Journal of Alzheimer's Disease 46 (2015) 901–912
DOI 10.3233/JAD-142943
IOS Press

Neuropsychological Testing Predicts Cerebrospinal Fluid Amyloid-β in Mild Cognitive Impairment

Benjamin M. Kandel[a,*], Brian B. Avants[b], James C. Gee[b], Steven E. Arnold[c], David A. Wolk[d] and for the Alzheimer's Disease Neuroimaging Initiative[1]

[a]*Penn Image Computing and Science Laboratory and Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA*

[b]*Penn Image Computing and Science Laboratory and Department of Radiology, Perelman School of Medicine of the University of Pennsylvania, Philadelphia, PA, USA*

[c]*Department of Psychiatry, Perelman School of Medicine of the University of Pennsylvania, Philadelphia, PA, USA*

[d]*Department of Neurology, Perelman School of Medicine of the University of Pennsylvania, Philadelphia, PA, USA*

Handling Associate Editor: Andrew Saykin

Accepted 26 March 2015

Abstract.

Background: Psychometric tests predict conversion of mild cognitive impairment (MCI) to probable Alzheimer's disease (AD). Because the definition of clinical AD relies on those same psychometric tests, the ability of these tests to identify underlying AD pathology remains unclear.

Objective: To determine the degree to which psychometric testing predicts molecular evidence of AD amyloid pathology, as indicated by cerebrospinal fluid (CSF) amyloid-β $(A\beta)_{1\text{-}42}$, in patients with MCI, as compared to neuroimaging biomarkers.

Methods: We identified 408 MCI subjects with CSF Aβ levels, psychometric test data, FDG-PET scans, and acceptable volumetric MR scans from the Alzheimer's Disease Neuroimaging Initiative (ADNI). We used psychometric tests and imaging biomarkers in univariate and multivariate models to predict Aβ status.

Results: The 30-min delayed recall score of the Rey Auditory Verbal Learning Test was the best predictor of Aβ status among the psychometric tests, achieving an AUC of 0.67 ± 0.02 and odds ratio of 2.5 ± 0.4. FDG-PET was the best imaging-based biomarker (AUC 0.67 ± 0.03, OR 3.2 ± 1.2), followed by hippocampal volume (AUC 0.64 ± 0.02, OR 2.4 ± 0.3). A multivariate analysis based on the psychometric tests improved on the univariate predictors, achieving an AUC of 0.68 ± 0.03 (OR 3.38 ± 1.2). Adding imaging biomarkers to the multivariate analysis did not improve the AUC.

Conclusion: Psychometric tests perform as well as imaging biomarkers to predict presence of molecular markers of AD pathology in MCI patients and should be considered in the determination of the likelihood that MCI is due to AD.

Keywords: Alzheimer's disease, magnetic resonance imaging, mild cognitive impairment, positron emission tomography

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

*Correspondence to: Benjamin M. Kandel, B.A., Penn Image Computing and Science Laboratory and Department of Bioengineering, University of Pennsylvania, 3600 Market St, Ste. 370, Philadelphia, PA 19104, USA. Tel.: +1 314 610 7256; E-mail: bkandel@seas.upenn.edu.

INTRODUCTION

Recent guidelines for diagnosing mild cognitive impairment (MCI) due to Alzheimer's disease (AD) have emphasized the importance of psychometric testing for establishing the existence of MCI, and subsequently relying on biomarkers based on imaging and biofluids to assess the likelihood that the existing cognitive impairment is "due to AD" relative to a different cause [1]. In particular, cognitive testing is a component of the "core clinical criteria" for MCI, which requires that impairment greater than expected for age must be present in at least one cognitive domain. Once clinical categorization of MCI is established, the guidelines suggest that the likelihood that the cognitive phenotype is "due to AD" should rely on various imaging and molecular biomarkers (each classified as either a biomarker of neurodegeneration or cerebral amyloid), without specifically taking into account the severity of the cognitive deficit within the MCI category.

Although imaging-derived biomarkers for diagnosis of AD and prediction of conversion from MCI to AD have been the subject of intensive research [2–4], how these biomarkers can be used most effectively in the presence of alternative sources of clinical information about a subject's status, such as cognitive testing, is still not settled. Several recent studies have examined the relative utility of cognitive testing, imaging, or molecular biomarkers for predicting conversion from MCI to AD [5–9], These studies have generally found that cognitive testing performs similarly to other biomarkers, but a potential criticism of these study designs is that using psychometric measurements to predict conversion to AD is circular, as the diagnosis of AD is itself determined in large part based on psychometric tests that are the same as or similar to those used to predict conversion.

To avoid this circularity, we sought to determine if cognitive testing with standard psychometric measures can predict the presence of cerebral amyloid based on a well-established cerebrospinal fluid (CSF) molecular biomarker, the detection of which is independent of cognitive scores, unlike clinical diagnosis of conversion to AD. Although postmortem histology remains the gold standard for establishing AD pathology, measures of CSF $A\beta_{1\text{-}42}$ and amyloid positron emission tomography (PET) imaging are the closest currently available surrogate [10, 11]. For the present study, we used CSF Aβ as a marker for AD pathology given its higher uniform availability in the studied cohort. We choose CSF Aβ in isolation, as opposed to tau/Aβ ratio, because we were specifically comparing the relationship between cognitive and neuroimaging neurodegenerative biomarkers and evidence of AD molecular pathology; thus, incorporating a molecular neurodegenerative marker like tau may confound the results. Moreover, we wanted to determine the relative and combined predictive value of psychometric testing with neuroimaging biomarkers of neuronal injury or neurodegeneration.

In particular, we examined several cognitive measures, including verbal memory, given their putative sensitivity to prodromal AD. We used diverse imaging-derived biomarkers to accurately represent both standard and developing measurement approaches. Further, we chose structural magnetic resonance imaging (MRI) and FDG-PET measures given their emphasis in the MCI guidelines. For MRI data, we used an automated hippocampal volume measurement, several cortical-thickness measurements including a summary measure of several regions associated with AD-related tissue loss [12, 13], and multivariate analysis of voxelwise measurements of cortical thickness [14, 15]. Hippocampal volume is considered to be one of the most established biomarkers of AD with numerous studies demonstrating its predictive value in MCI. We also used FDG-PET data from a set of regions (meta-region of interest, ROI) previously determined to be sensitive to early AD and predictive of clinical conversion to AD in MCI cohorts [16]. To obtain such a wide variety of clinical data in a sufficiently large population, we utilized the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. If cognitive measures perform similarly to both more standard and developing imaging biomarkers in prediction of AD pathology with MCI patients, they can provide a cost-effective and easily accessible method for assessing the likelihood of prodromal AD in patients with MCI.

METHODS

Clinical data

Subjects

This study was a retrospective analysis of data obtained from the ADNI database (http://adni.loni.usc.edu). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies, and non-profit organizations, as a $60 million, 5-year public-private partnership. The primary goal of ADNI has

been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials.

Data used in this article were downloaded from the ADNI website in January 2014. We included only MCI subjects with complete datasets for the current analysis, including CSF Aβ levels, all neuropsychological tests examined, and FDG-PET. Only those subjects with Freesurfer cortical and hippocampal segmentations of acceptable quality, as determined by the publicly available Freesurfer dataset available through ADNI, were included.

In the ADNI study, MCI is split into two groups, early MCI (EMCI) and late MCI (LMCI). Diagnostic criteria for both EMCI and LMCI subjects were as follows: Mini-Mental State Examination (MMSE) scores between 24–30 (inclusive), a subjective memory concern reported by subject, informant, or clinician, a Clinical Dementia Rating of 0.5, absence of significant levels of impairment in other cognitive domains, essentially preserved activities of daily living, and an absence of dementia. They also were required to have objective memory loss measured by education adjusted scores on delayed recall of one paragraph from Wechsler Memory Scale Logical Memory II, which further determined EMCI ($\geq$16 years: 9–11; 8–15 years: 5–9; 0–7 years: 3–6) or LMCI ($\geq$16 years: $\leq$8; 8–15 years: $\leq$4; 0–7 years: $\leq$2) status. In this manuscript, MCI refers to both EMCI and LMCI.

The ADNI study includes a variety of collection sites around the United States and Canada, and a full list is available at http://adni.loni.usc.edu/about/centers-cores/study-sites/. Recruitment for the ADNI study aimed to achieve a balance of normal controls, MCI, and AD subjects. For ADNI 1, a random subsample of subjects was selected for FDG imaging; in ADNI 2/GO, all subjects enrolled received FDG imaging. For up-to-date information on specific inclusion and exclusion criteria, please see http://www.adni-info.org.

Psychometric testing

We aimed to include a battery of psychometric tests that would cover a broad range of cognitive domains, with special focus on memory due to its importance in AD. For memory, we included components of the Rey Auditory Verbal Learning Test (AVLT) [17] given its richness of measures for various aspects of mnemonic processing (e.g., immediate versus delayed recall versus delayed recognition); for assessment of cognitive speed, sequencing, and executive function, the Trail Making Test [18] (Trails A and Trails B) was used; for language/semantics, category fluency [19] (Animals) and the Boston Naming Test [20] were examined; and as a measure of global cognition, the MMMSE was utilized [21]. We examined several of the AVLT measures, which depend on differential aspects of episodic and working memory [22]. The AVLT consists of five learning trials in which a list of 15 words is read and the subject is asked to immediately recall as many items as possible. After an interference list of 15 novel words is read and recalled, subjects are then asked to recall words from the initial list (5-min delayed recall). A 30-min delayed recall trial and recognition test follow. For the recognition test, subjects are presented with a list of the 15 studied words and 15 nonstudied foils and are asked to circle all words previously studied. To account for false alarms (FA) to nonstudied items, we calculated a measure of discriminability, d-prime (d′), in a standard fashion based on classic signal detection theory [23]. Because d' is undefined when either proportion is 0 or 1, we used standard formulas to convert these values: Hits = (no. of hits+0.5)/(no. of studied items+1) and FA = (no. of FA+0.5)/(no. of unstudied items+1). For the current study, we analyzed performance on the fifth immediate memory trial (AVLT Trial 5 Recall), 5- and 30-min delayed recall (AVLT 5-min Recall, AVLT 30-min Recall), and recognition memory discrimination (AVLT Recognition Discrimination). In addition, we calculated a retention score, which is the number of items remembered after a 30-min delay (AVLT 30-min Recall) divided by the number of items remembered during the last immediate memory trial (AVLT Trial 5 Recall).

Determination of amyloid and ApoE status

CSF-based molecular biomarkers were processed by the University of Pennsylvania/ADNI Biomarker Core Laboratory as previously described [10, 24]. An $A\beta_{1\text{-}42}$ value of less than or equal to 191 pg/ml was considered to be "positive" for the presence of amyloid pathology based on a prior autopsy-based study performed at the University of Pennsylvania [10]. For analyses involving ApoE status, subjects were dichotomized into ApoE ε4 positive and negative groups. ApoE ε4 positive status is defined as having at least one ApoE ε4 allele.

Neuroimaging measures

Processing of neuroimaging data included both analyses made publicly available by ADNI and in-house image processing. The following analyses were based on preprocessed data downloaded from the ADNI website: FDG-PET scans were acquired and analyzed in accordance with a standard protocol [16]. Mean FDG uptake was averaged over 5 ROIs that are sensitive to AD-related changes in metabolism, including right and left angular gyri, right and left inferior temporal regions, and bilateral posterior cingulate. These regions were averaged into a meta-ROI and normalized to an ROI focused on the pons and cerebellar vermis to give a summary FDG PET measure. Cortical thickness and hippocampal measurement of the MRI scans were performed according to the standard ADNI Freesurfer [25] processing pipeline, and downloaded from the ADNI website. Only images that passed ADNI quality control for the temporal, occipital, temporal, and parietal lobe were included. Cortical thickness in the caudal portion of the middle frontal gyrus, medial portion of the orbital frontal cortex, inferior parietal lobule, lateral portion of the occipital cortex, inferior temporal gyrus, entorhinal cortex, temporal pole, and the isthmus of the cingulate cortex were averaged to form a meta-ROI thought sensitive to early AD related neurodegeneration, as previously suggested [26].

Image analysis

In addition to the image analysis performed by various ADNI investigators, we ran additional analyses of MR images to supplement standard approaches with a state of the art multivariate analysis technique. 1.5T and 3T non-accelerated T1-weighted MPRAGE and SPGR MRI scans of all MCI subjects from ADNI1 and ADNI2/GO were downloaded from http://adni.loni.usc.edu. We computed an alternative measure of cortical thickness using DiReCT [12, 13], and used the AAL label set [27] to define medial temporal and precuneal ROIs, as these areas are known to atrophy in early AD. We performed a singular value decomposition (SVD) analysis of the whole-brain cortical thickness data, as this analysis has proven useful in differentiating AD from frontotemporal dementia and predicting CSF-based biomarkers in this population [28, 29]. The SVD was performed using the princomp function in R, and we retained the top 10 components. A grid search strategy using bootstrapping with 100 repetitions, with half the subjects left out for a validation cohort, was used to determine the optimal number of components to retain.

Statistical analysis

All statistical analysis was performed using the R programming language, version 3.1.0. For predictive studies, we randomly split the subjects 5 times into training and testing cohorts, retaining half the subjects for training and using the other half for testing in a 5×2 cross-validation scheme [30]. All area under the curve (AUC), odds ratios, and positive and negative predictive values are on the testing cohorts. Two-tail *t*-tests were used to compare AUC values between testing cohorts of different models to calculate a *p*-value for differences in mean AUC; false discovery rate (FDR) correction was applied to correct for multiple comparisons. For all analyses, patient age, gender, and education were used as additional predictors; for all MR-based imaging analyses, magnet field strength (1.5 or 3T) was included as a covariate. In addition to univariate predictions of Aβ status from psychometrics and imaging modalities, we performed principal component regression, using three principal components, on all the psychometric scores, as well as the psychometric and imaging values combined. AUC analysis was performed using the ROCR package in R [31].

RESULTS

Subject demographics

Subject data was collected between January 2006 and January 2013. A total of 622 MCI subjects with CSF-derived Aβ values were identified, and 407 of those were Aβ positive. Of these, 547 (350 Aβ positive) had FDG scans; 450 (286 Aβ positive) had complete Freesurfer segmentations without failures; 433 (273 Aβ positive) had intracranial volume available; and 408 subjects (257 Aβ positive) had complete psychometric scores available. There was a mean difference of 15 days between the psychometric tests and imaging studies, with 95% of subjects having the imaging and psychometric tests done within 55 days of each other. The maximum time difference was 138 days. A total of 62 adverse events were reported from the lumbar punctures, most of which were headaches (25 cases) or pain (23 cases), with two subjects reporting nausea and a few reporting a variety of other effects, including bruising, tenderness, and swelling. One adverse event, transient procedural anxiety, occurred during the imaging.

A summary of the demographics of the study population, including the psychometric and imaging information, is given in Table 1. We computed

Table 1
Summary of demographics, psychometric scores, and imaging data for subjects

	All subjects (mean ± standard deviation)	Aβ+	Aβ−
Number of subjects	408	257	151
Number of males	232	151	81
Number of ApoE ε4+	207	178	29
Age	71.61 ± 7.16	72.66 ± 6.76	69.79 ± 7.47
Education	16.24 ± 2.71	16.14 ± 2.79	16.41 ± 2.59
Mini-Mental Status Examination	28.0 ± 1.74	27.7 ± 1.80	28.4 ± 1.54
AVLT Trial 5 Recall	9.03 ± 3.00	8.35 ± 2.85	10.19 ± 2.90
AVLT 5-min Recall	5.65 ± 3.74	4.82 ± 3.42	7.05 ± 3.87
AVLT 30-min Recall	4.27 ± 3.92	3.30 ± 3.33	5.92 ± 4.29
AVLT Recognition Discrimination	2.31 ± 1.21	2.07 ± 1.18	2.72 ± 1.14
Retention	0.41 ± 0.31	0.34 ± 0.29	0.53 ± 0.31
Trail Making Test A	39.00 ± 16.71	41.64 ± 18.21	34.50 ± 12.63
Trail Making Test B	105.70 ± 57.60	116.30 ± 62.47	87.66 ± 42.69
Boston Naming Test	26.92 ± 3.28	26.73 ± 3.20	27.26 ± 3.39
Category fluency (animals)	18.05 ± 4.93	17.44 ± 4.88	19.08 ± 4.84
Hippocampal volume	3497.62 ± 577.07	3386.02 ± 537.17	3687.56 ± 537.17
Medial Temporal Thickness	3.83 ± 0.60	3.78 ± 0.61	3.93 ± 0.57
Precuneus Thickness	1.54 ± 0.39	1.52 ± 0.39	1.58 ± 0.37
Mean Cortical Thickness of AD Meta-ROI	2.64 ± 0.17	2.61 ± 0.17	2.68 ± 0.16
Mean FDG-PET SUVR of AD Meta-ROI	1.26 ± 0.14	1.23 ± 0.15	1.31 ± 0.11

Table 2
Summary of univariate logistic regressions predicting Aβ status from each psychometric test and imaging biomarker. Age, gender, and education level (in years) were included as covariates. All data were scaled before regression to facilitate inspection of regression coefficients

	β Estimate	Std. Error	Zval	*p* val
Mini-Mental State Examination	−0.36	0.12	−3.11	1.9E-3
AVLT Trial 5 Recall	−0.57	0.12	−4.946	7.6E-7
AVLT 5-min Recall	−0.55	0.11	−4.83	1.3E-6
AVLT 30-min Recall	−0.63	0.11	−5.47	4.4E-8
Trail Making Test A	0.44	0.14	3.18	1.5E-3
AVLT Recognition Discrimination	−0.50	0.11	−4.45	8.7E-6
Retention	−0.59	0.11	−5.25	1.5E-7
Trail Making Test B	0.52	0.15	3.57	3.6E-4
Boston Naming Test	−0.08	0.11	−0.76	4.5E-1
Category fluency (animals	−0.26	0.11	−2.39	1.7E-2
Hippocampal volume	−0.43	0.13	−3.44	5.9E-4
Medial Temporal Thickness	−0.12	0.11	−1.01	3.1E-1
Precuneal Thickness	−0.01	0.12	−0.05	9.6E-1
Mean Cortical Thickness of AD Meta-ROI	−0.23	0.12	−1.88	6.1E-2
Mean FDG-PET SUVR of AD Meta-ROI	−0.54	0.12	−4.57	4.9E-6

a logistic regression relating each psychometric test and modality with Aβ status, while covarying for age, gender, and education (Table 2). The logistic regression results indicated that the psychometric tests and imaging modalities were predictive of Aβ status, even when included in a univariate model.

Predictive models

The associations between the various psychometric scores and Aβ status were strong enough to predict Aβ status when the data used to train the model was separate from the data used for evaluation. While many of the psychometric measures displayed predictive value, varying in range of AUCs from 0.59 to 0.67, immediate and delayed recall measures performed particularly well, reaching an AUC of 0.65 and 0.67 respectively, corresponding to odds ratios of 3.0 and 2.5 (Fig. 1, Table 3). The 30-min delayed recall test was significantly better than both Trails tests, the Boston Naming Test, category fluency, and MMSE. The standard imaging modalities were similar to each other and the individual psychometric tests in prediction of Aβ status with FDG-PET displaying the highest AUC at 0.67, followed by hippocampal volume at 0.64. Delayed recall performed significantly better than all of the cortical thickness-based measurements and trended

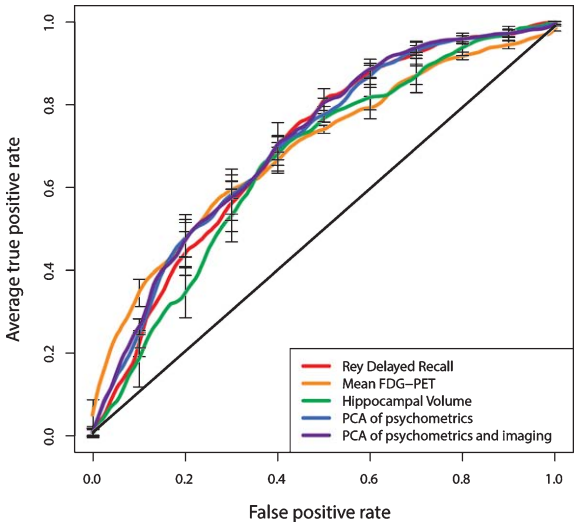

Fig. 1. ROC curves for predicting Aβ status from psychometric scores, imaging biomarkers, and principal components analysis of a collection of psychometric scores, and principal components of psychometric and imaging biomarkers.

better, but was not statistically significantly better, than hippocampal volume. Delayed recall performed similarly to FDG-PET. Despite the prior evidence of SVD analysis of the whole-brain cortical thickness data in prediction of CSF Aβ measures in a cohort of AD and FTD patients, this approach did not appear to enhance prediction (AUC = 0.59) versus more standard structural MRI measures. Performing a principal components analysis on the psychometric scores and using the resulting components boosted the AUC slightly to 0.68 with an odds ratio of 3.38; adding the imaging modalities to that model increased the AUC to 0.69, but the increase was not significant (Table 4). The multivariate analysis of the cognitive tests, however, was statistically significantly better than hippocampal volume, which was not true for any individual cognitive test. Repeating the analysis using only subjects with 3T MR scans did not significantly change the results.

Effect of ApoE allele

Because of the tight link between ApoE ε4 and Aβ pathology, we sought to determine, as a secondary analysis, whether the observed effects are modulated by ε4 status. We divided the subjects into ε4 positive and ε4 negative groups and performed the analyses in the same way as before (Table 5). The results were broadly the same in that imaging did not significantly improve diagnostic accuracy over psychometric tests. Nearly all psychometric and neuroimaging biomarkers were more predictive of Aβ status in ε4 negative as compared to ε4 positive subjects. This trend was highly statistically significant ($p < 0.001$ using a paired t-test).

DISCUSSION

Impact

The results shown here indicate that a psychometric evaluation can be as useful as FDG-PET or quantitative MR imaging in predicting whether or not a given amnestic MCI patient likely has underlying

Table 3
Area under the curve (AUC), odds ratios, and positive (PPV) and negative predictive values (NPV) predicting Aβ status from biomarkers

	AUC	Odds Ratio	PPV	NPV
Mini-Mental Status Examination	0.61 ± 0.03	1.94 ± 0.60	0.71 ± 0.05	0.43 ± 0.05
AVLT Trial 5 Recall	0.65 ± 0.03	3.01 ± 0.36	0.75 ± 0.04	0.50 ± 0.05
AVLT 5-min Recall	0.65 ± 0.02	2.50 ± 0.44	0.73 ± 0.05	0.47 ± 0.04
AVLT 30-min Recall	0.67 ± 0.02	2.46 ± 0.52	0.73 ± 0.05	0.48 ± 0.06
AVLT Recognition Discrimination	0.64 ± 0.03	2.44 ± 0.55	0.73 ± 0.02	0.48 ± 0.07
Retention	0.67 ± 0.03	2.48 ± 0.48	0.73 ± 0.03	0.47 ± 0.06
Trail Making Test A	0.62 ± 0.02	2.13 ± 0.46	0.73 ± 0.04	0.44 ± 0.05
Trail Making Test B	0.63 ± 0.02	2.49 ± 0.48	0.75 ± 0.05	0.45 ± 0.05
Boston Naming Test	0.59 ± 0.02	1.66 ± 0.17	0.70 ± 0.03	0.42 ± 0.04
Category fluency (animals)	0.60 ± 0.02	1.88 ± 0.43	0.71 ± 0.05	0.42 ± 0.03
Hippocampal volume	0.64 ± 0.02	2.41 ± 0.34	0.74 ± 0.04	0.46 ± 0.04
Medial Temporal Thickness	0.59 ± 0.01	1.67 ± 0.07	0.70 ± 0.04	0.42 ± 0.04
Precuneal Thickness	0.59 ± 0.02	1.83 ± 0.25	0.71 ± 0.03	0.43 ± 0.05
Mean Cortical Thickness of AD Meta-ROI	0.61 ± 0.02	1.90 ± 0.31	0.71 ± 0.04	0.43 ± 0.04
Mean FDG-PET SUVR of AD Meta-ROI	0.67 ± 0.03	3.19 ± 1.22	0.76 ± 0.05	0.49 ± 0.08
PCA of psychometric scores	0.68 ± 0.02	3.38 ± 1.16	0.71 ± 0.03	0.56 ± 0.10
PCA of psychometric scores and imaging biomarkers	0.69 ± 0.02	3.18 ± 0.76	0.71 ± 0.03	0.55 ± 0.08
PCA of cortex-wide cortical thickness	0.59 ± 0.03	1.57 ± 0.21	0.67 ± 0.04	0.43 ± 0.02

PCA, principal components analysis.

Table 4

Table of *p*-values of AUC's for each variable compared with every other variable (FDR corrected). *p*-values of less than 0.05 are color-coded to indicate which measure is better: Blue indicates that the test indicated in the row name is better, whereas green indicates that the test indicated in the column name is better

AVLT Trial 5 recall	AVLT 5-min recall	AVLT 30-min recall	Trails A	Trails B	Boston Naming Test	Category Fluency (animals)	MMSE	Discrimi-nation	Retention	Medial Temporal Thickness	Precuneus Thickness	Mean FDG	Hippo-campal Volume	Thickness of Meta-ROI	PCA of psychometrics	PCA of psychometrics and imaging
	1.00	0.37	0.23	0.39	0.02	0.04	0.12	0.77	0.25	0.02	0.02	0.29	0.54	0.05	0.12	0.10
1.00		0.24	0.17	0.33	0.01	0.03	0.10	0.77	0.12	0.00	0.01	0.23	0.50	0.03	0.07	0.05
0.37	0.24		0.03	0.05	0.00	0.01	0.03	0.23	0.82	0.00	0.00	0.76	0.10	0.01	0.38	0.28
0.23	0.17	0.03		0.62	0.07	0.24	0.62	0.44	0.02	0.05	0.08	0.04	0.52	0.38	0.01	0.01
0.39	0.33	0.05	0.62		0.03	0.09	0.31	0.73	0.03	0.01	0.03	0.07	0.87	0.12	0.02	0.02
0.02	0.01	0.00	0.07	0.03		0.50	0.25	0.03	0.00	1.00	0.86	0.01	0.03	0.23	0.00	0.00
0.04	0.03	0.01	0.24	0.09	0.50		0.62	0.09	0.01	0.38	0.62	0.01	0.08	0.69	0.00	0.00
0.12	0.10	0.03	0.62	0.31	0.25	0.62		0.24	0.02	0.20	0.30	0.03	0.27	0.87	0.01	0.01
0.77	0.77	0.23	0.44	0.73	0.03	0.09	0.24		0.16	0.03	0.04	0.19	0.86	0.13	0.09	0.07
0.25	0.12	0.82	0.02	0.03	0.00	0.01	0.02	0.16		0.00	0.00	0.87	0.06	0.00	0.50	0.35
0.02	0.00	0.00	0.05	0.01	1.00	0.38	0.20	0.03	0.00		0.78	0.01	0.02	0.11	0.00	0.00
0.02	0.01	0.00	0.08	0.03	0.86	0.62	0.30	0.04	0.00	0.78		0.01	0.03	0.27	0.00	0.00
0.29	0.23	0.76	0.04	0.07	0.01	0.01	0.03	0.19	0.87	0.01	0.01		0.10	0.02	0.79	0.65
0.54	0.50	0.10	0.52	0.87	0.03	0.08	0.27	0.86	0.06	0.02	0.03	0.10		0.12	0.03	0.03
0.05	0.03	0.01	0.38	0.12	0.23	0.69	0.87	0.13	0.00	0.11	0.27	0.02	0.12		0.00	0.00
0.12	0.07	0.38	0.01	0.02	0.00	0.00	0.01	0.09	0.50	0.00	0.00	0.79	0.03	0.00		0.86
0.10	0.05	0.28	0.01	0.02	0.00	0.00	0.01	0.07	0.35	0.00	0.00	0.65	0.03	0.00	0.86	

Table 5
AUC values for prediction of Aβ status from cognitive tests when stratifying patients by ApoE ε4 status. Cognitive tests were overall more predictive of Aβ status in ε4 negative subjects than ε4 positive subjects.

	AUC	
	ε4+	ε4−
AVLT Trial 5 recall	0.72 ± 0.03	0.71 ± 0.04
AVLT 5-min recall	0.70 ± 0.04	0.72 ± 0.04
AVLT 30-min recall	0.70 ± 0.03	0.74 ± 0.03
Trails A	0.68 ± 0.03	0.75 ± 0.03
Trails B	0.67 ± 0.02	0.76 ± 0.03
Boston Naming Test	0.67 ± 0.02	0.72 ± 0.06
Category Fluency (animals)	0.70 ± 0.03	0.72 ± 0.05
MMSE	0.68 ± 0.03	0.73 ± 0.02
Discrimination	0.69 ± 0.04	0.72 ± 0.02
Retention	0.70 ± 0.03	0.73 ± 0.03
Medial Temporal Thickness	0.68 ± 0.03	0.72 ± 0.04
Precuneus Thickness	0.68 ± 0.03	0.70 ± 0.05
Mean FDG	0.70 ± 0.02	0.75 ± 0.03
Hippocampal Volume	0.70 ± 0.04	0.74 ± 0.04
Thickness of Meta-ROI	0.67 ± 0.03	0.69 ± 0.03
PCA of psychometrics	0.69 ± 0.03	0.74 ± 0.03
PCA of psychometrics and imaging	0.69 ± 0.03	0.73 ± 0.04

PCA, principal components analysis.

AD pathology. The low cost and ready availability of psychometric batteries as compared to imaging studies makes them an attractive and useful alternative to specialized imaging techniques in clinical evaluation. Although the psychometric batteries do not approach perfect classification between Aβ-positive and Aβ-negative subjects, they can be useful in clinical practice to broadly estimate risk of prodromal AD and, perhaps, guide the process of obtaining additional studies, including molecular biomarkers. For situations in which obtaining an accurate measure of Aβ is paramount, such as evaluating appropriateness of a future anti-amyloid therapy, direct molecular imaging or CSF measurement of Aβ is still necessary, perhaps after initial screening with psychometrics to enrich with amyloid positive patients.

One intriguing finding of this study is that multivariate analysis using principal components analysis of the psychometric scores only marginally improved on the single best psychometric test, and the difference in AUC was not statistically significant at the $p < 0.05$ level. At the same time, the modest boost in AUC achieved by a multivariate analysis was sufficient to give a statistically significant improvement over hippocampal volume, but not over FDG-PET. These results suggest that improvements in diagnostic capability by using a multivariate cognitive profile as opposed to a single test offer only marginal improvements while at the same time suffering from less interpretability than a single test. Adding the imaging biomarkers to the multivariate analysis did not significantly improve the AUC, suggesting that imaging offers little added value over a cognitive profile when screening for underlying AD pathology.

Further, the fact that even the "standard" cognitive measures examined here displayed some success in determining the likelihood of AD pathology suggests that more research is warranted on designing and evaluating psychometric tests optimized for detection of early AD-related cognitive decline. In particular, measures guided by the cognitive neuroscience literature may be particularly useful in this regard [32]. Finally, the results here indicate that the ability of psychometric scores to identify patients who will progress to AD is not due solely to the fact that those same scores are used to establish presence of probable AD. Instead, it appears that the predictive value of psychometric tests are due, at least in part, to their ability to separate MCI patients into sub-populations with higher and lower prevalence of AD pathology.

Limitations

Although this study does indicate that a psychometric battery should be an important component of the evaluation of MCI subjects beyond initial categorization to the MCI designation, there are several factors that may influence the relative ability of imaging to predict AD pathology. First, this study focused exclusively on cross-sectional imaging studies. Longitudinal imaging may provide a more reliable representation of disease progression. Nevertheless, longitudinal imaging may not be feasible for many care settings, so evaluating the diagnostic power of cross-sectional imaging is also important. It is worth noting that this study is meant to help guide providers caring for patients with MCI, not to detect AD pathology in presymptomatic patients. By the time cognitive scores become clearly abnormal, significant neurodegeneration has likely already occurred while this may be more variable in the preclinical phase. Thus, it is unclear whether the same relative predictive value of cognitive versus neuroimaging methods would hold in that context. The patient selection criteria also may limit the applicability of the findings presented here to a broader range of patients. This study focused on amnestic MCI subjects. It is possible that in a broader selection of MCI subjects, the memory tests proposed may provide even greater capability in prediction of amyloid status. On the other hand, in non-amnestic MCI populations, these tests may be less predictive due to differences in the loci of neurodegenerative change in amnestic

versus non-amnestic prodromal AD. In addition, the ADNI study population is enriched in AD or AD-like pathology. In a more general clinical setting, providers must also consider the possibility of other sources of cognitive impairment, such as depression or stroke. It is uncertain how this greater heterogeneity would impact the predictive value of both cognitive and neuroimaging measures. Another drawback to the current study is the sampling procedure. We excluded subjects who did not have all the biomarkers examined here, including those for whom the automated hippocampal segmentation failed. As such, the subset in this study would, if anything, overestimate the ability of hippocampal segmentation to track AD pathology; had we not excluded patients with unreliable segmentations, the predictive ability of hippocampal volumes would likely be lower.

It is also possible that advances in image processing techniques may improve the diagnostic capability of neuroimaging data. Although it is impossible to rule out such advances, the variety of imaging modalities and image processing techniques used here make it less likely that new analytic approaches would improve the predictive power of imaging data enough to supplant psychometric measures as a key method for characterization of MCI patients. Indeed, the current work did use a promising analytic approach involving singular value decomposition across the entire cortical mantle, which had previously demonstrated good predictive value of CSF t-tau/Aβ in patients with AD and frontotemporal dementia [28]. Nonetheless, this approach did not display significant advantages over more traditional measures (e.g., hippocampal volume) or psychometric tests. In any case, psychometric tests are more accessible than sophisticated image processing techniques, especially to physicians who do not work in academic medical centers.

An obvious limitation of this study is the use of CSF-derived Aβ status as a gold standard in the prediction models, as CSF Aβ does not perfectly reflect brain AD pathology. While we took this approach to avoid the circularity of longitudinal studies of conversion, a better design would have autopsy-confirmed AD pathology for comparison with the other biomarkers. Nonetheless, CSF Aβ, along with amyloid PET, are the closest surrogates to histopathologic evaluation presently available and have displayed high sensitivity in autopsy studies [10, 11].

Finally, the limited accuracy for prediction of amyloid status of even the most accurate models indicates that caution should be exercised when using values from these models to guide clinical decision-making and, at most, they should be considered another piece in the overall assessment of risk in MCI patients. Fundamentally, the main conclusion of this study is that psychometric scores provide as much information as neurodegenerative imaging biomarkers in prediction of underlying amyloid pathology, not that either imaging or cognitive biomarkers should be regarded as having perfect diagnostic accuracy. This conclusion strengthens the argument made in previous studies that cognitive tests are a crucial component in multivariate predictive models for conversion from MCI to AD by demonstrating that cognitive scores predict molecular AD pathology, not just cognition-based diagnoses of AD. Therefore, cognitive tests should be considered just as important a biomarker for AD pathology as other neurodegenerative biomarkers, which have already been recognized by the National Institute on Aging – Alzheimer's Association (NIA-AA) work group for MCI diagnosis. While the AUC values are relatively modest, the odds ratios suggest that poorer performance on the best cognitive predictors are associated with approximately a three-fold risk of underlying AD pathology, which may influence counseling of patients.

Effect of ApoE

One intriguing result in this study is the marked difference in prediction accuracy in ApoE ε4 positive versus ε4 negative subjects. This finding is consistent with previous work showing that cognitive function is more closely linked to Aβ status within ε4 negative than within ε4 positive subjects [33, 34]. The mechanism behind this effect is not clear, but may be that the effects of Aβ on cognitive function are modulated by ApoE isoforms. However, an important confounding factor is the highly unbalanced nature of the samples: The ε4 negative group had 79 Aβ+ and 120 Aβ− subjects, whereas the ε4 positive group had 178 Aβ+ and only 29 Aβ− subjects. The relative paucity of ε4 positive but Aβ− subjects may contribute to the lower performance of the predictive model in the ε4 positive group. Thus, it is possible that the strong association of Aβ with ε4 status obscures the association with cognitive measures.

Psychometric scores as functional biomarkers

It is worth pointing out that the current algorithm for determining the likelihood of "MCI due to AD" in the recently proposed criteria treats neurodegenerative and molecular markers as dissociable modalities of evidence. In a sense, psychometric tests can be considered

another type of downstream neurodegenerative measure. Thus, it may seem somewhat odd to use one type of biomarker (neurodegenerative) to predict another (molecular) in this context if these measures provide orthogonal information. However, these measures are obviously related and multiple studies have demonstrated the significant predictive value for conversion to clinical AD in patients either with "positive" CSF or PET amyloid studies or neurodegenerative markers [1, 35, 36].

Nonetheless, one reason for the modest ability of cognitive measures to predict amyloid status is that MCI Aβ+ likely is associated with variable levels of impairment. This is almost certainly an issue for any neurodegenerative biomarker given the range of disease severity within the MCI category. Indeed, neurodegenerative biomarkers, in addition to providing some currency on the underlying pathology (e.g., cerebral amyloid), also are informative on disease stage and enhance prediction of the timing of transitions to dementia, as has been suggested in the literature [37–39]. Thus, relatively poor performance on cognitive measures within the MCI category increases both the likelihood that the underlying process is AD and that progression to dementia is more likely to occur in the near future, which may help provide additional context for clinicians in their assessment of these patients.

The choice of CSF Aβ as the proxy or standard for AD pathology in the present analysis also reflects the notion that it is a more specific measure of AD pathology than neurodegenerative markers given the defining nature of cerebral amyloid in the pathologic criteria for AD. Indeed, more and more therapeutic trials, including in MCI, are using a positive amyloid study as inclusion criteria [40]. Thus, examination of psychometric measures within the MCI category may contribute to increasing the likelihood that a given patient may qualify for such a study on that basis.

CONCULSION

In an MCI population, psychometric scores predict presence of CSF-based amyloid pathology that overlaps with predictions obtainable from FDG-PET and structural MR images. Thus, psychometric measures may be preferable in the cross-sectional context to provide initial screening on the likelihood of prodromal AD. The ability of cognitive scores to predict the existence of underlying AD pathology indicates that in addition to using cognitive test cutoffs to establish the existence of MCI, the severity of the test scores is as reliable an indicator as imaging biomarkers of neurodegeneration that the cognitive impairment is due to AD pathology. Thus, these measures could be included in the MCI algorithm as a type of neurodegenerative marker that could further help clinicians prognosticate in the clinical setting.

ACKNOWLEDGMENTS

This research was supported by NIH grants T32-EB009384, AG010124, and AG040271. The funding agencies had no say in the design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, and approval of the manuscript.

Authors' disclosures available online (http://j-alz.com/manuscript-disclosures/14-2943r2).

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

REFERENCES

[1] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carrillo MC, Thies B, Phelps CH (2011) The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 270-279.

[2] Chételat G, Desgranges B, De la Sayette V, Viader F, Eustache F, Baron J-C (2003) Mild cognitive impairment: Can FDG-PET predict who is to rapidly convert to Alzheimer's disease? *Neurology* **60**, 1374-1377.

[3] Davatzikos C, Bhatt P, Shaw LM, Batmanghelich KN, Trojanowski JQ (2011) Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiol Aging* **32**, 2322.e19-2322.e27.

[4] Killiany RJ, Gomez-Isla T, Moss M, Kikinis R, Sandor T, Jolesz F, Tanzi R, Jones K, Hyman BT, Albert MS (2000) Use of structural magnetic resonance imaging to predict who will get Alzheimer's disease. *Ann Neurol* **47**, 430-439.

[5] Eckerström C, Olsson E, Bjerke M, Malmgren H, Edman A, Wallin A, Nordlund A (2013) A combination of neuropsychological, neuroimaging, and cerebrospinal fluid markers predicts conversion from mild cognitive impairment to dementia. *J Alzheimers Dis* **36**, 421-431.

[6] Ewers M, Brendel M, Rizk-Jackson A, Rominger A, Bartenstein P, Schuff N, Weiner MW (2014) Reduced FDG-PET brain metabolism and executive function predict clinical progression in elderly healthy subjects. *Neuroimage Clin* **4**, 45-52.

[7] Ewers M, Walsh C, Trojanowski JQ, Shaw LM, Petersen RC, Jack Jr. CR, Feldman HH, Bokde ALW, Alexander GE, Scheltens P, Vellas B, Dubois B, Weiner M, Hampel H (2012) Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. *Neurobiol Aging* **33**, 1203-1214.e2.

[8] Gomar JJ, Bobes-Bascaran MT, Conejero-Goldberg C, Davies P, Goldberg TE, 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. *Arch Gen Psychiatry* **68**, 961-969.

[9] Runtti H, Mattila J, Van Gils M, Koikkalainen J, Soininen H, Lötjönen J (2014) Quantitative evaluation of disease progression in a longitudinal mild cognitive impairment cohort. *J Alzheimers Dis* **39**, 49-61.

[10] Shaw LM, Vanderstichele H, Knapik-Czajka M, Clark CM, Aisen PS, Petersen RC, Blennow K, Soares H, Simon A, Lewczuk P, Dean R, Siemers E, Potter W, Lee VM-Y, Trojanowski JQ (2009) Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Ann Neurol* **65**, 403-413.

[11] Clark CM, Schneider JA, Bedell BJ, Beach TG, Bilker WB, Mintun MA, Pontecorvo MJ, Hefti F, Carpenter AP, Flitter ML, Krautkramer MJ, Kung HF, Coleman RE, Doraiswamy PM, Fleisher AS, Sabbagh MN, Sadowsky CH, Reiman EP, Zehntner SP, Skovronsky DM, AV45-A07 Study, Group (2011) Use of florbetapir-pet for imaging β-amyloid pathology. *JAMA* **305**, 275-283.

[12] Das SR, Avants BB, Grossman M, Gee JC (2009) Registration based cortical thickness measurement. *Neuroimage* **45**, 867-879.

[13] Tustison NJ, Cook PA, Klein A, Song G, Das SR, Duda JT, Kandel BM, van Strien N, Stone JR, Gee JC, Avants BB (2014) Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. *Neuroimage* **99**, 166-179.

[14] Avants BB, Libon DJ, Rascovsky K, Boller A, McMillan CT, Massimo L, Coslett HB, Chatterjee A, Gross RG, Grossman. M Sparse canonical correlation analysis relates network-level atrophy to multivariate cognitive measures in a neurodegenerative population. *Neuroimage* **84**, 698-711.

[15] Dhillon PS, Wolk DA, Das SR, Ungar LH, Gee JC, Avants BB (2014) Subject-specific functional parcellation via Prior Based Eigenanatomy. *Neuroimage* **99**, 14-27.

[16] Landau SM, Harvey D, Madison CM, Koeppe RA, Reiman EM, Foster NL, Weiner MW, Jagust WJ (2011) Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI. *Neurobiol Aging* **32**, 1207-1218.

[17] Rey A (1964) *L'examen clinique en psychologie*, Presses universitaires de France, Paris.

[18] Reitan RM (1958) Validity of the trail making test as an indicator of organic brain damage. *Percept Mot Skills* **8**, 271-276.

[19] Butters N, Granholm E, Salmon DP, Grant I, Wolfe J (1987) Episodic and semantic memory: A comparison of amnesic and demented patients. *J Clin Exp Neuropsychol* **9**, 479-497.

[20] Kaplan E, Goodglass H, Weintraub S, Goodglass H (1983) *Boston Naming Test*, Lea & Febiger, Philadelphia.

[21] Folstein MF, Folstein SE, McHugh PR (1975) Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* **12**, 189-198.

[22] Wolk DA, Dickerson BC (2011) Fractionating verbal episodic memory in Alzheimer's disease. *Neuroimage* **54**, 1530-1539.

[23] Snodgrass JG, Corwin J (1988) Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *J Exp Psychol Gen* **117**, 34-50.

[24] Shaw LM, Vanderstichele H, Knapik-Czajka M, Figurski M, Coart E, Blennow K, Soares H, Simon AJ, Lewczuk P, Dean RA, Siemers E, Potter W, Lee VM-Y, Trojanowski JQ (2011) Qualification of the analytical and clinical performance of CSF biomarker analyses in ADNI. *Acta Neuropathol (Berl)* **121**, 597-609.

[25] Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* **9**, 179-194.

[26] Desikan RS, Sabuncu MR, Schmansky NJ, Reuter M, Cabral HJ, Hess CP, Weiner MW, Biffi A, Anderson CD, Rosand J, Salat DH, Kemper TL, Dale AM, Sperling RA, Fischl B, the Alzheimer's Disease Neuroimaging Initiative (2010) Selective disruption of the cerebral neocortex in Alzheimer's disease. *PLoS One* **5**, e12853.

[27] Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15**, 273-289.

[28] McMillan CT, Avants B, Irwin DJ, Toledo JB, Wolk DA, Van Deerlin VM, Shaw LM, Trojanoswki JQ, Grossman M (2013) Can MRI screen for CSF biomarkers in neurodegenerative disease? *Neurology* **80**, 132-138.

[29] McMillan CT, Avants BB, Cook P, Ungar L, Trojanowski JQ, Grossman M (2014) The power of neuroimaging biomarkers for screening frontotemporal dementia. *Hum Brain Mapp* **35**, 4827-4840.

[30] Dietterich TG (1998) Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Comput* **10**, 1895-1923.

[31] Sing T, Sander O, Beerenwinkel N, Lengauer T (2005) ROCR: Visualizing classifier performance in R. *Bioinformatics* **21**, 3940-3941.

[32] Rentz DM, Parra Rodriguez MA, Amariglio R, Stern Y, Sperling R, Ferris S (2013) Promising developments in neuropsychological approaches for the detection of preclinical Alzheimer's disease: A selective review. *Alzheimers Res Ther* **5**, 58.

[33] Vemuri P, Wiste HJ, Weigand SD, Knopman DS, Shaw LM, Trojanowski JQ, Aisen PS, Weiner M, Petersen RC, Jack CR (2010) Effect of apolipoprotein E on biomarkers of amyloid load and neuronal pathology in Alzheimer disease. *Ann Neurol* **67**, 308-316.

[34] Kantarci K, Lowe V, Przybelski SA, Weigand SD, Senjem ML, Ivnik RJ, Preboske GM, Roberts R, Geda YE, Boeve BF, Knopman DS, Petersen RC, Jack CR (2012) APOE modifies the association between Aβ load and cognition in cognitively normal older adults. *Neurology* **78**, 232-240.

[35] Rowe CC, Bourgeat P, Ellis KA, Brown B, Lim YY, Mulligan R, Jones G, Maruff P, Woodward M, Price R, Robins P, Tochon-Danguy H, O'Keefe G, Pike KE, Yates P, Szoeke C, Salvado O, Macaulay SL, O'Meara T, Head R, Cobiac L, Savage G, Martins R, Masters CL, Ames D, Villemagne VL (2013) Predicting Alzheimer disease with β-amyloid imaging: Results from the Australian imaging, biomarkers, and lifestyle study of ageing. *Ann Neurol* **74**, 905-913.

[36] Heister D, Brewer JB, Magda S, Blennow K, McEvoy LK (2011) Predicting MCI outcome with clinically available MRI and CSF biomarkers. *Neurology* **77**, 1619-1628.

[37] Jack CR Jr, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, Petersen RC, Trojanowski JQ (2010) Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol* **9**, 119-128.

[38] Buchhave P, Minthon L, Zetterberg H, Wallin AK, Blennow K, Hansson O (2012) Cerebrospinal fluid levels of β-amyloid 1-42, but not of tau, are fully changed already 5 to 10 years before the onset of Alzheimer dementia. *Arch Gen Psychiatry* **69**, 98-106.

[39] Dickerson BC, Wolk D (2013) Biomarker-based prediction of progression in MCI: Comparison of AD signature and hippocampal volume with spinal fluid amyloid-β and tau. *Front Aging Neurosci* **5**, 55.

[40] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack CR Jr, Kaye J, Montine TJ, Park DC, Reiman EM, Rowe CC, Siemers E, Stern Y, Yaffe K, Carrillo MC, Thies B, Morrison-Bogorad M, Wagster MV, Phelps CH (2011) Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 280-292.