Experimental Brain Research
https://doi.org/10.1007/s00221-021-06182-w

RESEARCH ARTICLE

Harnessing forgetfulness: can episodic-memory tests predict early Alzheimer's disease?

Samuel L. Warren[1] · Ahmed A. Moustafa[1,2] · Hany Alashwal[3] · for the Alzheimer's Disease Neuroimaging Initiative

Received: 25 March 2021 / Accepted: 16 July 2021

Abstract

A rapid increase in the number of patients with Alzheimer's disease (AD) is expected over the next decades. Accordingly, there is a critical need for early-stage AD detection methods that can enable effective treatment strategies. In this study, we consider the ability of episodic-memory measures to predict mild cognitive impairment (MCI) to AD conversion and thus, detect early-stage AD. For our analysis, we studied 307 participants with MCI across four years using data from the *Alzheimer's Disease Neuroimaging Initiative* (ADNI). Using a binary logistic regression, we compared episodic-memory tests 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). We also combined variables to test the accuracy of mixed-predictor models. Our results indicated that the best predictors of MCI to AD conversion were the following: a combined episodic-memory and neuroimaging model in year one (59.8%), the Rey Auditory Verbal Learning Test in year two (71.7%), a mixed episodic-memory predictor model in year three (77.7%) and the Logical Memory Test in year four (77.2%) of ADNI. Overall, we found that individual episodic-memory measure and mixed models performed similarly when predicting MCI to AD conversion. Comparatively, individual neuroimaging measures predicted MCI conversion worse than chance. Accordingly, our results indicate that episodic-memory tests could be instrumental in detecting early-stage AD and enabling effective treatment.

Keywords Alzheimer's disease (AD) · Episodic memory · Mild cognitive impairment (MCI) · Alzheimer's disease neuroimaging initiative (ADNI) · Disease prediction · Big data

Introduction

Alzheimer's disease (AD) is the second leading cause of death in Australia and is, consequentially, one of the greatest medical threats of our time (Australian Bureau of Statistics 2017). AD is a progressive degenerative disorder that can lead to severe dementia. There is a critical need for treatments that can counteract AD; however, most treatments are ineffective due to the late stage of AD diagnoses, the complex multifaceted nature of the disorder, and a fixation on singular theories of AD progression (e.g., the amyloid cascade hypothesis) (Banik et al. 2015; Rasmussen and Langerman 2019). Accordingly, the creation of new early-stage diagnostic methods could help enable effective treatments for AD. The Alzheimer's Association estimates that the formulation of early-stage diagnostic methods could revolutionise AD treatment saving the US alone $7.9 trillion (Alzheimer's Association 2018). Moreover, early-stage diagnoses and treatment are predicted to be instrumental in improving the quality of life of those living with AD (Chu 2012). Therefore, it is crucial to develop new diagnostic methods that can detect early-stage AD and, thus, combat the disease.

In research, neuroimaging, biomarker, and cognitive measures are commonly used to detect early-stage AD. Specifically, these measures are used to predict mild cognitive impairment (MCI) to AD conversion and, thus, discover

Communicated by Bill J Yates.

✉ Samuel L. Warren
19466949@student.westernsydney.edu.au

[1] School of Psychology, Western Sydney University, Sydney, Australia

[2] MARCS Institute for Brain and Behaviour, Western Sydney University, Sydney, Australia

[3] College of Information Technology, United Arab Emirates University, Al-Ain 15551, United Arab Emirates

Published online: 27 July 2021

markers of early-stage AD (Gainotti et al. 2014; Ottoy et al. 2019). There is evidence to suggest that neuroimaging and cognitive measures can detect MCI to AD conversion; however, these markers also have significant limitations. For example, neuroimaging and biomarker methods are often invasive or unaffordable to patients and, therefore, lack widespread clinical use (Jack et al. 2011). Moreover, cognitive tests (specifically, general multi-domain tests), which are commonly used to diagnose mid-stage AD, cannot accurately diagnose early-stage AD and are prone to misdiagnosing forms of dementia (e.g., diagnosing vascular dementia as AD) (Bak et al. 2005; Arevalo-Rodriguez et al. 2015; Larner 2019). In the literature, it is common for studies to combine neuroimaging, biomarker, demographic, and cognitive measures to predict MCI to AD conversion; however, although these combined models boast a high accuracy, they rarely overcome the limitations of each method discussed above. Accordingly, each category of AD research requires significant innovation to effectively diagnose early-stage AD. In this study, we investigate how cognitive tests can be adapted to predict MCI to AD conversion, and thus detect early-stage AD.

Research has suggested that specific tests, that measure the initial cognitive symptoms of AD, could be used to diagnose the early stages of the disorder (Bastin and Salmon 2014; Brown 2015). For example, the Rey Auditory Verbal Learning Test (RAVLT; a specific episodic-memory test) is theorised to detect early-stage AD because episodic-memory decline is one of the first symptoms of the disease (Eckerström et al. 2013; El Haj et al. 2016). In theory, specific tests are uniquely equipped to detect early-stage AD as they can assess a singular cognitive domain in detail. The detail of specific tests is distinctly opposite to current general multi-domain tests which rely on broad assessments and late-stage symptoms of AD. Episodic memory tests are also thought to perform better than other types of specific tests (e.g., visual memory) because episodic-memory decline is one of the earliest cognitive symptoms of AD (Eckerström et al. 2013; El Haj et al. 2016). Consequently, specific episodic-memory tests may be able to overcome the limitations of general multi-domain tests. Furthermore, specific episodic-memory tests (e.g., RAVLT) may also be able to overcome some of the limitations of neuroimaging and biomarker measures as they are easier to administer. Accordingly, this study investigates the ability of episodic-memory tests to predict MCI to AD conversion.

Various studies have found that episodic-memory measures can predict MCI to AD conversion as a part of a mixed-predictor model. For example, episodic-memory measures have been found as strong predictors of MCI to AD conversion when combined with cortical thickness, hippocampal atrophy, neuron connectivity, and amyloid-beta measures (Gomar et al. 2011; Cai et al. 2015; Nathan et al. 2016; Moradi et al. 2017; Russo et al. 2017; Ihara et al. 2018). However, because of the prominence of mixed-predictor models, episodic-memory measures are often studied with little focus**.** Accordingly, the accuracy and ability of individual episodic-memory tests are not fully understood when predicting MCI to AD conversion. Some preliminary research has suggested that episodic-memory measures could individually predict MCI to AD conversion with a high accuracy (Chapman et al. 2011; Irish et al. 2011; Derby et al. 2013; Gomar et al. 2014; De Simone et al. 2019). However, in the current literature, studies often neglect to research multiple episodic-memory tests and compare episodic-memory measures to other prominent markers of early-stage AD (e.g., general multi-domain tests and neuroimaging markers). Accordingly, it is unclear how accurate individual episodic-memory measures are, how they compare to prominent predictors of early-stage AD, and if they perform better than mixed predictive models. Consequently, there is a critical need for researchers to further evaluate and understand episodic-memory predictors of early-stage AD.

In this study, we investigate the ability of episodic-memory measures to predict early-stage AD. To our knowledge, no study has sought to evaluate and compare the predictive ability of multiple episodic-memory measures in-depth (e.g., the comparison of multiple episodic-memory tests, multivariate models, and neuroimaging variables in a big dataset). Moreover, there is a lacking understanding of how episodic-memory tests compare to one another and other predictors of early-stage AD (e.g., MRI). Subsequently, in this study, we evaluate the accuracy of multiple episodic-memory tests when predicting MCI to AD conversion. We also evaluate episodic-memory tests both individually and in mixed models to understand the optimal method for early-stage AD detection. Finally, we seek to understand the viability of episodic-memory measures by comparing them to other prominent neuroimaging markers of early-stage AD (volumetric MRI). Accordingly, in this study, we aim to do the following:

(1) Determine the accuracy of each episodic-memory test when predicting MCI to AD conversion.
(2) Create and assess mixed-predictor models of MCI to AD conversion.
(3) Investigate whether neuroimaging predictors can outperform or improve prior episodic-memory and mixed predictive models.

Methods

Data source and acquisition

The data for this project was acquired from the *Alzheimer's disease neuroimaging initiative* (ADNI) database. ADNI is a

research organisation (led by Principal Investigator Michael W. Weiner, MD) that investigates neuroimaging, biomarker, and neuropsychological markers of AD and MCI. We chose to use ADNI for this study as the initiative has an extensive longitudinal database that includes numerous episodic-memory and neuroimaging variables. Moreover, ADNI is consistently used throughout the literature to study MCI to AD conversion. The only prerequisite for obtaining ADNI data was a data access application and institutional ethics approval. Our study gained ethics approval from *Western Sydney Universities* Human Research Ethics Committee in March 2019.

Participants

ADNI recruits all their participants through traditional media (e.g., newspapers), new media (e.g., social media), third-party health providers, and their own website. All data collection is performed at ADNI sites or partnering organisations in North America. At these sites, ADNI collects neuroimaging, biomarker, neuropsychological, and demographic data using trained clinicians. These professionals also conduct and confirm participant diagnoses. The diagnoses for MCI and AD are performed using measures of subjective memory concern, the Logical Memory test (LMT; only the delayed recall task), Mini-mental state exam (MMSE), Clinical dementia rating scale (CDR), and the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association Alzheimer's Criteria (NINCDS-ADRDA). For more information on ADNI's diagnostic methods and criteria see the ADNI2 Procedures Manual (Alzheimer's Disease Neuroimaging Initiative 2008, p. 27).

For this project, we studied 307 participants with MCI from the ADNI2 cohort. We specifically chose to research the ADNI2 cohort as, out of all four ADNI cohorts, ADNI2 is the most recently completed study. However, we only included new MCI participants from the ADNI2 cohort as old participants carrying over from previous initiatives had different baselines and measures that could not be easily compared. Our MCI participants formed two groups dictated by their diagnosis over the first four years of ADNI2. Specifically, participants were sorted into a converter group if they converted from MCI to AD and into a non-converter group if they remained stable over the first four years of ADNI2. These groups were made to accurately address whether disease markers could predict MCI to AD conversion. The resulting population contained 95 converters and 212 non-converters at baseline.

Across the years of ADNI2, there was a gradual increase in participant drop out 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. In our sample, some participants also sporadically missed testing dates across the four years of ADNI2 (e.g., some participants measured at baseline missed the year one follow-up but attended the year two assessment). Accordingly, while there was a total of 307 participants studied across the four years of ADNI2, no individual year measured 307 participants (besides baseline). It is also important to note that we excluded the fifth years of ADNI2 from our analysis. As, after running a sample size analysis using the software G*Power (critical $z = -1.96$ and actual power $= 0.953$), we concluded that the fifth year did not have enough participants to warrant analysis (Faul et al. 2007, 2009).

Episodic memory variables

All variables were obtained from the ADNIMERGE.csv dataset, which incorporates the most common measures of AD across all ADNI participants and cohorts. From the ADNIMERGE dataset, we selected the LMT, RAVLT, Alzheimer's Disease Assessment Scale (ADAS-cog), and the ADNI composite memory score (ADNIMEM) as measures of episodic memory. We chose the LMT and RAVLT as they are specific tests that are widely used to measure episodic memory (Alzheimer's Disease Neuroimaging Initiative 2016, p. 21). The LMT is a subtest of the Weschler Memory Scale, which assesses episodic-memory formulation and recollection using a short story (Abikoff et al. 1987). In contrast, the RAVLT assesses episodic memory using a list-learning strategy that measures delayed word recall (Vakil and Blachstein 1993). We also chose the ADAS-cog which is a general multi-domain test with an episodic-memory component (the ADASQ4) (Gomar et al. 2011; Crane et al. 2012). Finally, we included the ADNIMEM measure as studies have suggested that a combined total episodic-memory score from all the tests above can strongly predict AD (Seo et al. 2016). ADNI calculates ADNIMEM using specific items from the RAVLT, ADAS-cog, LMT and MMSE.

It should be noted that the ADNIMERGE dataset reports multiple measures for some of the tests used. Specifically, the RAVLT contains measures of forgetting, immediate recall, learning, and percent forgetting (the percentage of words forgotten). Conversely, the ADAS-cog contains variables that represent different tests. The ADAS13 is the thirteen-question version of the ADAS-cog, the ADAS11 is the eleven-question variant of the ADAS-cog, and the ADASQ4 is the episodic memory (delayed word recall) portion of the ADAS-cog. The remaining LMT and ADNIMEM tests were only reported as single measures. This is appropriate as ADNIMEM is only a total score and because one portion of the LMT was used for diagnoses. Specifically, the LMT can be split into two immediate recall and delayed recall measures. Our study only used the LMT immediate recall

measure because the delayed recall measure was used by ADNI to diagnose MCI and AD. For more information on the cognitive measures used by ADNI see the ADNI2 Procedures Manual (Alzheimer's Disease Neuroimaging Initiative 2008, p. 119).

Neuroimaging variables

Our neuroimaging data was also acquired from the ADNIMERGE dataset which contains volumetric magnetic resonance imaging (MRI) measures for most ADNI participants. The University of California San Francisco (UCSF) processes and analyses all the neuroimaging data used in this study as part of a partnership with ADNI. Specifically, the UCSF use FreeSurfer software to clean and analyse raw MRI data, visually reconstruct the cortex, and segment brain regions into volumes for analysis. For more information about FreeSurfer, see the Athinoula (2019) the Athinoula 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 ADNIMERGE dataset were determined via ADNI according to the literature. The specific variables we obtained from the ADNIMERGE dataset were MRI ventricle, hippocampal, whole brain, entorhinal, fusiform, and medial temporal lobe volume.

Missing data

The ADNIMERGE dataset contained missing data that complicated our analysis. Accordingly, we ran a missing values pattern analysis to visualise the data and inform the best way to deal with the missing values. Usually, a study would omit missing values if under five per cent; however, this was not a comprehensive solution. In our dataset, all the episodic-memory measures only had approximately 2% of their data missing, meaning that we could omit the missing data. However, there was a severe amount of missing neuroimaging data in years three and four of ADNI2 that could not be omitted. 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. The missing data was so large for the latter two years of ADNI2 that 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 remaining 20% of missing data in the first two years was fixed using multiple imputations. Using SPSS 26, 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 Sterne et al. (2009) and Manly and Wells (2015).

Statistical analysis

Our analysis individually assessed the first four years of ADNI2. We thought that it was safest to compare measures within years rather than across the whole initiative because of the problems with missing data and participant dropout. Moreover, we wanted to observe the predictive ability of each episodic-memory measure during different stages of disease progression. Within each year, we individually assessed all three research questions derived from the aims specified in the introduction. Specifically, our research questions are as follows: (1) Can episodic-memory measures predict MCI to AD conversion, (2) Can mixed-predictor models outperform individual episodic-memory tests, and (3) Can neuroimaging measures outperform or improve episodic-memory and mixed-predictor models of MCI to AD conversion?

For the first research question, we tested which episodic-memory measures (RAVLT, ADNIMEM, LMT, ADAS-cog) could best predict MCI to AD conversion. Specifically, we used a binary logistic regression to compare the predictive accuracy of each episodic-memory measures against the observed conversion in the ADNI database. For the second research question, we constructed mixed-predictor models using our episodic-memory measures. Specifically, we used a block-wise hierarchical binary logistic regression to create mixed-predictor models of MCI to AD conversion. We used a forced entry hierarchical method to input our data into the model because alternative methods (stepwise data entry) are open to suppression effects that bias variables. Accordingly, we entered variables in the order of the ADNIMEM, LMT, RAVLT, and ADAS-cog. The best predictive model must have contained the largest accounted variance, highest accuracy, only significant predictors, and could not violate the assumptions of a logistic regression. For the third research question, we determined the predictive ability of neuroimaging measures and then created a mixed-predictor model (using both episodic memory and neuroimaging variables). For this analysis, we used the same logistic regression methods as noted above. In the mixed-predictor models, variables were entered in order of their individual predictive accuracy.

Results

Descriptive statistics

Before the primary analysis, we ran descriptive statistics and frequency analyses to understand our population. Table 1

below summarises the baseline age, education, gender, MMSE, and CDR statistics of all participants. Both the conversion and non-conversion groups contained similar means and standard deviations with some variation in CDR scores and group sizes. At baseline, our participants had an approximate mean age of 71 and a standard deviation of seven years. Both groups had the same educational values with a mean education of 16 years and a standard deviation of 2.6 years. These education levels are considered high and may affect the generalisability of the results.

We also plotted participant diagnosis across the four years of ADNI2 (see Fig. 1 below). Approximately 20–25% of our participants converted from MCI to AD. This conversion coincides with the MCI to AD conversion rate observed across the literature (Mitchell and 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. However, it should be noted that the conversion rate increased above expectations in year four. It is unknown whether the observed increase in conversions is due to disease progression or participant dropout. Overall, 21% of our participants converted from MCI to AD and 79% remained stable with MCI.

Episodic memory predictors of MCI to AD conversion

ADNI2 year one

In this section, we examined the ability of episodic-memory measures to predict MCI to AD conversion. Specifically, we used binary logistic regressions to understand the predictive odds and accuracy of each episodic-memory measure. Following the assessment of individual measures, we combined the episodic-memory variables using a hierarchical binary logistic regression to form a mixed predictive model of MCI to AD conversion for comparison.

In the first year of ADNI2, all 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 episodic-memory measure are detailed in Table 2 below. It is important to note that the results for the LMT, ADASQ4, RAVLT Forgetting and RAVLT Percent Forgetting could not be interpreted because they all violated the assumptions of a binary logistic regression (linearity of the logit). Out of the interpretable results, the ADNIMEM composite memory score had the best accuracy and accounted variance when predicting MCI to AD conversion [C&S $R^2 = 0.286$, Nagelkerke $R^2 = 0.402$, Conv-Acc. = 56.4%, $p < 0.001$]. The next best predictors of MCI to AD conversion were the ADAS13 [Conv-Acc. = 50.2%], ADAS11 [Conv-Acc. = 42.6%], and RAVLT Immediate [Conv-Acc. = 41.4%].

Table 1 Participant descriptive statistics at baseline

Diagnosis	Age	Gender (F/M)	Education	MMSE	CDR
MCI converter	71.2 (±7.3)	98/125	16 (±2.67)	28 (±1.65)	1.3 (±0.75)
MCI non-converter	72.4 (±7.3)	44/50	16 (±2.59)	27 (±1.77)	2.2 (±1.00)

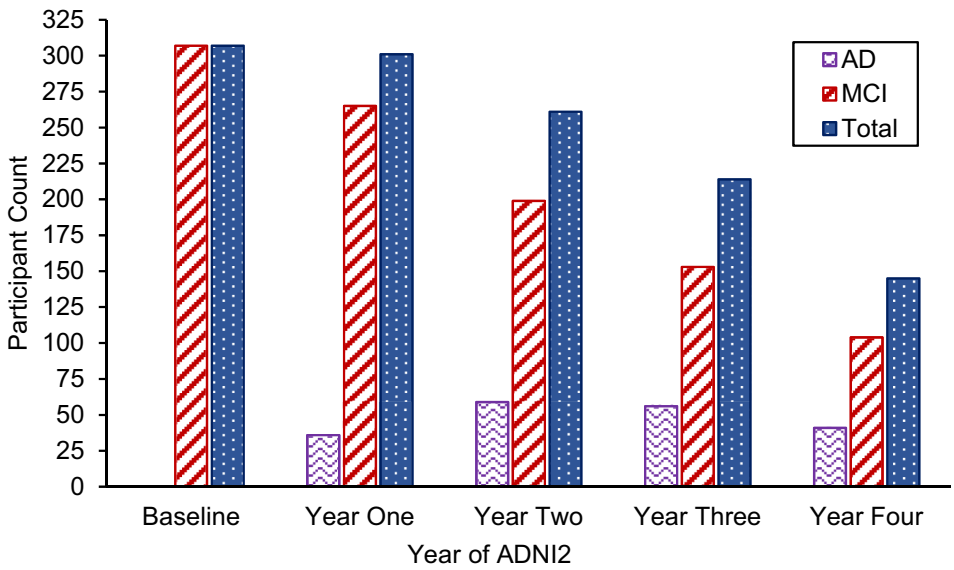

Fig. 1 Participant diagnosis across ADNI2

Table 2 Episodic memory predictors of MCI to AD conversion in year one and two of ADNI2

ADNI year	EM measures	*B*	S.E.	Wald	Sig.	Exp (*B*)	C&S R^2	N R^2	Conv acc.	Model acc.
1	ADNI_MEM	− 1.950	0.096	412.857	< 0.001	0.142	0.286	0.402	56.4	78.1
	RAVLT forgetting	*0.058*	*0.019*	*9.308*	*0.002*	*1.059*	*0.004*	*0.006*	*0.0*	*68.8*
	RAVLT immediate	− 0.122	0.007	331.591	< 0.001	0.885	0.227	0.319	41.4	72.4
	Logical memory test	*− 0.310*	*0.015*	*433.988*	*< 0.001*	*0.733*	*0.269*	*0.378*	*60.6*	*77.7*
	RAVLT percent forgetting	*0.030*	*0.002*	*267.456*	*< 0.001*	*1.031*	*0.154*	*0.217*	*63.8*	*71.7*
	RAVLT learning	− 0.371	0.023	256.490	< 0.001	0.690	0.142	0.200	38.1	75.0
	ADAS 13	0.177	0.009	423.350	< 0.001	1.193	0.274	0.385	50.2	77.5
	ADAS Q4	*0.450*	*0.023*	*386.623*	*< 0.001*	*1.569*	*0.226*	*0.318*	*53.2*	*75.7*
	ADAS 11	0.237	0.012	382.434	< 0.001	1.267	0.247	0.348	42.6	77.4
2	ADNI_MEM	− 2.240	0.112	402.790	< 0.001	0.106	0.358	0.501	63.9	81.6
	RAVLT forgetting	0.100	0.019	27.591	< 0.001	1.105	0.016	0.022	0.0	66.9
	RAVLT immediate	− 0.124	0.007	313.531	< 0.001	0.883	0.255	0.356	51.8	76.4
	Logical memory test	*− 0.396*	*0.018*	*460.720*	*< 0.001*	*0.673*	*0.382*	*0.533*	*70.7*	*81.1*
	RAVLT percent forgetting	0.039	0.002	277.774	< 0.001	1.039	0.213	0.297	71.7	75.6
	RAVLT learning	− 0.362	0.025	208.274	< 0.001	0.696	0.139	0.194	34.9	71.1
	ADAS 13	0.207	0.010	436.523	< 0.001	1.230	0.386	0.539	69.8	84.3
	ADAS Q4	0.564	0.027	431.223	< 0.001	1.758	0.324	0.452	66.3	82.0
	ADAS 11	0.290	0.014	404.279	< 0.001	1.337	0.366	0.511	60.2	82.4

Italic variables violated assumption checks and could not be interpreted

B regression coefficient used to calculate Exp(*B*), *Exp(B)* predictive odds ratio, *C&S* R^2 Cox & Snell *R*-squared, *N*-R^2 nagelkerke *R*-squared, *Conv acc.* conversion accuracy, *S.E.* standard error, *Model acc.* model accuracy

In year one, the best mixed episodic-memory predictor model consisted of the ADASQ4 and RAVLT Immediate (see Table 3 below). However, the mixed episodic-memory model [C&S $R^2 = 0.266$, N-$R^2 = 0.374$, Conv. Accuracy = 53.1%, $p < 0.001$] predicted MCI to AD conversion worse than the ADANIMEM model (but only marginally worse).

ADNI2 year two

In the second year of ADNI2, all episodic-memory measures could predict MCI to AD conversion besides the RAVLT Forgetting. Overall, episodic-memory measures performed better in the second year than they did in the first year. However, the LMT violated the assumption of linearity of the logit and could not be interpreted (see Table 2). The best predictors of conversion were the RAVLT Percent Forgetting [C&S $R^2 = 0.213$, N-$R^2 = 0.297$, Conv-Acc. = 71.7%, $p < 0.001$], ADAS 13 [Conv-Acc. = 69.8%], ADASQ4 [Conv-Acc. = 66.3%], and ADNIMEM [Conv-Acc. = 63.9%].

When constructing a mixed model, no significant model could be found that matched or outperformed our 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.

Table 3 The best mixed episodic-memory model in year one of ADNI2

Model	C&S R^2	N-R^2	Sig.	Conv acc.	Model acc.
Mixed EM Model Y1	0.266	0.374	< 0.001	53.1	77.1
Components	*B*	S.E.	Wald	Sig.	Exp(*B*)
RAVLT immediate	− 0.078	0.008	98.489	< 0.001	0.925
ADAS Q4	0.281	0.028	101.285	< 0.001	1.324
Constant	0.011	0.363	0.001	0.976	1.011

B regression coefficient used to calculate Exp(*B*), *Exp(B)* predictive odds ratio, *S.E.* standard error, C&S R^2 Cox & Snell *R*-squared, N-R^2 Nagelkerke *R*-squared, *Conv acc.* conversion accuracy, *Model acc.* model accuracy

ADNI2 year three

In the third year of ADNI2, the predictive accuracy of the episodic-memory measures continued to improve (see the conversion accuracy in Table 4 below). Unlike prior years, all measures met the assumptions of the linearity of the logit and could, therefore, be interpreted. Once again, all episodic-memory measures were significant besides the RAVLT forgetting variable. The LMT [C&S $R^2 = 0.374$, N-$R^2 = 0.521$, Conv-Acc. = 76.9%, $p < 0.001$] was the single best predictor followed by the ADASQ4 [Conv-Acc. = 71.6%], ADNIMEM [Conv-Acc. = 67.6%], and the ADAS13 [Conv-Acc. = 67.3%].

In the third year of ADNI2, the best mixed episodic-memory predictive model of MCI to AD conversion contained the LMT, ADNIMEM, and ADASQ4 measures (see Table 5 below). The mixed episodic-memory model [Conv-Acc. = 77.7%, $p < 0.001$] performed marginally better than the LMT model. However, it could be argued that the disparity in accuracy between the mixed model and LMT is not large enough to justify using three measures (including the LMT) compared to just using the LMT. Nonetheless, year three provided strong evidence for the viability of mixed predictive models.

Table 4 Episodic memory predictors of MCI to AD conversion in year three and four of ADNI2

ADNI year	EM measures	*B*	S.E.	Wald	Sig.	Exp(*B*)	C&S R^2	N R^2	Conv acc.	Model acc.
3	ADNI_MEM	− 2.445	0.134	333.229	< 0.001	0.087	0.409	0.571	67.6	83.2
	RAVLT forgetting	0.025	0.023	1.248	0.264	1.026	0.001	0.001	0.0	67.4
	RAVLT immediate	−0.151	0.009	283.746	< 0.001	0.860	0.321	0.447	55.6	76.9
	Logical memory test	− 0.367	0.019	358.583	< 0.001	0.693	0.374	0.521	76.9	82.9
	RAVLT percent forgetting	0.044	0.003	262.105	< 0.001	1.045	0.248	0.346	67.1	75.3
	RAVLT learning	− 0.405	0.027	228.073	< 0.001	0.667	0.200	0.279	51.2	74.0
	ADAS 13	0.233	0.012	351.517	< 0.001	1.262	0.453	0.631	67.3	85.5
	ADAS Q4	0.608	0.031	383.384	< 0.001	1.837	0.364	0.508	71.6	82.1
	ADAS 11	0.311	0.017	340.292	< 0.001	1.365	0.425	0.593	65.8	84.5
4	ADNI_MEM	− 2.284	0.149	233.721	< 0.001	0.102	0.400	0.563	71.1	85.6
	RAVLT forgetting	0.018	0.029	0.372	0.542	1.018	0.000	0.001	0.0	69.4
	RAVLT immediate	− 0.142	0.010	199.219	< 0.001	0.868	0.301	0.425	57.1	78.6
	Logical memory test	− 0.311	0.021	220.538	< 0.001	0.733	0.327	0.461	77.2	84.1
	RAVLT percent forgetting	0.037	0.003	148.148	< 0.001	1.037	0.191	0.270	65.3	72.2
	RAVLT learning	− 0.402	0.037	121.157	< 0.001	0.669	0.144	0.204	40.1	74.1
	ADAS 13	0.198	0.012	252.225	< 0.001	1.219	0.431	0.608	75.5	88.3
	ADAS Q4	0.637	0.039	266.060	< 0.001	1.891	0.363	0.513	75.0	85.5
	ADAS 11	0.262	0.017	238.778	< 0.001	1.299	0.406	0.573	73.2	87.6

B regression coefficient used to calculate Exp(*B*), *Exp(B)* predictive odds ratio, *C&S* R^2 Cox & Snell *R*-squared, *N*-R^2 Nagelkerke *R*-squared, *Conv acc.* conversion accuracy, *S.E.* standard error, *Model acc.* model accuracy

Table 5 The best mixed episodic-memory model in year three of ADNI2

EM measures	C&S R^2	N-R^2	Sig.	Conv acc.	Model acc.
Mixed EM Model Y3	0.428	0.598	< 0.001	77.7	85.1
Components	*B*	S.E.	Wald	Sig.	Exp(*B*)
ADNI_MEM	− 1.250	0.219	32.535	< 0.001	0.287
Logical memory	− 0.157	0.024	41.174	< 0.001	0.855
ADAS Q4	0.136	0.053	6.644	0.010	1.145
Constant	− 1.007	0.364	7.652	0.006	0.365

B regression coefficient used to calculate Exp(*B*), *Exp(B)* predictive odds ratio, *S.E.* standard error, *C&S* R^2 cox & Snell *R*-squared, *N*-R^2 Nagelkerke *R*-squared, *Conv acc.* conversion accuracy, *Model acc.* model accuracy

ADNI2 year four

In the fourth and final year of ADNI2, our analysis confirmed the trend that the accuracy of episodic-memory measures improves over time. As observed in prior years, RAVLT Forgetting could not predict MCI to AD conversion and was not significant. The best episodic-memory predictor was the LMT which predicted MCI to AD conversion with an accuracy of 77.2% (see Table 4 above). 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 $R^2 = 0.327$, N-$R^2 = 0.461$] and was greatly surpassed by the ADAS13 [C&S $R^2 = 0.431$, N-$R^2 = 0.608$] and the ADASQ4 [C&S $R^2 = 0.363$, N-$R^2 = 0.513$]. Accordingly, the LMT, the ADASQ4 or the ADAS13 could be considered the best predictive measures of MCI to AD conversion depending on research preferences. 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 measures and meet the assumption of multicollinearity.

Neuroimaging and mixed-predictor models of MCI to AD conversion

After we determined the best episodic-memory predictors of MCI to AD conversion, we compared them to neuroimaging measures to find the best measure of early-stage AD. Specifically, we assessed the predictive ability of each neuroimaging variable using a binary logistic regression. Following the individual analysis, we used a hierarchical binary logistic regression to determine whether a combined neuroimaging and episode memory predictive model could outperform prior models. It is important to note that the third and fourth years of ADNI2 neuroimaging data could not be assessed due to problems with missing data that are discussed in the Methods section (p.6).

In the first year of ADNI2, the neuroimaging markers were tested for their predictive odds and accuracy. All variables were statistically significant yet had problematic odds ratios [approximately Exp(B) = 1.00]. This meant that we could not interpret the odds ratios. All the neuroimaging variables performed very poorly when predicting MCI to AD conversion, as can be seen in Table 6 below. The best 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, neuroimaging variables also performed worse than episodic-memory measures in the first year of ADNI2.

In the first year of ADNI2, the best mixed neuroimaging and episodic-memory model contained the ADAS11, the LMT and MRI entorhinal variables. The mixed model predicted MCI to AD conversion with a higher accuracy than both the individual and composite episodic-memory measures (see Table 7 below). Specifically, the mixed model predicted disease conversion with an accuracy of 59.8% compared to the 56.4% conversion accuracy of the ADNI-MEM model.

In the second year of ADNI2, all our variables generally improved in predictive ability compared to the first year; however, the odds ratios were still problematic. Specifically, almost all variables have a positive odds ratio of Exp(B) = 1.00 and could not be interpreted; however,

Table 6 MRI predictors of MCI to AD conversion in year one and two of ADNI2

ADNI year	MRI measures	*B*	S.E.	Wald	Sig.	Exp(*B*)	C&S R^2	N R^2	Conv acc.	Model acc.
1	UCSF ventricles	0.000	0.000	5.736	0.038	1.000	0.020	0.027	1.95	68.15
	UCSF hippocampus	−0.001	0.000	30.879	< 0.001	0.999	0.115	0.161	27.32	71.58
	UCSF whole brain	0.000	0.000	6.733	0.025	1.000	0.023	0.032	1.07	68.85
	UCSF entorhinal	− 0.001	0.000	32.048	< 0.001	0.999	0.120	0.169	27.85	71.26
	UCSF fusiform	0.000	0.000	15.833	0.001	1.000	0.057	0.081	8.17	69.17
	UCSF med temp	0.000	0.000	19.031	0.001	1.000	0.067	0.095	14.17	69.55
2	UCSF ventricles	0.000	0.000	6.912	0.022	1.000	0.027	0.038	4.4	67.97
	UCSF hippocampus	− 0.001	0.000	37.006	< 0.001	0.999	0.169	0.235	42.4	75.08
	UCSF whole brain	0.000	0.000	7.680	0.016	1.000	0.031	0.044	4.2	68.17
	UCSF entorhinal	− 0.001	0.000	38.019	< 0.001	0.999	0.174	0.243	43.4	75.17
	UCSF fusiform	0.000	0.000	17.822	0.002	1.000	0.077	0.107	19.68	71.07
	UCSF med temp	0.000	0.000	22.626	< 0.001	1.000	0.097	0.135	30.3	73.27

B regression coefficient used to calculate Exp(*B*), *Exp(B)* predictive odds ratio, *C&S* R^2 Cox & Snell *R*-squared, *N*-R^2 Nagelkerke *R*-squared, *Conv acc.* conversion accuracy, *S.E.* standard error, *Model acc.* model accuracy

Table 7 The best mixed MRI and episodic-memory models in years one and two of ADNI2

Mixed Models	C&S R^2	N-R^2	Sig.	Conv acc.	Model acc.
Mixed MRI + EM Model Y1	0.323	0.454	< 0.001	59.8	80.4
Mixed MRI + EM Model Y2	0.408	0.570	< 0.001	70.5	83.9
Model	Components	*B*	S.E.	Sig.	Exp(*B*)
Mixed MRI + EM Model Y1	ADAS 11	0.130	0.037	< 0.001	1.139
	Logical memory test	− 0.190	0.046	< 0.001	0.827
	UCSF entorhinal	0.000	0.000	0.044	1.000
	Constant	0.528	0.978	0.590	1.695
Mixed MRI + EM Model Y2	ADAS 13	0.188	0.027	< 0.001	1.207
	UCSF entorhinal	− 0.001	0.000	0.018	0.999
	Constant	− 1.930	1.176	0.103	0.145

B regression coefficient used to calculate Exp(*B*), *Exp(B)* predictive odds ratio, *S.E.* standard error, *C&S* R^2 Cox & Snell *R*-squared, *N*-R^2 nagelkerke *R*-squared, *Conv acc.* conversion accuracy, *Model Acc.* model accuracy

the Wald statistics could still be used to indicate predictive odds. We theorised that the problematic odds ratios occurred due to the multiple imputations, however, this interaction is beyond the scope of this study. 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 6 above). All remaining variables performed poorly in comparison when predicting disease conversion.

In year two of ADNI2, the best mixed model contained the ADAS13 and MRI entorhinal variables. 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 Tables 7 above). However, the model performed worse by a negligible amount. Accordingly, the best predictor of MCI to AD conversion in the second year of ADNI 2 was the RAVLT Percent Forgetting episodic-memory measure.

When assessing individual predictors of MCI to AD conversion, we found that individual episodic-memory measures were the best predictors of MCI to AD conversion (see Fig. 2 above). Comparatively, individual neuroimaging markers could not predict MCI to AD conversion above chance averages. When assessing mixed-predictor models (both episodic memory and neuroimaging), we found that all models could predict MCI to AD conversion similarly to individual episodic-memory tests. Specifically, each mixed model varied approximately ± 2.2% in accuracy from the

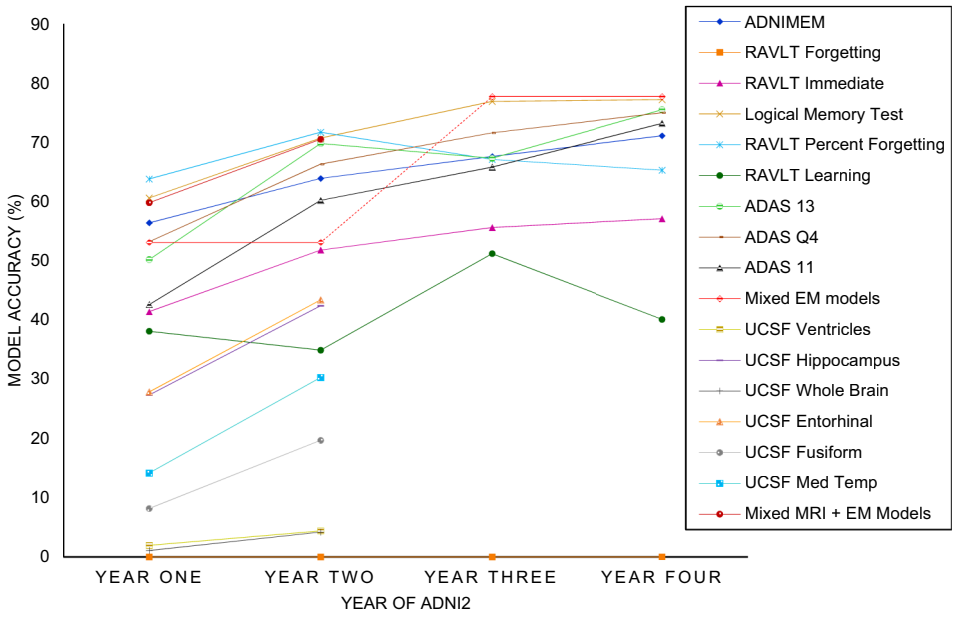

Fig. 2 The accuracy of all models over ADNI2. The dotted lines indicate that no mixed episodic memory (EM) models were created in years two and four of ADNI2

best episodic-memory measure each year. Accordingly, in some circumstances, mixed models may be a viable alternative to episodic-memory tests; however, when factoring in the logistical complexity of multiple measures, episodic-memory tests remain the most reliable and clinically viable option.

Discussion

In this study, we investigated the ability of several episodic-memory measures to predict early-stage AD. Specifically, we used episodic-memory tests to predict MCI to AD conversion, and thus detect markers of early-stage AD. While other studies have examined episodic memory as a predictor of MCI to AD conversion, this is the first study that assesses all ADNI episodic-memory measures in depth and compares them to common neuroimaging predictors of AD. For example, we evaluated episodic-memory measures individually and in mixed models. The following section discusses our results and the implications of our findings.

In research question one, we evaluate the ability of individual episodic-memory tests to predict MCI to AD conversion. We found that, across the first four years of ADNI2, all episodic-memory tests could predict disease conversion, apart from the RAVLT Forgetting (which was consistently not significant). Specifically, the following individual episodic-memory measures were the best predictors of MCI to AD conversion: 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%). Across the whole of ADNI2, the LMT, the ADASQ4, the ADAS13, ADNIMEM and the RAVLT Percent Forgetting were consistently strong predictors of disease conversion. We also found that over the years, all episodic-memory tests improved in predictive accuracy. When assessing the best type of cognitive test to use, we found that specific episodic-memory measures, general multi-domain tests, and composite memory tests performed similarly. However, specific episodic-memory measures were commonly the best individual predictors of MCI to AD conversion as we theorised. We were surprised to find that some forms of general multi-domain tests could strongly predict MCI to AD conversion. Namely, the ADAS13 but not the ADAS11 had a good predictive ability. We theorise that the better performance of the ADAS13 is due to the inclusion of the episodic-memory component (ADASQ4).

In research question two, we assessed the predictive ability of mixed episodic-memory predictor models. We found that all mixed models performed similarly or marginally worse than individual predictor tests. Specifically, the mixed episodic-memory models had an accuracy of 53.1% in year one and 77.7% in year three. In comparison, the best individual episodic-memory models had an accuracy of 56.4% in year one and 76.9% in year three. These results align with De Simone et al. (2019), who found that individual and mixed episodic-memory models are both equally viable and tend to have similar accuracies. However, we found that even when mixed models outperformed individual measures, the difference was so negligible that it did not make logistical sense to use multiple tests over a single measure. Moreover, we found that in some years of ADNI2, mixed-predictor models were incredibly hard to compute and often violated the assumption of multicollinearity. Accordingly, our results indicated that individual episodic-memory tests were the most versatile measures of early-stage AD.

In research question three, we compared our episodic-memory measures to neuroimaging markers commonly used in the literature. In the first and second years of ADNI2, 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 volume 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 (Moradi et al. 2017; Ihara et al. 2018). Other studies using the ADNI2 cohort have found that MRI volumetric measures can predict MCI to AD conversion with an accuracy ranging from 60 to 85% depending on the specific measure (Korolev et al. 2016; Sun et al. 2017). We theorise that our results are weaker than previous research because we were restricted to only studying two years of ADNI2 and because neuroimaging measures improve with disease progression. However, it is important to note that many neuroimaging studies combine measures to create mixed models with higher accuracy. If we combined different neuroimaging measures (e.g., PET) to create a mixed-predictor model, maybe we could have observed similar results to the models in the literature. However, the refinement of neuroimaging models was beyond the scope of this paper.

When assessing mixed variable models, we found that 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%. In comparison, our other models achieved an accuracy of 56.4% (ADNIMEM) in year one and 71.7% (RAVLT percent forgetting) in year two. 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 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.

In summary, our study determined that, out of several variables, individual episodic-memory measures such as the LMT are strong predictors of MCI to AD conversion and therefore of early-stage AD. While neuroimaging measures and mixed models can increase the overall accuracy of predictive models by a small amount, there is no significant advantage when factoring in the logistical complexity and cost of the technology. 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. Evidence also suggests that general multi-domain tests should be added to inform diagnoses later during AD progression (e.g., ADAS13). However, it is important to note that researchers and clinicians 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). In turn, it is imperative that multiple episodic-memory tests be repeatedly and consistently administered throughout disease progression to confirm MCI and AD diagnoses.

It is important to note that this study had some limitations that should be considered. Namely, we had problems with missing data and a restricted methodology. Researchers should be aware of our limited neuroimaging data, high participant education, and problems with multicollinearity (in mixed models) when generalising our results. It is also apparent that our methodology could have been more robust. For example, while splitting the ADNI years was beneficial for observing MCI progression and test characteristics, we fear that it may have overcomplicated the results. Our logistic regression also limited our ability to have a control group. In turn, future studies should use a survival analysis to include a control group and mediate for participant characteristics (e.g., Aβ, neuropsychiatric symptoms, APOE4). Future research should also seek to investigate the predictive ability of other episodic-memory tests. In turn, we hope that future research can overcome our limitations and continue to apply episodic-memory tests to early-stage AD research. We believe that, with further research, episodic-memory tests can be used to help diagnose early-stage AD and enable sufficient treatment for individuals with AD.

Funding This study would like to acknowledge funding received from 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.

Availability of data and materials This data is available via application to the Alzheimer's Disease Neuroimaging Initiative (ADNI). Data collection and sharing for this project was funded by ADNI (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defence award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organisation is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Code availability Not applicable.

Declarations

Conflict of interest The authors of this study declare no conflicts of interest.

Ethical approval Our study gained ethics approval from *Western Sydney Universities* Human Research Ethics Committee in March 2019 (H13176).

Consent to participate Not applicable.

Consent for publication Not applicable.

References

Abikoff H, Alvir J, Hong G et al (1987) Logical memory subtest of the wechsler memory scale: age and education norms and

alternate-form reliability of two scoring systems. J Clin Exp Neuropsychol 9:435–448. https://doi.org/10.1080/01688638708405063

Alzheimer's Association (2018) 2018 Alzheimer's disease facts and figures. Alzheimers Dement 14:367–429

Alzheimer's Disease Neuroimaging Initiative (2008) ADNI2 Procedures Manual. In: Alzheimers Dis. Neuroimaging Initiat. https://adni.loni.usc.edu/wp-content/uploads/2008/07/adni2-procedures-manual.pdf. Accessed 30 Jan 2021

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

Alzheimer's Disease Neuroimaging Initiative (2017) ADNI MRI Tool. http://adni.loni.usc.edu/methods/mri-tool/. Accessed 11 Dec 2019

Arevalo-Rodriguez I, Smailagic N, Roqué I, Figuls M et al (2015) Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). Cochrane Database Syst Rev. https://doi.org/10.1002/14651858.CD010783.pub2

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

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

Bak TH, Rogers TT, Crawford LM et al (2005) Cognitive bedside assessment in atypical parkinsonian syndromes. J Neurol Neurosurg Psychiatry 76:420–422. https://doi.org/10.1136/jnnp.2003.029595

Banik A, Brown RE, Bamburg J et al (2015) Translation of pre-clinical studies into successful clinical trials for Alzheimer's disease: what are the roadblocks and how can they be overcome? J Alzheimers Dis JAD 47:815–843. https://doi.org/10.3233/JAD-150136

Bastin C, Salmon E (2014) Early neuropsychological detection of Alzheimer's disease. Eur J Clin Nutr 68:1192–1199. https://doi.org/10.1038/ejcn.2014.176

Brown J (2015) The use and misuse of short cognitive tests in the diagnosis of dementia. J Neurol Neurosurg Psychiatry 86:680–685. https://doi.org/10.1136/jnnp-2014-309086

Cai S, Huang L, Zou J et al (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:e0115573. https://doi.org/10.1371/journal.pone.0115573

Chapman RM, Mapstone M, McCrary JW et al (2011) Predicting conversion from mild cognitive impairment to Alzheimer's disease using neuropsychological tests and multivariate methods. J Clin Exp Neuropsychol 33:187–199. https://doi.org/10.1080/13803395.2010.499356

Chu LW (2012) Alzheimer's disease: early diagnosis and treatment. Hong Kong Med J Xianggang Yi Xue Za Zhi 18:228–237

Crane PK, Carle A, Gibbons LE et al (2012) Development and assessment of a composite score for memory in the Alzheimer's Disease Neuroimaging Initiative (ADNI). Brain Imaging Behav Indianap 6:502–16. https://doi.org/10.1007/s11682-012-9186-z

De Simone MS, Perri R, Fadda L et al (2019) Predicting progression to Alzheimer's disease in subjects with amnestic mild cognitive impairment using performance on recall and recognition tests. J Neurol 266:102–111. https://doi.org/10.1007/s00415-018-9108-0

Derby CA, Burns LC, Wang C et al (2013) Screening for predementia AD: time-dependent operating characteristics of episodic memory tests. Neurology 80:1307–1314. https://doi.org/10.1212/WNL.0b013e31828ab2c9

Eckerström C, Olsson E, Bjerke M et al (2013) A combination of neuropsychological, neuroimaging, and cerebrospinal fluid markers predicts conversion from mild cognitive impairment to dementia. J Alzheimers Dis JAD 36:421–431. https://doi.org/10.3233/JAD-122440

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

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. Behav Res Methods 39:175–191. https://doi.org/10.3758/BF03193146

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

Gainotti G, Quaranta D, Vita MG, Marra C (2014) Neuropsychological predictors of conversion from mild cognitive impairment to Alzheimer's disease. J Alzheimers Dis JAD 38:481–495. https://doi.org/10.3233/JAD-130881

Gomar JJ, Bobes-Bascaran MT, Conejero-Goldberg C et al (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. Arch Gen Psychiatry 68:961–969. https://doi.org/10.1001/archgenpsychiatry.2011.96

Gomar JJ, Conejero-Goldberg C, Davies P et al (2014) Extension and refinement of the predictive value of different classes of markers in ADNI: four-year follow-up data. Alzheimers Dement J Alzheimers Assoc 10:704–712. https://doi.org/10.1016/j.jalz.2013.11.009

Ihara R, Iwata A, Suzuki K et al (2018) Clinical and cognitive characteristics of preclinical Alzheimer's disease in the Japanese Alzheimer's Disease Neuroimaging Initiative cohort. Alzheimers Dement Transl Res Clin Interv 4:645–651. https://doi.org/10.1016/j.trci.2018.10.004

Irish M, Lawlor BA, Coen RF, O'Mara SM (2011) Everyday episodic memory in amnestic mild cognitive impairment: a preliminary investigation. BMC Neurosci 12:80. https://doi.org/10.1186/1471-2202-12-80

Jack CR, Albert M, Knopman DS et al (2011) Introduction to revised criteria for the diagnosis of Alzheimer's disease: national institute on aging and the Alzheimer association workgroups. Alzheimers Dement J Alzheimers Assoc 7:257–262. https://doi.org/10.1016/j.jalz.2011.03.004

Korolev IO, Symonds LL, Bozoki AC, Alzheimer's Disease Neuroimaging Initiative (2016) Predicting progression from mild cognitive impairment to Alzheimer's dementia using clinical, MRI, and plasma biomarkers via probabilistic pattern classification. PLoS ONE 11:e0138866. https://doi.org/10.1371/journal.pone.0138866

Larner AJ (2019) Evaluating cognitive screening instruments with the "likelihood to be diagnosed or misdiagnosed" measure. Int J Clin Pract 73:e13265. https://doi.org/10.1111/ijcp.13265

Manly CA, Wells RS (2015) Reporting the use of multiple imputation for missing data in higher education research. Res High Educ 56:397–409. https://doi.org/10.1007/s11162-014-9344-9

Mitchell AJ, Shiri-Feshki M (2009) Rate of progression of mild cognitive impairment to dementia–meta-analysis of 41 robust inception cohort studies. Acta Psychiatr Scand 119:252–265. https://doi.org/10.1111/j.1600-0447.2008.01326.x

Moradi E, Hallikainen I, Hänninen T et al (2017) Rey's Auditory Verbal Learning Test scores can be predicted from whole brain MRI in Alzheimer's disease. NeuroImage Clin 13:415–427. https://doi.org/10.1016/j.nicl.2016.12.011

Nathan PJ, Abbott R, Lim YY et al (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. Int J Neuropsychopharmacol 19:18–19. https://doi.org/10.1093/ijnp/pyw044.593

Ottoy J, Niemantsverdriet E, Verhaeghe J et al (2019) Association of short-term cognitive decline and MCI-to-AD dementia conversion with CSF, MRI, amyloid- and 18F-FDG-PET imaging. NeuroImage Clin 22:101771. https://doi.org/10.1016/j.nicl.2019.101771

Rasmussen J, Langerman H (2019) Alzheimer's disease—why we need early diagnosis. Degener Neurol Neuromuscul Dis 9:123–130. https://doi.org/10.2147/DNND.S228939

Russo MJ, Campos J, Vázquez S et al (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. Front Aging Neurosci. https://doi.org/10.3389/fnagi.2017.00046

Seo EH, Choo IH, Alzheimer's Disease Neuroimaging Initiative (2016) Amyloid-independent functional neural correlates of episodic memory in amnestic mild cognitive impairment. Eur J Nucl Med Mol Imaging 43:1088–1095. https://doi.org/10.1007/s00259-015-3261-9

Sterne JAC, White IR, Carlin JB et al (2009) Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 338:b2393. https://doi.org/10.1136/bmj.b2393

Sun Z, van de Giessen M, Lelieveldt BPF, Staring M (2017) Detection of conversion from mild cognitive impairment to Alzheimer's disease using longitudinal brain MRI. Front Neuroinformatics. https://doi.org/10.3389/fninf.2017.00016

Vakil E, Blachstein H (1993) Rey auditory-verbal learning test: structure analysis. J Clin Psychol 49:883–890. https://doi.org/10.1002/1097-4679(199311)49:6%3c883::AID-JCLP2270490616%3e3.0.CO;2-6