geriatric Depression Scale score of normal range.

2.2. Neuroimaging acquisition and processing

Participants underwent high-resolution MRI brain scans on 1.5 (N = 889) or 3 T (N = 872) scanners from General Electric, Siemens, or Philips (Milwaukee, WI, USA; Germany; the Netherlands, respectively) using a standardized ADNI acquisition protocol for 3D MP-RAGE sequence (Jack et al., 2008). Images which had undergone specific ADNI preprocessing correction steps to standardize images from different sites and platforms were obtained for this study: (1) Grad wrap: a specific correction of image geometry distortion due to nonlinearity, (2) B1 nonuniformity: B1 calibration to correct the image intensity nonuniformity that results when RF transmission is performed with a more uniform body coil while reception is performed with a less-uniform head coil, (3) N3 correction: a histogram peak sharpening algorithm applied after grad wrap and B1 correction. For MCI participants, only images acquired before conversion to AD or reversion to CN were included. The MRI scans of individual participants were acquired on the same scanner with the same parameters throughout the follow-up.

All scans were segmented with FreeSurfer version 5.3 (http://surfer.nmr.mgh.harvard.edu/), processed with the longitudinal pipeline. For each participant, all scans were initially processed by the default workflow. Then an unbiased template (an average template) was created from all time points. The unbiased template was used as a base for registering all the time point scans to reduce the random within-subject variation in the processing procedure of the longitudinal analysis. Finally, all time points were longitudinally processed.

The output-segmented images were visually checked. The criterion was a clear segmentation error as assessed by an experienced neuroscientist. Scans with segmentation errors were reprocessed and would only be excluded if the error could not be corrected. Six scans with error were not correctable and excluded from the study.

2.3. Measurements

Four brain volumes were considered as regions of interest (ROIs) in this study: (1) total whole brain volume (sum of the total gray and white matter), (2) total ventricular volume (sum of the lateral, third, and fourth ventricular volumes), (3) total hippocampal volume (sum of the left and right hippocampus), and (4) total entorhinal cortex volume (sum of the left and right entorhinal cortex). Baseline volume and annual change rate (atrophy rate for the whole brain, hippocampus, and entorhinal cortex and enlargement rate for the ventricles) of each ROI were investigated as the measures of interest for predicting time to conversion from MCI to AD.

The annual change rate for each ROI was computed by the least square linear regression method for each individual separately: brain volume (at each time point) was used as the dependent variable, with age at each time point (centered at 55, the minimum age at the baseline) as the independent variable. The regression coefficient for age was considered as the volume change for each year increase in age in mm^3. The regression coefficient was used to compute the annual change rate in percentage using the formula

$$100 \times (\textit{the regression coefficient for age} / \textit{baseline volume})$$

Because the results from our previous study revealed that the baseline scores on the MMSE, the Alzheimer's Disease Assessment Scale (ADAS cognitive version), the Functional Assessment Questionnaire, and the Rey Auditory Learning Test (immediate memory subtype) were predictive of time to conversion from MCI to AD when also taking into account hippocampal volume (Tabatabaei-Jafari et al., 2019), the annual change rates of these measures were also evaluated to better characterize the participants.

While CSF level of amyloid β 1-42 and total and phosphorylated tau were only available for a subsample of participants (236 for amyloid β, 232 for total tau, and 236 for phosphorylated tau of the

332 participants), they could not be included in analyses but are reported to better characterize the sample investigated.

2.4. Statistical analysis

Statistical analyses were performed using the R statistical software (version 3.3.2). Data were checked for missing values and for univariate and multivariate outliers using Mahalanobis distance. There were no missing values or outliers. Group differences in demographic variables were assessed by *t*-test for continuous variables and chi-square tests for categorical variables. The alpha level was set at < 0.05.

Cox regression analysis (package survival; version 2.40–1) using time to event as time metric was used to investigate the predictive value of brain ROIs for time to conversion from MCI to AD. The event in the model was specified as happened if the individual converted to AD, thus MCIc were coded as 1 and MCIs and MCIr were coded as 0 in the model. For MCIs, the time to event was the time from the baseline to the last scan, whereas in MCIc and MCIr, it was the time from the baseline to diagnosis change (change to AD for MCIc and change to CN for MCIr). One-sided Wald tests were used to test associations because only increase in the risk of conversion to AD was predicted. Baseline volume and annual change rate were considered as predictors of time to conversion and were standardized to reduce the variance inflation factor in the model. Baseline volumes were adjusted for age, sex, field strength, and intracranial volume using the residual method before adopted in the models (Pintzka et al., 2015).

Table 1
Participants characteristic and measurements

Diagnostic group	MCIr	MCIs	MCIc	Significant pair difference ($p < 0.05$)
Sample size	38	96	198	-
Age; year, mean (SD)	69.30 (8.23)	71.65 (7.48)	74.25 (7.16)	MCIc vs. MCIr and MCIs
Age range, year	55–87	57–88	55–89	-
Male sex; N (%)	18 (47.37)	58 (60.42)	121 (61.42)	No difference
Education; year, mean (SD)	16.68 (2.52)	15.88 (3.04)	16.01 (2.78)	No difference
APOE e4; N (%)				
One allele	17 (45)	22 (23)	102 (51.51)	MCIc vs. MCIr and MCIs
Two alleles	1 (3)	6 (6)	32 (16.33)	MCIr vs. MCIs
Number of scan points	11.4 (2.7)	7.7 (2.5)	5.9 (1.8)	MCIc vs. MCIr and MCIs MCIr vs. MCIs
Follow-up, range; day	1082 - 3662	1850 - 3927	343 - 3690	-
Follow-up, mean (SD)	1704 (676)	2381 (686)	1790 (869)	-
Time to diagnosis change, range; day	184–1583	-	357–3714	-
Time to DX change, mean (SD)	762 (411)	-	1041 (603)	-
Brain measures				
Whole brain				
Baseline, mm^3	1,081,597 (35,162)	1,095,415 (38,165)	1,070,628 (42,231)	MCIc vs. MCIs MCIr vs. MCIs
Annual change rate, %/y	−0.15 (1.20)	−0.55 (0.37)	−0.73 (1.26)	MCIc vs. MCIr MCIr vs. MCIs
Ventricles				
Baseline, mm^3	37,782 (14,261)	38,808 (15,115)	44,744 (17,982)	MCIc vs. MCIr and MCIs
Annual change rate, %/y	2.42 (3.94)	3.93 (2.66)	7.62 (5.74)	MCIc vs. MCIr and MCIs MCIr vs. MCIs
Hippocampus				
Baseline, mm^3	7229 (794)	7035 (953)	6127 (912)	MCIc vs. MCIr and MCIs
Annual change rate, %/y	0.13 (3.34)	−1.29 (1.10)	−3.12 (2.86)	MCIc vs. MCIr and MCIs MCIr vs. MCIs
Entorhinal cortex				
Baseline, mm^3	3787 (693)	3645 (644)	3224 (698)	MCIc vs. MCIr and MCIs
Annual change rate, %/y	−0.11 (5.05)	−1.75 (1.56)	−3.62 (5.95)	MCIc vs. MCIr and MCIs
Cognitive/functional measures				
MMSE				
Baseline	28.53 (1.50)	28.22 (1.42)	27.09 (1.78)	MCIc vs. MCIr and MCIs
Annual change, u/y	0.64 (1.94)	−0.15 (0.29)	−0.93 (1.94)	MCIc vs. MCIr and MCIs MCIr vs. MCIs
ADAS cog				
Baseline	10.66 (4.24)	12.08 (4.63)	19.94 (5.81)	MCIc vs. MCIr and MCIs
Annual change, u/y	−1.97 (4.31)	0.24 (0.57)	1.50 (3.98)	MCIc vs. MCIr and MCIs MCIr vs. MCIs
RAVLT immediate				
Baseline	43.55 (10.21)	39.70 (10.96)	29.64 (7.98)	MCIc vs. MCIr and MCIs
Annual change, u/y	−1.44 (7.88)	−0.46 (1.44)	−2.06 (6.14)	MCIc vs. MCIs
FAQ				
Baseline	0.87 (1.73)	1.60 (3.07)	4.61 (4.54)	MCIc vs. MCIr and MCIs
Annual change, u/y	−0.16 (2.35)	0.22 (0.69)	1.61 (3.55)	MCIc vs. MCIr and MCIs
CSF measures (baseline)				
Amyloid β level, pg/mL	211.54 (50.51)	198.23 (48.10)	143.19 (42.85)	MCIc vs. MCIr and MCIs
TAU	60.24 (29.31)	75.21 (49.97)	115.45 (55.86)	MCIc vs. MCIr and MCIs
P-TAU	27.69 (14.45)	32.21 (49.97)	49.61 (26.13)	MCIc vs. MCIr and MCIs

Baseline measures adjusted for age, sex, field strength, and intracranial volume.
Key: MCIc, mild cognitive impairment converted to Alzheimer's disease; MCIs, mild cognitive impairment remained stable for more than five years; MCIr, mild cognitive impairment reverted to cognitively; APOE e4, apolipoprotein E allele 4; MMSE, Mini–Mental State Examination; ADAS cog, Alzheimer's Disease Assessment Scale (cognitive subscale); RAVLT, Rey Auditory Verbal Learning Test; FAQ, functional assessment questionnaire; CSF, cerebrospinal fluid; Aβ, amyloid-beta 1–42; TAU, total tau protein; P-TAU, phosphorylated tau protein.

Univariate models were used to investigate the association between brain measures and time to conversion. Four separate bivariate models, each consisting of standardized baseline volume, standardized annual change rate, and their interaction, were conducted for the whole brain, ventricles, hippocampus, and entorhinal cortex. Hazard ratios with 95% confidence intervals for a 1-SD different in baseline volume and 1-SD change in annual change rate were used to quantify the magnitude of the effect. In the case of significant interaction between baseline volume and annual change rate (as continuous variables), to better conceptualize the interaction, participants were categorized into 3 groups based on their baseline volume (for each brain area separately) and the bivariate analyses were repeated with categorical baseline volume in the model. Categorization was based on the standard deviation (SD, round values): (1) small category: smaller than 1 SD below the mean, (2) medium category: 1 SD below and above the mean, and (3) large category: larger than one SD above the mean. In addition, to better visualize the contribution of baseline volume and annual change rate in predicting conversion from MCI to AD, the density of those converted to AD over time was plotted across different stratified annual change rate for each baseline volume category separately.

3. Results

3.1. Participants' characteristics

Three hundred thirty-two participants (59% male), who were followed up for up to ten years (5.35 ± 2.31 years), were categorized into MCIr, MCIs, and MCIc (Table 1). Individuals with MCIc were about three years older than other individuals with MCI. There was no significant difference in education across the groups, but the proportion of males was somewhat lower in MCIr (47.37%) than in MCIs (60.42%) and MCIc (60.42%). The proportion of individuals carrying the APOE e4 allele was significantly larger in MCIc than others, and more so for those with two e4 alleles (Table 1, Supplementary table 1).

3.2. Baseline brain volumes and annual changes

For all ROIs, baseline volumes and annual change rates were different between MCIc and other MCI types. Differences were most pronounced in the hippocampus, entorhinal cortex, and ventricles and followed the direction MCIr > MCIs > MCIc for volumes, and MCIr < MCIs < MCIc for change rates (Table 1, Supplementary table 1).

Despite significant group differences, the distribution of the brain measures revealed a large overlap across the groups (Supplementary Fig. 1). When considered across the whole sample, there was no significant correlation between baseline volume and annual change rate for the whole brain and the ventricles. A moderate correlation was detected for the hippocampus (r = 0.27), and a smaller correlation for the entorhinal cortex (r = 0.12). However, when computed separately in each group, associations between baseline volume and annual change rate were only significant in MCIs for the hippocampus and entorhinal cortex as well as for the ventricles in MCIc (r = −0.19) (Supplementary table 2).

3.3. Cognitive/functional measures

Similar to brain measures, cognitive/functional measures were significantly different between MCIc and other MCI types. Differences were most pronounced in baseline volumes following the order MCIr <MCIs <MCIc. While annual changes were significantly different between MCIc and other MCI types, differences between MCIr and MCIs did not follow a constant pattern (Table 1, Supplementary table 1).

3.4. CSF measures

The pattern in CSF differences was consistent (in the subsample that data were available) across the groups. Amyloid β was significantly lower in MCIc than MCIs and MCIr, and total tau and phosphorylated tau were significantly greater in MCIc than MCIs and MCIr. These measures were not different between MCIr and MCIs (Table 1, Supplementary table 1).

3.5. Prediction of time to AD conversion

Baseline volume and annual change rate for each brain area significantly predicted time to AD conversion (Z > 5, $p < 0.01$) when they were evaluated separately (univariate model). When baseline volume and annual change rate were tested in the same model (bivariate model) both measures remained significantly predictive in all ROIs. In addition, an interaction between annual change rate and baseline volume was detected. It means, in

Table 2
Cox proportional hazard

Diagnostic group	Coef.	SE	HR (95% CI)	Z
Whole brain				
Baseline volume	0.43	0.08	1.53 (1.31–1.79)	**5.344, $p < 0.0001$**
Annual atrophy rate	0.32	0.09	1.38 (1.15–1.65)	**4.024, $p < 0.0001$**
Interaction	0.21	0.07	1.24 (1.08–1.41)	**3.110, $p < 0.01$**
Ventricles				
Baseline volume	0.25	0.07	1.29 (1.14–1.46)	**3.919, $p < 0.0001$**
Annual enlargement rate	0.46	0.06	1.58 (1.41–1.77)	**7.988, $p < 0.0001$**
Interaction	−0.16	0.08	0.85 (0.73–0.99)	**−2.133, $p < 0.05$**
Hippocampus				
Baseline volume	0.63	0.08	1.87 (1.60–2.19)	**7.840, $p < 0.0001$**
Annual atrophy rate	0.66	0.10	1.94 (1.61–2.33)	**7.001, $p < 0.0001$**
Interaction	0.45	0.09	1.56 (1.30–1.88)	**4.755, $p < 0.0001$**
Entorhinal cortex				
Baseline volume	0.44	0.08	1.55 (1.33–1.80)	**5.671, $p < 0.0001$**
Annual atrophy rate	0.56	0.10	1.75 (1.45–2.12)	**5.768, $p < 0.0001$**
Interaction	0.30	0.10	1.35 (1.11–1.64)	**3.061, $p < 0.01$**

All measures have been adjusted for age, field strength, and intracranial volume.
Bolded values represents the significance of $p < 0.05$.
Key: HR, hazard ratio for 1-SD decrease in whole brain, hippocampal volume and entorhinal cortex volume and their annual rates as well as 1-SD increase in ventricular volume and its annual ventricular volume enlargement.

addition to a constant increase in the risk for each 1-SD decrease in ROIs' baseline volume (1-SD increase in ventricular volume) and 1-SD increase in annual volume loss, there was an additional risk for each measure, which was dependent on the other measure (Table 2). To better conceptualize this interactive effect between the 2 measures, analyses were repeated with a categorical baseline volume (small, medium and large) and annual change rate in percent in the model (Table 3). Following are brief reports for each ROI separately.

Because APOE e4 carrier prevalence was significantly higher in MCIc than other MCI types, post hoc analyses were carried out to investigate the effect of APOE e4 on the predictive values of baseline volumes and annual change rates and their interactions. The result showed that APOE e4 genotype had no effect on the predictive values of these measures as well as their interactions.

Fig. 1 demonstrates the distribution of individuals converted from MCI to AD within ten years using Cox analysis to estimate probability in each separate baseline category across stratified annual atrophy rates (enlargement rate for the ventricles). It reveals that at hippocampal baseline volumes less than 5500 mm^3, conversion within 3 years occurs regardless of the atrophy rate. A similar but somewhat weaker pattern was observed for an entorhinal cortex volume smaller than 2800 mm^3, a whole brain volume smaller than 1,040,000 mm^3, and a ventricular volume larger than 55,000 mm^3. By contrast, atrophy rate (enlargement rate for the ventricles) is determinant of probability of conversion over time at medium to large baseline brain volumes (medium to small for the ventricles). It is especially noticeable for the hippocampus with atrophy rate more than the average.

3.5.1. Whole brain

Atrophy rate did not predict time to conversion in whole brain baseline volumes less than 1,040,000 mm^3, whereas it had significant predictive value at higher volumes. Medium to large whole brain volumes were associated with 61% and 72% lower risk of conversion from MCI to AD compared with small volumes. An additional 35% and 43% decrease in the risk of conversion were demonstrated for every 1 percent lower atrophy rate in medium and large volumes.

3.5.2. Ventricles

Enlargement rate did not predict time to conversion in ventricular baseline volumes larger than 55,000 mm^3, whereas it had significant predictive value at small volumes (lower than 28,000 mm^3). Medium to small volumes were, respectively, associated with 48% and 83% lower risk of conversion from MCI to AD compared with large volumes. An additional 14% increase in the risk of conversion was demonstrated for 1 percent greater enlargement rate in small volumes.

3.5.3. Hippocampus

Atrophy rate did not predict time to conversion in hippocampal baseline volumes less than 5500 mm^3, whereas it had significant predictive value at higher volumes. Medium to large volumes were associated with 69% and 95% lower risk of conversion from MCI to AD compared with small volumes. An additional 15% and 50% decrease in the risk of conversion were demonstrated for every 1 percent lower atrophy rate in medium and large volumes.

3.5.4. Entorhinal cortex

Atrophy rate did not predict time to conversion in entorhinal cortex baseline volumes less than 2800 mm^3, whereas it had significant predictive value at large volumes (larger than 4000 mm^3). Medium to large entorhinal cortex volumes were, respectively, associated with 47% and 86% lower risk of conversion from MCI to AD compared with small volumes. An additional 24% decrease in the risk of conversion was demonstrated for 1 percent lower atrophy rate in large entorhinal cortex baseline volumes.

Table 3
Risk of conversion from MCI to AD over time (Cox proportional hazard ratios) in medium and large brain volume categories (small and medium categories in the ventricles) compared with the small brain volume category (large category in the ventricles)

Diagnostic group	Coef.	SE	HR (95% CI)	Z, *p*-value
Whole brain				
Medium whole brain	**−0.96**	**0.21**	**0.39 (0.25–0.58)**	**−4.488, $p < 0.0001$**
Large whole brain	**−1.28**	**0.32**	**0.28 (0.15–0.53)**	**−3.949, $p < 0.0001$**
Atrophy rate[a]	0.08	0.18	1.08 (0.76–1.54)	0.438, $p = 0.66$
Medium whole brain: atrophy rate	**−0.44**	**0.21**	**0.65 (0.43–0.98)**	**−2.072, $p < 0.05$**
Large whole brain: atrophy rate	**−0.56**	**0.34**	**0.57 (0.34–0.96)**	**−2.132, $p < 0.05$**
Ventricles				
Medium ventricles	**−0.65**	**0.32**	**0.52 (0.28–0.97)**	**−2.068, $p < 0.05$**
Small ventricles	**−1.76**	**0.48**	**0.17 (0.07–0.42)**	**−3.664, $p < 0.001$**
Enlargement rate[b]	0.05	0.05	1.05 (0.96–1.15)	1.000, $p = 0.32$
Medium ventricles: atrophy rate	0.04	0.05	1.04 (0.94–1.14)	0.757, $p = 0.45$
Small ventricles: atrophy rate	**0.13**	**0.06**	**1.14 (1.01–1.29)**	**2.065, $p < 0.05$**
Hippocampus				
Medium hippocampus	**−1.16**	**0.28**	**0.31 (0.18–0.54)**	**−4.122, $p < 0.0001$**
Large hippocampus	**−3.09**	**0.50**	**0.05 (0.02–0.12)**	**−6.252, $p < 0.0001$**
Atrophy rate[a]	−0.04	0.05	0.96 (0.86–1.06)	−0.818, $p = 0.41$
Medium hippocampus: atrophy rate	**−0.16**	**0.06**	**0.85 (0.75–0.97)**	**−2.472, $p < 0.05$**
Large hippocampus: atrophy rate	**−0.69**	**0.14**	**0.50 (0.38–0.66)**	**−5.022, $p < 0.0001$**
Entorhinal cortex				
Medium entorhinal	**−0.64**	**0.22**	**0.53 (0.34–0.81)**	**−2.954, $p < 0.01$**
Large entorhinal	**−1.97**	**0.35**	**0.14 (0.07–0.28)**	**−5.600, $p < 0.0001$**
Atrophy rate[a]	−0.05	0.03	0.96 (0.91–1.01)	−1.623, $p = 0.11$
Medium entorhinal: atrophy rate	−0.03	0.04	0.98 (0.90–1.04)	−0.900, $p = 0.37$
Large entorhinal: atrophy rate	**−0.27**	**0.07**	**0.76 (0.67–0.87)**	**−3.941, $p < 0.0001$**

Small category, smaller than one SD below the mean; medium category, one SD below and above the mean; and large category, larger than one SD above the mean.
All measures have been adjusted for age, sex, field strength, and intracranial volume.
Bolded values represents the significance of $p < 0.05$.
Key: AD, Alzheimer's disease; MCI, mild cognitive impairment; Coef, coefficient; SE, standard error; HR, hazard ratio.
[a] Atrophy rate in the small volume category.
[b] Enlargement rate in the large ventricular volume category.

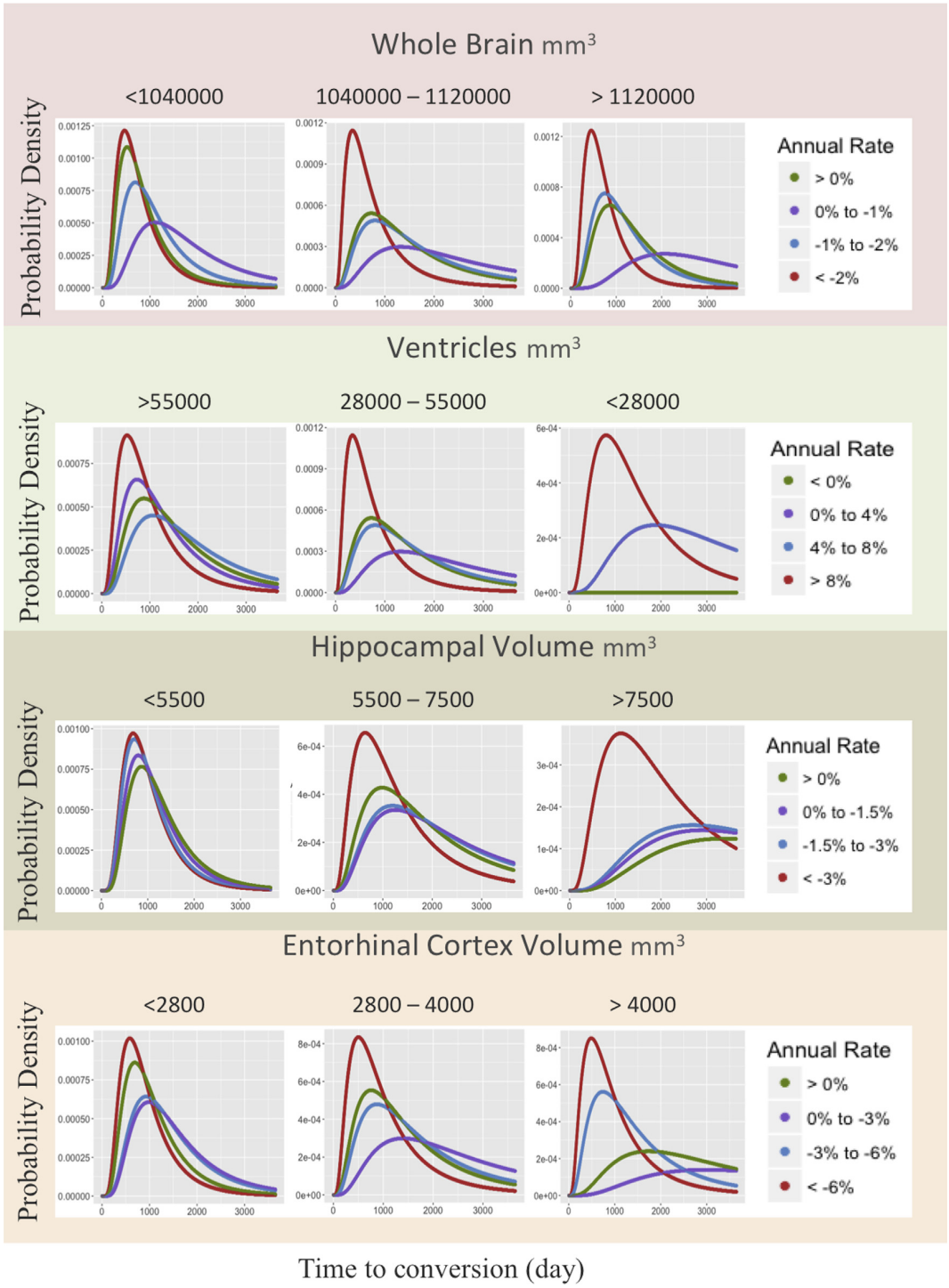


Fig. 1. Distribution of probability of conversion over time: Separate illustration of probability density measured by Cox proportional models in 4 brain areas at 3 baseline categories across stratified respected annual atrophy rate (enlargement rate for the ventricles) within ten years. The figure shows that at hippocampal baseline volumes less than 5500 mm^3, conversions mostly happen within three years regardless of atrophy rate. Similar patterns but relatively less determinant are noticeable at entorhinal cortex volumes lower than 2800 mm^3, at whole brain baseline volumes lower than 1,040,000 mm^3, and at ventricle baseline volumes larger than 55,000 mm^3. By contrast, atrophy rate (enlargement rate for the ventricles) has an impact on probability of conversion over time at medium to large baseline brain volumes (medium to small for the ventricles), especially noticeable for the hippocampus with atrophy rate more than the average.

4. Discussion

This study aimed to investigate of the volume or change in volume over time of different brain regions could predict the time to conversion from MCI to AD. The main finding was that the baseline volumes of the whole brain, ventricles, hippocampus, and entorhinal cortex and their respective atrophy rates (enlargement rate for ventricles) were all significant predictors of earlier conversion. However, the predictive value of these ROIs' atrophy rates was highly dependent on their baseline volume.

Although volume and change in volume over time are predictive across all ROIs, the effect of baseline volume on the predictive value of volume change over time is more distinctive in the hippocampus than other ROIs (Fig. 1). Individuals with hippocampal volumes smaller than 5500 mm^3 mostly convert to AD within three years regardless of atrophy rate. This has an important implication for clinical trials aiming to delay AD conversion by reducing atrophy rate. In these trials, any treatment effects on brain atrophy rate should be interpreted in light of baseline volumes because at small hippocampal volumes, any reduction in atrophy rate is less likely to be associated with delay in disease progression. Indeed, it may be better for clinical trials to exclude individuals with small hippocampal volumes to identify interventions that can really delay the conversion by reducing volume loss. In addition, hippocampal volume can be used as a simple heuristics to identify those at risk of early conversion in clinical practice. However, it is important to note that the baseline brain volumes in this study were normalized for age, sex, field strength, and ICV, and therefore, hippocampal threshold for small volume (i.e., 5500 mm^3) for any individual must be corrected with the provided formula

$$\text{Hippocampal threshold} = \begin{array}{l}\text{Male} \geq 3141 + 74.5 * \text{Age} - 477.5 * \text{field strength} - 0.0014 * \text{ICV} \\ \text{Female} \geq 3438 + 74.5 * \text{Age} - 477.5 * \text{field strength} - 0.0014 * \text{ICV}\end{array}$$

Although we cannot shed light on specific reasons for this hippocampal threshold, we speculate that volumes below this value are indicative of an accumulation of pathology, which makes conversion to AD all but inevitable. Regional accumulation of pathology is associated with concomitant spread of pathology to the adjacent brain areas. At early stages of the disease, neuropathology and brain atrophy is mainly limited to the medial and inferior temporal lobes (including hippocampus and entorhinal cortex) particularly in relation to tauopathy. As the disease progresses, degeneration spreads into more posterior regions of the temporal lobe and starts to spread to the parietal lobe. By the time of conversion to AD, atrophy has become more severe in the areas first affected and has spread further into the frontal lobes (Braak and Braak, 1991; Thal et al., 2002; Whitwell et al., 2007). Therefore, hippocampal volume below a certain threshold is not only indicative of pathology accumulation in this structure but also of spreading neurodegeneration in adjacent regions, which together indicate poorer prognosis.

By contrast, in those with larger ROI volumes, atrophy rate is a predictor of the time to conversion but is dependent on baseline volume. The pattern in larger volumes is also somewhat more distinctive in the hippocampus than other ROIs. Atrophy rate in those with medium to large hippocampal baseline volume (5500 mm^3–7500 mm^3) is determinant of the risk of AD conversion, whereas at volume larger than 7500 mm^3, atrophy rate more than the average, i.e., more than 3%/y, is determinant. This can also be explained by the contribution of previous (reflected in baseline volume) and ensuing (reflecting in atrophy rate) neurodegeneration in prediction of progression. It is likely that at medium baseline volumes, there is a balance between previous and ensuing neurodegeneration; thus both measures are determinant of the time to conversion. While at volume larger than 7500 mm^3, because of low level of previous degeneration, only a large atrophy rate (more than the average of 3%/y) can be determinant of time to AD conversion.

The present results are particularly significant because they provide a guide on how structural imaging measures can assist in predicting conversion to AD as recommended by the National Institute on Aging and the Alzheimer's Association although to date they have been unable to advise on how this should or could be done (Jack et al., 2018). This approach also aligns with our understanding of AD's pathological progression, which recognizes MCI as a clinical stage of the disease continuum, rather than a distinct clinical entity with a higher risk of AD conversion (Albert et al., 2011; Dubois et al., 2016).

It is noteworthy that the selection of the brain ROIs in this study was based on the typical spread of the neurofibrillary tangles and neurodegeneration in the course of the disease. Typically, AD's neurofibrillary tangles aggregation and subsequent neurodegeneration originate in the transentorhinal cortex and spread through the hippocampus to subcortical structures and the lateral temporal, parietal, and frontal association and primary cortices (Braak and Braak, 1991). However, there is some evidence demonstrating the presence of at least 2 atypical subtypes of AD that do not follow the typical pattern, i.e., limbic-predominant AD and hippocampus-sparing AD (Byun et al., 2015; Ferreira et al., 2017; Whitwell et al., 2012). In the limbic-predominant AD, fibrillary tangles and degeneration remain restrictively in medial temporal lobe and cortical areas remain relatively preserved. The hippocampus and entorhinal cortex are severely involved and progression to the final stages of the disease is faster than the other subtypes (Ferreira et al., 2017; Murray et al., 2011). Thus, hippocampal atrophy would be expected to remain predictive of time to conversion, and to be consistent with the present findings. By contrast, in the hippocampus-sparing subtype, the pathology originates in the lateral cortical areas and the medial temporal lobe including the hippocampus remains preserved and hippocampal atrophy is in line with that found in normal aging (Ferreira et al., 2017). Thus, hippocampal atrophy is not expected to be predictive of time to AD conversion. Of relevance to the present findings, the possible presence of this subtype in the sample investigated—it affects approximately 10% of all AD cases in the population—may have negatively impacted the predictive value of the measures investigated, although probably only to a small extent.

To our knowledge, the present study is the first investigation of interaction between brain volume and annual change rate in predicting the time to conversion from MCI to AD. In addition, unlike previous studies, which investigated the prediction of conversion from MCI to AD within follow-ups of 1 to three years (Jack et al., 2005; Liu et al., 2017; McEvoy et al., 2011; Risacher et al., 2009), the follow-up time of the present study was up to ten years. However, these findings need replications in other population before their usefulness in clinical practice can ascertained.

5. Conclusion

These findings are among the first to demonstrate that simple structural imaging measures can make a useful contribution in predicting disease progression from MCI to AD. Importantly, they provide specific guidance on volumetric thresholds in specific brain structures, which can be used to inform clinical assessment. However, while this is an important first step, further investigation in different, more diverse, and larger populations is needed before recommendation for their routine use in clinical trials and clinical practice can be confidently made.

Disclosure

None.

Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI; National Institutes of Health Grant U01 AG024904). 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 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 co-ordinated 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 partly undertaken on the National Computational Infrastructure (NCI) facility in Canberra, Australia, which is supported by the Australian Commonwealth Government and by an Australian Government Research Training Program (RTP) Scholarship.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neurobiolaging.2019.08.033.

References

Albert, M.S., DeKosky, S.T., Dickson, D., Dubois, B., Feldman, H.H., Fox, N.C., Gamst, A., Holtzman, D.M., Jagust, W.J., Petersen, R.C., Snyder, P.J., Carrillo, M.C., Thies, B., Phelps, C.H., 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. Alzheimers Dement. 7, 270–279.

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

Braak, H., Braak, E., 1997. Frequency of stages of Alzheimer-related lesions in different age categories. Neurobiol. Aging 18, 351–357.

Byun, M.S., Kim, S.E., Park, J., Yi, D., Choe, Y.M., Sohn, B.K., Choi, H.J., Baek, H., Han, J.Y., Woo, J.I., Lee, D.Y., Alzheimer's Disease Neuroimaging, I, 2015. Heterogeneity of regional brain atrophy patterns associated with distinct progression rates in Alzheimer's disease. PLoS One 10, e0142756.

Douaud, G.I., Refsum, H., de Jager, C.A., Jacoby, R., Nichols, T.E., Smith, S.M., Smith, A.D., 2013. Preventing Alzheimer's disease-related gray matter atrophy by B-vitamin treatment. Proc. Natl. Acad. Sci. U. S. A. 110, 9523–9528.

Dubois, B., Chupin, M., Hampel, H., Lista, S., Cavedo, E., Croisile, B., Louis Tisserand, G., Touchon, J., Bonafe, A., Ousset, P.J., Ait Ameur, A., Rouaud, O., Ricolfi, F., Vighetto, A., Pasquier, F., Delmaire, C., Ceccaldi, M., Girard, N., Dufouil, C., Lehericy, S., Tonelli, I., Duveau, F., Colliot, O., Garnero, L., Sarazin, M., Dormont, D., Hippocampus Study, G., Hippocampus Study, G., 2015. Donepezil decreases annual rate of hippocampal atrophy in suspected prodromal Alzheimer's disease. Alzheimers Dement. 11, 1041–1049.

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 Jr., C.R., Proceedings of the Meeting of the International Working, G., the American Alzheimer's Association on "The Preclinical State of, A.D., July, Washington Dc, U.S.A, 2016. Preclinical Alzheimer's disease: definition, natural history, and diagnostic criteria. Alzheimer's & dementia. J. Alzheimers Assoc. 12, 292–323.

Falahati, F., Ferreira, D., Muehlboeck, J.S., Eriksdotter, M., Simmons, A., Wahlund, L.O., Westman, E., 2017. Monitoring disease progression in mild cognitive impairment: associations between atrophy patterns, cognition, APOE and amyloid. Neuroimage Clin. 16, 418–428.

Ferreira, D., Verhagen, C., Hernandez-Cabrera, J.A., Cavallin, L., Guo, C.J., Ekman, U., Muehlboeck, J.S., Simmons, A., Barroso, J., Wahlund, L.O., Westman, E., 2017. Distinct subtypes of Alzheimer's disease based on patterns of brain atrophy: longitudinal trajectories and clinical applications. Sci. Rep. 7, 46263.

Fjell, A.M., Westlye, L.T., Grydeland, H., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., Dale, A.M., Walhovd, K.B., 2014. Accelerating cortical thinning: unique to dementia or universal in aging? Cereb. Cortex 24, 919–934.

Gavidia-Bovadilla, G., Kanaan-Izquierdo, S., Mataro-Serrat, M., Perera-Lluna, A., Alzheimer's Disease Neuroimaging, I, 2017. Early prediction of Alzheimer's disease using null longitudinal model-based classifiers. PLoS One 12, e0168011.

Jack Jr., C.R., Bennett, D.A., Blennow, K., Carrillo, M.C., Dunn, B., Haeberlein, S.B., Holtzman, D.M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J.L., Montine, T., Phelps, C., Rankin, K.P., Rowe, C.C., Scheltens, P., Siemers, E., Snyder, H.M., Sperling, R., Contributors, 2018. NIA-AA Research Framework: toward a biological definition of Alzheimer's disease. Alzheimers Dement. 14, 535–562.

Jack Jr., C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., J, L.W., Ward, C., Dale, A.M., Felmlee, J.P., Gunter, J.L., Hill, D.L., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C.S., Krueger, G., Ward, H.A., Metzger, G.J., Scott, K.T., Mallozzi, R., Blezek, D., Levy, J., Debbins, J.P., Fleisher, A.S., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., Weiner, M.W., 2008. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. J. Magn. Reson. Imaging 27, 685–691.

Jack Jr., C.R., Shiung, M.M., Weigand, S.D., O'Brien, P.C., Gunter, J.L., Boeve, B.F., Knopman, D.S., Smith, G.E., Ivnik, R.J., Tangalos, E.G., Petersen, R.C., 2005. Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic MCI. Neurology 65, 1227–1231.

Kile, S., Au, W., Parise, C., Rose, K., Donnel, T., Hankins, A., Chan, M., Ghassemi, A., 2017. IVIG treatment of mild cognitive impairment due to Alzheimer's disease: a randomised double-blinded exploratory study of the effect on brain atrophy, cognition and conversion to dementia. J. Neurol. Neurosurg. Psychiatry 88, 106–112.

Kobe, T., Witte, A.V., Schnelle, A., Lesemann, A., Fabian, S., Tesky, V.A., Pantel, J., Floel, A., 2016. Combined omega-3 fatty acids, aerobic exercise and cognitive stimulation prevents decline in gray matter volume of the frontal, parietal and cingulate cortex in patients with mild cognitive impairment. Neuroimage 131, 226–238.

Li, Y., Wang, Y., Wu, G., Shi, F., Zhou, L., Lin, W., Shen, D., 2012. Discriminant analysis of longitudinal cortical thickness changes in Alzheimer's disease using dynamic and network features. Neurobiol. Aging 33, 427.e15–427.e30.

Liu, K., Chen, K., Yao, L., Guo, X., 2017. Prediction of mild cognitive impairment conversion using a combination of independent component analysis and the Cox model. Front Hum. Neurosci. 11, 33.

McEvoy, L.K., Holland, D., Donald, J., Hagler, J., Fennema-Notestine, C., Brewer, J.B., Dale, A.M., 2011. Mild cognitive impairment: baseline and longitudinal structural MR imaging measures improve predictive Prognosis1. Radiology 259, 834–843.

Murray, M.E., Graff-Radford, N.R., Ross, O.A., Petersen, R.C., Duara, R., Dickson, D.W., 2011. Neuropathologically defined subtypes of Alzheimer's disease with distinct clinical characteristics: a retrospective study. Lancet Neurol. 10, 785–796.

Petersen, R.C., Aisen, P.S., Beckett, L.A., Donohue, M.C., Gamst, A.C., Harvey, D.J., Jack, C.R., Jagust, W.J., Shaw, L.M., Toga, A.W., Trojanowski, J.Q., Weiner, M.W., 2010. Alzheimer's disease neuroimaging initiative (ADNI) clinical characterization. Neurology 74, 201–209.

Pintzka, C.W., Hansen, T.I., Evensmoen, H.R., Haberg, A.K., 2015. Marked effects of intracranial volume correction methods on sex differences in neuroanatomical structures: a HUNT MRI study. Front Neurosci. 9, 238.

Prins, N.D., van der Flier, W.A., Knol, D.L., Fox, N.C., Brashear, H., Nye, J.S., Barkhof, F., Scheltens, P., 2014. The effect of galantamine on brain atrophy rate in subjects with mild cognitive impairment is modified by apolipoprotein E genotype: post-hoc analysis of data from a randomized controlled trial. Alzheimers Res. Ther. 6, 47.

Risacher, S.L., Saykin, A.J., West, J.D., Shen, L., Firpi, H.A., McDonald, B.C., Alzheimer's Disease Neuroimaging, I, 2009. Baseline MRI predictors of conversion from MCI to probable AD in the ADNI cohort. Curr. Alzheimer Res. 6, 347–361.

Risacher, S.L., Shen, L., West, J.D., Kim, S., McDonald, B.C., Beckett, L.A., Harvey, D.J., Jack Jr., C.R., Weiner, M.W., Saykin, A.J., Alzheimer's Disease Neuroimaging, I, 2010. Longitudinal MRI atrophy biomarkers: relationship to conversion in the ADNI cohort. Neurobiol. Aging 31, 1401–1418.

Tabatabaei-Jafari, H., Shaw, M.E., Cherbuin, N., 2015. Cerebral atrophy in mild cognitive impairment: a systematic review with meta-analysis. Alzheimers Dement. 1, 487–504.

Tabatabaei-Jafari, H., Shaw, M.E., Walsh, E., Cherbuin, N., Alzheimer's Disease Neuroimaging, I, 2019. Cognitive/functional measures predict Alzheimer's disease, dependent on hippocampal volume. J. Gerontol. B Psychol. Sci. Soc. Sci.

Teipel, S.J., Kurth, J., Krause, B., Grothe, M.J., Alzheimer's Disease Neuroimaging, I, 2015. The relative importance of imaging markers for the prediction of Alzheimer's disease dementia in mild cognitive impairment - beyond classical regression. Neuroimage Clin. 8, 583–593.

Thal, D.R., Rub, U., Orantes, M., Braak, H., 2002. Phases of A beta-deposition in the human brain and its relevance for the development of AD. Neurology 58, 1791–1800.

Thung, K.H., Yap, P.T., Adeli, E., Lee, S.W., Shen, D., Alzheimer's Disease Neuroimaging, I, 2018. Conversion and time-to-conversion predictions of mild cognitive impairment using low-rank affinity pursuit denoising and matrix completion. Med. Image Anal. 45, 68–82.

Whitwell, J.L., Dickson, D.W., Murray, M.E., Weigand, S.D., Tosakulwong, N., Senjem, M.L., Knopman, D.S., Boeve, B.F., Parisi, J.E., Petersen, R.C., Jack, C.R., Josephs, K.A., 2012. Neuroimaging correlates of pathologically defined subtypes of Alzheimer's disease: a case-control study. Lancet Neurol. 11, 868–877.

Whitwell, J.L., Przybelski, S.A., Weigand, S.D., Knopman, D.S., Boeve, B.F., Petersen, R.C., Jack Jr., C.R., 2007. 3D maps from multiple MRI illustrate changing atrophy patterns as subjects progress from mild cognitive impairment to Alzheimer's disease. Brain. 130 (Pt 7), 1777–1786.

Zhang, Y.P., Miao, R., Li, Q., Wu, T., Ma, F., 2017. Effects of DHA supplementation on hippocampal volume and cognitive function in older adults with mild cognitive impairment: a 12-month randomized, double-blind, placebo-controlled trial. J. Alzheimers Dis. 55, 497–507.

Supplementary table 1: Statistics of participants' characteristics (pair comparison)

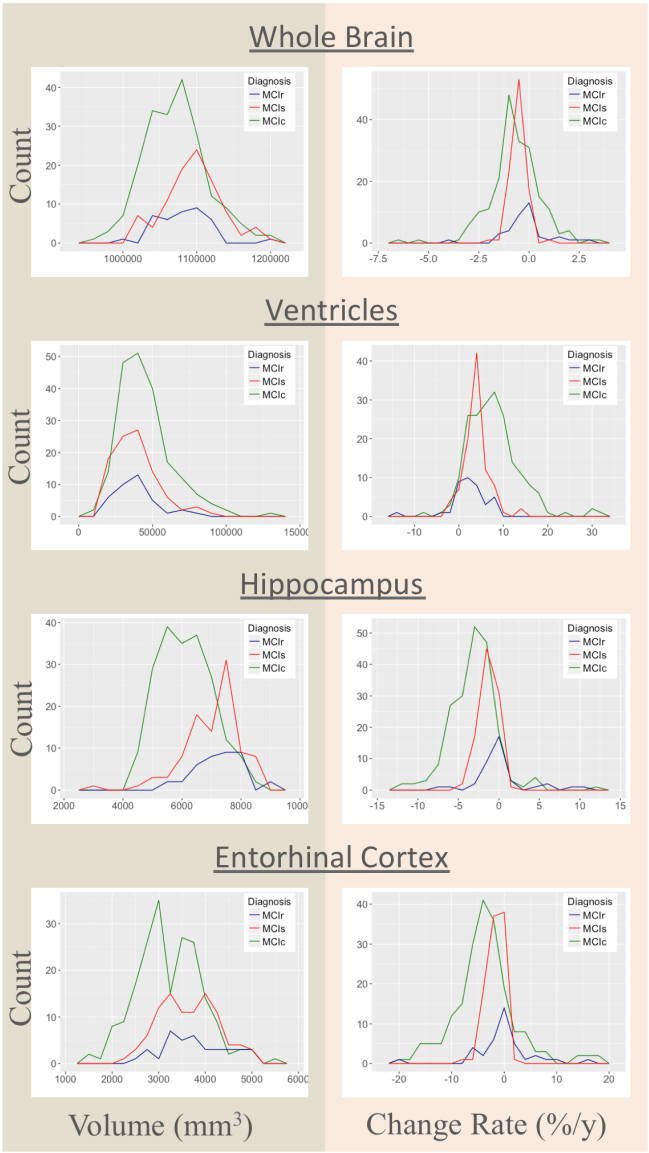
	MCIc vs. MCIs	MCIc vs. MCIr	MCIs vs. MCIr
Age	t(df=180.93)= 2.83, p<0.01	t(df=48.34)=3.46, p<0.01	t(df=62.56)=-1.53, p=0.13
Sex	X^2(df=1)=0, p=1	X^2(df=1)=1.95, p=0.16	X^2(df=1)=1.39, p=0.24
APOEe4	X^2(df=2)=39.05, p<0.0001	X^2(df=2)=8.29, p<0.05	X^2(df=2)=6.51, p<0.05
Education	t(df=173.95)=0.35, p=0.72	t(df=55.83)=-1.50, p=0.14	t(df=81.55)=1.58, p=0.12
Number of scan point	t(df=146.82)=-6.16, p<0.0001	t(df=43.65)=-11.94, p<0.0001	t(df=62.75)=7.35, p<0.0001
Brain Measures			
Whole Brain			
Baseline, mm^3	t(df=206.23)=-5.04, p<0.0001	t(df=59.47)=-1.70, p=0.09	t(df=73.34)=-2.00, p<0.05
Annual change rate, %/y	t(df=255.60)=-1.83, p=0.07	t(df=54.07)=-2.71, p<0.01	t(df=39.79)=-2.03, p<0.05
Ventricles			
Baseline, mm^3	t(df=220.14)=2.96, p<0.01	t(df=61.95)=2.63, p<0.05	t(df=71.70)=-0.37, p=0.71
Annual change rate, %/y	t(df=291.55)=7.54, p<0.0001	t(df=71.11)=6.87, p<0.0001	t(df=50.93)=-2.18, p<0.05
Hippocampus			
Baseline, mm^3	t(df=180.89)=-7.77, p<0.0001	t(df=57.39)=-7.64, p<0.0001	t(df=80.93)=1.20, p=0.23
Annual change rate, %/y	t(df=280.98)=-7.89, p<0.0001	t(df=47.99)=-5.62, p<0.0001	t(df=40.20)=2.56, p<0.05
Entorhinal cortex			
Baseline, mm^3	t(df=202.35)=-5.12, p<0.0001	t(df=52.45)=-4.58, p<0.0001	t(df=63.75)=1.09, p=0.28
Annual change rate, %/y	t(df=246.61)=-4.15, p<0.0001	t(df=54.98)=-3.56, p<0.001	t(df=39.38)=1.81, p=0.08
Cognitive/functional measures			
MMSE			
Baseline	t(df=230.58)=-5.91, p<0.0001	t(df=58.73)=-5.25, p<0.0001	t(df=64.49)=1.09, p=0.28
Annual change, u/y	t(df=211.03)=-5.40, p<0.0001	t(df=40.91)=-4.19, p<0.001	t(df=31.40)=2.35, p<0.05
ADAS cog			

Baseline	t(df=230.87)=12.50, p<0.0001	t(df=64.46)=11.44, p<0.0001	t(df=70.99)=-1.69, p=0.10
Annual change, u/y	t(df=223.32)=3.35, p<0.001	t(df=35.97)=3.75, p<0.001	t(df=30.48)=-2.84, p<0.01
RAVLT immediate			
Baseline	t(df=145.52)=-8.03, p<0.0001	t(df=46.06)=-7.95, p<0.0001	t(df=72.53)=1.93, p=0.06
Annual change, u/y	t(df=218.50)=-3.28, p<0.01	t(df=40.94)=-0.84, p=0.41	t(df=31.59)=-0.41, p=0.68
FAQ			
Baseline	t(df=261.49)=6.67, p<0.0001	t(df=150.96)=8.73, p<0.0001	t(df=116.37)=-1.75, p=0.08
Annual change, u/y	t(df=223.71)=5.28, p<0.0001	t(df=57.02)=3.65, p<0.001	t(df=32.80)=-0.92, p=0.37
CSF measures (baseline)			
Amyloid β level, pg/ml	t(df=129.55)=-8.08, p<0.0001	t(df=42.37)=-7.07, p<0.0001	t(df=57.29)=1.26, p=0.21
TAU	t(df=158.41)=5.22, p<0.0001	t(df=93.64)=7.73, p<0.0001	t(df=94)=-1.09, p=0.06
P-TAU	t(df=184.35)=5.47, p<0.0001	t(df=86.41)=6.42, p<0.0001	t(df=76.81)=-1.33, p=0.19

MCIc= mild cognitive impairment converted to Alzheimer's disease; MCIs= mild cognitive impairment remained stable for more than five years; MCIr= mild cognitive impairment reverted to cognitively. APOE e4; Apolipoprotein E allele 4;MMSE = mini-mental state examination; ADAS cog= Alzheimer Disease Assessment Scale (cognitive subscale); RAVLT = Rey Auditory Verbal Learning Test; FAQ = functional assessment questionnaire. CSF= cerebrospinal fluid; Aβ= amyloid-beta 1–42; TAU=total tau protein, P-TAU=phosphorylated tau protein. Baseline measures adjusted for age, sex, field strength and intracranial volume

Supplementary table 2: Correlation between baseline brain volume and annual change rate in the whole sample and across the groups

	All groups	MCIr	MCIs	MCIc
Whole brain	r=-0.03, p=0.54	r=0.23, p=0.17	r=-0.04, p=0.69	r=-0.11, p=0.13
Ventricles	r=-0.06, p=0.31	r=0.08, p=0.62	r=-0.01, p=0.99	**r=-0.19, p<0.01**
Hippocampus	**r=0.27, p<0.0001**	r=-0.09, p0.57	**r=0.24, p<0.05**	r=0.12, p=0.09
Entorhinal cortex	**r=0.12, p<0.05**	r=-0.11, p=0.50	**r=0.32, p<0.01**	r=0.05, p=0.53



Supplementary figure 1: Frequency of baseline volumes and atrophy rates across the groups. Left column shows the overlap of baseline volumes and right column shows the overlap of atrophy rates (enlargement rate for the ventricles) across MCI groups.

MCIc= mild cognitive impairment who convert to Alzheimer's disease,

MCIs= mild cognitive impairment who remain stable for at least five years, and

MCIr= mild cognitive impairment who revert to cognitively normal.