# Investigating Episodic-Memory Predictors of Early-Stage Alzheimer's Disease

**Samuel L. Warren**

Supervisor: A/Prof. Ahmed Moustafa

School of Social Science and Psychology

Western Sydney University

Sydney, Australia

*A thesis submitted in fulfilment of the requirements*

*for the degree of Master of Research*

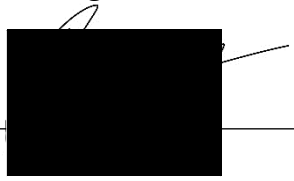
2019

# Table of Contents

# Statement of Originality

---

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution

(Signed)  Date: 20/12/19

# Acknowledgements

To my partner, thank you for always being my foundation.

To my family and friends, thank you for your ongoing support.

To my supervisor, thank you for your guidance throughout this project and my degree.

# Abstract

A rapid increase in Alzheimer's disease (AD) patients is expected over the next 30 years. Accordingly, there is a critical need for early-stage AD detection methods that can enable professionals to treat the disease adequately. The present study considers the ability of episodic-memory measures to predict mild cognitive impairment (MCI) to AD conversion and thus, detect early-stage AD. Using a binary logistic regression, episodic-memory tests were compared to each other and to prominent neuroimaging methods in MCI converter (MCI participants who developed AD) and MCI non-converter groups (MCI participants who did not develop AD). Standard tests for AD (e.g., MMSE) were also compared to specific episodic-memory tests—using a principal component analysis—to test if standard tests can measure episodic memory. Our study acquired all data from the *Alzheimer's Disease Neuroimaging Initiative* (ADNI) and tested participants over four years. Our results indicate that individual episodic-memory measures could predict MCI to AD conversion better than episodic memory, neuroimaging, and mixed models. We theorised that mixed models were worse than individual tests, as mixed episodic-memory models increase multicollinearity and neuroimaging measures had poor accuracy. Specifically, the most accurate predictors were the ADNI memory score in year one (56.4%), the RAVLT percent forgetting measure in year two (71.7%), and the logical memory test in years three (76.9%) and four (77.2%). Our results also indicated that standard tests could be used to measure episodic memory. In conclusion, our study highlighted the ability of episodic-memory tests to predict disease conversion and thus detect early-stage AD.

# Abbreviation List

- Alzheimer's Disease = AD
- Mild Cognitive Impairment = MCI
- Amnestic MCI= aMCI
- Non-Amnestic MCI= naMCI
- Alzheimer's Disease Neuroimaging Initiative = ADNI
- Amyloid Beta = Aβ
- Apolipoprotein E4 = APOE4
- Magnetic Resonance Imaging = MRI
- Positron Emission Tomography = PET
- Medial Temporal Lobe = MTL
- Logical Memory Test = LMT
- Rey Auditory Verbal Learning Test = RAVLT
- Alzheimer's Disease Assessment Scale-Cognitive = ADAS-cog
- Mini-Mental State Exam = MMSE
- Cerebrospinal Fluid = CSF
- ADNI Composite Memory Score = ADNIMEM
- Variance Inflation Factor = VIF

# Variable Dictionary

- UCSF Ventricles = UCSF_VENT
- UCSF Hippocampus = UCSF_HIP
- UCSF Whole Brain = UCSF_WB
- UCSF entorhinal = UCSF_ENT
- UCSF Fusiform = UCSF_FUSI
- UCSF Medial Temporal = UCSF_MED
- RAVLT Immediate = RAVLT_IMM
- RAVLT Learning = RAVLT_LEAR
- RAVLT Percent Forgetting = RAVLT_%_FOR
- RAVLT Forgetting = RAVLT_FOR
- Alzheimer's Disease Assessment Scale-Cognitive (13 task version) = ADAS13
- Alzheimer's Disease Assessment Scale-Cognitive (11 task version) = ADAS11
- ADAS Delayed Recall Task = ADASQ4
- Logical Memory Test (Delayed Recall) = LMTOTAL
- Mini-Mental State Exam = MMSE
- ADNI Composite Memory Score = ADNI_MEM
- MCI Conversion and non-conversion = Conv/Non-Conv

# Chapter One: Background and Literature Review

## 1.1 Introduction

Alzheimer's disease (AD) is the second leading cause of death in Australia and, consequentially, is one of the greatest medical threats of our time (Australian Bureau of Statistics, 2017). Specifically, AD is the most common form of dementia, which, in turn, currently affects 436,000 Australians (Australian Bureau of Statistics, 2017; Dementia Australia, 2018). Worryingly, dementia diagnoses are expected to increase drastically from 50 million to 132 million global patients in the next 30 years (Alzheimer's Disease International, 2019). The rapid increase in AD patients is problematic, as the cause of AD is currently unknown (Alzheimer's Association, 2018). Furthermore, current AD diagnoses occur at a later stage of the disease, when neurological damage is presumed to be irreversible and most medications are largely ineffective (Kalat, 2018, p. 206). Accordingly, current symptomatic treatments such as cholinesterase inhibitors and aged care can mostly treat physical and cognitive decline (e.g., memory and motor function loss) but cannot reverse or prevent the disease (Edvardsson, Winblad, & Sandman, 2008; Tan et al., 2014). Clearly, there is a need for significant innovations in AD diagnosis and treatment to overcome what is expected to be a global epidemic.

Specifically, innovations in early-stage AD diagnosis are required to adequately treat and understand disease mechanisms. Research suggests that current diagnostic methods, such as standard paper-and-pencil tests (e.g., Clinical Dementia Rating Scale), cannot reliably diagnose the early stages of AD. Moreover, diagnostic methods which could potentially categorise early-stage AD (e.g., MRI) are often unavailable to patients as they are vastly unaffordable and inaccessible. Alzheimer's Disease must be diagnosed as quickly as possible,

as early-stage treatment can reduce the rate of neurodegeneration and prolong the onset of severe cognitive decline (Chu, 2012). Furthermore, early-stage AD detection enables researchers to examine the onset of the disease longitudinally and, therefore, understand disease pathogenesis and progression (Reiman et al., 2012). Accordingly, innovations in early-stage AD diagnosis could lead to the detection of the disease at a significantly manageable and reversible stage (Frisoni et al., 2017). The Alzheimer's Association predicts that innovations in early-stage AD diagnostic methods could save the US alone $7.9 trillion in future treatment costs (Alzheimer's Association, 2018). Clearly, there is a critical need for new AD diagnostic methods which can detect the early stages of the disease.

Thus, this study investigates methods of predicting early-stage AD progression. Specifically, we use longitudinal data from episodic-memory tests to predict AD development in participants with MCI (an interim disorder which commonly converts to AD). Cognitive tests are used to predict disease conversion, as episodic-memory decline is one of the first symptoms of early-stage AD (Perry, Watson, & Hodges, 2000). Moreover, we examine the predictive ability of episodic-memory tests in comparison to magnetic-resonance-imaging (MRI) measures of brain volume, which are commonly used to predict early-stage AD. Accordingly, we seek to predict early-stage AD and to evaluate the predictive ability of episodic-memory tests compared to the currently standard neuroimaging measures. By predicting early-stage AD and evaluating current disease measures, this study hopes to contribute to the formulation of better diagnostic measures and, thus, improve disease detection and treatment capabilities.

## 1.2 Theories of Alzheimer's Disease Pathogenesis

Many aspects of AD, such as the development and progression of the disease (i.e., pathogenesis), are not fully understood. Accordingly, theories such as the amyloid cascade and cholinergic hypotheses provide great insight into AD pathogenesis and can be used to inform early-stage AD diagnostic research.

The amyloid cascade hypothesis suggests that AD development is related to the aggregation and formulation of neurotoxic matter known as plaques and tangles (Bloom, 2014). Specifically, high concentrations of the peptide amyloid beta (Aβ) result in the formulation of plaques and the protein tau in neurofibrillary tangles. Why Aβ and tau aggregate to the point of neurotoxicity is unknown; however, both Aβ and tau aggregation correlate with AD progression (Murphy & LeVine, 2010). Research has shown that high concentrations of Aβ interrupt the dendritic function and synaptic ability of neurons (Yankner, Duffy, & Kirschner, 1990). Accordingly, the loss of synaptic ability leads to communication breakdowns, neuronal failure, the formulation of further senile plaques, and thus cascading neurodegeneration. In addition, research has shown that the gene apolipoprotein E (APOE) assists in the regulation of Aβ in the brain (Michaelson, 2014). The apolipoprotein E4 (APOE4) allele is subsequently a risk factor for AD development, as the APOE gene variant lacks the ability to sufficiently clear Aβ build up in the brain; however, it is essential to note that APOE4 is only a risk factor for AD, as the gene has no known causal relation to the disease (Uddin et al., 2019).

Tau is a protein found in the axons of neurons. Tau is pivotal in maintaining cell structure and integrity (Weingarten, Lockwood, Hwo, & Kirschner, 1975). When senile plaques compromise the brain, Aβ modulates tau (in a process known as hyperphosphorylation), thereby causing the protein to spread from the axon to the wider neuron (Iqbal, Liu, Gong, & Grundke-Iqbal, 2010). Hyperphosphorylated tau forms into

clumps known as neurofibrillary tangles which impale neurons, thus compromising the structural integrity of the cell and causing further neurodegeneration. Accordingly, the assaults of both Aβ and tau are both malicious and meticulous, with tau compromising neurons internally and Aβ dismantling them externally. The amyloid cascade hypothesis states that Aβ and tau derived neurodegeneration starts in the hippocampus and progressively expands into the wider medial temporal lobe (MLT) and cortex (El Haj, Antoine, Amouyel, et al., 2016). Research shows that spreading neurodegeneration is correlated with symptoms of AD progression. For example, neurodegeneration in the cerebral cortex correlates with increased aggression and the loss of fine motor skills in AD (Kalat, 2018, p. 204). Accordingly, the amyloid cascade hypothesis is the current theoretical basis for the majority of AD progression and biomarker research; however, one common critique of the theory is that the hypothesis lacks specific information regarding the manifestation of higher-order deficits such as cognitive decline in AD.

The cholinergic hypothesis originally stated that the depletion and dysfunction of the neurotransmitter acetylcholine causes cognitive decline in AD (Bartus, Dean, Beer, & Lippa, 1982). However, modern iterations of the theory consider cholinergic decline to be a risk factor or contributory process towards cognitive decline in AD (Craig, Hong, & McDonald, 2011). The cholinergic system of the brain is instrumental in managing attention, neuroplasticity, and complex cognitive processes such as memory (Teipel, Grinberg, Hampel, & Heinsen, 2009). Accordingly, there is a clear link between the higher-order symptoms of AD and tasks which involve the cholinergic system. For example, neurodegeneration in the basal forebrain is related to the loss of acetylcholinic neurons and memory function in AD (Ferreira-Vieira, Guimaraes, Silva, & Ribeiro, 2016; Francis, Palmer, Snape, & Wilcock, 1999). In healthy aging, it is typical for cholinergic neurons to degenerate slowly (Davidson & Winocur, 2010); however, modern iterations of the cholinergic hypothesis state that in AD,

cholinergic decline is drastically increased by senile plaques and neurofibrillary tangles (Craig et al., 2011). Accordingly, the cholinergic hypothesis provides evidence to suggest that higher-order decline in AD occurs beyond the mechanisms of the amyloid cascade hypothesis.

It is essential to note that the cholinergic hypothesis, while providing valuable information on cognitive decline, does not provide as cohesive an explanation of AD as once thought. Cholinergic decline is only seen as a risk factor for AD, as the theory lacks widespread experimental support (Terry & Buccafusco, 2003). However, to date, cholinesterase inhibitors are the only medications which viably treat and delay cognitive decline in AD. Accordingly, theories such as the amyloid cascade hypothesis regard cholinergic decline as a symptom of neurodegeneration in AD (Armstrong, 2013). Nonetheless, while the cholinergic hypothesis is lacking, the theory does highlight the need for a deeper understanding of AD progression and cognitive decline.

Moreover, while current theories such as the amyloid cascade hypothesis provide insight into AD, they do not sufficiently explain the disease (e.g., cognitive decline). For example, biological markers of AD which are centric to current theories (e.g., APOE4 and Aβ) are not strong predictors of early-stage AD (Liu, Kanekiyo, Xu, & Bu, 2013). Accordingly, there is a gap in the literature concerning AD markers which can both detect the early stages of the disease and inform theories of pathogenesis and cognitive decline in AD. Clearly, there is a critical need for further research which can inform current theories and early-stage AD detection methods. Thus, by investigating cognitive decline (specifically episodic-memory decline), this research project seeks to inform gaps in early-stage AD progression theories and predict early-stage AD.

## 1.3 Early-Stage Alzheimer's Disease Research

### 1.3.1 The Alzheimer's Disease Continuum

Early-stage AD research often aims to predict disease development and observe the progression of pathological markers (e.g., Aβ aggregation). In turn, these disease predictors and markers inform early-stage AD diagnosis strategies and potential treatment areas. Researchers commonly predict early-stage AD and identify pathological markers by longitudinally studying precursor disorders which transition to AD. Specifically, the Alzheimer's Association and the National Institute on Aging recognise AD as a continuum with three different stages known as preclinical AD, MCI, and AD dementia (Jack et al., 2011). Accordingly, preclinical AD and MCI are often used to study early-stage AD. Preclinical AD is the initial stage of AD development and is an asymptomatic state which can last decades before transitioning to MCI or AD (Dubois et al., 2016). Due to the nature of preclinical AD, MCI is preferred in early-stage AD research as MCI is easier to predict, recognise, and study.

*MCI* refers to the nondemented impairment of at least one cognitive ability, such as language or memory (Csukly et al., 2016). It is important to note that these impairments are not debilitating in MCI. Debilitating impairments indicate another disorder such as dementia. Moreover, MCI is not exclusive to AD development, yet MCI does share many characteristics with AD and reliably converts to the disease (Albert et al., 2011). Specifically, individuals with MCI convert to AD at a rate of 10-15% while only 2-4% of the general population progress to AD (Roberts et al., 2008). Two types of MCI are characterised by the specific domain impaired. If memory is impaired, the disorder is known as amnestic MCI (aMCI). If any other cognitive domain is impaired (e.g., language), the disorder is known as non-amnestic MCI (naMCI) (Tannenbaum, Paquette, Hilmer, Holroyd-Leduc, & Carnahan, 2012). Accordingly, the classification of MCI is important, as aMCI reliably converts to AD

while naMCI converts to other disorders such as Parkinson's Disease and non-Alzheimer's dementia (Ferman et al., 2013; Goldman, Weis, Stebbins, Bernard, & Goetz, 2012). Consequently, most if not all of early-stage AD research investigates aMCI participants.

### 1.3.2 Overcoming Barriers in AD Research

Clearly, understanding aMCI and memory loss is essential when investigating early-stage AD. However, studying AD and MCI participants can be incredibly complicated.

AD and MCI research is incredibly complex and is often limited by three main factors. First, the logistical complexity of organising and recruiting neurological populations (e.g., individuals with dementia) can limit MCI and AD participant counts. For example, neurological participants, such as participants with AD, are likely to drop out of longitudinal studies as the symptoms of their disease worsen (Watson, Ryan, Silverberg, Cahan, & Bernard, 2014). Second, high monetary and administrative costs restrict the use of AD detection methods. For example, MRI technology can measure neurodegeneration with high accuracy, yet the device is incredibly expensive to use and hard to access (Stites, Milne, & Karlawish, 2018). Finally, the resources required to study AD progression and MCI conversion longitudinally are immense and beyond the scope of some labs. In any case, the resources required to study early-stage AD are beyond the scope of this study. However, these temporal, logistical and recruitment barriers to AD research can sometimes be overcome through joint research projects such as the *Alzheimer's Disease Neuroimaging Initiative* (ADNI).

ADNI is the largest and most commonly used research database in the AD literature. Most importantly, for researchers, ADNI provides complex longitudinal data with numerous clinical and experimental measures of MCI and AD participants. As an organisation, ADNI seeks to detect and observe AD progression using disease biomarkers. For more information see the Alzheimer's Disease Neuroimaging Initiative (2017a). Specifically, ADNI collects

neuroimaging, traditional biomarker, neuropsychological, and demographic measures in a cohort of control (neurologically healthy), MCI, and AD participants. For more information see the Alzheimer's Disease Neuroimaging Initiative (2017b). Most importantly, for this project, ADNI operates on a policy of information sharing which enables any study with an institutional ethics approval to use their data for research. Accordingly, researchers can study topics such as early-stage AD by using the ADNI database while overcoming the logistical, temporal and economic limitations common to AD research. Thus, this project uses the ADNI database to predict MCI to AD conversion and, therefore, to detect early-stage AD.

### 1.3.3 Common Methods of Early-Stage AD Research

Alzheimer's disease research is commonly categorised into the three fields of neuroimaging, biomarker, and neuropsychological research. In the case of ADNI, each of these fields seeks to investigate and predict AD progression using pathological markers or measures. Importantly, ADNI chooses its markers and methods according to the surrounding literature; however, the popularity and accessibility of the ADNI database also significantly affect the popularity of some measures and methods.

**Biomarker research**

Traditional biomarker research assesses biochemical changes in the brain which are associated with AD. For example, AD biomarker research commonly investigates genes, proteins, and peptides such as APOE4, tau, Aβ42, and presenilin-1 (Natelson Love et al., 2017; Sharma & Singh, 2016). Unlike neuroimaging and neuropsychological methods, traditional biomarkers require the extraction of biological material, which can often involve invasive procedures (e.g., lumbar punctures). For example, ADNI collects blood, cerebral spinal fluid (CSF), and urine samples to study plasma, enzymes, proteins, amino acids, and genes which are biomarkers for AD. Overall, biomarkers provide great insight into the

progression and symptoms of AD; however, traditional biomarkers are weak predictors of early-stage AD compared to neuroimaging and neuropsychological measures (Cui et al., 2011). Research has shown that traditional biomarkers cannot strongly predict and detect the early stages of AD, as they often rely on symptomatic measures and risk factors (Liu et al., 2013; Yip, Brayne, Easton, & Rubinsztein, 2002). Consequently, traditional biomarkers are only useful predictors of early-stage AD when they form parts of a mixed predictive model (e.g., combined with neuroimaging measures) (Vemuri et al., 2009). Moreover, biomarkers also find little use in clinical diagnoses, as procedures are often uncomfortable or invasive and biomarker samples requiring lab processing. Comparatively, neuropsychological pen-and-paper tests are often preferred because they require no adverse procedures and can be conducted quickly on site.

**Neuroimaging research**

Neuroimaging research commonly uses electromagnetic signals to noninvasively measure neurodegeneration and metabolic changes in AD (e.g., MRI measures of hippocampal atrophy). Neuroimaging research is, in part, a form of biomarker research but is often separated from traditional biomarkers due to the size and unique methodology of the field. Magnetic resonance imaging (MRI) and positron emission tomography (PET) are the most commonly used methods in neuroimaging research. In ADNI, MRI measures broadly assess brain atrophy, volume, and neuron connectivity while PET measures assess brain metabolism and Aβ pathology (Davatzikos, Bhatt, Shaw, Batmanghelich, & Trojanowski, 2011; Ewers et al., 2014; Landau et al., 2012). Neuroimaging technology is highly accurate when used to predict early-stage AD, especially when MRI measures of brain volume and atrophy are used. For example, MRI measures of hippocampal atrophy and entorhinal volume are some of the best predictors of MCI to AD conversion and, therefore, of early-stage AD (Brueggen et al., 2015; Tapiola et al., 2008; Velayudhan et al., 2013). Furthermore,

neuroimaging measures provide in-depth analyses of neuronal health and brain connectivity in AD (Cai et al., 2015).

Clearly, neuroimaging methods are at the forefront of AD research; however, these measures have little application to clinical practise as the technology has a high temporal, monetary, and logistical cost. Accordingly, the Alzheimer's Association and the National Institute on Aging suggest that biomarkers (including neuroimaging methods) can inform diagnoses when possible but cannot be used to diagnose the disease without further technological development (Jack et al., 2011).

**Neuropsychological research**

Neuropsychological research focuses primarily on measuring cognitive and functional ability (e.g., motor skills and memory). Accordingly, pen-and-paper tests are commonly used to diagnose AD and to monitor at-risk patients (e.g., MCI patients). For example, ADNI currently uses 11 cognitive, and 10 functional and behavioural tests to monitor all AD, MCI and control participants. In neuropsychology, cognitive tests can be further divided into three categories depending on the nature and use of the specific test (Brown, 2015). In AD research, cognitive tests can broadly be categorised as the following: Short questionnaires used for screening AD (e.g., abbreviated mental test), highly specific tests used to discriminate between similar diseases such as vascular dementia and AD (e.g., the clock-drawing test), and general multi-domain tests commonly used for AD diagnosis (e.g., the Mini-mental-state Exam) (Brown, 2015; Kato et al., 2013; Swain, O'Brien, & Nightingale, 1999).

General multi-domain tests are the most commonly used neuropsychological measures in both clinical practice and research. Accordingly, these general multi-domain tests are often referred to as standard tests. Standard tests boast a high accuracy in AD research and are the current standard for clinically diagnosing AD. Standard tests are widely

used throughout the literature, as they are easily administrable and affordable in comparison to other measures (e.g., MRI); however, standard tests lack the ability to diagnose early-stage AD and are prone to misdiagnosing forms of dementia (e.g., diagnosing vascular dementia as AD) (Arevalo-Rodriguez et al., 2015; Bak et al., 2005; Larner, 2019). Accordingly, early-stage AD research does not commonly focus on cognitive predictors of early disease development or MCI conversion. Nonetheless, the literature has suggested that highly specific tests which target the early cognitive symptoms of AD (e.g., episodic-memory decline) could be used to diagnose the early stages of the disease (Bastin & Salmon, 2014; Brown, 2015). However, due to the dominance of standard tests and preference for neuroimaging and biomarker measures in the literature, research is lacking on the predictive ability of specific cognitive tests in early-stage AD research. Thus, further research is required to validate the predictive ability of specific cognitive tests in early-stage AD research.

Clearly, neuroimaging, biomarker, and neuropsychological research have unique strengths which, with further innovation, could enable each field to detect early-stage AD. However, each field also has weaknesses which currently limit early-stage AD prediction and diagnosis. Accordingly, there is a need for neuroimaging and biomarker research to become more clinically reliable and accessible. Moreover, neuropsychological and biomarker research must overcome flaws in accuracy, reliability, and validity to warrant use in early-stage AD diagnosis. Thus, this study investigates the ability of specific cognitive tests to predict MCI to AD conversion and, therefore, early-stage AD. This project assesses cognitive tests over other measures (e.g., neuroimaging) for the following reasons: A) studying cognitive tests enables us to address the theoretical gaps concerning the development of cognitive decline in early-stage AD, B) cognitive tests are standard measures of AD yet lack research concerning early-stage AD, and C) specific cognitive tests could overcome some of

the limitations of general tests for AD diagnosis (e.g., poor early-stage AD detection). Accordingly, this project uses episodic-memory tests to predict MCI to AD conversion in the ADNI database and, thus, detect early-stage AD.

## 1.4 Episodic Memory in Alzheimer's Disease Research

### 1.4.1 Episodic Memory and Decline

Episodic-memory decline is one of the first indicators of AD progression and, thus, is an optimal candidate for early-stage AD detection and prediction. Episodic memory is an archival form of memory that encodes events with temporal and positional details to create complex situational memories (Tulving, 1972, 1985, 2002). Episodic memory is an incredibly complex form of memory; consequentially, our understanding of it remains mostly theoretical. Accordingly, current theory states that episodic memory function occurs using multiple neurological structures and systems; this is the component process model of episodic memory (Morris Moscovitch, Cabeza, Winocur, & Nadel, 2016, p. 2). It is widely agreed that episodic-memory processes occur predominantly in the medial temporal lobe (MTL) and cortex. Moreover, the component process model regards the hippocampus as the centrepiece of episodic-memory processing (Eichenbaum, 2017).

The component process model states that the hippocampus is a hub that combines and translates information from across the brain to construct episodic memories. It is not clear what specific relationships and executive processes directly result in the formulation of episodic memory; however, theory suggests that sensory, object, context, temporal, spatial, and emotional information are all incorporated in the hippocampus to develop an episodic-memory engram (Barker et al., 2017; Morris Moscovitch et al., 2016). Furthermore, this multitude of information is previously processed in and then received from the MTL and neocortex (M. Moscovitch, 1992). Because episodic memories form through the complex integration of multiple systems, research suggests that the brain uses electronic signals known as neural oscillations (also known as brain waves) to coordinate memory formation and recollection (Nyhus & Curran, 2010). Accordingly, electroencephalographic studies have observed theta-wave neural oscillations that are indicative of coordinated episodic memory

function (Hot et al., 2011). However, the structural complexity of the episodic memory system and its reliance on complex forms of communication also makes episodic memory a prime target for decline in AD.

It is, therefore, no coincidence that episodic-memory decline is one of the first symptoms of early-stage AD, as both amyloid cascade and neurodegeneration are theorised to start in the hippocampus. Thus, the hippocampal system is pivotal to both episodic memory function and neurodegeneration in AD (El Haj, Antoine, Nandrino, & Kapogiannis, 2016). Accordingly, in AD, episodic-memory decline is theorised to occur because of communication breakdowns in neural oscillations due to interference from senile plaques and neurofibrillary tangles. Moreover, episodic-memory decline is also theorised to occur because of neurodegeneration in the MTL (El Haj, Antoine, Amouyel, et al., 2016; El Haj, Antoine, Nandrino, et al., 2016; Morris Moscovitch et al., 2016).

As amyloid and neurodegenerative processes are poorly understood in AD, the pathogenesis and progression of episodic-memory decline also remain enigmatic. Contemporary theories of episodic memory focus more on the structures that decay than on the causes and progression of memory decline. However, with further research, contemporary theories of episodic-memory decline (e.g., the component process model) could address some of the theoretical gaps in the amyloid cascade hypothesis such as the onset of cognitive decline. If episodic-memory decline can be further understood, researchers can use it as a tool to predict and detect early-stage AD. Thus, this study seeks to predict early-stage AD and to understand episodic memory using the ADNI database.

### 1.4.2 Episodic-Memory Tests in ADNI

ADNI collects a vast amount of neuropsychological information for research through cognitive and psychological tests. ADNI uses 11 cognitive tests. The Logical Memory Test (LMT) and Rey Auditory Verbal Learning Test (RAVLT) are used to measure episodic

memory (Alzheimer's Disease Neuroimaging Initiative, 2016, p. 21). The LMT is a revised form of the Weschler Memory Scale, which assesses episodic memory formulation and recollection using a short story (Abikoff et al., 1987). Similarly, the RAVLT assesses episodic memory by using a list-learning strategy that measures delayed word recall (Vakil & Blachstein, 1993). Evidence also suggests that the cognitive variant of the Alzheimer's Disease Assessment Scale (ADAS-cog) and the Mini-mental-state Exam (MMSE) can predict episodic memory performance even though they do not explicitly assess episodic memory (Crane et al., 2012; Gomar et al., 2011). The MMSE and ADAS-cog are argued to test episodic memory, as they incorporate list learning and recall tasks. Some studies have argued that a combined total episodic-memory score from all the tests listed above best predicts AD (Seo, Choo, & Alzheimer's Disease Neuroimaging Initiative, 2016); however, to our knowledge, no study has comprehensively assessed all episodic-memory measures in the ADNI database. All past and present ADNI data collection has incorporated the 11 tests, yet only a small literature examines episodic memory performance in the initiative's cohort.

Most of the ADNI studies that investigate episodic memory use the RAVLT alone as a measure of episodic memory. Only using the RAVLT could be problematic, as the accuracy and ability of most, if not all, episodic-memory tests are not fully understood in early-stage AD research. Moreover, many tests seek to quantify and measure constructs of episodic memory; however, there is no single, all-encompassing episodic-memory test (Cheke & Clayton, 2015; Humphreys, Smith, Pachana, Tehan, & Byrne, 2010). A study by Cheke and Clayton (2013) examined multiple episodic-memory tests and found that no single test fully measured the construct of episodic memory. Variations in memory tests are understandable, as episodic memory is an incredibly complex mental process for which no cohesive understanding appears in the literature (Tulving, 2002). Accordingly, cognitive tests are both limited by our understanding of episodic memory and by their ability to measure the

construct of episodic memory wholly and directly. Therefore, relying on a single episodic-memory test such as the RAVLT may not wholly and directly measure episodic memory compared to a battery of tests, as suggested by Seo et al. (2016). The ADNI database has also used two versions of the RAVLT throughout the initiative's history. Evidence suggests that a difficulty is encountered in changing between RAVLT versions (Crane et al., 2012). The difficulty is confounding to studies that use only the RAVLT and do not control for variation differences when longitudinally assessing participants across ADNI cohorts. Accordingly, there is a critical need to comprehensively examine and understand episodic-memory measures in the ADNI database so that future research can use cognitive measures to their full potential. Moreover, there is a need to critically examine the ability of all ADNI episodic-memory measures to predict early-stage AD.

### 1.4.3 Episodic-Memory Research

The ADNI episodic-memory literature is relatively small compared to other prominent early-stage AD fields such as MRI and biomarker research. Accordingly, most ADNI studies often assess episodic memory as part of a mixed predictive model with other AD markers (e.g., episodic memory and hippocampal atrophy model). Although the mixed-predictive-model approach makes sense methodologically, in practice, most studies address episodic memory with little focus. Moreover, it is typical for studies that seek to predict early-stage AD to include episodic memory as a variable in a long list of many predictors; they thus do not study it comprehensively. The remaining few studies that do focus specifically on episodic memory often study only the correlational or predictive ability of one episodic-memory measure with one other disease marker, thereby limiting the ability of researchers to understand episodic memory in AD. While it is essential to study the interaction of episodic memory with other disease markers (e.g., Aβ), the mixed model approach has the result that episodic memory lacks specific and focused research. However,

mixed predictor model research still provides helpful information regarding the potential of episodic-memory measures as predictors of MCI to AD conversion.

**Mixed Neuroimaging and Episodic-Memory Models**

Accordingly, most if not all ADNI episodic memory studies combine episodic-memory scores with neuroimaging measures or biomarkers. Specifically, MRI brain-atrophy-and-volume measures are the most prominent markers assessed in conjunction with episodic-memory tests. Specifically, researchers have found strong relationships between MTL atrophy (hippocampal and entorhinal volume/atrophy) and episodic-memory decline, as theorised in the component process model. For example, one study investigating an array of predictors discovered that neuroimaging measures of cortical thickness and episodic-memory scores best predicted the conversion from MCI to AD; however, the study concluded that neurobiological measures were overall worse than episodic-memory tests in predicting MCI to AD conversion (Gomar et al., 2011). In a follow-up study three years later, the same researchers assessed more predictors of conversion across the whole of the ADNI database and concluded that episodic-memory measures were the best predictors of MCI to AD conversion (Gomar, Conejero-Goldberg, Davies, Goldberg, & Alzheimer's Disease Neuroimaging Initiative, 2014). Accordingly, there is contention concerning which neuroimaging and episodic-memory measure should be used to accurately predict early-stage AD.

Ihara et al. (2018) have investigated episodic memory, hippocampal atrophy and Aβ aggregation in ADNI to understand the relationship and predictive ability of each measure. Ihara et al. (2018) reasoned that episodic-memory decline, hippocampal atrophy and Aβ accumulation are all linked, as theorised in the amyloid cascade hypothesis (see Section 1.2, p.3). However, the study also found that, while Aβ and hippocampal atrophy were related, episodic memory and Aβ were not (Ihara et al., 2018). The researchers reasoned that

episodic-memory decline in AD derives from hippocampal atrophy that is mediated by Aβ concentration. Moreover, Ihara et al. (2018) have further suggested that there is no direct relationship between Aβ concentration and episodic-memory decline. The relationship between MTL atrophy and episodic memory ability was further reinforced by a study which showed that MRI measures of atrophy could predict episodic memory performance in ADNI (Moradi, Hallikainen, Hänninen, Tohka, & Alzheimer's Disease Neuroimaging Initiative, 2017). However, there is a need for more research, as it is clear that no decisive episodic memory and neuroimaging model consistently predicts MCI to AD conversion. Thus, this project builds on past research by examining multiple neuroimaging and episodic-memory measures together to assess the predictive accuracy of a mixed model.

Beyond atrophy, other neuroimaging studies have sought to determine how neurological connectivity affects episodic memory in ADNI by using functional MRI. Breakdowns in the communicative ability of the frontal temporal and thalamic regions in amnestic MCI participants are indicative of cognitive decline (Cai et al., 2015). However, little literature investigates neural connectivity in ADNI even though it is a fundamental aspect of the component process model of episodic memory. We theorise that the lack of clarity and consensus concerning neural connectivity (neural oscillations) occurs due to the lack of measures in the ADNI database. Specifically, ADNI only uses fMRI data to assess neuronal connectivity and not electroencephalographic measures which are preferred in the wider literature. Accordingly, this topic is beyond the scope of this research project and our acquired ADNI dataset.

**Mixed Biomarker and Episodic-Memory Models**

Traditional biomarkers, such as cerebrospinal fluid (CSF) and genes, are the other prominent AD markers which are studied alongside episodic memory in the ADNI database. The most prominent biomarkers studied in the ADNI database are the Aβ and tau CSF

markers. There are varying opinions regarding the viability of biomarkers in predicting AD and correlating with episodic memory. Some ADNI studies have found no relationship between Aβ and episodic memory (Gomar et al., 2011, 2014; Mormino et al., 2009); conversely, other studies have found evidence that Aβ and episodic memory correlate and can accurately predict MCI to AD conversion (Lin et al., 2017; Nathan et al., 2016; María J. Russo, Campos, Vázquez, Sevlever, & Allegri, 2017; María Julieta Russo et al., 2016; Seo et al., 2016). Thus, further research is required to determine the relationship between Aβ and episodic memory in AD.

As with CSF biomarkers, genetic risk factors are also often assessed with episodic-memory decline to predict early-stage AD. A small number of studies show positive results in predicting MCI to AD conversion and episodic-memory decline (Nathan et al., 2016; Ramanan et al., 2012); however, the literature is too small to conclude that genes can predict early-stage AD in the ADNI cohort. Furthermore, in the broader AD literature, genetic markers such as APOE4 have been rigorously investigated and shown to be mere risk factors of disease development (Povova et al., 2012). More research is required to understand the interaction of genes and CSF biomarkers in AD pathogenesis and their role in AD derived episodic-memory decline. Accordingly, current research would suggest that episodic memory best predicts MCI to AD conversion as a single measure or when combined with neuroimaging biomarkers. Consequently, this project focuses on predicting MCI to AD conversion using neuroimaging and episodic-memory measures.

**Individual Episodic-memory Predictors**

Outside of ADNI, a small number of studies have examined the ability of episodic-memory measures to predict MCI to AD conversion and, thus, early-stage AD. A study by Chapman et al. (2011) investigated the ability of neuropsychological tests to predict AD and found that episodic memory could predict MCI to AD conversion. Specifically, they found

that episodic memory, when combined with recognition, visuospatial memory, and executive functioning measures, could predict disease conversion with an accuracy of 84%. Chapman et al. (2011) primarily used the LMT, RAVLT, MMSE, brief visuospatial memory test revised, and Hopkins verbal learning test to study episodic memory. The statistical analysis incorporated neuropsychological test scores into total scores (e.g., episodic-memory scores) using a principal component analysis. However, the individual predictive accuracy of each episodic-memory test was not thoroughly discussed or compared. Moreover, Chapman et al. (2011) concluded that a single episodic-memory measure could not predict MCI to AD conversion and so recommended using a mixed model containing various neuropsychological tests (e.g., visuospatial memory, executive functioning and episodic-memory measures).

A more recent study by De Simone et al. (2019) examined the accuracy of specific episodic-memory tests when predicting MCI to AD conversion. De Simone et al. (2019), unlike prior researchers, analysed multiple episodic-memory measures; however, De Simone et al. (2019) combined them into total scores and did not report the accuracy of specific tests and measures. Nonetheless, De Simone et al. (2019) highlighted the potential of specific episodic-memory tests to predict MCI to AD conversion with a high accuracy. Accordingly, preliminary research has suggested that episodic-memory measures can predict MCI to AD conversion both individually and in a mixed model (Chapman et al., 2011; De Simone et al., 2019; Gomar et al., 2011). However, in the literature, there are discrepancies in research results, as some studies do not evaluate each contained episodic-memory test or report accuracy measures. There is also contention concerning the relationship between episodic memory and neuroimaging measures in mixed predictive models. Moreover, some key areas of episodic memory remain unexplored, such as the relationship between specific and general episodic-memory tests.

Thus, building on the above research, this project seeks to assess ADNI participants across four years by using multiple general and specific episodic-memory measures (i.e., the ADAS-cog, MMSE, ADNIMEM, RAVLT, and LMT). Moreover, through ADNI, this project has access to more participants and variables (neuroimaging and biomarker) than prior studies such as Chapman et al. (2011). Furthermore, this project can also use a more sophisticated analysis to rigorously and longitudinally compare episodic-memory predictors of MCI to AD conversion. Like the study of De Simone et al. (2019), this study longitudinally and comprehensively examines the predictive accuracy of various episodic-memory tests when predicting MCI to AD conversion. However, unlike De Simone et al. (2019), this project also does the following: A) tests the episodic-memory measures contained in the ADNI database (which to our knowledge has not been done before), B) uses a more complex logistic regression based predictive analysis, C) assesses neuroimaging predictors of conversion and mixed episodic-memory models, D) examines the predictive ability of general and specific episodic-memory measures, and, most importantly E) reports the accuracy of individual episodic-memory measures when predicting MCI to AD conversion.

## 1.5 The Present Study

Past studies have examined the ability of various cognitive measures and biomarkers to predict MCI to AD conversion and, subsequently, early-stage AD. However, there is disagreement amongst researchers regarding the reliability, validity, and accuracy of the varying individual measures of early-stage AD. For example, traditional biomarkers have a poor predictive accuracy which impedes early-stage AD diagnosis. Cognitive measures (e.g., standard tests) are often regarded as ineffective measures of early-stage AD, as they mostly assess the later symptoms of the disease; however, little research has been conducted on specific tests that target the initial cognitive symptoms of early-stage AD, such as episodic memory loss. Across the AD literature, episodic memory is consistently regarded as a strong predictor of AD development; yet episodic memory loss is not widely understood in early-stage AD.

Accordingly, this study seeks to investigate the predictive ability of episodic-memory measures in the ADNI database to inform current early-stage AD theory and diagnostic methods. Most ADNI studies rely heavily on the RAVLT as a single measure of episodic memory and do not comprehensively assess episodic memory. Moreover, to our knowledge, no study has sought to test and evaluate all the episodic-memory measures in the ADNI database for their predictive ability. Thus, this study tests and evaluates all measures of episodic memory in the ADNI database for their accuracy in predicting MCI to AD conversion. The evaluation of episodic-memory tests involved studying each measure both individually and in mixed cognitive and neuroimaging marker models. Specifically, this study seeks to create the best episodic-memory predictor model and compare it to neuroimaging predictor models, which are the preferred measures in the literature. Moreover, by comparing the accuracy of episodic memory and neuroimaging models, we can theorise how these measures interact and determine whether significant memory decline occurs before

or after neurodegeneration. Thereby, informing theories of episodic-memory decline and early-stage AD progression, such as the amyloid cascade hypothesis and component processing model.

Importantly, this study assesses each episodic-memory measure across each year of ADNI's second cohort. It was essential to study each year separately, as theory states that different episodic-memory tests optimally predict AD at different points in time during conversion; however, episodic-memory tests were also evaluated for their overall predictive accuracy across the whole cohort. It was also important that general episodic-memory measures (e.g., standard tests that can measure episodic memory such as the MMSE) be compared to specific episodic-memory measures, thus determining the ability of specific tests compared to standard tests in the ADNI database and early-stage AD research.

Accordingly, this project aims to do the following:

- Assess whether general cognitive assessment tests such as the MMSE and ADAS-Cog can measure episodic memory.
- Determine which episodic-memory tests can best predict MCI to AD conversion individually and in a mixed model.
- Investigate whether neuroimaging predictors can outperform or improve episodic-memory predictive models.

We hypothesise that hippocampal atrophy and specific episodic-memory scores (e.g., LMT) are the best predictors of MCI to AD conversion. We further hypothesise that an aggregate score of all episodic-memory tests performs better than each memory test individually. Moreover, we predict that a mixed episodic-memory and neuroimaging model will best predict MCI to AD conversion in the ADNI cohort.

It is important to note that this study does not seek to understand every facet of episodic-memory decline in early-stage AD. This study only seeks to understand early-stage AD using episodic-memory measures in the confines of the ADNI cohort. Moreover, this study is not a dictionary for episodic-memory measures but rather an evaluation of the potential of episodic-memory measures in early AD research. This report comprises an original contribution to the discussion and investigation of pathological markers in AD and a much needed discussion on the state of cognitive research in neuroscience. We expect that this project will result in a deeper understanding of episodic-memory decline in early-stage AD and the theoretical processes of episodic memory. Furthermore, we expect that the information gathered from this study will contribute to the pool of knowledge that will, in the future, improve cognitive tests and their ability to detect early-stage AD. This study seeks to provide an original contribution to the understanding of episodic memory and AD detection and prediction strategies. Consequently, our long-term goal is to contribute to the improvement of early-stage AD detection and, subsequently, to improve the quality of life of those with AD.

With the above background, aims, and literature review in mind, the following chapters explore the associated methodology, analysis, and discussion. Specifically, Chapter 2 outlines the methods and measures that are used to address the above aims and hypotheses. Next, Chapter 3 details the statistical analysis and answers our research questions. Finally, Chapter 4 discusses the outcomes of this study as well as detailing future directions for early-stage AD and episodic-memory research.

# Chapter Two: Methods

## 2.1 Design and Methodology

As discussed in the introduction (p. 22), this study seeks to predict MCI to AD conversion using measures of episodic memory. Furthermore, this study aims to evaluate current episodic-memory tests and determine which memory measures best predict early-stage AD. This study also seeks to investigate episodic-memory measures in comparison to neuroimaging measures and discusses the subsequent theoretical implications of our results. This chapter details the sample population, dataset, variables, methodology and statistical analysis used in this study. Particularly, this chapter seeks to outline the reasoning, tools, and methods that were used to transform the aims from Chapter 1 into the statistical analysis and results detailed in Chapters 3 and 4.

Accordingly, a quantitative design is used to compare control (MCI stable) and experimental groups (MCI to AD converters) as well as evaluate the predictive nature of episodic-memory and neuroimaging measures. Moreover, a predictive analysis is used to detect AD conversion over the four years of the ADNI2 cohort. This study assessed participants at yearly intervals, as most, if not all, of the measures contained were recorded annually. Participants came from a volunteer sample population and were initially sorted into control, MCI and AD diagnosis groups by ADNI. For this project, participants were further classified into MCI stable and MCI (to AD) converting groups. All data in this project was received from ADNI. The quality and complexity of the information received from the ADNI database surpass that of any information that can be collected during a one-year project and is in agreement with recent advances in the health sciences on the analysis of large datasets.

**Statistical Analyses**

For the statistical analysis, a principal component factor analysis was used to assess the construct validity of episodic-memory tests. Moreover, the factor analysis also determined whether the MMSE and ADAS-cog could measure episodic memory. For the main analysis, a binary logistic regression was used to predict MCI to AD conversion in MCI participants. The logistic regression tests the predictive accuracy of conversion of individual and mixed models against the observed conversion accuracy in the ADNI database. Binary logistic regressions were also used to determine the accuracy and predictive odds of neuroimaging measures and mixed models. All assumptions of a binary logistic regression were checked using tests of linearity and multicollinearity (For example see Field (2013), p. 794).

## 2.2 Data Source and Acquisition

The Alzheimer's Disease Neuroimaging Initiative was formed in 2004 as a publicly and privately funded entity that explores AD biomarkers and disease prediction strategies. Most importantly, ADNI provides all of its data to researchers for free, as discussed in Section 1.3.2 above (p. 7). The initiative has studied four cohorts over its history, with participants both carrying over from prior research and new participants being added at the beginning of each new initiative. ADNI studies typically last five years to allow sufficient time for longitudinal participant assessment. Upon the initiation of a new ADNI cohort, new biomarkers and research methodologies are added to keep the initiative up to date with the current AD literature. However, because of the gradual iteration of the ADNI cohorts, it is sometimes hard to examine measures across the cohorts. In order, the ADNI cohorts thus far have been the following: the ADNI1, the ADNIGO, the ADNI2 and the ADNI3 cohort.

The ADNI2 cohort is the chosen sample for this study because, out of all four ADNI cohorts, the ADNI2 study is the one most recently finished. The current ADNI3 cohort is still in process, thus limiting sample numbers and the availability of longitudinal participant results. The ADNI2 cohort went for five years from 2011 to 2016 with participants' data recorded at annual or biannual intervals depending on the measures in question. The ADNI2 cohort consists of 700 participants from previous initiatives and 150 cognitively healthy controls, 100 early MCI (EMCI), 150 late MCI (LMCI), 150 AD participants, and a new criterion of 107 participants with a significant memory concern (SMC). The SMC participants are, categorically, control subjects at a higher risk of disease conversion appearing on the AD spectrum between healthy ageing and MCI individuals.

The specific dataset we used was the ADNIMERGE.csv file, which incorporates the most common AD measures in all participants from all initiatives. The measures contained in the dataset include the most commonly used demographic, neuroimaging, biomarkers, and

neuropsychological measures, according to the literature. In the ADNIMERGE datafile participants data is recorded annually; however, if a participant misses a study interval, his or her data is omitted. The only prerequisite for obtaining ADNI data was institutional ethics approval. Accordingly, this study gained ethics approval from *Western Sydney Universities* Human Research Ethics Committee in March of 2019. All statistical analysis was performed using IBM SPSS Version 26.

## 2.3 Participants and Groups

ADNI2 consists of two main streams of participants: those overlapping from prior ADNI cohorts and newly recruited participants. The prior ADNI1 and ADNIGO cohorts carried over approximately 700 participants to ADNI2 out of their original 1000 recruits. As previously noted, the new ADNI2 recruits consist of approximately 150 CN, 100 SMC, 100 EMCI, 150 LMCI and 150 AD participants. ADNI recruited all participants for ADNI2 via print and new media advertising, third-party healthcare providers or previous initiatives. Since the ADNI data is open to the general science community, all participant data is de-identified for ethical reasons. Our participants formed two groups dictated by their diagnosis over the whole of ADNI2. Participants were sorted into a conversion group if they converted from MCI to AD over the five years and into a non-converter group if they remained stable over the five years. Subsequently, CN and SMC participants were not studied, as this study is only interested in MCI to AD conversion. Moreover, it was important to maintain only two groups throughout the study so that binary logistic regressions could be used for the analysis. The resulting population contained 95 converters and 212 non-converters overall with a total population of N = 307 participants. These groups were made to accurately address whether disease markers could predict MCI to AD conversion. When working with atypical neurological populations and the elderly, it is natural to experience mortality and participant drop out, due to the vulnerability of the population. Because of the nature of our population, corrections needed to be made in the form of multiple imputations to reduce the bias created by the exponential and longitudinal increase in missing data. Missing data and multiple imputations are discussed further in Section 2.6.2 below (see p. 37).

## 2.4 Variables and Measures

All available episodic memory and neuroimaging variables for ADNI2 were extracted from the ADNIMERGE dataset. The ADNIMEM composite memory score was also included from an auxiliary dataset, as this score can measure episodic memory (it contains items from the RAVLT and LMT) and is computed from data contained in the ADNIMERGE dataset. Although our research questions do not concern demographical measures such as age or education, these measures were kept as descriptive statistics to enable us to better understand our population. The specific measures contained in this study are detailed in what follows.

### 2.4.1 Episodic-Memory Measures

The RAVLT is a specific episodic-memory tests and is the most commonly used test for assessing episodic memory in the ADNI literature. The RAVLT is a list learning task that tests word recall using multiple trials after as time delay. Specifically, participants are given a list of words and then, after a break, they are asked to identify those words in a larger word list. Most of the episodic-memory tests used by ADNI contain multiple measures that examine different facets of episodic memory (e.g., ADAS-cog and RAVLT). The ADNI and ADNIMERGE data sets contain four measures of episodic memory which use the RAVLT. First, the RAVLT immediate score measures participant word recall after the first list learning trial. Second, the RAVLT learning score measures the number of words remembered across all trials. Next, the RAVLT forgetting score measures the number of words from the original word list missed over all trials. Finally, the RAVLT percent forgetting score is a different quantification of the prior RAVLT forgetting score. In the literature, RAVLT immediate and forgetting variables are the most common measures of episodic memory used and tend to have the best accuracy when predicting MCI to AD conversion.

The LMT is the second specific episodic-memory tests used by ADNI and is based on the Weschler Memory Scale. As discussed in Section 1.4.2 above (see p.14), the test uses a recall task centred around a short story to test episodic memory. The LMT is often used in clinical practice to provide evidence for an AD diagnosis or to monitor at-risk patients such as individuals with MCI. The ADNI database records the LMT as a single score of delayed recall; however, few of the papers which assess episodic memory in the ADNI database have used the LMT and have instead relied too much on the RAVLT. Outside of the ADNI database and early-stage AD research, the LMT commonly has a high accuracy when detecting AD.

The ADAS-cog is a general multi-domain test or standard test that is widely used in AD diagnosis. The ADAS-cog involves various tasks in areas such as language, delayed word recall, comprehension and word recognition. The specific tasks vary depending on the version of the ADAS-cog. Accordingly, the ADNI2 cohort and the ADNIMERGE datasets contain the 11(ADAS 11) and 13 (ADAS 13) task versions of the ADAS-cog. The ADNI database also contains the ADASQ4 measure, which is the fourth section of the ADAS-cog, which measures delayed recall. The ADASQ4 is included in the ADNIMERGE dataset, as it is a specific measure of delayed recall and thus of episodic memory.

The MMSE is also a general multi-domain test (standard test) that is widely used in dementia research and clinical diagnosis. Moreover, the MMSE is often used outside of dementia research to assess cognitive ability. The MMSE is a 30-item questionnaire that assesses multiple cognitive domains and tests abilities such as recall, naming and orientation. The ADNI database records MMSE scores as a total score.

The final episodic-memory measure is the ADNI composite memory score (ADNIMEM). ADNI calculates ANIMEM by using specific items from the RAVLT, ADAS-cog, LMT and MMSE scores. The ADNIMEM scores are calculated using item response

theory methods across all four of the tests (see Crane et al., 2012, for more information). The ADNIMEM score is available to all researchers with access to the ADNI database and is theorised to be better than individual measures of episodic memory (Crane et al., 2012). The ADNIMEM score is not contained in the ADNIMERGE dataset; however, the data used to calculate it is contained therein. It is unknown why ADNIMEM is not initially contained in the ADNIMERGE dataset, as that is the file its data originates from. Thus, we treated it as an ADNIMERGE measure and included it in our dataset.

### 2.4.2 Alzheimer's Disease Neuroimaging Markers

The ADNIMERGE dataset contains volumetric neuroimaging measures for the majority of participants throughout all cohorts. Participant data was initially collected by ADNI on site or through certified partner organisations such as universities or the Mayo Clinic. ADNI often outsources the collection and management of its neuroimaging data to partner companies and institutions, and the Mayo Clinic handles most of the neuroimaging data. The University of California San Francisco (UCSF) processes and analyses all of the neuroimaging data used in this study as part of a partnership with ADNI. ADNI2 and UCSF use FreeSurfer software to clean and analyse raw MRI data. In this case, FreeSurfer was explicitly used to visually reconstruct the cortex and segment brain regions into volumes for analysis. For more information about FreeSurfer, see the Athinoula A. Martinos Center for Biomedical Imaging (2019), and for more information about UCSF methods, see the Alzheimer's Disease Neuroimaging Initiative (2017). The volumetric neuroimaging variables contained in the analysis were determined via ADNI according to the literature. The specific measures are ventricle, hippocampal, whole brain, entorhinal, fusiform, and medial temporal lobe volume.

Initially, our analysis sought to contain more neuroimaging variables; however, much of the wider ADNI neuroimaging data available is either not for ADNI2, did not contain enough participants or did not contain the measures we needed to address our hypothesis. Furthermore, when trying to assimilate new neuroimaging data into our file, there was too much missing data to warrant adding new neuroimaging variables. The problems with missing data are discussed further in Section 2.5.2 below.

## 2.5 Data Pre-processing and Population

The ADNI data is broad. It contains approximately 100 different variables across all participants at every temporal stage of all ADNI cohorts. Due to the broad and general nature of the dataset, the data needed to be cleaned to fit our research project. We first removed all participants from all cohorts outside of the ADNI2 population. ADNI2 was the best population available, as it contains the most recent measures of the completed ADNI studies. Next, we removed all participants who originated from initiatives outside of the ADNI2 cohort. The varying ADNI initiatives used different measures and variables that are not interchangeable across cohorts, thereby limiting comparisons. Subsequently, this study removed participants who had baseline measures before the initiation of ADNI2, as their measures differed from those of new subjects. Other studies have attempted to compare variations in ADNI measures across initiatives and have found inconsistency. Consequently, we thought it best to control for ADNI cohort variations by only using participants originating in ADNI2.

After we had defined our population, the two conversion groups were created using a filter command in SPSS. Participants were extracted into the conversion group if they were diagnosed with MCI at baseline and developed AD at any yearly reporting session after (Select participant if [Base Line Diagnosis ~= AD & Diagnosis = AD]). Participants who remained stable were selected if they were diagnosed with MCI at baseline and remained with MCI throughout the study (Select participant if [Base Line Diagnosis = EMCI | Base Line Diagnosis = LMCI] & [Diagnosis = MCI]). For some unknown reason, the ADNIMERGE dataset reports MCI participants as EMCI or LMCI at baseline; during all follow up sessions, however, it only reports these participants as having MCI. Thus, in the analysis, we had to treat all EMCI and LMCI participants as MCI due to measurement changes made by ADNI. Next, because this filter only pulled out participants' data in the year

they converted or were stable, we extracted participants' data in each group using their ID number and cross-referenced the lists to make sure there was no participant overlap in our groups. Finally, the data was filtered by year because of participant mortality and for the analysis. As the study progressed, participants increasingly missed annual recordings, so the data was moderated by yearly intervals so that participant mortality would not confound conversion statistics over time. For a more detailed example of our data pre-processing procedure, see Gomar et al. (2011 & 2014).

Upon completion of the data cleaning, we merged the ADNI composite memory score from another file. The merge data sets command in SPSS was used to include the composite memory variable and match it to our participant rosters at the appropriate annual recording intervals. It was essential to include ADNIMEM to test whether a combined episodic-memory score could better predict AD than individual cognitive tests, as seen in (Crane et al., 2012). Thus, inclusion of the composite memory score was necessary for completing our research goal, which was to assess the viability and predictive ability of all ADNI episodic-memory measures. Other potential biomarker and neuroimaging measures in the ADNI database were cross-referenced with our sample to see if we could include more variables; however, none both, applied to the ADNI2 dataset and did not have severe levels of missing data. Once our population was defined and our dataset was sufficiently cleaned, we started the preliminary analysis.

## 2.6 Preliminary Analyses

### 2.6.1 Sample Size and Descriptive Statistics

Before conducting the initial analysis, descriptive statistics and a G*Power analysis were run so we could better understand our population and required sample size. Specifically, we used measures of diagnosis, participant counts, and group sizes to better understand our sample. Furthermore, descriptive statistics were run on participants' age, sex and education to help us understand our participant sample compared to other studies. In total, our analysis contained five descriptive, ten episodic-memory, and six neuroimaging variables. A sample size analysis was run using the software, G*Power, developed by the University of Düsseldorf, Germany (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007). Since we are using logistic regression, the z-test *a priori* settings were used to compute a recommended sample size from an alpha error probability of 0.05, a statistical power of 0.95, a conversion probability of 0.2 and a populations parameter of 0.45 in a binomial distribution. The recommended sample size was 37 participants, which we exceeded in all years of ADNI2 besides year five. The low sample size cut off occurred primarily because we were testing the dichotomous group fit, which is easy to compute due to the limited outcomes. It is important to report, for replication's sake, that our G*Power analysis resulted in a critical z of -1.96 and an actual power of 0.953. As previously mentioned, the final year of ADNI2 (year five) did not reach the recommended sample size and was therefore excluded from the analysis. Accordingly, in total, our analysis contained 307 participants with 927 cases recorded across the first four years of ADNI2.

## 2.6.2 Missing Data

Early on in the draft analysis, we discovered a significant portion of participants' data had been omitted due to data missing across multiple variables. We ran a missing values pattern analysis to visualise the data and inform the best way to fix the problem of missing data. Usually, a study would omit missing values if under five per cent; however, this was not a comprehensive solution. All of the episodic-memory measures contain approximately 2% missing data, which means that we could omit the missing data. Figure 1 below shows the percentage of missing data relative to all participants and variables. While there was only a small number of missing values, the study kept filtering out participant data in a listwise fashion. Incorrectly, if a participant was missing a single entry for a variable, the analysis filtered him or her out of the study.

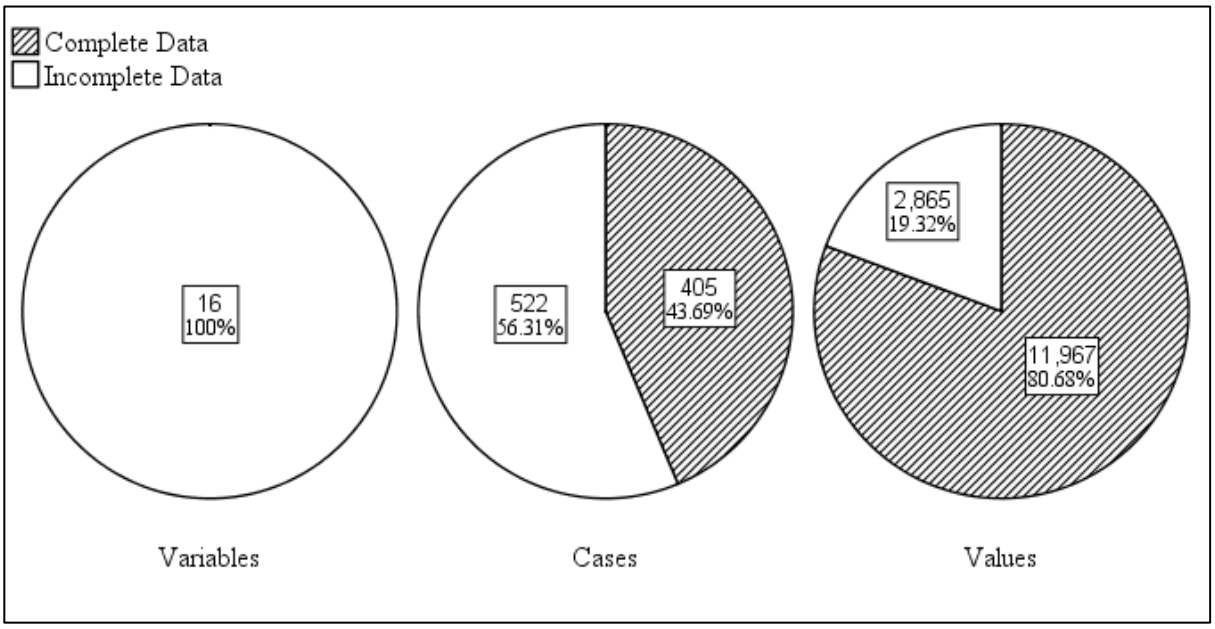



*Figure 1.* Visual representation of missing values across ADNI2 where: Variable = number of missing variables, Cases = percentage of participants with missing data, Values = amount of missing data across all variables and participants

Upon further investigation, we determined that there was a severe amount of missing neuroimaging data in years three and four of ADNI2. To be precise, an average of 85% of all neuroimaging data was missing from these two years. Comparatively, an average of 20% of participants' neuroimaging data was missing in years one and two of the ADNI2 cohort. As

the missing data was so large for the latter two years of ADNI2, there was no option but to remove the neuroimaging data for those years and to study neuroimaging markers only in the first two years of the ADNI2 cohort. The 20% of missing data in the first years was fixed using multiple imputations.

**Table 1**

*Missing Data in Years One and Two of ADNI2*

| *Variables* | *Missing* | | *Valid N* | *Mean* | *SD* |
|---|---|---|---|---|---|
| | *N* | *Per cent* | | | |
| *UCSF Med Temp* | 150 | 26.7% | 412 | 19476.70 | 3193.474 |
| *UCSF Fusiform* | 150 | 26.7% | 412 | 17669.77 | 3287.481 |
| *UCSF Entorhinal* | 150 | 26.7% | 412 | 3397.98 | 888.682 |
| *UCSF Ventricles* | 114 | 20.3% | 448 | 40511.12 | 25597.710 |
| *UCSF Whole Brain* | 79 | 14.1% | 483 | 1031415.44 | 114270.456 |
| *UCSF Hippocampus* | 74 | 13.2% | 488 | 6575.64 | 1412.750 |

*Note*. SD = Standard deviation, N = Participant cases

**Table 2**

*Missing Data in Years Three and Four of ADNI2*

| *Variables* | *Missing* | | *Valid N* | *Mean* | *SD* |
|---|---|---|---|---|---|
| | *N* | *Per cent* | | | |
| *UCSF Med Temp* | 322 | 88.2% | 43 | 19632.00 | 3792.481 |
| *UCSF Fusiform* | 322 | 88.2% | 43 | 17541.09 | 3637.192 |
| *UCSF Entorhinal* | 322 | 88.2% | 43 | 3328.56 | 868.462 |
| *UCSF Ventricles* | 320 | 87.7% | 45 | 42865.89 | 24493.282 |
| *UCSF Whole Brain* | 315 | 86.3% | 50 | 1018675.72 | 126639.129 |
| *UCSF Hippocampus* | 308 | 84.4% | 57 | 6474.85 | 1409.024 |

*Note*. SD = Standard deviation, N = Participant cases

To remedy the missing values, we used multiple imputation corrections to infer new values from the existing data (Manly & Wells, 2015). Multiple imputations work by computing probable values for missing data based on the existing patterns of variables in the dataset and a random number generator. The imputations for each variable are calculated multiple times orthogonally from each other, then pooled together and averaged to get the final imputed statistic. Using SPSS, multiple imputations were formulated using the Mersenne Twister random number generator and the monotonic method set to six imputations. We followed the guidelines and methodology for running multiple imputations suggested by Manly and Wells (2015) and Sterne et al., (2009). All variables were used to

calculate multiple imputations; however, in the analysis, imputations were used only for neuroimaging variables, as episodic-memory variables were missing only 2% of data. As a general rule, multiple imputations work best for variables which are missing less than 50% of their data. As the two years of neuroimaging data was missing only approximately 20%, this rule of thumb was met. Originally, we wanted to use CSF biomarkers (Aβ and Tau) in this analysis; however, too much data was missing (approximately 90%) to run multiple imputations and use the variables.

We hypothesize that missing data has resulted from problems with participant dropout, participant comfort and the flexibility of our data file. First, with a neurologically vulnerable population, it is understandable to have a gradual drop off of missing data as observed across the years of our ADNI sample. Second, uncomfortable and invasive procedures such as MRI or the lumbar punctures used to obtain participant information work on an opt-in/opt-out approach. Participants can choose when they want to partake in research and the attached medical procedures due to the ethical rules put in place to protect participant consent and wellbeing. Subsequently, when participants can choose which medical procedures to undertake, there is a chance to obtain less data. For example, it is not surprising to see that CSF data has the most missing values, as it is the most invasive of the procedures and would deter the most participants. Lastly, we hypothesize that a large section of the missing data comes from our manipulation of the original data file. It is apparent that the ADNI data was not collected explicitly for our study and accordingly does not fit perfectly; however, there is still ample data available with which to address our research questions, and without the help of ADNI, this project would not be possible.

## 2.7 Analysis Outline

The analysis contained four parts, one for each year of ADNI2 studied. Each year individually assessed all three research questions derived from the aims specified in Section 1.5 (p.23). Initially, we wanted to include all five years of ADNI2; however, the last year did not have enough participants to warrant analysis when accounting for participant drop-out and missing data. For the first research question, we sought to evaluate the construct validity of general AD diagnostic tests to see if they could measure episodic memory. For the second research question, we wanted to test which episodic-memory measure best predicted MCI to AD conversion and to determine whether a mixed episodic-memory model could predict conversion better than individual tests. Finally, we sought to create a mixed predictor model of MCI to AD conversion using both episodic memory and neuroimaging variables. Specifically, we sought to test if a mixed predictor model could predict MCI to AD conversion better than prior episodic-memory models and compare the predictive ability of neuroimaging measures to episodic-memory tests.

### 2.7.1 Research Question One Outline

*Research Question One: Can general cognitive assessment tests for AD, such as the MMSE and ADAS-Cog, measure episodic memory?*

Before conducting the majority of the analyses and evaluating the tests, we wanted to double check the construct validity of our episodic-memory measures. Specifically, in the literature, it is theorised that the standard dementia diagnostic tests such as the ADAS-cog and MMSE can be used to assess episodic memory (Crane et al., 2012; Gomar et al., 2011); however, we wanted to test the construct validity of these tests before assuming that they are episodic-memory tests. To test the construct validity of these AD tests, we ran a factor analysis to see if the general AD assessment tests function using the same underlying factors

as the specialised episodic-memory tests. We specifically used a principal component analysis with a varimax rotation to easily separate variables into potential factors. The results of the factor analysis and subsequent construct validity of the ADAS-cog and MMSE were taken into account when we performed the primary analysis.

### 2.7.2 Research Question Two Outline

*Research Question Two: Which episodic-memory test best predicts MCI to AD conversion, and can a mixed episodic-memory model predict conversion better in each year of the ADNI2 cohort?*

A binary logistic regression was used to formulate a model that could predict MCI to AD conversion in converter and non-converter groups. Usually, a linear regression would be favoured for predictive analyses; however, because this is secondary (previously collected) and longitudinal data, a logistic regression fit best due to its retrospective nature. Moreover, a logistic regression also fit our analysis best, as our dependant variable was a dichotomous categorical variable (disease conversion). Due to the secondary nature of our data, participants could be categorised into conversion groups based on their diagnosis history in the cohort, which means that we could test the accuracy of our predictive models against the reality of participant conversion in the cohort. Accordingly, we inferred the conversion variable from ADNI's measure of participant diagnosis across the four years of the initiative.

The first series of logistic regressions was used to determine the predictive ability of each episodic-memory test. Accordingly, we performed a logistic regression on each episodic-memory variable and tabulated the results. Each episodic-memory variable was ranked in order of predictive ability according to measures of accuracy, odds ratio, accounted variance and significance, which were determined by chi-squared, r-squared and Pearson's p-values. Accuracy was tested for by cross-referencing our model's classification of conversion in participants with their actual conversion in the initiative, as recorded by ADNI. This

analysis ran automatically as part of a logistic regression in IBM SPSS 26. We also used the Hosmer-Lemeshow test of sphericity as an indicator of model fit. While researchers debate the usefulness of the Hosmer-Lemeshow test in logistic regressions, we included the test as an indicator of model fit and did not use it as a statistical cut off for our analysis. We evaluated all individual episodic-memory predictors for predictive ability by using binary logistic regressions in all four years of ADNI2. Once the individual episodic-memory variables had their predictive ability determined, we put them together to formulate a mixed episodic-memory predictive model of MCI to AD conversion.

To create the best mixed model, we expanded upon the binary logistic regression design used for the individual predictors above. A block-wise hierarchical binary logistic regression was used to find the best predictive model. In the literature, multiple studies use stepwise regression methods to construct mixed predictive episodic-memory models, but these methods are often open to suppression effects that create preference biases for variables. To avoid suppression effects, we used a forced entry hierarchical method that enters variables individually in blocks and in order of importance, which is determined by the researcher. The block hierarchical method allows predictors that have historically performed best to go first while also lessening the suppression effects of latter variables by adding them in individually and in accord with their importance. We entered the ADNIMEM score first, as composite memory scores are theorised to be better than individual episodic-memory scores in AD research. Next, we gradually entered the RAVLT and LMT episodic-memory tests due to their reputation and prominence in the literature. Finally, we added in the ADAS-cog and MMSE tests, as they are general assessment tests that do not explicitly test episodic memory but are theorised to do so. Variables in the mixed model were removed or kept depending on significance, chi-squared, accuracy and R-squared statistics, as was done previously. To be counted the best predictive model, the model must have had the best balance of accounted

variance and accuracy while containing only significant predictors. It is important to note that these episodic-memory mixed models were run without imputations, as they can lead to high correlations between variables and effect multicollinearity.

It was imperative to check the assumptions of a binary logistic regression before interpreting our results. All of the assumptions—dichotomous dependent variable, multiple continuous or categorical independent variables and independence of observations—were passed by our basic research design and methodology. The assumption of the linearity of the logit was tested by creating log versions of our variables and testing the interaction between the original variables and its log form in logistic regression. If the results of the interaction were nonsignificant in the regression, then the assumption was met. We did not assess the assumption of multicollinearity during the individual predictor's stage, as only one variable was studied at a time; however, we assessed the assumption of multicollinearity during the mixed model analysis by running collinearity tests contained in linear regression analysis in SPSS. Multicollinearity was assessed using tests of tolerance and variance inflation factor (VIF). When assessing multicollinearity, to meet the assumption of multicollinearity, the general rule is that tolerance should be above 0.1 and VIF less than 10. The results of the assumption checks dictated the analysis and interpretation of our results.

### 2.7.3 Research Question Three Outline

*Research Question Three: Which neuroimaging markers best predict MCI to AD conversion, and can a mixed model predict conversion better than prior models in each year of the ADNI2 cohort?*

For the third research question, we ran an analysis similar to the one we ran for Research Question 2; however, this time, the binary logistics regression contained the neuroimaging variables. We first analysed each predictor individually for predictive ability by using a binary logistic regression. We then checked the assumption of the logit with log

forms of each variable, as in Research Question 2. Next, we created a mixed model using a hierarchical binary logistic regression. This time variables were entered in blocks of two, organised by the type of variable and its importance to the literature. It was not possible to enter variables individually as done in the mixed episodic-memory analysis, as there were too many variables that would have made the analysis cumbersome and unwieldy. The software IBM SPSS statistics also does not allow for the individual addition of so many variables during a logistic regression; thus, we were forced to use a multiple variable block-wise method. We used the individual predictive ability of each variable to dictate the order within blocks and variables history in the literature. Accordingly, episodic-memory variables were entered in first followed by neuroimaging measures. Variables were once again removed or kept in the model depending on their R-squared measure, significance and accuracy in predicting disease conversion. We then checked the assumptions of multicollinearity and linearity of the logit on the best predictive model.

Following the completion of the assumption checks, the best mixed predictor model was compared with the best individual and mixed episodic-memory models in the appropriate year to determine which measure best predicted MCI to AD conversion overall. It is important to note that all of the variables in the mixed model analysis, except for the episodic-memory variables, contained multiple imputations. Though the use of multiple imputation corrections for data is permissible and widely accepted in the literature, we were careful in interpreting the results; however, not using any statistical corrections would have resulted a worse circumstance with mostly biased results. The interpretation of multiple imputations is discussed further in Chapter 4 (p.72).

In conclusion, this chapter detailed the sample population, variables, data, and statistical methods used in this study. Firstly, the ADNI2 sample and MCI-converter and MCI non-converter groups were discussed. Next, we discussed the ADNI database and

ADNIMERGE dataset, including a list of episodic memory and neuroimaging variables. Subsequently, the G*Power analysis, pattern analysis, multiple imputation corrections, and problems with missing data were also detailed. Lastly, the statistical analysis and methodology were outlined, and research questions were developed from the aims detailed in Chapter 1.

# Chapter Three: Results

## 3.1 Descriptive Statistics

This study investigates the predictive ability of episodic-memory measures in the ADNI database to inform current early-stage AD theory and detection methods. Specifically, this study compares general-multidomain tests and specific episodic-memory tests. Moreover, this study also predicts MCI to AD conversion using various episodic-memory and neuroimaging markers.

**Table 3**
*Participant Count, Age, and Education Descriptive Statistics*

| ADNI2 Year | Variable | N | Minimum | Maximum | Mean | SD |
|---|---|---|---|---|---|---|
| 1 | Age | 301 | 55.0 | 91.4 | 71.693 | 7.3329 |
| | Education | 301 | 9 | 20 | 16.32 | 2.654 |
| 2 | Age | 261 | 55.0 | 91.4 | 71.372 | 7.4096 |
| | Education | 261 | 9 | 20 | 16.28 | 2.648 |
| 3 | Age | 214 | 55.0 | 88.6 | 71.217 | 7.1362 |
| | Education | 214 | 9 | 20 | 16.28 | 2.653 |
| 4 | Age | 151 | 55.0 | 88.4 | 70.621 | 6.7987 |
| | Education | 151 | 12 | 20 | 16.11 | 2.551 |
| 5 | Age | 27 | 55.0 | 82.2 | 68.878 | 7.4641 |
| | Education | 27 | 12 | 20 | 16.26 | 2.566 |

*Note*. SD = Standard deviation, N = Participant count

Before the primary analysis, descriptive statistics and frequency analyses were run to understand our ADNI2 sample population. Table 3 above summarises the age and educational statistics of all participants organised by each year of ADNI2. Each year of ADNI2 in the analysis contained participants with similar age distributions and educational values. Specifically, in the five years of the ADNI2 cohort, participants had an approximate mean age of 70, an age range of 55-91 and a standard deviation of seven years. Participants' educational levels remained constant with a mean of 16, a range of 9-20 and a standard

deviation of 2.6 across all four years of ADNI2. This meant that, although the ADNI2 cohort had gradual participant drop out, our sample stayed relatively consistent.

Our study contained 307 participants (N = 307) across the first four years of ADNI2. However, there was a gradual drop off of participants in our sample, with year one containing 301 participants, year two containing 261 participants, year three containing 214 participants, year four containing 151 participants, and year five containing 27 participants. Accordingly, year five was the only sample that we could not study, as the subject count did not reach the 37 participants recommended by the G*Power analysis (see Section 2.6.1 p.36). In our sample, MCI participants remained the largest group throughout, consisting of approximately 74% of all participants. Accordingly, AD participants made up approximately 26% of our sample (see Table 4 below).

**Table 4**

*Participant Diagnosis Across ADNI2*

| Year | Diagnosis | Frequency | Percent |
|---|---|---|---|
| 1 | AD | 36 | 12.0 |
| | MCI | 265 | 88.0 |
| | Total | 301 | 100.0 |
| 2 | AD | 59 | 22.6 |
| | MCI | 199 | 76.2 |
| | Total | 261 | 100.0 |
| 3 | AD | 56 | 26.2 |
| | MCI | 153 | 71.5 |
| | Total | 214 | 100.0 |
| 4 | AD | 41 | 39.4 |
| | MCI | 104 | 60.6 |
| | Total | 145 | 100.0 |

*Note*. Year five was removed due to lack of participants

From the diagnosis statistics observed across ADNI2, we were able to determine participant conversion and group participants. The AD conversion and non-conversion groups were pivotal in testing our model's accuracy in comparison to the observed conversion in ADNI2. Because our sample consisted of only MCI participants at baseline, the conversion statistics were the same as the percentage of AD participants. Thus, approximately 20-25% of our

participants converted from MCI to AD, which is in agreement with the MCI to AD conversion rate observed across the literature (Mitchell & Shiri-Feshki, 2009). Fewer participants converted to AD in the first year of ADNI2 (12%), which is to be expected, as AD progression is exponentially related to age. On average, 48 participants converted to AD from MCI in each year of ADNI2, and 180 remained stable with MCI.

It is also important to note that, in the ADNIMERGE datafile, all MCI participants were split into EMCI and LMCI groups at baseline but only categorised as MCI participants in follow-up sessions. It was therefore impossible to analyse the different forms of MCI, as EMCI and LMCI were not distinguished in the dataset. Consequentially all instances of LMCI and EMCI were treated as purely MCI. The same scenario occurred for SMC participants who became normal controls. The operationalisation of MCI participants is further discussed in Section 2.5 above (p. 34).

Lastly, before progressing to the main analysis, it is important to explore the assumption that episodic-memory declines over time in AD participants. If episodic memory does not decline over time in our population, we cannot use episodic-memory measures to predict the disease. Accordingly, we ran descriptive statistics for each cognitive test and tabulated the means across the four years of ADNI2 in MCI converters and non-converters. Specifically, standardised averages (Z-scores) were used to compare the episodic-memory tests, as various tests had different base scales that initially hindered comparison. Furthermore, note that higher scores on the ADAS 11, 13, Q4, and RAVLT percent forgetting tests indicate poor memory performance while lower scores on the MMSE, LMT, ADNIMEM and all other RAVLT measures indicate poor memory performance. Accordingly, cognitive ability worsened across all episodic-memory measures across the four years of ADNI2 (see Figure 2 below).

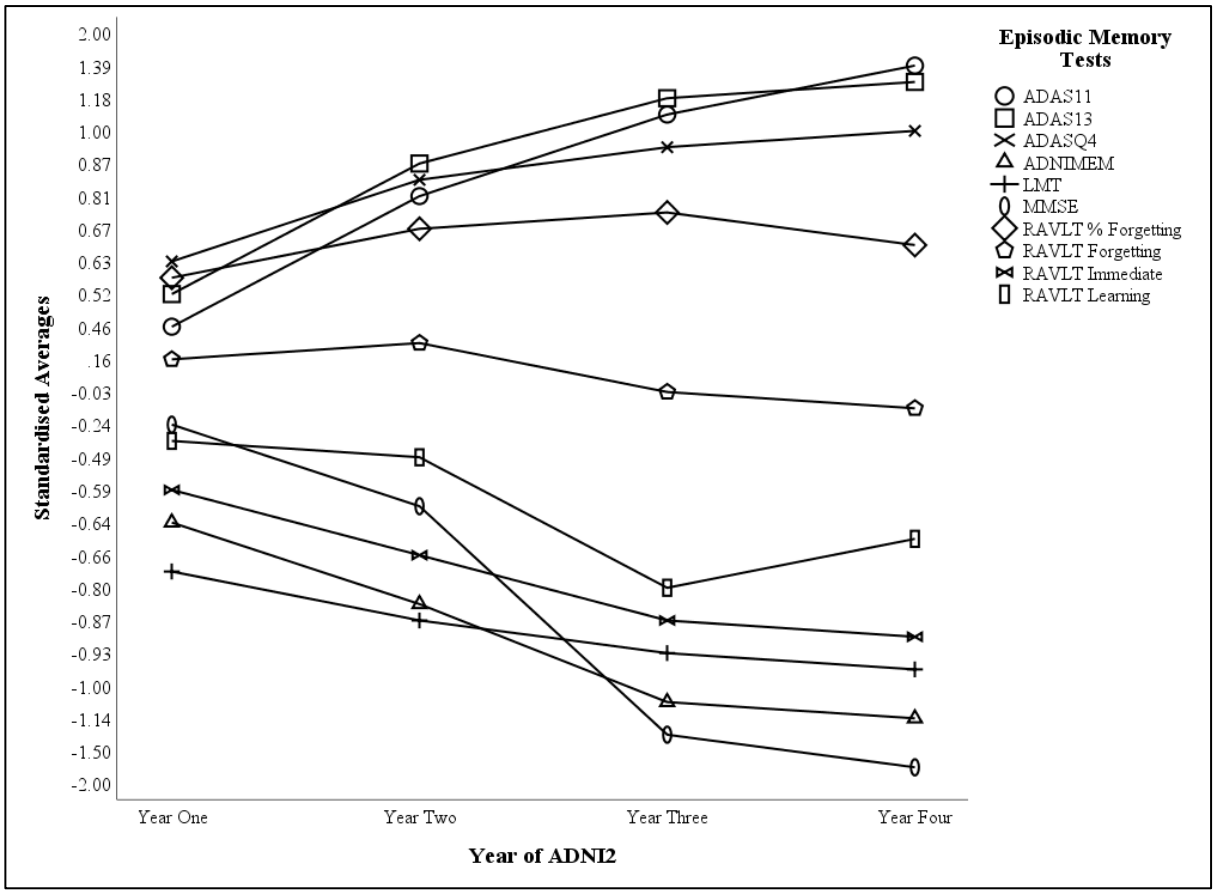



*Figure 2.* Standardised averages of episodic-memory scores in MCI converting participants across the first four years of ADNI2.

Comparatively, stable MCI participants remained relatively the same, which is to be expected (see Figure 3, p.50). There was little variation is scores over time, with most standardised averages varying only within a hundredth of their initial score in the first year of ADNI2. Therefore, it is safe to assume that episodic memory worsens over time in AD participants (MCI converters) and remains comparatively stable in MCI non-converters. Furthermore, the observed patterns in episodic-memory decline in this study align with those observed in the literature.

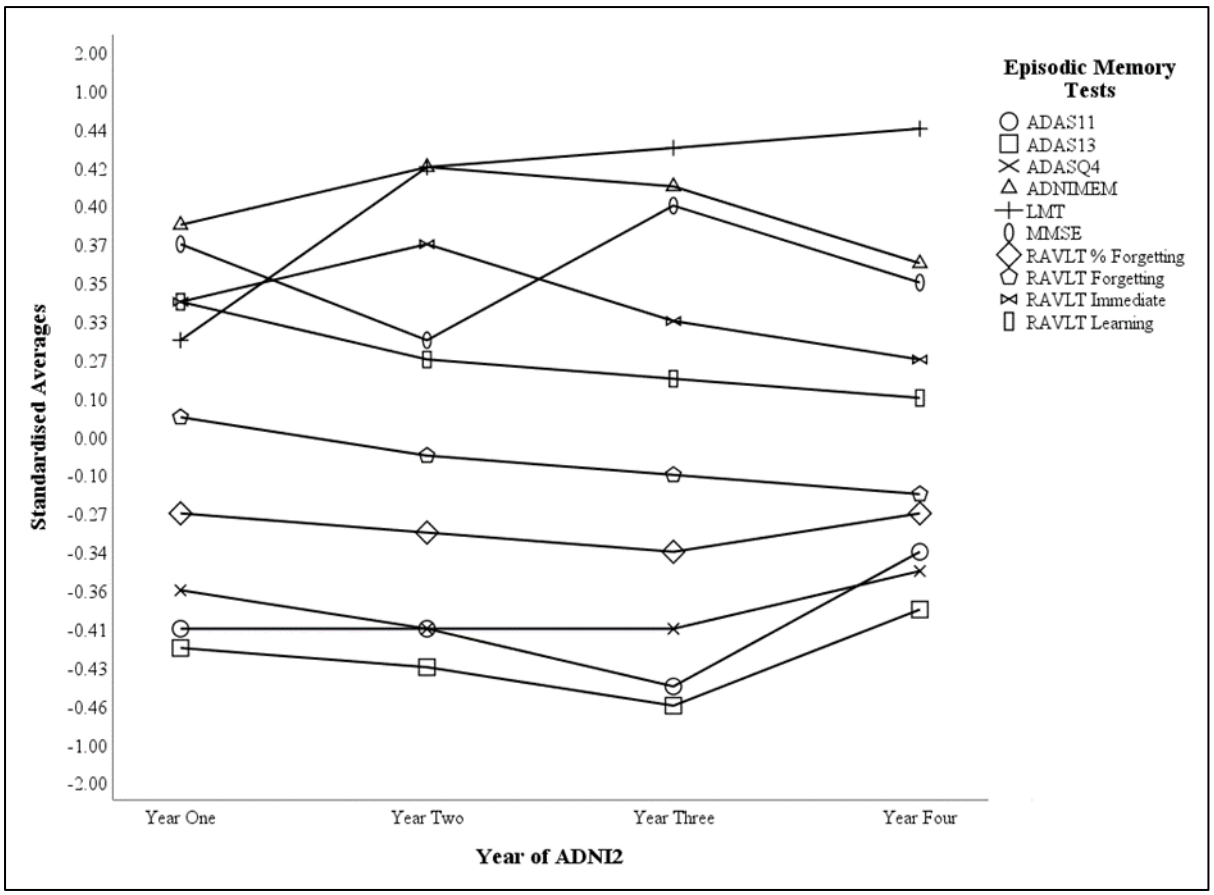



*Figure 3.* Standardised averages of episodic-memory scores in MCI stable participants across the first four years of ADNI2.

## 3.2 Research Question One

A principal component factor analysis was used to test the assumption that general AD diagnostic tests can measure episodic memory. Specifically, the factor analysis assessed construct validity between episodic-memory tests and general AD tests in hope that they would all measure episodic memory in AD. Ten episodic-memory measures were assessed originating from four main groups of tests: the RAVLT, MMSE, LMT and ADAS-cog (see Section 2.2 for more information). Before running the full analysis, the Kaiser-Myer-Olkin (KMO) Measure of Sampling Adequacy and Bartlett's Test of Sphericity were used to check the viability of running a factor analysis on these variables.

**Table 5**
*Principal Component Analysis ADNI2 Year One*

| | *Loadings* | | |
|---|---|---|---|
| *EM Measures* | *Factor One:* | *Factor Two:* | *Communality* |
| *ADAS13* | **-.904** | .147 | .838 |
| *ADNI_MEM* | **.928** | -.225 | .912 |
| *ADAS11* | **-.843** | .048 | .712 |
| *RAVLT Immediate* | **.802** | -.019 | .643 |
| *ADASQ4* | **-.718** | .415 | .689 |
| *MMSE* | **.665** | .062 | .446 |
| *LMT* | **.549** | -.486 | .538 |
| *RAVLT Learning* | **.606** | -.133 | .385 |
| *RAVLT Forgetting* | .286 | **.900** | .892 |
| *RAVLT % Forgetting* | -.438 | **.840** | .898 |
| *Eigenvalue* | 4.932 | 2.022 | |
| *% of Total variance* | 49.316 | 20.219 | |
| *Total Variance* | | 69.534% | |

*Note*. Only MCI converter cases were used in the principal component analysis

In year one of ADNI2, we were cleared to run our analysis because Bartlett's test was significant [p= <.001] and because KMO = .672, which is above the required minimum of 0.6. A varimax rotation was used to clarify factor loadings, as there was some overlap in factors. Our principal component analysis resulted in two factors being extracted according to an eigenvalue of one, which resulted in an explained total variance of 70% amongst all variables (see Table 5 above). Factor 1 confirmed our assumption that the general AD tests

were highly related to the episodic-memory tests and could independently measure episodic memory. Factor 2 appeared only to contain the RAVLT forgetting measures, which is understandable, as they are both variants of the same measure. Accordingly, in the first year of ADNI2, all cognitive measures, including standard tests, could measure episodic memory.

**Table 6**
*Principal Component Analysis ADNI2 Year Two*

| | *Loadings* | | |
|---|---|---|---|
| *EM Measures* | *Factor One:* | *Factor Two:* | *Communality* |
| *ADAS13* | **-.895** | .244 | .861 |
| *ADNI_MEM* | **.904** | -.300 | .908 |
| *ADAS11* | **-.864** | .139 | .766 |
| *RAVLT Immediate* | **.902** | -.020 | .813 |
| *ADASQ4* | **-.613** | .601 | .737 |
| *MMSE* | **.717** | -.126 | .530 |
| *LMT* | .496 | **-.630** | .643 |
| *RAVLT Learning* | **.663** | -.137 | .459 |
| *RAVLT Forgetting* | .657 | **.682** | .896 |
| *RAVLT % Forgetting* | -.141 | **.909** | .846 |
| *Eigenvalue* | 5.205 | 2.253 | |
| *% of Total variance* | 52.054 | 22.534 | |
| *Total Variance* | | 74.588% | |

*Note*. Only MCI converter cases were used in the principal component analysis

The second year of ADNI yielded results similar to those obtained the first year. We were cleared to run our analysis, as the KMO test was above the 0.6 threshold [KMO = .740] and because Bartlett's test was significant [p= <.001]. Once again, a varimax rotation was used to separate factors into more distinct groups, and an eigenvalue of one was used to define the number of factors. The principal component analysis resulted in an explained total variance of 75% (see Table 6 above). As in the first year, Factor 1 confirmed our assumption that the general AD tests were highly related to the episodic-memory tests and could independently measure episodic memory. Again as in the first year, Factor 2 appeared to contain the RAVLT forgetting measures due to their similarity but also contained the LMT;

however, the LMT and RAVLT forgetting measures were still highly prominent in the first factor. Accordingly, the analysis once again suggested that all tests were related enough to measure episodic memory due to sharing the same primary factor.

**Table 7**
*Principal Component Analysis ADNI2 Year Three*

| | *Loadings* | | |
|---|---|---|---|
| *EM Measures* | *Factor One:* | *Factor Two:* | *Communality* |
| *ADAS13* | **-.849** | .394 | .877 |
| *ADNI_MEM* | **.770** | -.596 | .947 |
| *ADAS11* | **-.863** | .293 | .831 |
| *RAVLT Immediate* | **.857** | -.311 | .831 |
| *ADASQ4* | -.280 | **.840** | .785 |
| *MMSE* | **.819** | -.235 | .726 |
| *LMT* | .171 | **-.903** | .845 |
| *RAVLT Learning* | .235 | -.**508** | .313 |
| *RAVLT Forgetting* | **.783** | .317 | .714 |
| *RAVLT % Forgetting* | -.023 | **.925** | .857 |
| *Eigenvalue* | 4.241 | 3.485 | |
| *% of Total variance* | 42.409 | 34.852 | |
| *Total Variance* | | 77.261% | |

*Note*. Only MCI converter cases were used in the principal component analysis

In the third year of ADNI2, the principle components analysis contained two factors similar to those observed in prior years; however, in the third year, the factors had larger differences in values. The third-year analysis used the same rotation and eigenvalue for the extraction, but we had to limit the analysis to two factors. The factors were limited to two, as the analysis was extracting a third factor that was relatively useless and only just over the eigenvalue cut-off. As in prior years, the KMO [KMO = .711] and Bartlett's Test of Sphericity [p= <.001] indicated that the data met the assumptions of a principal component analysis. The two factors explained a total variance of 77% but were more evenly distributed compared to prior years (see Table 7 above). Factor 1 contained the general tests and the RAVLT Immediate and forgetting. Factor 2 was comprised of all the delayed recall measures and contained the ADASQ4, LMT, RAVLT Learning and RAVLT percent forgetting. While

there was still an overlap between specific episodic-memory tests and general standard tests, it was strange to see a greater divide later on in ADNI2. Accordingly, the divide between tests may provide evidence for the theory that different tests optimally measure episodic memory during different stages of AD progression.

**Table 8**

*Principal Component Analysis ADNI2 Year Four*

| | *Loadings* | | |
|---|---|---|---|
| *EM Measures* | *Factor One:* | *Factor Two:* | *Communality* |
| *ADAS13* | **-.892** | .381 | .941 |
| *ADNI_MEM* | **.877** | -.407 | .935 |
| *ADAS11* | **-.882** | .299 | .867 |
| *RAVLT Immediate* | **.921** | -.149 | .869 |
| *ADASQ4* | **-.660** | .622 | .822 |
| *MMSE* | **.859** | -.134 | .756 |
| *LMT* | .384 | **-.759** | .723 |
| *RAVLT Learning* | **.448** | -.227 | .252 |
| *RAVLT Forgetting* | **.726** | .602 | .890 |
| *RAVLT % Forgetting* | -.183 | **.916** | .872 |
| *Eigenvalue* | 5.273 | 2.655 | |
| *% of Total variance* | 52.727 | 26.550 | |
| *Total Variance* | | 79.277% | |

*Note*. Only MCI converter cases were used in the principal component analysis

In the final year of ADNI2, our analysis returned to the two factor loadings observed in the first two years. A KMO of .790 and a significant Bartlett's test [p= <.001] indicated that we were clear to run our analysis. In the fourth year, the analysis found two factors, which means that there was no need to restrict the factor loadings as in the previous year. The two factors explained a total variance of 79%, with factor one accounting for the most variance (52%) (see Table 8 above). All measures were related in Factor 1 with the exception of the RAVLT percentage forgetting score. Factor 2 contained the delayed recall tests like the third year, yet in year four, the factor only had a small contribution. Thus, all tests appear to measure the same episodic memory construct.

In conclusion, across all four years studied in ADNI2, general memory tests and specific episodic-memory tests were related enough to assume that general memory tests can

assess episodic memory. This means that the rest of the analysis can investigate how well all these tests can measure episodic memory in MCI and AD. Accordingly, general and specific tests can also be compared for their ability to predict MCI to AD conversion and thus early-stage AD.

## 3.3 Research Question Two

A binary logistic regression was run on each episodic-memory predictor in each year of ADNI2 to determine their sole predictive abilities. Following the analysis of individual predictors, all episodic-memory measures were combined by using a hierarchical binary logistic model to form a mixed predictive model of MCI to AD conversion.

**Table 9**

*Individual Episodic Memory Predictors of MCI to AD Conversion in Year One of ADNI2*

| | | | | | | *95% C.I. for Exp(B)* | |
|---|---|---|---|---|---|---|---|
| *EM Measures* | *B* | *S.E.* | *Wald* | *Sig.* | *Exp(B)* | *Lower* | *Upper* |
| *ADNI_Mem* | -1.950 | .096 | 412.857 | .000 | .142 | .118 | .172 |
| *RAVLT Forgetting* | .058 | .019 | 9.308 | .002 | 1.059 | 1.021 | 1.099 |
| *RAVLT Immediate* | -.122 | .007 | 331.591 | .000 | .885 | .873 | .897 |
| *Logical Memory Test* | -.310 | .015 | 433.988 | .000 | .733 | .712 | .755 |
| *RAVLT Percent Forgetting* | .030 | .002 | 267.456 | .000 | 1.031 | 1.027 | 1.034 |
| *RAVLT Learning* | -.371 | .023 | 256.490 | .000 | .690 | .660 | .722 |
| *ADAS 13* | .177 | .009 | 423.350 | .000 | 1.193 | 1.173 | 1.213 |
| *MMSE* | -.381 | .023 | 271.007 | .000 | .683 | .653 | .715 |
| *ADAS Q4* | .450 | .023 | 386.623 | .000 | 1.569 | 1.500 | 1.641 |
| *ADAS 11* | .237 | .012 | 382.434 | .000 | 1.267 | 1.238 | 1.298 |

*Note*. Highlighted variables violated assumption checks and could not be interpreted
B = Regression coefficient used to calculate Exp(B)
Exp(B) = predictive odds ratio

### 3.3.1 ADNI2 Year One

In the first year of ADNI2, all of the episodic-memory variables could significantly predict MCI to AD conversion except for the RAVLT Forgetting. The predictive odds, significance and confidence intervals of each individual episodic-memory measure is detailed in Table 9 above. Initially, the ADASQ4 had an odds ratio of Exp(β) = 1.569 and was positively predicting conversion with better odds than any other variable [$\beta = 0.450$, $z = 386.623$, $p < .001$]; however, the LMT, ADASQ4, RAVLT Forgetting and RAVLT Percent Forgetting all violated the assumptions of a binary logistic regression. Therefore, out of the interpretable results, the ADAS11 had the best predictive odds [Exp(β) = 1.267]. The odds of

the ADAS11 were followed by the ADAS13 [Exp(β) = 1.193], RAVLT Immediate [Exp(β) = 0.690], and RAVLT Learning [Exp(β) = 0.683].

**Table 10**

*Accuracy and Accounted Variance of Individual Episodic Memory Predictors in Year One of ADNI2*

| *EM Measures* | *C&S $R^2$* | *N-$R^2$* | *Chi-squared* | *Sig.* | *Conversion predictive Acc.* | *Model Accuracy* |
|---|---|---|---|---|---|---|
| *ADNI_Mem* | .286 | .402 | 709.297 | .000 | 56.4 | 78.1 |
| *RAVLT Forgetting* | .004 | .006 | 9.372 | .002 | 0.0 | 68.8 |
| *RAVLT Immediate* | .227 | .319 | 541.876 | .000 | 41.4 | 72.4 |
| *Logical Memory Test* | .269 | .378 | 658.881 | .000 | 60.6 | 77.7 |
| *RAVLT Percent Forgetting* | .154 | .217 | 352.418 | .000 | 63.8 | 71.7 |
| *RAVLT Learning* | .142 | .200 | 322.400 | .000 | 38.1 | 75.0 |
| *ADAS 13* | .274 | .385 | 672.581 | .000 | 50.2 | 77.5 |
| *MMSE* | .149 | .210 | 340.442 | .000 | 23.4 | 71.4 |
| *ADAS Q4* | .226 | .318 | 540.637 | .000 | 53.2 | 75.7 |
| *ADAS 11* | .247 | .348 | 598.100 | .000 | 42.6 | 77.4 |

*Note*. Highlighted variables violated assumption checks and could not be interpreted
*C&S $R^2$ = Cox & Snell R-squared*
N-$R^2$ = *Nagelkerke R-squared*

After taking the odds ratios into account, the best individual predictive episodic-memory measure was determined according to accuracy, chi-squared, accounted variance and significance, as seen in Table 10 above. The accounted variance was reported using Cox and Snell's R-squared (C&S $R^2$) and Nagelkerke's R-squared (N-$R^2$) measures. Initially, the specific episodic-memory measures (ADASQ4, LMT etc.) had the best predictive accuracy, yet they had to be omitted because they violated the assumptions of linearity of the logit. Of the remaining measures, the ADNIMEM composite memory score had the best accuracy and accounted variance when predicting MCI to AD conversion [C&S $R^2$ = .286, Nagelkerke $R^2$ = .402, Conv-Acc. = 56.4%, $p < .001$]. The next best individual episodic-memory predictors of MCI to AD conversion were the ADAS13 [Conv-Acc. = 50.2%], ADAS11 [Conv-Acc. = 42.6%], RAVLT Immediate [Conv-Acc. = 41.4%]. Interestingly, the RAVLT Forgetting, which is the most common episodic-memory measure in the literature, could not predict MCI to AD conversion. Accordingly, ADNIMEM was the best individual episodic-memory predictor of MCI to AD conversion in the first year of ADNI2. It should also be noted that

some researchers may opt to use the ADAS13 over the ANIMEM because it has a better odds ratio with a similar accuracy.

**Mixed Episodic-Memory Predictor Model**

Whilst multiple predictors of MCI conversion were found, a true accuracy as fair as 56% and 50% is not good enough for clinical or academic use. We therefore investigated whether a mixed episodic-memory model could predict MCI to AD conversion better than our individual episodic-memory models. We used a hierarchical binary logistic regression to create a mixed episodic-memory predictor model for the first year of ADNI2. The best model consisted of the ADASQ4 and RAVLT Immediate, as seen in Table 11 below. The odds ratios of the mixed model were better than those of the ADNIMEM and the other interpretable individual predictors. The model odds when predicting conversion for the ADASQ4 and RAVLT Immediate were Exp(β) = 0.925 and Exp(β) = 1.324, respectively. It is important to note that, when part of a mixed model, the ADASQ4 did not violate any assumptions of a logistic regression and thus could be used in this stage of the analysis.

**Table 11**

*The Best Mixed Episodic-Memory Model Predicting MCI to AD Conversion in ADNI2 Year One*

| | | | | | | *95% C.I. for Exp(B)* | |
|---|---|---|---|---|---|---|---|
| *EM Measures* | *B* | *S.E.* | *Wald* | *Sig.* | *Exp(B)* | *Lower* | *Upper* |
| *RAVLT Immediate* | -.078 | .008 | 98.489 | .000 | .925 | .911 | .939 |
| *ADAS Q4* | .281 | .028 | 101.285 | .000 | 1.324 | 1.254 | 1.399 |
| *Constant* | .011 | .363 | .001 | .976 | 1.011 | | |

B = Regression coefficient used to calculate Exp(B)
Exp(B) = predictive odds ratio
S.E. = Standard Error

For the predictive accuracy and accounted variance of the mixed episodic-memory predictor model, see Table 12 below. The mixed episodic-memory model [C&S $R^2$ = .266, N-$R^2$ = .310, Conv. Accuracy = 53.1%, $p < .001$] accurately predicted MCI to AD conversion worse than the ADANIMEM individual model (but only marginally worse). On all counts besides conversion accuracy, the mixed model was better than ADNIMEM; however, the

differences between each model was so marginal that the models could be considered equal. Once again, the accuracy of the mixed model was not close to research or clinical standards for diagnostic and predictive tests.

**Table 12**
*Accuracy and Accounted Variance of the Mixed Episodic Memory Model in Year One of ADNI2*

| *EM Measures* | *C&S $R^2$* | *N-$R^2$* | *Chi-squared* | *Sig.* | *Conv Acc.* | *Total Model Acc.* |
|---|---|---|---|---|---|---|
| *Mixed EM Model Y1* | .266 | .374 | 650.207 | .000 | 53.1 | 77.1 |

*C&S $R^2$ = Cox & Snell R-squared*
N-$R^2$ = *Nagelkerke R-squared*

## 3.3.2 ADNI2 Year Two

**Table 13**
*Individual Episodic Memory Predictors of MCI to AD Conversion in Year Two of ADNI2*

| | | | | | | *95% C.I. for Exp(B)* | |
|---|---|---|---|---|---|---|---|
| *EM Measures* | *B* | *S.E.* | *Wald* | *Sig.* | *Exp(B)* | *Lower* | *Upper* |
| *ADNI_Mem* | -2.240 | .112 | 402.790 | .000 | .106 | .086 | .132 |
| *RAVLT Forgetting* | .100 | .019 | 27.591 | .000 | 1.105 | 1.065 | 1.148 |
| *RAVLT Immediate* | -.124 | .007 | 313.531 | .000 | .883 | .871 | .896 |
| *Logical Memory Test* | -.396 | .018 | 460.720 | .000 | .673 | .649 | .698 |
| *RAVLT Percent Forgetting* | .039 | .002 | 277.774 | .000 | 1.039 | 1.035 | 1.044 |
| *RAVLT Learning* | -.362 | .025 | 208.274 | .000 | .696 | .663 | .731 |
| *ADAS 13* | .207 | .010 | 436.523 | .000 | 1.230 | 1.206 | 1.254 |
| *MMSE* | -.427 | .024 | 325.185 | .000 | .652 | .623 | .683 |
| *ADAS Q4* | .564 | .027 | 431.223 | .000 | 1.758 | 1.667 | 1.854 |
| *ADAS 11* | .290 | .014 | 404.279 | .000 | 1.337 | 1.299 | 1.375 |

*Note*. Highlighted variables violated assumption checks and could not be interpreted
B = Regression coefficient used to calculate Exp(B)
Exp(B) = predictive odds ratio
S.E. = Standard Error

In the second year of ADNI2, all individual episodic-memory measures could predict MCI to AD conversion besides RAVLT Forgetting, just as in the first year. However, overall, episodic-memory measures performed better in the second year than they did in the first year. Both the LMT and MMSE violated the assumption of linearity of the logit and could, consequentially, not be interpreted. The ADASQ4 [Exp(β) = 1.758, $\beta$ = .564, $z$ = 431.223, $p$ < .001] had the best odds ratio when predicting MCI to AD conversion out of all individual

episodic-memory measures, as seen in Table 13 above. The ADAS 11 [Exp(β) = 1.337], ADAS13 [Exp(β) = 1.230], and RAVLT Forgetting [Exp(β) = 1.105] followed, with similar increases in disease conversion odds.

When assessing accounted variance and predictive accuracy with the above odds, the best individual predictor was the RAVLT Percent Forgetting [C&S R2 = .213, N- R2 = .297, Conv-Acc. = 71.7%, p < .001]. The RAVLT Percent Forgetting was followed in predictive accuracy by the ADAS 13 [Conv-Acc. = 69.8%], ADASQ4 [Conv-Acc. = 66.3%] and the ADNIMEM [Conv-Acc. = 63.9%] in predictive accuracy. Although the RAVLT Percent Forgetting did have a lover accounted variance compared to most predictors, the ADAS13 had the best balance of odds, accounted variance and accuracy; however, with cognitive tests, a higher predictive accuracy is often preferred; thus, the RAVLT Percent Forgetting would remain the best predictor of MCI to AD conversion in the second year of ADNI2.

**Table 14**

*Accuracy and Accounted Variance of Individual Episodic Memory Predictors in Year Two of ADNI2*

| *EM Measures* | *C&S $R^2$* | *N- $R^2$* | *Chi-squared* | *Sig.* | *Conversion predictive Acc.* | *Model Accuracy* |
|---|---|---|---|---|---|---|
| *ADNI_Mem* | .358 | .501 | 797.755 | .000 | 63.9 | 81.6 |
| *RAVLT Forgetting* | .016 | .022 | 28.184 | .000 | 0.0 | 66.9 |
| *RAVLT Immediate* | .255 | .356 | 527.911 | .000 | 51.8 | 76.4 |
| *Logical Memory Test* | .382 | .533 | 862.725 | .000 | 70.7 | 81.1 |
| *RAVLT Percent Forgetting* | .213 | .297 | 429.404 | .000 | 71.7 | 75.6 |
| *RAVLT Learning* | .139 | .194 | 268.613 | .000 | 34.9 | 71.1 |
| *ADAS 13* | .386 | .539 | 875.889 | .000 | 69.8 | 84.3 |
| *MMSE* | .233 | .326 | 477.476 | .000 | 44.6 | 76.3 |
| *ADAS Q4* | .324 | .452 | 699.946 | .000 | 66.3 | 82.0 |
| *ADAS 11* | .366 | .511 | 817.625 | .000 | 60.2 | 82.4 |

*Note*. Highlighted variables violated assumption checks and could not be interpreted
*C&S $R^2$ = Cox & Snell R-squared*
N-$R^2$ = *Nagelkerke R-squared*

To determine whether we could improve upon our individual predictor models, we used a hierarchical binary logistic regression to discover the best mixed episodic-memory model in the second year of ADNI2; however, no significant mixed episodic-memory model could be found that matched or outperformed our individual RAVLT Percent Forgetting or

ADAS13 models. This was, in part, due to high amounts of multicollinearity between variables. It is unclear why multicollinearity was such a large issue in the second year and not in the first year of ADNI2. Consequently, the RAVLT Percent Forgetting remained the best episodic-memory predictor of MCI to AD conversion in the second year of ADNI2.

### 3.3.3 ADNI2 Year Three

In the third year of ADNI2, individual episodic-memory predictors generally improved in predictive accuracy. Unlike prior years, all measures met the assumptions of the linearity of the logit and could therefore be interpreted. All individual episodic-memory measures were significant with the exception of the RAVLT forgetting variable. Accordingly, the best odds ratio of all individual episodic-memory predictors in the third year of ADNI2 was once again the ADASQ4 (see Table 15 below). Specifically, the ADASQ4 [$\beta = 0.608$, $z = 383.384$, $p < .001$] had a positive odds increase of Exp(β) = 1.837 when predicting conversion. TheADASQ4 was followed by the ADAS11 [Exp(β) = 1.365], ADAS13 [Exp(β) = 1.262] and the RAVLT Percent Forgetting [Exp(β) = 1.045] in predictive odds.

**Table 15**
*Individual Episodic Memory Predictors of MCI to AD Conversion in Year Three of ADNI2*

| | | | | | | *95% C.I. for Exp(B)* | |
|---|---|---|---|---|---|---|---|
| *EM Measures* | *B* | *S.E.* | *Wald* | *Sig.* | *Exp(B)* | *Lower* | *Upper* |
| *ADNI_Mem* | -2.445 | .134 | 333.229 | .000 | .087 | .067 | .113 |
| *RAVLT Forgetting* | .025 | .023 | 1.248 | .264 | 1.026 | .981 | 1.072 |
| *RAVLT Immediate* | -.151 | .009 | 283.746 | .000 | .860 | .845 | .875 |
| *Logical Memory Test* | -.367 | .019 | 358.583 | .000 | .693 | .667 | .720 |
| *RAVLT Percent Forgetting* | .044 | .003 | 262.105 | .000 | 1.045 | 1.039 | 1.050 |
| *RAVLT Learning* | -.405 | .027 | 228.073 | .000 | .667 | .633 | .703 |
| *ADAS 13* | .233 | .012 | 351.517 | .000 | 1.262 | 1.232 | 1.293 |
| *MMSE* | -.597 | .032 | 338.741 | .000 | .551 | .517 | .587 |
| *ADAS Q4* | .608 | .031 | 383.384 | .000 | 1.837 | 1.728 | 1.952 |
| *ADAS 11* | .311 | .017 | 340.292 | .000 | 1.365 | 1.320 | 1.410 |

B = Regression coefficient used to calculate Exp(B)
Exp(B) = predictive odds ratio
S.E. = Standard Error

The predictive accuracy and accounted variance of each individual episodic-memory predictor can be found in Table 16 below. Overall, episodic-memory predictors exhibited a better accuracy predicting conversion in the third year of ADNI2 than in prior years. The LMT [C&S $R^2$ = .374, N-$R^2$ = .521, Conv-Acc. = 76.9%, $p$ < .001] was the single best predictor followed by the ADASQ4 [Conv-Acc. = 71.6%], the ADNIMEM [Conv-Acc. = 67.6%], and the ADAS13 [Conv-Acc. = 67.3%]. The RAVLT Forgetting was the only episodic-memory measure that could not predict MCI to AD conversion, as in prior years. The LMT did not contain the best accounted variance or odds ratio; however, when its strong predictive accuracy is factored in, it is clearly the best predictor.

**Table 16**

*Accuracy and Accounted Variance of Individual Episodic Memory Predictors in Year Three of ADNI2*

| *EM Measures* | *C&S $R^2$* | *N-$R^2$* | *Chi-squared* | *Sig.* | *Conversion predictive Acc.* | *Model Accuracy* |
|---|---|---|---|---|---|---|
| *ADNI_Mem* | .409 | .571 | 766.886 | .000 | 67.6 | 83.2 |
| *RAVLT Forgetting* | .001 | .001 | 1.247 | .264 | 0.0 | 67.4 |
| *RAVLT Immediate* | .321 | .447 | 561.585 | .000 | 55.6 | 76.9 |
| *Logical Memory Test* | .374 | .521 | 679.139 | .000 | 76.9 | 82.9 |
| *RAVLT Percent Forgetting* | .248 | .346 | 413.650 | .000 | 67.1 | 75.3 |
| *RAVLT Learning* | .200 | .279 | 324.256 | .000 | 51.2 | 74.0 |
| *ADAS 13* | .453 | .631 | 876.677 | .000 | 67.3 | 85.5 |
| *MMSE* | .389 | .543 | 716.585 | .000 | 59.1 | 83.3 |
| *ADAS Q4* | .364 | .508 | 655.406 | .000 | 71.6 | 82.1 |
| *ADAS 11* | .425 | .593 | 805.531 | .000 | 65.8 | 84.5 |

*C&S $R^2$ = Cox & Snell R-squared*
N-$R^2$ = *Nagelkerke R-squared*

**Mixed Episodic-Memory Predictor Model**

In the third year of ADNI2, the best mixed episodic-memory predictive model of MCI to AD conversion contained the LMT, the ADNIMEM, and the ADASQ4 measures. The LMT [Exp(β) = 0.855], the ADNIMEM [Exp(β) = 0.287], and the ADASQ4 [Exp(β) = 1.145] all predicted an increase in conversion odds (see Table 17 below). These predictive odds were worse than most of the individual predictors in the third year.

**Table 17**
*The Best Mixed Episodic Memory Predictor Model in Year Three of ADNI2*

| | | | | | | 95% C.I. for Exp(B) | |
|---|---|---|---|---|---|---|---|
| *EM Measures* | *B* | *S.E.* | *Wald* | *Sig.* | *Exp(B)* | *Lower* | *Upper* |
| *ADNI_MEM* | -1.250 | .219 | 32.535 | .000 | .287 | .186 | .440 |
| *Logical Memory* | -.157 | .024 | 41.174 | .000 | .855 | .815 | .897 |
| *ADAS Q4* | .136 | .053 | 6.644 | .010 | 1.145 | 1.033 | 1.269 |
| *Constant* | -1.007 | .364 | 7.652 | .006 | .365 | | |

B = Regression coefficient used to calculate Exp(B)
Exp(B) = predictive odds ratio
S.E. = Standard Error

Overall, the mixed episodic-memory model [Conv-Acc. = 77.7%, p < .001] performed like the LMT [Conv-Acc. = 76.9%, p < .001], which was the best individual episodic-memory predictor in terms of conversion predictive accuracy. However, the mixed model accounted for more variance [C&S R2 = .428, N-R2 = .598] when predicting MCI to AD conversion compared to the LMT [C&S R2 = .374, N-R2 = .521] (see Table 18 below). Moreover, the conversion predictive accuracy [Conv-Acc. = 77.7%] and total predictive accuracy [Total Acc. = 85.3%] of the third-year mixed models are more in line with research and clinical standards. The mixed model as a whole is only just better than individual episodic-memory predictors of conversion; however, the disparity between models is arguably not large enough to justify using three measures (including the LMT) compared to just using the LMT.

**Table 18**
*The Accounted Variance and Accuracy of the Best Mixed episodic Memory Model in Year Three of ADNI2*

| *EM Measures* | *C&S $R^2$* | *N-$R^2$* | *Chi-squared* | *Sig.* | *Conv Acc.* | *Total Model Acc.* |
|---|---|---|---|---|---|---|
| *Mixed EM Model Y3* | .428 | .598 | 808.244 | .000 | 77.7 | 85.1 |

*C&S $R^2$ = Cox & Snell R-squared*
N-$R^2$ = *Nagelkerke R-squared*

### 3.3.4 ADNI2 Year Four

In the final year of ADNI2, our analysis confirmed the trend that episodic-memory measure accuracy improves over time. The odds ratios were almost identical to those of the third year with the multiple ADAS-cog measures and RAVLT Percent Forgetting having the best predictive odds (see Table 19 below). Specifically, the ADASQ4 [Exp(β) = 1.891] had the best predictive odds followed by the ADAS11 [Exp(β) = 1.299], ASAD13 [Exp(β) = 1.219] and the RAVLT Percentage Forgetting [Exp(β) = 1.037].

**Table 19**
*Individual Episodic Memory Predictors of MCI to AD Conversion in Year Four of ADNI2*

| | | | | | | *95% C.I. for Exp(B)* | |
|---|---|---|---|---|---|---|---|
| *EM Measures* | *B* | *S.E.* | *Wald* | *Sig.* | *Exp(B)* | *Lower* | *Upper* |
| *ADNI_Mem* | -2.284 | .149 | 233.721 | .000 | .102 | .076 | .136 |
| *RAVLT Forgetting* | .018 | .029 | .372 | .542 | 1.018 | .962 | 1.077 |
| *RAVLT Immediate* | -.142 | .010 | 199.219 | .000 | .868 | .851 | .885 |
| *Logical Memory Test* | -.311 | .021 | 220.538 | .000 | .733 | .703 | .763 |
| *RAVLT Percent Forgetting* | .037 | .003 | 148.148 | .000 | 1.037 | 1.031 | 1.043 |
| *RAVLT Learning* | -.402 | .037 | 121.157 | .000 | .669 | .623 | .719 |
| *ADAS 13* | .198 | .012 | 252.225 | .000 | 1.219 | 1.189 | 1.249 |
| *MMSE* | -.547 | .035 | 248.935 | .000 | .579 | .541 | .619 |
| *ADAS Q4* | .637 | .039 | 266.060 | .000 | 1.891 | 1.751 | 2.041 |
| *ADAS 11* | .262 | .017 | 238.778 | .000 | 1.299 | 1.257 | 1.343 |

B = Regression coefficient used to calculate Exp(B)
Exp(B) = predictive odds ratio
S.E. = Standard Error

As observed in prior years, RAVLT Forgetting could not predict MCI to AD conversion and was not significant. In the fourth year, the best individual episodic-memory predictor was the LMT. It predicted MCI to AD conversion with an accuracy of 77.2% (see Table 20 below). The LMT was followed by the ADAS13 [Conv-Acc. = 75.5%] and the ADASQ4 [Conv-Acc. = 75.0%] in predictive ability. The LMT did not have the best accounted variance [C&S $R^2$ = .327, N-$R^2$ = .461] and was greatly surpassed by the ADAS13 [C&S $R^2$ = .431, N-$R^2$ = .608] and the ADASQ4 [C&S $R^2$ = .363, N-$R^2$ = .513]. Accordingly, the ADASQ4 is equally viable, as it has the best balance of high accounted variance, predictive ability and odds ratio. Either the LMT, the ADASQ4 or the ADAS13 could be

considered the best predictive measures of MCI to AD conversion depending on research preferences. We consider the ADASQ4 to be the best predictor of MCI to AD conversion due to its superior predictive odds balanced with its accuracy and accounted variance.

**Table 20**

*Accuracy and Accounted Variance of Individual Episodic Memory Predictors in Year Four of ADNI2*

| *EM Measures* | *C&S $R^2$* | *N-$R^2$* | *Chi-squared* | *Sig.* | *Conversion predictive Acc.* | *Model Accuracy* |
|---|---|---|---|---|---|---|
| *ADNI_Mem* | .400 | .563 | 521.400 | .000 | 71.1 | 85.6 |
| *RAVLT Forgetting* | .000 | .001 | .372 | .542 | 0.0 | 69.4 |
| *RAVLT Immediate* | .301 | .425 | 364.931 | .000 | 57.1 | 78.6 |
| *Logical Memory Test* | .327 | .461 | 403.133 | .000 | 77.2 | 84.1 |
| *RAVLT Percent Forgetting* | .191 | .270 | 216.294 | .000 | 65.3 | 72.2 |
| *RAVLT Learning* | .144 | .204 | 158.846 | .000 | 40.1 | 74.1 |
| *ADAS 13* | .431 | .608 | 574.623 | .000 | 75.5 | 88.3 |
| *MMSE* | .415 | .586 | 548.715 | .000 | 73.3 | 87.7 |
| *ADAS Q4* | .363 | .513 | 457.133 | .000 | 75.0 | 85.5 |
| *ADAS 11* | .406 | .573 | 531.548 | .000 | 73.2 | 87.6 |

*C&S $R^2$ = Cox & Snell R-squared*
N-$R^2$ = *Nagelkerke R-squared*

Note that we attempted to formulate a mixed episodic-memory predictor model in the fourth year of ADNI2; however, as with the second year, no model was found that could match or surpass our individual episodic-memory models and meet the assumption of multicollinearity.

## 3.4 Research Question Three

After we had formulated our episodic-memory prediction models, we sought to assess the predictive ability of the neuroimaging predictors of MCI to AD conversion in the first two years of ADNI2. We sought to assess the individual predictive ability of each neuroimaging variable by using a binary logistic regression. Following the individual analysis, we used a hierarchical binary logistic regression to determine whether neuroimaging measures could improve prior episode memory predictive models. The third and fourth years of ADNI2 could not be assessed due to problems with missing data that are discussed in Section 2.6.2 (p. 37).

**ADNI2 Year One**

In the first year of ADNI2, the individual neuroimaging markers were tested for their predictive odds and accuracy. All variables were statistically significant yet had problematic odds ratios [approximately Exp(β) = 1.00] (see Table 21 below). This meant that we could not interpret the odds ratios; however, we could infer from the Wald statistics that the entorhinal [z = 32.048] and hippocampal [z = 30.879] variables were the best models.

**Table 21**
*Individual Neuroimaging Predictor Odds of MCI to AD Conversion in Year One of ADNI2*

| | | | | | | *95% C.I. for Exp(B)* | |
|---|---|---|---|---|---|---|---|
| *AD Markers* | *B* | *S.E.* | *Wald* | *Sig.* | *Exp(B)* | *Lower* | *Upper* |
| *UCSF Ventricles* | .000 | .000 | 5.736 | .038 | 1.000 | 1.000 | 1.000 |
| *UCSF Hippocampus* | -.001 | .000 | 30.879 | .000 | .999 | .999 | 1.000 |
| *UCSF Whole Brain* | .000 | .000 | 6.733 | .025 | 1.000 | 1.000 | 1.000 |
| *UCSF Entorhinal* | -.001 | .000 | 32.048 | .000 | .999 | .998 | .999 |
| *UCSF Fusiform* | .000 | .000 | 15.833 | .001 | 1.000 | 1.000 | 1.000 |
| *UCSF Med Temp* | .000 | .000 | 19.031 | .001 | 1.000 | 1.000 | 1.000 |

B = Regression coefficient used to calculate Exp(B)
Exp(B) = predictive odds ratio
S.E. = Standard Error

All of the variables performed very poorly when predicting MCI to AD conversion, as can be seen in Table 22 below. The best individual predictor of conversion was the entorhinal variable [Conv-Acc. = 27.85%], followed by hippocampal volume [Conv-Acc. = 27.32%].

All models performed worse than chance when predicting MCI to AD conversion. Moreover, individual neuroimaging variables also performed worse than episodic-memory measures in the first year of ADNI2. The accounted variance of entorhinal volume [C&S R2 = .120, N-R2 = .169], the best neuroimaging predictor, was also far worse than the ANIMEM [C&S $R^2$ = .286, N-$R^2$ = .402, Conv-Acc. = 56.4%, $p < .001$], which was the best episodic-memory predictor.

**Table 22**

*Accuracy and Accounted Variance of Individual Neuroimaging Predictors in Year One of ADNI2*

| *AD Markers* | *C&S $R^2$* | *N-$R^2$* | *Chi-squared* | *Sig.* | *Conversion Acc.* | *Total Accuracy* |
|---|---|---|---|---|---|---|
| *UCSF Ventricles* | .020 | .027 | 5.79 | .030 | 1.95 | 68.15 |
| *UCSF Hippocampus* | .115 | .161 | 36.601 | .000 | 27.32 | 71.58 |
| *UCSF Whole Brain* | .023 | .032 | 7.005 | .017 | 1.07 | 68.85 |
| *UCSF Entorhinal* | .120 | .169 | 38.455 | .000 | 27.85 | 71.26 |
| *UCSF Fusiform* | .057 | .081 | 17.783 | .000 | 8.17 | 69.17 |
| *UCSF Med Temp* | .067 | .095 | 21.133 | .000 | 14.17 | 69.55 |

*C&S $R^2$ = Cox & Snell R-squared*
N-$R^2$ = *Nagelkerke R -squared*

## Mixed predictor Model

In the first year of ADNI2, the best mixed neuroimaging and episodic-memory model contained the ADAS11, the LMT and the entorhinal variables (see Table 23). The odds ratio positively predicted disease conversion similarly to each variable in their individual models.

**Table 23**

*The Best Mixed Predictor Model of MCI to AD Conversion in Year One of ADNI2*

| | | | | | *95% C.I. for Exp(B)* | |
|---|---|---|---|---|---|---|
| *Mixed Measures* | *B* | *S.E.* | *Sig.* | *Exp(B)* | *Lower* | *Upper* |
| *ADAS 11* | .130 | .037 | .000 | 1.139 | 1.060 | 1.225 |
| *Logical Memory* | -.190 | .046 | .000 | .827 | .756 | .905 |
| *UCSF Entorhinal* | .000 | .000 | .044 | 1.000 | .999 | 1.000 |
| *Constant* | .528 | .978 | .590 | 1.695 | .247 | 11.615 |

B = Regression coefficient used to calculate Exp(B)
Exp(B) = predictive odds ratio
S.E. = Standard Error

The mixed model predicted MCI to AD conversion with a higher accuracy than both the individual and composite episodic-memory models. Specifically, the mixed model predicted disease conversion with an accuracy of 59.8% compared to the 56.4% conversion accuracy of the ADNIMEM model (see Table 24 below). Accordingly, the mixed neuroimaging and episodic-memory predictive model was the best model in the first year of ADNI2. However, it could be argued that the individual episodic-memory models are more practical while containing similar accuracies.

**Table 24**
*The Accuracy and Accounted Variance of the Best Mixed Model in Year One of ADNI2*

| *Mixed Measures* | *C&S $R^2$* | *N-$R^2$* | *Chi-squared* | *Sig.* | *Conv Acc.* | *Total Model Acc.* |
|---|---|---|---|---|---|---|
| *Mixed Model Y1* | .323 | .454 | 117.315 | .000 | 59.8 | 80.4 |

*C&S $R^2$ = Cox & Snell R-squared*
N-$R^2$ = *Nagelkerke R-squared*

## ADNI2 Year Two

**Table 25**
*The Best Mixed Predictor Model of MCI to AD Conversion in Year Two of ADNI2*

| | | | | | | *95% C.I. for Exp(B)* | |
|---|---|---|---|---|---|---|---|
| *AD Markers* | *B* | *S.E.* | *Wald* | *Sig.* | *Exp(B)* | *Lower* | *Upper* |
| *UCSF Ventricles* | .000 | .000 | 6.912 | .022 | 1.000 | 1.000 | 1.000 |
| *UCSF Hippocampus* | -.001 | .000 | 37.006 | .000 | .999 | .999 | .999 |
| *UCSF Whole Brain* | .000 | .000 | 7.680 | .016 | 1.000 | 1.000 | 1.000 |
| *UCSF Entorhinal* | -.001 | .000 | 38.019 | .000 | .999 | .998 | .999 |
| *UCSF Fusiform* | .000 | .000 | 17.822 | .002 | 1.000 | 1.000 | 1.000 |
| *UCSF Med Temp* | .000 | .000 | 22.626 | .000 | 1.000 | 1.000 | 1.000 |

B = Regression coefficient used to calculate Exp(B)
Exp(B) = predictive odds ratio
S.E. = Standard Error

When assessing the predictive ability of individual neuroimaging variables in the second year, an improvement in predictive ability was achieved compared to the first year; however, the odds ratios were still problematic. Specifically, almost all variables have a positive odds ratio of Exp(β) = 1.00 and could thus not be interpreted; however, the Wald

statistic indicated that the entorhinal [z = 38.019], hippocampal [z = 37.006] and MTL [z = 22.626] models outperform the others (see Table 25 above).

The average accounted variance and predictive accuracy of the neuroimaging variables in the second year of ADNI2 was generally worse than the episodic-memory measures in the same year. The entorhinal [Conv-Acc. = 49.1%], hippocampal [Conv-Acc. = 47.4%] and MTL [Conv-Acc. = 30.3%] models had the best accuracy when predicting MCI to AD conversion (See Table 26 below). All remaining variables performed poorly in comparison when predicting disease conversion. The entorhinal [C&S R2 = .174, N-R2 = .243], hippocampal [C&S R2 = .169, N-R2 = .235] and MTL [C&S R2 = .097, N-R2 = .135] models also had the best accounted variance.

**Table 26**

*The Accuracy and Accounted Variance of the Best Mixed Model in Year Two of ADNI2*

| *AD Markers* | *C&S $R^2$* | *N-$R^2$* | *Chi-squared* | *Sig.* | *Conversion Acc.* | *Total Accuracy* |
|---|---|---|---|---|---|---|
| *UCSF Ventricles* | .027 | .038 | 7.131 | .017 | 4.4 | 67.97 |
| *UCSF Hippocampus* | .169 | .235 | 47.432 | .000 | 42.4 | 75.08 |
| *UCSF Whole Brain* | .031 | .044 | 8.101 | .009 | 4.2 | 68.17 |
| *UCSF Entorhinal* | .174 | .243 | 49.142 | .000 | 43.4 | 75.17 |
| *UCSF Fusiform* | .077 | .107 | 20.539 | .000 | 19.68 | 71.07 |
| *UCSF Med Temp* | .097 | .135 | 26.151 | .000 | 30.3 | 73.27 |

*C&S $R^2$ = Cox & Snell R-squared*
N-$R^2$ = *Nagelkerke R-squared*

## Mixed predictor Model

A hierarchical binary logistic regression was run in year one of ADNI2 using both neuroimaging and episodic-memory variables. The aim was to see if a mixed predictor model could outperform our prior episodic-memory models. The best mixed model contained the ADAS13 and entorhinal variables (see Table 27 below). The odds ratios were similar to each

measure's predictive odds in their individual models observed above. Moreover, the entorhinal variable still had problematic odds ratios.

**Table 27**
*The Best Mixed Predictor Model of MCI to AD Conversion in Year Two of ADNI2*

| | | | | | *95% C.I. for Exp(B)* | |
|---|---|---|---|---|---|---|
| *Mixed Measures* | *B* | *S.E.* | *Sig.* | *Exp(B)* | *Lower* | *Upper* |
| *ADAS 13* | .188 | .027 | .000 | 1.207 | 1.145 | 1.273 |
| *UCSF Entorhinal* | -.001 | .000 | .018 | .999 | .999 | 1.000 |
| *Constant* | -1.930 | 1.176 | .103 | .145 | .014 | 1.483 |

B = Regression coefficient used to calculate Exp(B)
Exp(B) = predictive odds ratio
S.E. = Standard Error

The accuracy and accounted variance of the mixed predictor model performed slightly worse than the episodic-memory models observed in the second year of ADNI2 (see Table 28 below). However, the model performed worse by a negligible amount. Accordingly, the mixed neuroimaging model could be used to predict MCI to AD conversion, yet it would not realistically be used, as a single episodic-memory test can achieve the same results at a lower cost. Thus, the predictor of MCI to AD conversion in the second year of ADNI 2 was the RAVLT Percent Forgetting episodic-memory measure.

**Table 28**
*The Accuracy and Accounted Variance of the Best Mixed Model in Year Two of ADNI2*

| *Measures* | *C&S $R^2$* | *N-$R^2$* | *Chi-squared* | *Sig.* | *Conv Acc.* | *Total Model Acc.* |
|---|---|---|---|---|---|---|
| *Mixed Model Y2* | .408 | .570 | 134.789 | .000 | 70.5 | 83.9 |

*C&S $R^2$ = Cox & Snell R-squared*
N-$R^2$ = *Nagelkerke R-squared*

In conclusion, Chapter 3 investigated the ability of general-multidomain tests (i.e., MMSE and ADAS-cog) to measure episodic memory. Moreover, Chapter 3 investigated the ability of episodic-memory tests and neuroimaging measures to predict MCI to AD conversion. The results indicated that general-multidomain tests measured the same construct as specific tests and could thus, measure episodic memory. Furthermore, the results indicated

that individual episodic-memory measures were the best predictors of MCI to AD conversion. Comparatively, Individual neuroimaging markers could not predict MCI to AD conversion above chance averages and performed worse than episodic-memory measures. Mixed neuroimaging and episodic-memory models improved over time with these mixed models performing relatively worse or the same as mixed episodic-memory models or individual episodic-memory predictors of disease conversion.

# Chapter Four: Discussion and Conclusion

## 4.1 Interpretation of Results

The exponential increase in AD patients over the next 30 years is expected to create a global epidemic. Stronger treatment and diagnostic methods are required to actively combat AD; however, current measures cannot detect the disease at a significantly treatable or reversible stage. Accordingly, there is a critical need for early-stage AD detection methods. The present study sought to investigate the ability of episodic-memory measures to predict early-stage AD. Episodic-memory tests were used as they can specifically target the initial cognitive symptoms of AD during the early stages of the disease. Specifically, this study sought to predict MCI to AD conversion using multiple episodic-memory tests. The episodic-memory models were also compared to common neuroimaging measures of AD to find the best predictor of MCI to AD conversion and thus early-stage AD. While other studies have examined episodic memory as a predictor of MCI to AD conversion, this is the first to incorporate all the episodic-memory measures from ADNI and properly evaluate and compare these measures to common neuroimaging predictors of AD.

The results of this study are discussed in reference to the project research questions which are the following: Can general cognitive assessment tests used for AD, such as the MMSE and ADAS-Cog, be used to measure episodic memory? Which episodic-memory test best predicts MCI to AD conversion, and can mixed episodic-memory models predict conversion better than individual tests? Can individual neuroimaging measures predict MCI to AD conversion better than episodic-memory variables, and can a mixed model (neuroimaging and episodic-memory measures) predict conversion better than prior models?

### 4.1.1 Construct Validity in Episodic-Memory Tests

Before assessing the predictive ability of episodic-memory tests, we determined whether general multi-domain tests—namely, the MMSE and ADAS-cog—could be used to measure episodic memory. This was, in part, because prior research had suggested that these general diagnostic tests could measure episodic memory and because these tests were part of the ADNI composite memory score that we used to measure episodic memory (Crane et al., 2012). Theoretically, it is not difficult to infer that the MMSE and ADAS-cog can measure episodic memory, as both tests involve recall tasks similar to the RAVLT. Accordingly, we consistently found, in all four years of ADNI2, that the general diagnostic tests and specific episodic-memory tests measured the same construct. This means that, to a degree, the MMSE and ADAS-cog can be used to measure episodic memory. While these results should not justify the use of the MMSE or ADAS-cog over specific tests such as the RAVLT for the study of episodic memory, it does clarify that these general tests can measure episodic memory and are justified for use in our study. Moreover, this means that, in research and clinical practice, specific delayed recall sections of general tests can be used as measures of episodic memory (e.g., ADASQ4).

During the principle component analysis, the RAVLT Forgetting and RAVLT Percentage Forgetting were consistently in their own factor loading group. We theorised that the RAVLT forgetting group occurred because both measures are two quantifications of the same base statistic. Moreover, the second factor was not important, as the analysis was only concerned with general diagnostic tests and not with the relationship between RAVLT forgetting measures. In the third year of the ADNI2 cohort, there was more of a divide between general cognitive assessment tests and episodic-memory tests; however, overall, these two groups of tests were still related and could both measure episodic memory in MCI converters. The split between factors in the third year may be caused by the progression of

AD, as general multi-domain tests are often better later on during the progression of the disease, thus causing more of a split between types of tests. However, the disparity between specific and general tests reduced in the fourth year, and thus the third-year results, may also have just been an anomaly.

### 4.1.2 Episodic-Memory Predictive Models

The prime focus of this project was to evaluate the ability of episodic-memory tests to predict MCI to AD conversion both individually and in a battery. Across the first four years of ADNI2, all episodic-memory tests could predict MCI, with the exception on the RAVLT Forgetting variable. The predictive accuracy of individual episodic-memory tests also improved with time across the four years of ADNI2. The best episodic-memory tests predictive of MCI to AD conversion observed across all four years of ADNI2 were the LMT, the ADASQ4, the ADAS13, the ADNIMEM and the RAVLT Percent Forgetting. The following individual episodic-memory measures were the best predictors: ANIMEM in the first year (56.4%), the RAVLT Percent Forgetting in the second year (71.7%) and the LMT in the third (76.9%) and fourth years (77.2%). It is important to note that the RAVLT Percent Forgetting and LMT predicted MCI to AD conversion with a higher accuracy (63.8% and 60.6%) in the first year of ADNI2; however, the results of these tests had to be omitted, as they violated the assumptions of a logistic regression (specifically the assumption of linearity of the logit).

It was interesting to find that, unlike other general tests, the ADAS 13 had a good predictive ability similar to specific episodic-memory tests such as the LMT. Conversely, the MMSE was one of the worst predictors of MCI to AD conversion along with the RAVLT forgetting, the RAVLT immediate and the RAVLT learning measures. The results indicate that the MMSE is a poor predictor of MCI to AD conversion, which is consistent with findings in the literature (Arevalo-Rodriguez et al., 2015). The RAVLT forgetting measure

consistently could not predict MCI to AD conversion across all of the years of ADNI2. In the literature, this measure is often rated as one of the best episodic-memory measures of disease conversion. It is unknown whether prior studies have confused the RAVLT forgetting and RAVLT percentage forgetting measures or if our results are anomalous. Equally, our results have shown that the RAVLT forgetting measure can strongly predict MCI stability, so there might be some confusion with this statistic in the literature. Accordingly, the RAVLT Percentage Forgetting measure has also been found to be a strong predictor of AD in the literature (Moradi et al., 2017). It is important to note that the variability in RAVLT measures reinforce the notion that both clinicians and researchers should not use just the RAVLT to measure episodic memory.

When assessing the predictive ability of mixed episodic-memory predictor models, it was found that all mixed models performed similarly or worse than individual predictor tests. Moreover, in the second and fourth years of ADNI2, mixed episodic-memory predictive models could not be formulated due to assumption violations. In the first year of ADNI2, the best mixed episodic-memory predictor model contained the RAVLT Immediate and the ADAS Q4 (53.1%) and was worse than the ADNIMEM measure alone (56.4%). The predictive accuracy of this model did not justify the use of multiple tests in the assessment of episodic-memory decline and MCI to AD conversion. Conversely, in the third year of ADNI2, the best mixed predictor model (accuracy of 77.7%) contained the ADNIMEM, the LMT and the ADASQ4 measures, and it predicted disease conversion better than individual episodic-memory measures. However, the best individual predictor in the third year had an accuracy of 76.9%; thus, an increase in accuracy of 1.2% does not logistically justify the use of two extra tests. Even the specially created ADNI composite memory score was often superseded by individual episodic-memory tests. These results are in contrast to those found by Chapman et al. (2011), who suggested that cognitive batteries could predict MCI to AD

conversion better than individual episodic-memory tests. Our results align more with De Simone et al. (2019), who found that individual and mixed measure episodic-memory models are both equally viable and tend to have similar accuracies.

Statistically speaking, mixed episodic-memory predictive models can potentially provide small advantages in accuracy over individual episodic-memory models; however, clinically speaking, the logistical complexity gained through multiple tests is not worth the small increase in accuracy. In the case of this study, it is safe to say that more predictors and measures do not result in better models of MCI to AD conversion. Moreover, no mixed episodic-memory predictor models could be computed for the second and fourth years of the ADNI2 cohort, as all significant models violated the assumptions of linearity and multicollinearity, thus suggesting that model complexity can often diminish statistical power. We theorised that the assumption violations occurred because many of the episodic-memory measures share similar parent tests (e.g., ADAS 11, ADAS 13, ADASQ4) or because they all sought to measure episodic memory similarly (e.g., using list learning tasks).

In conclusion, individual episodic-memory tests were found to be the best predictors of MCI to AD conversion in Research Question 2. We also found that episodic-memory predictive models generally improved across the years of ADNI2. At the beginning of ADNI2, the most accurate episodic-memory tests performed just above chance. Moreover, the best predictive model was the LMT in the fourth year of ADNI2. The results also indicated that specific episodic-memory tests consistently outperformed general multidomain tests, with only the ADAS 11 and 13 consistently rivalling specific episodic-memory tests. We theorise that the ADAS-cog performed well as an episodic-memory test, as it includes a delayed recall task. Finally, the RAVLT generally performed poorly with only the RAVLT Percent Forgetting measure strongly predicting MCI to AD conversion.

### 4.1.3 Mixed model Predictors of Alzheimer's Disease

ADNI research commonly combines various cognitive, neuroimaging and biomarker measures together in models used to predict early-stage AD and MCI to AD conversion. For this reason, we wanted to utilise the biomarker and neuroimaging measures contained in the ADNI database to see if we could improve our episodic-memory predictive models. Unfortunately, we could not use any biomarker measures (e.g., Tau and Aβ), as these markers had severe amounts of missing data; however, this was not too discouraging, as biomarkers often perform poorly when predicting MCI to AD conversion as discussed in Section 1.3.3 above (p. 8). Similarly, neuroimaging markers contained a decent amount of missing data in the third and fourth years of ADNI2; however, in this circumstance, we were able to study the markers in the first two years of ADNI2.

Accordingly, in the first and second years of ADNI2, individual neuroimaging markers poorly predicted MCI to AD conversion in comparison to episodic-memory measures. In the first year, the best predictors of conversion were MRI atrophy measures of participants entorhinal (27.9%) and hippocampal (27.3%) regions. Similarly, in the second year, the entorhinal (43.4%) and hippocampal (42.4%) regions were the best neuroimaging predictors of MCI to AD conversion; however, these statistics were in no way good individual predictors of MCI to AD conversion. While we found that the entorhinal and hippocampal regions of the brain were the best neuroimaging predictors of disease conversion in concordance with the literature, our accuracy statistics were significantly weaker than those observed in other studies (Ihara et al., 2018; Moradi et al., 2017). We theorise that these weak results may have been obtained because neuroimaging measures improve with disease progression. As has been observed with the episodic-memory predictors mentioned above, neuroimaging measures may have improved in the third and fourth years of ADNI2; however, this cannot be confirmed due to the nature of our missing data.

Neuroimaging markers predicted MCI to AD conversion better when they were combined with episodic memory markers. Moreover, in both the first and second years of ADNI2, combined neuroimaging and episodic-memory mixed models predicted MCI to AD conversion the same as or better than all other models. In the first year, the ADAS11, LMT and UCSF entorhinal measures predicted disease conversion with an accuracy of 59.8%. In the second year, the ADAS13 and UCSF entorhinal measures predicted MCI to AD conversion with an accuracy of 70.5%. However, like the mixed episodic-memory models, these mixed models did not provide any significant increases in accuracy that would justify the use of multiple measures and MRI equipment over one single episodic-memory test. These results are similar to those obtained by Gomar et al. (2014), who found that adding neuroimaging measures to episodic-memory tests did not significantly improve predictive models. However, it would be interesting to see how episodic-memory models improved with neuroimaging measures in the later years of ADNI2, as there is a gradual increase in the accuracy of predictive measures observed in our data.

In conclusion, our study has determined that, out of our variables, individual episodic-memory measures such as the LMT are the best predictors of MCI to AD conversion and therefore of early-stage AD. While neuroimaging measures can increase the overall accuracy of predictive models, there is no advantage to using the technology when factoring in the logistical complexity and cost of the technology compared to the accessibility and accuracy of episodic-memory tests. Accordingly, when researching episodic memory in AD and MCI conversion, we recommend the use of the RAVLT Percent Forgetting, the ADASQ4 and the LMT for episodic-memory testing. It is also important to note that researchers need to be careful when using single episodic-memory measures, as they sometimes do not match the target population and perform differently at various stages of AD progression (e.g., tests performed with varying accuracies across ADNI2 in our study). Accordingly, we suggest that

multiple specific episodic-memory tests should be used to inform early-stage MCI to AD diagnosis. Furthermore, general multidomain tests should be added to inform diagnoses later on during AD progression. It is also imperative that multiple episodic-memory tests be repeatedly and consistently administered throughout disease progression to confirm MCI and AD diagnoses.

## 4.2 Clinical and Theoretical Implications

When concerning clinical research and practise, our study suggests that episodic-memory tests are viable measures of MCI to AD conversion and, subsequently, of early-stage AD. These results indicate that clinicians could use specific episodic-memory tests such as the LMT to observe and diagnose early-stage AD. An increase in early-stage AD diagnoses through episodic-memory tests could consequentially enable researchers and clinicians to treat and observe the disease at a more manageable and reversible stage. This early detection ability may be the first step in combating the global epidemic of AD expected in the next three decades. Moreover, our study has shown that neuroimaging technology may not be necessary in early-stage AD diagnosis, as individual episodic-memory tests could predict the disease with the same or higher accuracy. This means that early-stage AD diagnoses and MCI to AD prediction methods can be relatively affordable and accessible due to the simplistic nature of pen-and-pencil tests.

Furthermore, our study has confirmed that general multi-domain tests used in AD diagnosis (e.g., MMSE) are not strong predictors of early-stage MCI to AD conversion and thus should be used only when assessing mid to late-stage AD. However, our results have also indicated that multi domain tests that include delayed recall measures (e.g., ADAS13) may be more useful than other general tests (e.g., MMSE) in early-stage AD research. Our study has also found that all episodic memory and MRI predictors of disease conversion increase in accuracy over time. Accordingly, there is a need for clinicians to repeatedly observe tests over time to provide accurate diagnoses and avoid misdiagnosis. Moreover, using an incorrect test during the wrong stage of AD progression—such as using the MMSE in the early stages of MCI to AD conversion—will not provide accurate measures of AD. Generally, according to our results, we suggest the use of the LMT or ADAS Q4 when seeking to predict MCI to AD conversion. It is imperative that the correct test is used during

specific stages of the disease and, unfortunately, there is no consensus on which test is best suited for each stage of the disease or even how to determine which stage of AD an individual is in.

**Theoretical Implications**

As discussed in Section 1.2 (see p. 3), there is no clear theory of cognitive decline in AD. Moreover, theories of episodic memory function (see Section 1.4.1 p. 13) often do not discuss memory decline. Building on the component processing model of episodic memory, the amyloid cascade hypothesis, and our results, we theorise that episodic-memory decline in AD occurs due to a loss of connectivity in the MTL. Specifically, the component processing model states that the episodic-memory system formulates memories using a multitude of brain regions across the MLT and cortex. Accordingly, we theorize that episodic-memory decline in early-stage AD is derived from breakdowns in communication between brain regions relating to the aggregation of Aβ and tau pathology. We hypothesize that episodic-memory decline relates to the disruption of neurons in line with the amyloid cascade hypothesis and is not a by-product of neurodegeneration. Specifically, episodic-memory declines because of communication breakdowns relating to tau and Aβ aggregation and not due to the loss of neurons in the MLT and cortex, which happens at a later stage. We theorise this because, in our population, episodic-memory decline was a significant predictor of MCI to AD conversion before neuroimaging predictors were. This led us to believe that the episodic memory system is interrupted before the progression of mass neurodegeneration.

Like prior researchers, we also found that measures of the MLT such as hippocampal and entorhinal volume or atrophy are the best neuroimaging predictors of MCI to AD conversion. This agrees with the component processing model that regards the hippocampus and entorhinal cortex as pivotal to the aggregation and processing of episodic-memory

information. Conversely, we found that these predictors did not predict conversion with a high accuracy. For this reason, and due to our missing data, this topic requires further research and clarification.

## 4.3 Review and Future Direction

This study sought to predict disease conversion from MCI to AD as a way of detecting early-stage AD. Accordingly, we discovered that episodic-memory tests could be used to predict disease conversion with a moderate to strong accuracy. Thus, this study provides a unique outlook on early-stage AD prediction and has both strengths and limitations that should influence future research.

### 4.3.1 Strengths

This study excelled in its statistical analysis, theoretical background and population. First, the statistical analysis utilised both binary logistic regressions and multiple imputations which are not commonly used in this area of research; however, these statistical methods are highly revered in the statistical literature. A binary logistic regression gave us the unique opportunity to investigate a large number of variables at once while providing detailed analyses of each individual predictor of disease conversion. In the literature, it is common to use receiver operating characteristic curves to assess predictor accuracy and use linear regressions to formulate predictive models. By utilising a binary logistic regression, both the predictive accuracy and model construction can be performed at the same time while also obtaining detailed information on the odds ratios (increase or decrease in predictive ability) of individual variables. A binary logistic regression also fit our research questions and dataset best, as the analysis specialises in dichotomous categorical prediction (e.g., conversion and non-conversion) and secondary data (testing a model against observations). The multiple imputations also strengthened our dataset, as the corrections reduced the error rate and enabled us to correct missing data. Without the multiple imputations, this study would only have been able to consider episodic-memory variables and not neuroimaging measures.

Another strength of this study can be found in its theoretic background. A key topic in this study was the lack of theoretical information explaining processes of episodic-memory

decline in AD; however, the existing theories gave us a strong background from which to consider how episodic memory is affected in AD. Specifically the amyloid cascade hypothesis in the component processing model of episodic memory provided enough information for us to consider and test a potential reason for episodic-memory decline in AD. Uniquely, the lack of theory surrounding episodic-memory decline in AD drove this study and enabled us to make a unique contribution to research. Furthermore, we were able to draw on research from multiple disciplines such as neuroscience, cognitive science and psychology. We believe that the incorporation of interdisciplinary knowledge was a strong foundation for this project and analysis.

Lastly, the study population and, accordingly, the data acquired from the *Alzheimer's Disease Neuroimaging Initiative* was pivotal to the success of this project. As previously mentioned, without the ADNI database, this project would not be able to longitudinally study early-stage AD or research so many participant and variables. Accordingly, the large population and the multitude of variables contained in the ADNI database statistically and theoretically strengthened this study. Specifically, the large number of MCI participants enabled us to use a binary logistic regression and perform a detailed analysis of each of the first four years of ADNI2. Moreover, the multitude of cognitive tests enabled us to investigate the differences between general tests for AD diagnosis (e.g., MMSE) and specific episodic-memory tests (e.g., RAVLT). This was important, as cognitive tests are often avoided in early-stage AD research because of the poor accuracy of general tests for AD diagnosis, and specific episodic-memory tests are rarely compared or utilised to predict ability early-stage AD.

### 4.3.2 Limitations

It is important to note that this study had some limitations that should be considered in future studies. Primarily, this study was limited by missing data, the tests used by ADNI and the ability to assess disease misdiagnosis.

Missing data was the largest limitation of this study. We were unable to study biomarkers for two years of neuroimaging markers due to the amount of missing data in the ADNI database. Ultimately, missing data could not be avoided, as we did not collect any data ourselves and we utilised statistical methods to recover as much data as possible. Accordingly, this study can only comment on the predictive ability of episodic-memory tests in the first four years and on neuroimaging measures in the first two years of ADNI2. We would have liked to have commented on the performance of neuroimaging predictors in the later years of ADNI2 and AD, as this is a key area of interest in the literature. It is unknown whether patterns of participant drop out occur in other ADNI cohorts; however, participant drop out is common in AD research due to the vulnerable nature of neurological populations.

Next, this study was limited by the variables and tests available to use from the ADNI database. Accordingly, we had to fit the variables given to us to our research questions and tests. Subsequently, statistical models had problems with multicollinearity due to the similarity of the tests selected by ADNI. However, ADNI is in no way to blame, as the tests used are widely recommended in the literature. Ultimately, problems of multicollinearity and the need for more tests are consequences of using data not specifically collected for this study.

Lastly, this study was limited in its ability to assess misdiagnosis. In both research and clinical practice, general assessment tests used for AD such as the MMSE are often prone to misdiagnosing forms of dementia (e.g., diagnosing Lewy body dementia as AD). While this study used a control group of MCI stable participants to test against converting participants, it

would have been good to also test stable AD and healthy aging participants and those with various forms of dementia (e.g., vascular dementia). We could not do this because our dataset from ADNI did not contain all of these participant groups and because a binary logistic regression must have dichotomous groups. In the case of our study, it was more important to confirm that episodic-memory tests can predict conversion than test misdiagnosis; however, this is the next logical step for future research.

### 4.3.3 Future Direction

Both the results and limitations of this project should direct future research. Particularly, future research should build on the episodic-memory models contained in this study and seek to apply findings about such tests and predictors to clinical research.

Future research should continue to explore more episodic-memory tests and compare them to current mainstream predictors such as MRI measures. Specifically, future research should expand the pool of episodic-memory tests used to predict MCI to AD conversion to determine whether any other specific test can predict disease conversion with a high accuracy. Equally, new episodic-memory measures should be selected from different tests with unique methodologies to avoid problems of multicollinearity, such as those which occurred in this study. While all tests should be evaluated, it is also important to note that tests that are clinically assessable and easily administrable should be preferred over complex batteries. Our study suggests that complex measures involving many tests, such as the ADNINMEM measure, provide no advantage over individual tests. It is also important to note that future research should seek to explore specific tests that target episodic memory over general multidomain tests and also explore the relationships between said tests. Our study suggests that one reason for the lack of confidence in cognitive tests during early-stage AD diagnosis may be because researchers prefer general tests over specific tests.

Furthermore, it is important that future research further investigates the predictive ability of neuroimaging methods in detecting early-stage AD and, subsequently, MCI conversion. With respect to our study, future research should longitudinally assess a multitude of neuroimaging measures of MCI to AD conversion. While a study evaluating many neuroimaging measures requires a large amount of time and money, this may be possible once the current ADNI3 cohort is completed. Future research should also seek to evaluate episodic-memory measures in the ADNI3 cohort to replicate and confirm this study. This is important, as ADNI3 includes new biomarker and neuroimaging measures, such as PET measures, that may predict disease conversion better than current measures in the ADNI2 cohort.

## 4.4 Concluding Remarks

Overall, this study was both a successful study and a defining research experience that will affect the rest of my research career. The primary goal of this study was to determine whether episodic-memory tests could predict MCI to AD conversion and therefore detect early AD. Moreover, this study sought to find reasons for episodic-memory decline in AD. Academically, this project has greatly improved my writing, statistical and communicative skills. I have learnt countless lessons that have already started to define my academic career and work. Personally, this study has also influenced my understanding of myself and my work ethic. While this project is only approximately 100 pages to the reader, to me it is the culmination of two years of work and personal development. I am incredibly proud of where I am and the effort I have put into this project. I have learned the true meaning of stress and anxiety yet also patience, satisfaction and happiness.

I hope that this project has a positive effect on the AD literature and will be part of a wider movement that seeks to improve the lives of those with neurodegenerative diseases. It is my wish that AD is cured in my lifetime and that projects like this one can assist in getting the research one step closer to that goal. I think it is incredibly important that we continue to research and invest in early-stage detection methods for AD and MCI and promote research that can be translated into clinical practice. Everyone can do their part to help individuals with AD, and this project is the beginning of my contribution.

# Declarations

This study would like to acknowledge the *Alzheimer's Disease Neuroimaging Initiative* for their data which was crucial to the success of this study.

This study would also like to acknowledge funding received through the *Australian Government Research Training Program* and *Western Sydney Universities Master of Research Scholarship*. Without funding from the Australian Government or Western Sydney University this project would not have been possible.

The authors of this study declare no conflicts of interest.

# References


Abikoff, H., Alvir, J., Hong, G., Sukoff, R., Orazio, J., Solomon, S., & Saravay, S. (1987). Logical memory subtest of the wechsler memory scale: Age and education norms and alternate-form reliability of two scoring systems. *Journal of Clinical and Experimental Neuropsychology*, *9*(4), 435–448. https://doi.org/10.1080/01688638708405063

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. *Alzheimer's & Dementia*, *7*(3), 270–279. https://doi.org/10.1016/j.jalz.2011.03.008

Alzheimer's Association. (2018). 2018 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, *14*(3), 367–429.

Alzheimer's Disease International. (2019). Dementia Statistics. Retrieved April 5, 2019, from Alzheimer's Disease International website: https://www.alz.co.uk/research/statistics

Alzheimer's Disease Neuroimaging Initiative. (2016). ADNI3 Protocol. Retrieved April 11, 2019, from http://adni.loni.usc.edu/wp-content/themes/freshnews-dev-v2/documents/clinical/ADNI3_Protocol.pdf

Alzheimer's Disease Neuroimaging Initiative. (2017a). ADNI About. Retrieved December 11, 2019, from http://adni.loni.usc.edu/about/

Alzheimer's Disease Neuroimaging Initiative. (2017b). ADNI Data Types. Retrieved December 11, 2019, from http://adni.loni.usc.edu/data-samples/data-types/

Alzheimer's Disease Neuroimaging Initiative. (2017c). ADNI MRI Tool. Retrieved December 11, 2019, from http://adni.loni.usc.edu/methods/mri-tool/

Arevalo-Rodriguez, I., Smailagic, N., Roqué I Figuls, 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). *The Cochrane Database of Systematic Reviews*, (3), CD010783. https://doi.org/10.1002/14651858.CD010783.pub2

Armstrong, R. A. (2013). What causes alzheimer's disease? *Folia Neuropathologica*, *51*(3), 169–188. https://doi.org/10.5114/fn.2013.37702

Athinoula A. Martinos Center for Biomedical Imaging. (2019). Free Surfer Wiki. Retrieved December 11, 2019, from https://surfer.nmr.mgh.harvard.edu/fswiki

Australian Bureau of Statistics. (2017). *Causes of death, Australia, 2016*. Australian Bureau of Statistics Canberra.

Bak, T. H., Rogers, T. T., Crawford, L. M., Hearn, V. C., Mathuranath, P. S., & Hodges, J. R. (2005). Cognitive bedside assessment in atypical parkinsonian syndromes. *Journal of*

*Neurology, Neurosurgery & Psychiatry*, *76*(3), 420–422. https://doi.org/10.1136/jnnp.2003.029595

Barker, G. R. I., Banks, P. J., Scott, H., Ralph, G. S., Mitrophanous, K. A., Wong, L.-F., … Warburton, E. C. (2017). Separate elements of episodic memory subserved by distinct hippocampal-prefrontal connections. *Nature Neuroscience*, *20*(2), 242–250. https://doi.org/10.1038/nn.4472

Bartus, R. T., Dean, R. L., Beer, B., & Lippa, A. S. (1982). The cholinergic hypothesis of geriatric memory dysfunction. *Science*, *217*(4558), 408–414. https://doi.org/10.1126/science.7046051

Bastin, C., & Salmon, E. (2014). Early neuropsychological detection of Alzheimer's disease. *European Journal of Clinical Nutrition*, *68*(11), 1192–1199. https://doi.org/10.1038/ejcn.2014.176

Bloom, G. S. (2014). Amyloid-β and tau: The trigger and bullet in Alzheimer disease pathogenesis. *JAMA Neurology*, *71*(4), 505–508. https://doi.org/10.1001/jamaneurol.2013.5847

Brown, J. (2015). The use and misuse of short cognitive tests in the diagnosis of dementia. *Journal of Neurology, Neurosurgery & Psychiatry*, *86*(6), 680–685. https://doi.org/10.1136/jnnp-2014-309086

Brueggen, K., Dyrba, M., Barkhof, F., Hausner, L., Filippi, M., Nestor, P. J., … Teipel, S. J. (2015). Basal Forebrain and Hippocampus as Predictors of Conversion to Alzheimer's Disease in Patients with Mild Cognitive Impairment—A Multicenter DTI and Volumetry Study. *Journal of Alzheimer's Disease: JAD*, *48*(1), 197–204. https://doi.org/10.3233/JAD-150063

Cai, S., Huang, L., Zou, J., Jing, L., Zhai, B., Ji, G., … Ren, A. (2015). Changes in Thalamic Connectivity in the Early and Late Stages of Amnestic Mild Cognitive Impairment: A Resting-State Functional Magnetic Resonance Study from ADNI. *PLoS ONE*, *10*(2), e0115573. https://doi.org/10.1371/journal.pone.0115573

Chapman, R. M., Mapstone, M., McCrary, J. W., Gardner, M. N., Porsteinsson, A., Sandoval, T. C., … Reilly, L. A. (2011). Predicting Conversion from Mild Cognitive Impairment to Alzheimer's Disease Using Neuropsychological Tests and Multivariate Methods. *Journal of Clinical and Experimental Neuropsychology*, *33*(2), 187–199. https://doi.org/10.1080/13803395.2010.499356

Cheke, L. G., & Clayton, N. S. (2013). Do different tests of episodic memory produce consistent results in human adults? *Learning & Memory (Cold Spring Harbor, N.Y.)*, *20*(9), 491–498. https://doi.org/10.1101/lm.030502.113

Cheke, L. G., & Clayton, N. S. (2015). The six blind men and the elephant: Are episodic memory tasks tests of different things or different tests of the same thing? *Journal of Experimental Child Psychology*, *137*, 164–171. https://doi.org/10.1016/j.jecp.2015.03.006

Chu, L. W. (2012). Alzheimer's disease: Early diagnosis and treatment. *Hong Kong Medical Journal = Xianggang Yi Xue Za Zhi*, *18*(3), 228–237.

Craig, L. A., Hong, N. S., & McDonald, R. J. (2011). Revisiting the cholinergic hypothesis in the development of Alzheimer's disease. *Neuroscience & Biobehavioral Reviews*, *35*(6), 1397–1409. https://doi.org/10.1016/j.neubiorev.2011.03.001

Crane, P. K., Carle, A., Gibbons, L. E., Insel, P., Mackin, R. S., Gross, A., … Mungas, D. (2012). Development and assessment of a composite score for memory in the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Brain Imaging and Behavior; Indianapolis*, *6*(4), 502–516. http://dx.doi.org.ezproxy.uws.edu.au/10.1007/s11682-012-9186-z

Csukly, G., Sirály, E., Fodor, Z., Horváth, A., Salacz, P., Hidasi, Z., … Szabó, Á. (2016). The Differentiation of Amnestic Type MCI from the Non-Amnestic Types by Structural MRI. *Frontiers in Aging Neuroscience*, *8*. https://doi.org/10.3389/fnagi.2016.00052

Cui, Y., Liu, B., Luo, S., Zhen, X., Fan, M., Liu, T., … Alzheimer's Disease Neuroimaging Initiative. (2011). Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors. *PloS One*, *6*(7), e21896. https://doi.org/10.1371/journal.pone.0021896

Davatzikos, C., Bhatt, P., Shaw, L. M., Batmanghelich, K. N., & Trojanowski, J. Q. (2011). Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiology of Aging*, *32*(12), 2322.e19-27. https://doi.org/10.1016/j.neurobiolaging.2010.05.023

Davidson, P. S. R., & Winocur, G. (2010). Aging and Cognition. In G. F. Koob, M. L. Moal, & R. F. Thompson (Eds.), *Encyclopedia of Behavioral Neuroscience* (pp. 20–26). https://doi.org/10.1016/B978-0-08-045396-5.00031-2

De Simone, M. S., Perri, R., Fadda, L., Caltagirone, C., & Carlesimo, G. A. (2019). Predicting progression to Alzheimer's disease in subjects with amnestic mild cognitive impairment using performance on recall and recognition tests. *Journal of Neurology*, *266*(1), 102–111. https://doi.org/10.1007/s00415-018-9108-0

Dementia Australia. (2018). Dementia Prevalence Estimates 2018-2058. Retrieved April 5, 2019, from https://www.dementia.org.au/statistics

Edvardsson, D., Winblad, B., & Sandman, P. (2008). Person-centred care of people with severe Alzheimer's disease: Current status and ways forward. *The Lancet Neurology*, *7*(4), 362–367. https://doi.org/10.1016/S1474-4422(08)70063-2

Eichenbaum, H. (2017). Prefrontal-hippocampal interactions in episodic memory. *Nature Reviews. Neuroscience*, *18*(9), 547–558. https://doi.org/10.1038/nrn.2017.74

El Haj, M., Antoine, P., Amouyel, P., Lambert, J.-C., Pasquier, F., & Kapogiannis, D. (2016). Apolipoprotein E (APOE) ε4 and episodic memory decline in Alzheimer's disease: A review. *Ageing Research Reviews*, *27*, 15–22. https://doi.org/10.1016/j.arr.2016.02.002

El Haj, M., Antoine, P., Nandrino, J.-L., & Kapogiannis, D. (2016). Autobiographical memory decline in Alzheimer's Disease. *Ageing Research Reviews*, *27*, 15–22. https://doi.org/10.1016/j.arr.2016.02.002

Ewers, M., Brendel, M., Rizk-Jackson, A., Rominger, A., Bartenstein, P., Schuff, N., … Alzheimer's Disease Neuroimaging Initiative (ADNI). (2014). Reduced FDG-PET brain metabolism and executive function predict clinical progression in elderly healthy subjects. *NeuroImage. Clinical*, *4*, 45–52. https://doi.org/10.1016/j.nicl.2013.10.018

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Ferman, T. J., Smith, G. E., Kantarci, K., Boeve, B. F., Pankratz, V. S., Dickson, D. W., … Petersen, R. C. (2013). Nonamnestic mild cognitive impairment progresses to dementia with Lewy bodies. *Neurology*, *81*(23), 2032–2038. https://doi.org/10.1212/01.wnl.0000436942.55281.47

Ferreira-Vieira, T. H., Guimaraes, I. M., Silva, F. R., & Ribeiro, F. M. (2016). Alzheimer's Disease: Targeting the Cholinergic System. *Current Neuropharmacology*, *14*(1), 101–115. https://doi.org/10.2174/1570159X13666150716165726

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. SAGE.

Francis, P. T., Palmer, A. M., Snape, M., & Wilcock, G. K. (1999). The cholinergic hypothesis of Alzheimer's disease: A review of progress. *Journal of Neurology, Neurosurgery & Psychiatry*, *66*(2), 137–147. https://doi.org/10.1136/jnnp.66.2.137

Frisoni, G. B., Boccardi, M., Barkhof, F., Blennow, K., Cappa, S., Chiotis, K., … Winblad, B. (2017). Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers. *The Lancet. Neurology*, *16*(8), 661–676. https://doi.org/10.1016/S1474-4422(17)30159-X

Goldman, J. G., Weis, H., Stebbins, G., Bernard, B., & Goetz, C. G. (2012). Clinical differences among mild cognitive impairment subtypes in Parkinson's disease. *Movement Disorders*, *27*(9), 1129–1136. https://doi.org/10.1002/mds.25062

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*(9), 961–969. https://doi.org/10.1001/archgenpsychiatry.2011.96

Gomar, J. J., Conejero-Goldberg, C., Davies, P., Goldberg, T. E., & Alzheimer's Disease Neuroimaging Initiative. (2014). Extension and refinement of the predictive value of different classes of markers in ADNI: Four-year follow-up data. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *10*(6), 704–712. https://doi.org/10.1016/j.jalz.2013.11.009

Hot, P., Rauchs, G., Bertran, F., Denise, P., Desgranges, B., Clochon, P., & Eustache, F. (2011). Changes in sleep theta rhythm are related to episodic memory impairment in early Alzheimer's disease. *Biological Psychology*, *87*(3), 334–339. https://doi.org/10.1016/j.biopsycho.2011.04.002

Humphreys, M. S., Smith, S., Pachana, N. A., Tehan, G., & Byrne, G. J. (2010). Measuring episodic memory: A novel approach with an indefinite number of alternative forms. *Applied Cognitive Psychology*, *24*(8), 1080–1094. https://doi.org/10.1002/acp.1591

Ihara, R., Iwata, A., Suzuki, K., Ikeuchi, T., Kuwano, R., & Iwatsubo, T. (2018). Clinical and cognitive characteristics of preclinical Alzheimer's disease in the Japanese Alzheimer's

Disease Neuroimaging Initiative cohort. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, *4*, 645–651. https://doi.org/10.1016/j.trci.2018.10.004

Iqbal, K., Liu, F., Gong, C.-X., & Grundke-Iqbal, I. (2010). Tau in Alzheimer Disease and Related Tauopathies. *Current Alzheimer Research*, *7*(8), 656–664.

Jack, C. R., Albert, M., Knopman, D. S., McKhann, G. M., Sperling, R. A., Carillo, M., … Phelps, C. H. (2011). Introduction to Revised Criteria for the Diagnosis of Alzheimer's Disease: National Institute on Aging and the Alzheimer Association Workgroups. *Alzheimer's & Dementia : The Journal of the Alzheimer's Association*, *7*(3), 257–262. https://doi.org/10.1016/j.jalz.2011.03.004

Kalat, J. W. (2018). *Biological psychology* (13th edition.). Boston, MA, USA: Cengage.

Kato, Y., Narumoto, J., Matsuoka, T., Okamura, A., Koumi, H., Kishikawa, Y., … Fukui, K. (2013). Diagnostic performance of a combination of Mini-Mental State Examination and Clock Drawing Test in detecting Alzheimer's disease. *Neuropsychiatric Disease and Treatment*, *9*, 581–586. https://doi.org/10.2147/NDT.S42209

Landau, S. M., Mintun, M. A., Joshi, A. D., Koeppe, R. A., Petersen, R. C., Aisen, P. S., … Jagust, W. J. (2012). Amyloid deposition, hypometabolism, and longitudinal cognitive decline. *Annals of Neurology*, *72*(4), 578–586. https://doi.org/10.1002/ana.23650

Larner, A. J. (2019). Evaluating cognitive screening instruments with the "likelihood to be diagnosed or misdiagnosed" measure. *International Journal of Clinical Practice*, *73*(2), e13265. https://doi.org/10.1111/ijcp.13265

Lin, F., Ren, P., Mapstone, M., Meyers, S. P., Porsteinsson, A., & Baran, T. M. (2017). The cingulate cortex of older adults with excellent memory capacity. *Cortex*, *86*, 83–92. https://doi.org/10.1016/j.cortex.2016.11.009

Liu, C.-C., Kanekiyo, T., Xu, H., & Bu, G. (2013). Apolipoprotein E and Alzheimer disease: Risk, mechanisms and therapy. *Nature Reviews Neurology*, *9*(2), 106–118. https://doi.org/10.1038/nrneurol.2012.263

Manly, C. A., & Wells, R. S. (2015). Reporting the Use of Multiple Imputation for Missing Data in Higher Education Research. *Research in Higher Education*, *56*(4), 397–409. https://doi.org/10.1007/s11162-014-9344-9

Michaelson, D. M. (2014). APOE ε4: The most prevalent yet understudied risk factor for Alzheimer's disease. *Alzheimer's & Dementia*, *10*(6), 861–868. https://doi.org/10.1016/j.jalz.2014.06.015

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. https://doi.org/10.1111/j.1600-0447.2008.01326.x

Moradi, E., Hallikainen, I., Hänninen, T., Tohka, J., & Alzheimer's Disease Neuroimaging Initiative. (2017). Rey's Auditory Verbal Learning Test scores can be predicted from whole brain MRI in Alzheimer's disease. *NeuroImage. Clinical*, *13*, 415–427. https://doi.org/10.1016/j.nicl.2016.12.011

Mormino, E. C., Kluth, J. T., Madison, C. M., Rabinovici, G. D., Baker, S. L., Miller, B. L., … Jagust, W. J. (2009). Episodic memory loss is related to hippocampal-mediated β-amyloid deposition in elderly subjects. *Brain*, *132*(5), 1310–1323. https://doi.org/10.1093/brain/awn320

Moscovitch, M. (1992). Memory and Working-with-Memory: A Component Process Model Based on Modules and Central Systems. *Journal of Cognitive Neuroscience*, *4*(3), 257–267. https://doi.org/10.1162/jocn.1992.4.3.257

Moscovitch, Morris, Cabeza, R., Winocur, G., & Nadel, L. (2016). Episodic Memory and Beyond: The Hippocampus and Neocortex in Transformation. *Annual Review of Psychology*, *67*, 105–134. https://doi.org/10.1146/annurev-psych-113011-143733

Murphy, M. P., & LeVine, H. (2010). Alzheimer's Disease and the β-Amyloid Peptide. *Journal of Alzheimer's Disease : JAD*, *19*(1), 311. https://doi.org/10.3233/JAD-2010-1221

Natelson Love, M., Clark, D. G., Cochran, J. N., Den Beste, K. A., Geldmacher, D. S., Benzinger, T. L., … Roberson, E. D. (2017). Clinical, imaging, pathological, and biochemical characterization of a novel presenilin 1 mutation (N135Y) causing Alzheimer's disease. *Neurobiology of Aging*, *49*, 216.e7-216.e13. https://doi.org/10.1016/j.neurobiolaging.2016.09.020

Nathan, P. J., Abbott, R., Lim, Y. Y., Galluzzi, S., Marizzoni, M., Bagnoli, C., … Frisoni, G. B. (2016). PT593. CSF Beta-Amyloid and APOE E4 Related Decline in Episodic Memory over 12 months measured using the CANTAB in individuals with amnestic MCI: Results from the European-ADNI study. *International Journal of Neuropsychopharmacology*, *19*(Suppl_1), 18–19. https://doi.org/10.1093/ijnp/pyw044.593

Nyhus, E., & Curran, T. (2010). Functional Role of Gamma and Theta Oscillations in Episodic Memory. *Neuroscience and Biobehavioral Reviews*, *34*(7), 1023–1035. https://doi.org/10.1016/j.neubiorev.2009.12.014

Perry, R. J., Watson, P., & Hodges, J. R. (2000). The nature and staging of attention dysfunction in early (minimal and mild) Alzheimer's disease: Relationship to episodic and semantic memory impairment. *Neuropsychologia*, *38*(3), 252–271. https://doi.org/10.1016/S0028-3932(99)00079-2

Povova, J., Ambroz, P., Bar, M., Pavukova, V., Sery, O., Tomaskova, H., & Janout, V. (2012). Epidemiological of and risk factors for Alzheimer's disease: A review. *Biomedical Papers of the Medical Faculty of the University Palacky, Olomouc, Czechoslovakia*, *156*(2), 108–114. https://doi.org/10.5507/bp.2012.055

Ramanan, V. K., Kim, S., Holohan, K., Shen, L., Nho, K., Risacher, S. L., … for the Alzheimer's Disease Neuroimaging Initiative (ADNI). (2012). Genome-wide pathway analysis of memory impairment in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort implicates gene candidates, canonical pathways, and networks. *Brain Imaging and Behavior*, *6*(4), 634–648. https://doi.org/10.1007/s11682-012-9196-x

Reiman, E. M., Quiroz, Y. T., Fleisher, A. S., Chen, K., Velez-Pardo, C., Jimenez-Del-Rio, M., … Lopera, F. (2012). Brain imaging and fluid biomarker analysis in young adults at genetic risk for autosomal dominant Alzheimer's disease in the presenilin 1 E280A kindred: A case-

control study. *The Lancet Neurology*, *11*(12), 1048–1056. https://doi.org/10.1016/S1474-4422(12)70228-4

Roberts, R., Geda, Y., Knopman, D., Cha, R., Pankratz, V., Boeve, B., … Rocca, W. (2008). The Mayo Clinic Study of Aging: Design and Sampling, Participation, Baseline Measures and Sample Characteristics. *Neuroepidemiology*, *30*(1), 58–69. https://doi.org/10.1159/000115751

Russo, María J., Campos, J., Vázquez, S., Sevlever, G., & Allegri, R. F. (2017). Adding Recognition Discriminability Index to the Delayed Recall Is Useful to Predict Conversion from Mild Cognitive Impairment to Alzheimer's Disease in the Alzheimer's Disease Neuroimaging Initiative. *Frontiers in Aging Neuroscience*, *9*. https://doi.org/10.3389/fnagi.2017.00046

Russo, María Julieta, Cohen, G., Chrem Mendez, P., Campos, J., Nahas, F. E., Surace, E. I., … Sevlever, G. (2016). Predicting episodic memory performance using different biomarkers: Results from Argentina-Alzheimer's Disease Neuroimaging Initiative. *Neuropsychiatric Disease and Treatment*, *12*, 2199–2206. https://doi.org/10.2147/NDT.S107051

Seo, E. H., Choo, I. H., & Alzheimer's Disease Neuroimaging Initiative. (2016). Amyloid-independent functional neural correlates of episodic memory in amnestic mild cognitive impairment. *European Journal of Nuclear Medicine and Molecular Imaging*, *43*(6), 1088–1095. https://doi.org/10.1007/s00259-015-3261-9

Sharma, N., & Singh, A. N. (2016). Exploring Biomarkers for Alzheimer's Disease. *Journal of Clinical and Diagnostic Research : JCDR*, *10*(7), KE01–KE06. https://doi.org/10.7860/JCDR/2016/18828.8166

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., … Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *BMJ*, *338*, b2393. https://doi.org/10.1136/bmj.b2393

Stites, S. D., Milne, R., & Karlawish, J. (2018). Advances in Alzheimer's imaging are changing the experience of Alzheimer's disease. *Alzheimer's & Dementia : Diagnosis, Assessment & Disease Monitoring*, *10*, 285–300. https://doi.org/10.1016/j.dadm.2018.02.006

Swain, D. G., O'Brien, A. G., & Nightingale, P. G. (1999). Cognitive assessment in elderly patients admitted to hospital: The relationship between the Abbreviated Mental Test and the Mini-Mental State Examination. *Clinical Rehabilitation*, *13*(6), 503–508. https://doi.org/10.1191/026921599670895633

Tan, C.-C., Yu, J.-T., Wang, H.-F., Tan, M.-S., Meng, X.-F., Wang, C., … Tan, L. (2014). Efficacy and Safety of Donepezil, Galantamine, Rivastigmine, and Memantine for the Treatment of Alzheimer's Disease: A Systematic Review and Meta-Analysis. *Journal of Alzheimer's Disease*, *41*(2), 615–631. https://doi.org/10.3233/JAD-132690

Tannenbaum, C., Paquette, A., Hilmer, S., Holroyd-Leduc, J., & Carnahan, R. (2012). A Systematic Review of Amnestic and Non-Amnestic Mild Cognitive Impairment Induced by Anticholinergic, Antihistamine, GABAergic and Opioid Drugs. *Drugs & Aging*, *29*(8), 639–658. https://doi.org/10.1007/BF03262280

Tapiola, T., Pennanen, C., Tapiola, M., Tervo, S., Kivipelto, M., Hänninen, T., … Soininen, H. (2008). MRI of hippocampus and entorhinal cortex in mild cognitive impairment: A follow-

up study. *Neurobiology of Aging*, *29*(1), 31–38. https://doi.org/10.1016/j.neurobiolaging.2006.09.007

Teipel, S. J., Grinberg, L. T., Hampel, H., & Heinsen, H. (2009). Cholinergic System Imaging in the Healthy Aging Process and Alzheimer Disease. *Encyclopedia of Neuroscience*, 857–868. https://doi.org/10.1016/B978-008045046-9.02041-6

Terry, A. V., & Buccafusco, J. J. (2003). The Cholinergic Hypothesis of Age and Alzheimer's Disease-Related Cognitive Deficits: Recent Challenges and Their Implications for Novel Drug Development. *Journal of Pharmacology and Experimental Therapeutics*, *306*(3), 821–827. https://doi.org/10.1124/jpet.102.041616

Tulving, E. (1972). Episodic and semantic memory. In *Organization of memory* (pp. xiii, 423–xiii, 423). Oxford, England: Academic Press.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, *26*(1), 1–12. https://doi.org/10.1037/h0080017

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, 1–19. Retrieved from Academic OneFile.

Uddin, M. S., Kabir, M. T., Al Mamun, A., Abdel-Daim, M. M., Barreto, G. E., & Ashraf, G. M. (2019). APOE and Alzheimer's Disease: Evidence Mounts that Targeting APOE4 may Combat Alzheimer's Pathogenesis. *Molecular Neurobiology*, *56*(4), 2450–2465. https://doi.org/10.1007/s12035-018-1237-z

Vakil, E., & Blachstein, H. (1993). Rey auditory-verbal learning test: Structure analysis. *Journal of Clinical Psychology*, *49*(6), 883–890. https://doi.org/10.1002/1097-4679(199311)49:6<883::AID-JCLP2270490616>3.0.CO;2-6

Velayudhan, L., Proitsi, P., Westman, E., Muehlboeck, J.-S., Mecocci, P., Vellas, B., … dNeuroMed Consortium. (2013). Entorhinal cortex thickness predicts cognitive decline in Alzheimer's disease. *Journal of Alzheimer's Disease: JAD*, *33*(3), 755–766. https://doi.org/10.3233/JAD-2012-121408

Vemuri, P., Wiste, H. J., Weigand, S. D., Shaw, L. M., Trojanowski, J. Q., Weiner, M. W., … Alzheimer's Disease Neuroimaging Initiative. (2009). MRI and CSF biomarkers in normal, MCI, and AD subjects: Predicting future clinical change. *Neurology*, *73*(4), 294–301. https://doi.org/10.1212/WNL.0b013e3181af79fb

Watson, J. L., Ryan, L., Silverberg, N., Cahan, V., & Bernard, M. A. (2014). Obstacles And Opportunities In Alzheimer's Clinical Trial Recruitment. *Health Affairs (Project Hope)*, *33*(4), 574–579. https://doi.org/10.1377/hlthaff.2013.1314

Weingarten, M. D., Lockwood, A. H., Hwo, S. Y., & Kirschner, M. W. (1975). A protein factor essential for microtubule assembly. *Proceedings of the National Academy of Sciences of the United States of America*, *72*(5), 1858–1862.

Yankner, B. A., Duffy, L. K., & Kirschner, D. A. (1990). Neurotrophic and neurotoxic effects of amyloid beta protein: Reversal by tachykinin neuropeptides. *Science*, *250*(4978), 279–282. https://doi.org/10.1126/science.2218531

Yip, A. G., Brayne, C., Easton, D., & Rubinsztein, D. C. (2002). Apolipoprotein E4 is only a weak predictor of dementia and cognitive decline in the general population. *Journal of Medical Genetics*, *39*(9), 639–643.