Routledge
Taylor & Francis Group

Journal of Clinical and Experimental Neuropsychology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ncen20

Predicting conversion from mild cognitive impairment to Alzheimer's disease with multimodal latent factors

Minyu Chang & C. J. Brainerd

To cite this article: Minyu Chang & C. J. Brainerd (2022) Predicting conversion from mild cognitive impairment to Alzheimer's disease with multimodal latent factors, Journal of Clinical and Experimental Neuropsychology, 44:4, 316-335, DOI: 10.1080/13803395.2022.2115015

To link to this article: https://doi.org/10.1080/13803395.2022.2115015

JOURNAL OF CLINICAL AND EXPERIMENTAL NEUROPSYCHOLOGY
2022, VOL. 44, NO. 4, 316–335
https://doi.org/10.1080/13803395.2022.2115015

Predicting conversion from mild cognitive impairment to Alzheimer's disease with multimodal latent factors

Minyu Chang and C. J. Brainerd

Department of Psychology and Human Neuroscience Institute, Cornell University, Ithaca, New York, USA

ABSTRACT

Introduction: We studied the ability of latent factor scores to predict conversion from mild cognitive impairment (MCI) to Alzheimer's disease (AD) and investigated whether multimodal factor scores improve predictive power, relative to single-modal factor scores.

Method: We conducted exploratory factor analyses (EFAs) and confirmatory factor analyses (CFAs) of the baseline data of MCI subjects in the Alzheimer's Disease Neuroimaging Initiative (ADNI) to generate factor scores for three data modalities: neuropsychological (NP), magnetic resonance imaging (MRI), and cerebrospinal fluid (CSF). Factor scores from single or multiple modalities were entered in logistic regression models to predict MCI to AD conversion for 160 ADNI subjects over a 2-year interval.

Results: NP factors attained an area under the curve (AUC) of .80, with a sensitivity of .66 and a specificity of .77. MRI factors reached a comparable level of performance (AUC = .80, sensitivity = .66, specificity = .78), whereas CSF factors produced weaker prediction (AUC = .70, sensitivity = .56, specificity = .79). Combining NP factors with MRI or CSF factors produced better prediction than either MRI or CSF factors alone. Similarly, adding MRI factors to NP or CSF factors produced improvements in prediction relative to NP or CSF factors alone. However, adding CSF factors to either NP or MRI factors produced no improvement in prediction.

Conclusions: Latent factor scores provided good accuracy for predicting MCI to AD conversion. Adding NP or MRI factors to factors from other modalities enhanced predictive power but adding CSF factors did not.

ARTICLE HISTORY
Received 2 February 2022
Accepted 16 August 2022

KEYWORDS
ADNI; mild cognitive impairment; alzheimer's disease; factor scores; disease progression

According to the Alzheimer's Association's (2021) report, 11.3% of American adults age 65 or older develop Alzheimer's disease (AD), which is the fifth leading cause of death in this population. In recent years, mild cognitive impairment (MCI) was introduced as an intermediate state between normal cognition and AD (Petersen, 2004; Petersen, 2011), with roughly one-third of MCI patients converting to AD within three years following MCI diagnoses (Mitchell & Shiri-Feshki, 2009). Considering that only a subset of MCI patients eventually develops AD, identifying those patients in advance is critical for developing targeted interventions that reduce the MCI to AD conversion rate. In other words, it is crucial to develop methods to distinguish MCI patients who convert to AD (MCI_C) from patients who do not convert (MCI_{NC}) as early as possible. The current paper tests the ability of latent factors to predict MCI to AD conversion and examines whether multimodal factor scores improve prediction. In the following sections, we first explain the advantages of using latent factors as predictors and review recent studies that implemented the latent-variable approach. Next, we briefly summarize the relevant findings for predictors of MCI to AD conversion with the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, which we used in the current study.

Latent factors as predictors of MCI to AD conversion

In most prior studies that investigated the prediction of MCI to AD conversion, individual variable scores were used as predictors. However, a few recent studies have attempted to use latent variables that are extracted from individual variables as predictors (Chapman et al., 2011; Eckerström et al., 2013; Giraldo et al., 2017; Wilhalme et al., 2017). In such an approach, a dimension reduction method such as principal component analysis (PCA) or exploratory factor analysis (EFA)[1] is used to extract the latent components or factors that underlie

CONTACT Minyu Chang mc2674@cornell.edu Department of Psychology and Human Neuroscience Institute, Cornell University, G341 MVR Hall, Ithaca, NY 14853

for the Alzheimer's Disease Neuroimaging Initiative*

Supplemental data for this article can be accessed online at https://doi.org/10.1080/13803395.2022.2115015.

a set of individual predictors. Then, subjects' component or factor scores, rather than their scores on individual predictors, supply the variables that are entered into prediction models.

Psychometrically, using latent variables as predictors for MCI to AD conversion has several benefits over using individual variables. First, latent variable scores reduce error variance through statistical aggregation (Rushton et al., 1983), and hence, they are normally more reliable than subjects' scores on individual predictors (Crane et al., 2012; Wilhalme et al., 2017). Second, some predictors can be differentially sensitive to AD conversion in different sub-populations, and latent variables can minimize such bias (Gibbons et al., 2012). For instance, word list recognition can detect cognitive declines when functional impairment is still absent, whereas word list immediate recall cannot capture declines until mild functional impairment is present (Jutten et al., 2021). In such circumstances, incorporating these test scores into component or factor scores should be more appropriate for a total population with subjects at different clinical stages. Third, using latent variables can eliminate multicollinearity and overfitting problems without dropping predictors from the model. That is, by using latent variables as predictors, one can decrease the number of variables that are entered when fitting prediction models while simultaneously preserving information about the individual variables.

In line with these points, the latent-variable approach has demonstrated good potential for improving MCI to AD prediction. To illustrate, Wilhalme et al. (2017) conducted an EFA for the neuropsychological (NP) data of 71 MCI patients in the Imaging and Genetic Biomarkers for AD (ImaGene) study. They found that the area under the receiver operating characteristic (ROC) curve (AUC) for the individual NP factor scores were very promising, ranging from .73 to .89. Similarly, Chapman et al. (2011) used a PCA to generate component scores for the NP data of 43 MCI patients in a longitudinal study. They reported that the accuracy level of MCI_C-MCI_{NC} classifications using component scores was quite good (.84, sensitivity = .86, specificity = .83). Finally, Eckerström et al. (2013) conducted PCAs of the NP, magnetic resonance imaging (MRI), and cerebrospinal fluid (CSF) data of 42 MCI patients in the Gothenburg MCI study and reported that combining the component scores for the three modalities accounted for a near-perfect AUC of .98.

In brief, the latent-variable approach has several psychometric benefits, and a few studies have shown that it is a promising candidate for improving the prediction of MCI to AD conversions. However, the aforementioned studies were all conducted with very small sample sizes and mostly with a single data modality (i.e., NP data). Thus, it is important to verify the predictive ability of latent factors using larger sample sizes and multiple data modalities.

Major predictors of MCI to AD conversion in the ADNI

In the current paper, we used the ADNI dataset for testing the latent-variable approach. The ADNI is an ongoing, longitudinal study that collects cognitive, biochemical, and neuroimaging data from healthy controls, MCI patients, and AD patients at multiple sites in the United States and Canada (Mueller et al., 2005). Since the study began, the ADNI database has formed the basis for large numbers of studies in which potential predictors of AD have been evaluated, particularly NP assessments, MRI biomarkers, and CSF biomarkers (Weiner et al., 2017). Below, we briefly summarize the findings for these three modalities of predictors.

NP assessments are the least expensive, invasive, and time-consuming of AD predictors. Moreover, they have demonstrated comparable or higher accuracy relative to predictors from the other modalities (Cui et al., 2011; Devanand et al., 2012; Eckerström et al., 2013; Gomar et al., 2014; Li et al., 2017). For example, when classifying 143 MCI patients as MCI_C or MCI_{NC}, Cui et al. (2011) found that a cluster of five NP instruments reached an AUC of .76 (sensitivity = .91 and specificity = .48). Furthermore, Gomar et al. (2014) obtained an AUC of .78 (sensitivity = .58 and specificity = .74) for 318 MCI patients with only two NP assessments.

Turning to MRI predictors, many ADNI studies have provided supporting evidence that MRI biomarkers can be reliable predictors of MCI to AD conversion (Barnes et al., 2014; Cui et al., 2011; Cuingnet et al., 2011; Davatzikos et al., 2011; Devanand et al., 2012; Ewers et al., 2012; Li et al., 2017). For instance, Ewers et al. (2012) showed that right entorhinal cortical thickness was the best single predictor for a sample of 130 MCI patients, with an overall prediction accuracy of .69 (sensitivity = .53 and specificity = .77). In addition, Devanand et al. (2012) found that hippocampal and entorhinal volumes together yielded an AUC of .74 (sensitivity = .56 with specificity fixed at .80) for 282 MCI patients.

Continuing to CSF biomarkers, the most common CSF predictors of AD are the CSF concentration of total tau (t-tau), amyloid beta ($A\beta_{42}$), phosphorylated tau (p-tau), ratios of t-tau to $A\beta_{42}$, and ratios of t-tau to $A\beta_{42}$ (Cui et al., 2011; Davatzikos et al., 2011; Fjell et al., 2010; Gomar et al., 2011, 2014; Shaw et al., 2009;

Ye et al., 2012), although there is also increasing interest in novel non-amyloid and non-tau CSF predictors (Weiner et al., 2017). In predicting MCI to AD conversion, Gomar et al. (2011) and Ye et al. (2012) reported similar AUCs for t-tau/$A\beta_{42}$: .64 with 168 MCI patients and .63 with 319 MCI patients, respectively. In addition, Cui et al. (2011) found that a combination of t-tau/$A\beta_{42}$ and p-tau/$A\beta_{42}$ delivered a similar level of predictive performance (AUC = .64, sensitivity = .80, specificity = .48). Notably, the reported predictive power of CSF variables with ADNI subjects seems relatively weaker than that of NP or MRI variables.

Because abundant data are available for three distinct modalities in ADNI, it is natural to ask whether combining data across modalities can improve the prediction of future AD. However, prior studies have yielded mixed findings on the utility of multimodal predictors. On the one hand, some research suggests that combining predictors from different modalities can increase prediction accuracy (Chen et al., 2015; Cui et al., 2011; Devanand et al., 2012; Shaffer et al., 2013; Ye et al., 2012; Zhang et al., 2012). On the other hand, multimodal predictors have failed to outperform single-modal predictors in other studies (Ewers et al., 2012; Gomar et al., 2011, 2014; Richard et al., 2013).

The current study

Our study had two goals. First, although some research shows that latent variable scores are promising candidates for improving predictions of MCI to AD conversions, they were all conducted with small patient samples. In the current study, we built on this line of research by implementing the EFA and CFA approaches with the ADNI data, which is one of the most comprehensive AD databases. Second, as prior studies reported mixed findings on whether combining cognitive, neuroimaging, and biochemical variables can improve predictions of MCI to AD conversions, we examined whether combining latent factor scores from multiple modalities (NP, MRI, and CSF) yields better predictive accuracy. To isolate the latent factors for each modality and to generate factor scores, we conducted three separate EFAs and three separate CFAs. Then, single-modal and multimodal factor scores were entered into a series of logistic regression models that predicted MCI to AD conversions in the ADNI sample over a 2-year interval. Our working hypotheses were that latent factor scores should yield good predictive power and that combinations of NP, MRI, and CSF factor scores should improve predictive power, relative to single-modal factor scores.

Method

Subjects

The data that we analyzed were obtained from the ADNI 1 database (http://adni.loni.usc.edu/). The ADNI includes four grant periods: ADNI 1, ADNI GO, ADNI 2, and ADNI 3. The ADNI 1 includes the initial ADNI cohort of 819 ADNI subjects. In each phase following ADNI 1, new subjects were added and the subjects from the prior phases were followed up to the extent possible. In the current paper, we focused on the ADNI 1 data because the data for most non-amyloid and non-tau CSF variables were only available for this cohort. Specifically, we focused on ADNI 1 subjects with MCI diagnoses in the baseline session (N = 397) and tracked their clinical diagnoses in the following 24 months. Clinical diagnoses of MCI or AD were established at baseline and at each subsequent session (at 6, 12, 18, 24 months) based on the following five criteria: (a) both MCI and AD subjects must have memory complaints made by the subject or by a study partner; (b) education-adjusted Logical Memory II subscale [from the Wechsler Memory Scale-Revised (Wechsler, 1987)] scores: for ≥ 16 years of education, 9–11 for MCI and ≤ 8 for AD; for 8–15 years of education, 5–9 for MCI and ≤ 4 for AD; for 0–7 years of education, 3–6 for MCI and ≤ 2 for AD; (c) Mini Mental State Exam scores: 24–30 for MCI and 20–26 for AD; (d) Clinical Dementia Rating (CDR; Morris, 1993) scores: total score = 0.5 and memory box score ≥ 0.5 for MCI[2] and total score = 0.5 or 1 for AD; (e) MCI subjects must have relatively preserved general cognitive and functional performance that rule out an AD diagnosis, whereas AD subjects must have met the National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA) criteria for probable AD. Full details of the diagnostic criteria can be found in the ADNI 1 Procedures Manual (http://adni.loni.usc.edu/methods/documents/).

We included all 397 ADNI 1 baseline MCI subjects in the EFA and CFA of NP predictors, as NP data were available for all these subjects. For the same reason, we used all 397 MCI subjects' pre-processed and quality-controlled MRI data in the EFA and CFA of MRI predictors. Because CSF data were only collected from a subsample of 187 MCI subjects, the EFA and CFA of CSF predictors were confined to those subjects. A final group of logistic regressions was restricted to the MCI

subjects for whom NP, MRI, and CSF data were available for baseline, and MCI to AD conversion statuses were identifiable for the following 24 months (N = 160).[3] These subjects were split into two subgroups: MCI_C (N = 68) and MCI_{NC} (N = 92).[4]

Overview of predictors

NP predictors

We included baseline scores from the following NP tests in our analyses: Alzheimer's Disease Assessment Scale – Cognitive (ADAS-Cog; Rosen et al., 1984) delayed recall, recognition, naming, and cancelation, Rey Auditory Verbal Learning Test (RAVLT; Rey, 1964) trial 1–5 recall and forgetting, Clock Drawing Test (CDT) command and copy, Digit Span forward and backward, Categorical Fluency Test (CFT; Harrison et al., 2000) animals and vegetables, Boston Naming Test (Goodglass et al., 1983), and Trail Making Tests (TMT; Reitan & Wolfson, 1985) parts A and B.

MRI predictors

MRI data were acquired with 1.5 Tesla MRI scanners at multiple sites and were processed with FreeSurfer Version 4.3.0 at the University of California San Francisco.[5] We selected multiple regions of interest that were found to be most discriminating between MCI_C and MCI_{NC} groups based on Risacher et al.'s (2009) review, including entorhinal cortex, hippocampus, amygdala, temporal pole, middle temporal gyrus, inferior temporal gyrus, superior temporal gyrus, inferior parietal gyrus, superior parietal gyrus, inferior lateral ventricle, cerebral cortex, precuneus, supramarginal gyrus, inferior lateral ventricle, lateral ventricle, nucleus accumbens, and ventral dorsal column. For these regions of interest, we considered cerebral volume or cortical thickness or both. For cerebral volume, we used the summed values of the left and right hemispheres. In addition, to control for individual differences in brain size, we used a normalization procedure in which cerebral volume was divided by intracranial volume (ICV; Shi et al., 2009; Whitwell et al., 2001). For cortical thickness, we used mean values of the left and right hemispheres.

CSF predictors

CSF samples were collected via lumbar puncture. We considered 15 CSF predictors that had been identified as promising biomarkers based on Weiner et al.'s (2017) review and multiple other ADNI studies (Deming et al., 2016; Khan et al., 2015; Mattsson et al., 2014; Morenas-Rodríguez et al., 2016; Paterson et al., 2014; Portelius et al., 2015; Suárez-Calvet et al., 2016; Toledo et al., 2014), including CSF concentrations of t-tau, p-tau, $A\beta_{42}$, t-tau/$A\beta_{42}$, p-tau/$A\beta_{42}$, neurogranin, Interleukin 6, factor H, complement C3, progranulin, clusterin, chromogranin-A, cystatin-C, fatty acid-binding protein 3, and sTREM2.

Overview of statistical procedures

EFAs

The EFAs were aimed at identifying latent factors that could be used to predict MCI to AD conversion. Three separate EFAs were conducted for NP, MRI, and CSF predictors using the psych package (Revelle, 2016) in (R Core Team, 2019). We only included predictors with Kaiser-Meyer-Olkin (KMO; Kaiser, 1970; Kaiser & Rice, 1974) scores > .5, which is the conventional indicator of sampling adequacy in factor analyses (Williams et al., 2010; Yong & Pearce, 2013). Following Costello and Osborne's (2005) recommendations, we considered multiple sources of information when determining the number of factors to extract, including Kaiser's criterion (eigenvalue > 1; Kaiser, 1958), scree plots (Cattell & Vogelmann, 1977), parallel analysis (Horn, 1965) and *a priori* factor structure. Specifically, we ran multiple EFAs with the number of factors set at values above, equal to, and below the suggested numbers. After that, we chose the best-fitting factor solution by applying the following criteria: statistical interpretability, theoretical significance, parsimony, and absence of Heywood cases. The factors were then extracted using principal axis factoring and rotated using the varimax method. Finally, factor scores were estimated for these factors using the Bartlett algorithm (Hershberger, 2005).

CFAs[6]

The CFAs were also conducted to generate factor scores that can be used to predict MCI to AD conversion. All the CFA analyses were conducted with the lavaan package in R (Rosseel, 2012). All variables were standardized before being entered into the CFA models. As will be seen in the Results section, the factor solutions produced by EFAs are of high theoretical interpretability. Thus, we used the same factor solutions in CFAs as in EFAs. Additionally, we modified the models based on the modification indices (MIs)[7] by adding paths between latent factors and observed variables or adding error covariances between observed variables. For parsimony, we only considered modifications with MI > 10. To avoid over-fitting, we only made modifications that are theoretically meaningful. Meanwhile, we

avoid making any modifications that would introduce Heywood cases.

Logistic regressions

As mentioned, baseline MCI subjects were classified as MCI_C or MCI_{NC}. We used logistic regression to predict conversion status (MCI_C vs. MCI_{NC}). The logistic regressions and cross-validations were conducted using the stats and caret packages in R, respectively (R Core Team, 2019; Kuhn, 2008), and the accompanying ROC analyses were conducted with the pROC package in (Robin et al., 2011). In all the logistic regression models, we included age, education, gender, and APOE ε4 status as covariates. We first conducted logistic regressions using single-modal factors (NP, MRI or CSF) as predictors. Next, we used factors from two modalities (NP + MRI, NP + CSF, or MRI + CSF) as predictors, and last, we used factors from all three modalities (NP + MRI + CSF). By adopting such an approach, we were able to determine (a) whether factor scores had good predictive power, as measured by AUC, sensitivity, and specificity; (b) whether incorporating factor scores from multiple modalities improved predictive power compared to using single-modal factor scores. Finally, we used a leave-one-out cross-validation to estimate the robustness of our models. In the leave-one-out cross-validation, we left out one participant's data in each iteration, which served as testing data to which we applied the parameter estimates, and used the remaining data ($N - 1$, where N is the sample size) as training data from which we derived the parameter estimates. Thus, such a procedure was repeated N times, each for one particular participant's data.

Results

Descriptive statistics

We report the descriptive statistics for the key demographic characteristics, 16 NP predictors, 19 MRI predictors, and 15 CSF predictors, separately for MCI_C and MCI_{NC} subjects in Table 1. In addition, we conducted chi-squared tests or two-sample t-tests to compare those values between MCI_C and MCI_{NC} subjects and report the effect sizes and p values in Table 1.

Factor analyses

EFAs

The scree plots and parallel analyses for the EFAs of NP, MRI, and CSF predictors are displayed in Figure 1. As indicated there, the different tests for factor retention suggested 2–5 factors for NP predictors, 2–4 factors for MRI predictors, and 2–4 factors for CSF predictors. After careful comparison between these potential factor solutions using the aforementioned selection criteria, we ultimately extracted five factors in the EFA of NP predictors, three factors in the EFA of MRI predictors, and three factors in the EFA of CSF predictors. It should be noted that the five-factor solution for NP predictors aligns with the results of other recent factor analyses of the ADNI NP battery (Chang & Brainerd, 2021; Park et al., 2012).

The factor loadings in the EFAs of NP, MRI, and CSF predictors are reported in Tables 2, 3, and 4, respectively. A visual inspection of Table 2 reveals a highly interpretable NP factor structure. Factor 1 is clearly a memory factor as all memory measures loaded highly on it. Similarly, the specific tests that load highly on factors 2–5 indicated that they are language, executive function, attention, and visuospatial processing factors, respectively. Table 3 reveals a highly interpretable MRI factor structure, too. Factor 1 is a temporal-parietal factor because middle, inferior, and superior temporal gyrus, inferior and superior parietal cortex, precuneus (part of the superior parietal lobule), and supramarginal gyrus (part of the parietal lobe) all loaded highly on it. Factor 2 is an entorhinal-amygdala factor as entorhinal cortex loaded the highest on this factor, followed by amygdala.[8] Finally, factor 3 is a ventricle factor as inferior lateral ventricle and lateral ventricle loaded the highest on it.

Next, continuing to CSF factors in Table 4, factor 1 is clearly an amyloid-tau factor because t-tau, p-tau, $A\beta_{42}$, t-tau/$A\beta_{42}$, and p-tau/$A\beta_{42}$ all loaded highly on it. Factor 2 is a neuroprotection factor, as the variables that loaded the highest on it (clusterin and cystatin C) were shown to play a protective role against AD by modulating Aβ fibril formation and toxicity (Boggs et al., 1996; Kaur & Levy, 2012; Miners et al., 2017). Additionally, cystatin C and chromogranin-A were both negatively associated with brain atrophy rate (Paterson et al., 2014). Factor 3 is a neuroinflammation factor as it is characterized by the strong loadings of complement C3, factor H, progranulin, and Interleukin 6, which all contribute to or regulate neuroinflammation (Hampel et al., 1999; Hu et al., 2016; Kumar et al., 2015; Martens et al., 2012; Pogue et al., 2009).[9]

We generated factor scores from the three EFAs using the Bartlett method. The descriptive statistics of those factor scores are provided in Table 5. For the five NP factors, MCI_{NC} subjects had higher scores of factors 1 (memory) and 5 (visuospatial processing), and lower scores of factor 3 (executive function) than MCI_C subjects, while they did not differ in scores of the other two

Table 1. Descriptive data for the 160 ADNI 1 subjects with mild cognitive impairment at baseline.

		Group			
Variables		MCI_C (N = 68)	MCI_{NC} (N = 92)	Effect Size	p value
Demographic					
	Age	74.80 (7.16)	74.10 (7.69)	.04	.474
	Education	15.57 (2.98)	16.25 (2.74)	.24	.059
	Gender, Female (%)	35.29%	33.70%	.003	.966
	APOE ε4 carriers (%)	64.71%	46.73%	.17	.072
NP					
	ADAS Delayed Recall	.29 (.20)	.43 (.24)	.63	**< .001**
	ADAS Recognition	.56 (.24)	.64 (.22)	.37	**.041**
	RAVLT Trial 1–5 Recall	.35 (.09)	.44 (.12)	.83	**< .001**
	RAVLT Forgetting	.79 (.28)	.63 (.32)	.54	**< .001**
	CDT Command	3.84 (1.11)	4.36 (.86)	.52	**.006**
	CDT Copy	4.59 (.63)	4.64 (.78)	.08	.686
	Digit Span Forward	8.46 (2.02)	8.41 (2.05)	.02	.895
	Digit Span Backward	6.10 (1.79)	6.35 (2.40)	.12	.530
	CFT Animal	15.04 (4.64)	16.25 (4.54)	.26	.167
	CFT Vegetable	9.84 (2.91)	11.40 (3.55)	.48	**.008**
	Boston Naming Test	25.25 (3.96)	26.16 (4.02)	.23	.215
	ADAS Naming	4.75 (.47)	4.73 (.58)	.04	.817
	TMT Part A	49.26 (24.24)	41.26 (16.82)	.38	**.038**
	TMT Part B	154.12 (74.62)	108.14 (56.28)	.70	**< .001**
	ADAS Cancellation	3.93 (.86)	4.09 (.80)	.20	.295
MRI					
	Entorhinal Cortex CV	.0020 (.0005)	.0022 (.0005)	.45	**.013**
	Entorhinal Cortex TA	3.00 (.42)	3.22 (.46)	.52	**.005**
	Hippocampus SV	.0037 (.0006)	.0041 (.0007)	.63	**< .001**
	Amygdala SV	.0015 (.0002)	.0016 (.0002)	.50	**.007**
	Temporal Pole TA	3.33 (.33)	3.45 (.33)	.37	**.041**
	Middle Temporal Gyrus CV	.0110 (.0014)	.0123 (.0015)	.90	**< .001**
	Middle Temporal Gyrus TA	2.29 (.19)	2.43 (.15)	.85	**< .001**
	Inferior Temporal Gyrus TA	2.57 (.20)	2.70 (.19)	.71	**< .001**
	Superior Temporal Gyrus TA	2.36 (.20)	2.45 (.18)	.47	**.011**
	Inferior Parietal Gyrus TA	2.07 (.19)	2.21 (.14)	.80	**< .001**
	Superior Parietal Gyrus TA	1.86 (.18)	1.94 (.16)	.52	**.005**
	Cerebral Cortex SV	.2443 (.0181)	.2516 (.0201)	.38	**.036**
	Precuneus TA	2.96 (.17)	2.07 (.14)	.75	**< .001**
	Supramarginal Gyrus TA	2.17 (.17)	2.28 (.16)	.64	**< .001**
	Inferior Lateral Ventricle SV	.0018 (.0010)	.0012 (.0009)	.60	**<.001**
	Lateral Ventricle SV	.0283 (.0120)	.0234 (.1112)	.43	**.020**
	Nucleus Accumbens SV	.0005 (.0001)	.0006 (.0001)	.45	**.013**
	Cerebral Cortex White Matter SV	.2690 (.0237)	.2782 (.0234)	.39	**.034**
	Ventral Dorsal Column SV	.0045 (.0005)	.0046 (.0005)	.25	.177
CSF					
	t-tau	111.06 (50.14)	92.94 (51.64)	.36	**.045**
	p-tau	40.08 (16.88)	32.25 (15.30)	.49	**.008**
	$A\beta_{42}$	145.33 (37.37)	173.83 (56.88)	.59	**< .001**
	t-tau/$A\beta_{42}$	.82 (.41)	.63 (.51)	.37	**.037**
	p-tau/$A\beta_{42}$	.30 (.15)	.22 (.15)	.49	**.007**
	Neurogranin	529.91 (378.79)	471.34 (334.75)	.16	.386
	Interleukin 6	6.89 (6.07)	5.11 (6.17)	.13	.569
	Factor H	1545.39 (654.11)	1678.34 (639.41)	.21	.268
	Complement C3	3789.43 (2287.03)	4335.09 (2334.38)	.24	.206
	Progranulin	1721.28 (1266.81)	1541.31 (286.94)	.20	.386
	Clusterin	20.55 (.50)	20.69 (.46)	.29	.175
	Chromogranin-A	19.971(.83)	20.09 (1.16)	.17	.388
	Cystatin-C	34.10 (.39)	34.19 (.38)	.23	.268
	Fatty Acid-Binding Protein 3	14.76 (.39)	14.74 (.46)	.04	.817
	sTREM2	4362.28 (1992.74)	4575.78 (2841.07)	.09	.686

Standard deviations are included in parentheses. MCI_C = MCI subjects who converted to AD within 24 months; MCI_{NC} = MCI subjects who did not convert to AD within 24 months.

NP = neuropsychological; MRI = magnetic resonance imaging; CSF = cerebrospinal fluid.

ADAS = Alzheimer's Disease Assessment Scale; RAVLT = Rey Auditory Verbal Learning test; CDT = Clock Drawing Test; CFT = Categorical Fluency Test; TMT = Trail Making Test.

RAVLT Forgetting was measured as the percentage of recall decline between the 5th and the delayed recall test. Scores derived from the ADAS battery were all reversed-coded so higher scores indicate better performance.

SV = subcortical volume. CV = cortical volume. TA = cortical thickness average. Subcortical and cortical volumes were normalized by intracranial volume. The unit of measurement for cortical thickness is mm.

t-tau = total tau; p-tau = phosphorylated tau. sTREM2 = soluble triggering receptor expressed on myeloid cell 2.

The unit of measurement for t-tau, p-tau, and $A\beta_{42}$, Neurogranin, Interleukin 6, Progranulin, Fatty Acid-Binding Protein 3, and sTREM2 is pg/ml. The unit of measurement for Factor H, Complement C3, Cystatin-C, and Chromogranin-A is ng/ml. The unit of measurement for Clusterin is µg/ml.

The effect size column indicates Cramer's V of chi-squared tests or Cohen's d of two-sample t tests between MCI_C and MCI_{NC}.

The p value column indicates p values of chi-squared tests or of two-sample t tests between MCIC and MCINC, which were corrected for multiple comparisons using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995). P values < .05 were highlighted in bold font.

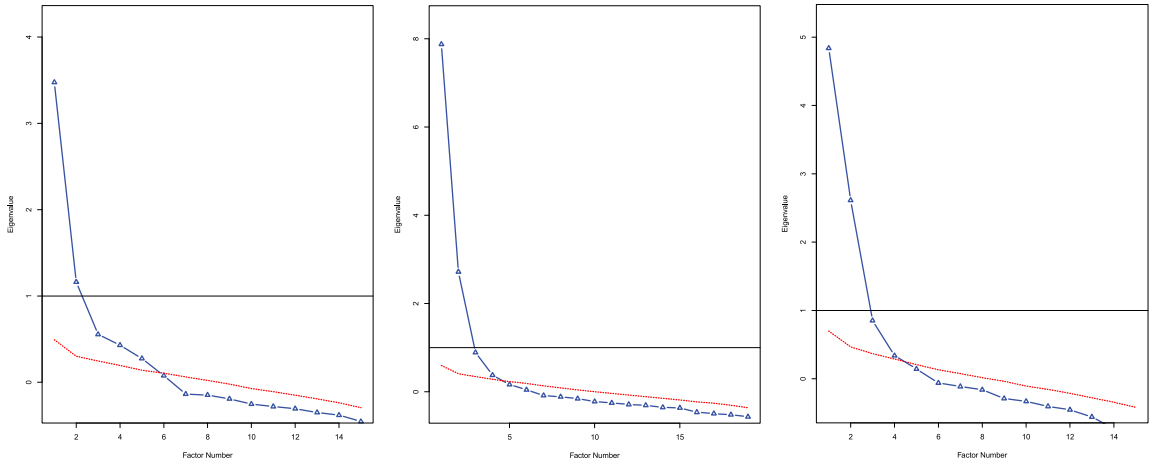

Figure 1. The left panel displays the scree plot with parallel analysis for the neuropsychological predictors. The middle panel displays the scree plot with parallel analysis for the magnetic resonance imaging predictors. The right panel displays the scree plot with parallel analysis for the cerebrospinal fluid predictors. The blue solid line with triangles indicates the observed eigenvalues, with each triangle representing one factor. The black solid line indicates eigenvalue = 1, and thus the number of triangles above the black line indicates the number of factors to retain, as suggested by the Kaiser's criterion. The red dotted line indicates the random eigenvalues of the simulated data. The number of triangles on the blue line that lies above the corresponding red line indicates the number of factors to retain, as suggested by parallel analysis.

factors (language and attention). Here, because both TMT variables loaded positively on factor 3 and their index variable is response time, lower factor 3 scores actually indicate better executive function. Thus, the results of factor scores suggest that MCI subjects who did not convert to AD had better memory, executive function, and visuospatial processing than those who converted to AD.

Continuing to MRI factors, Table 5 shows that MCI_{NC} subjects had higher scores of all three MRI factors (temporal-parietal, entorhinal-amygdala, and ventricle) than MCI_C subjects. For factor 3, because both inferior lateral ventricle and lateral ventricle loaded negatively on it, higher factor 3 scores suggest smaller ventricles. Thus, the results indicate that MCI_{NC} subjects had larger brain volumes and thicker cortex in the selected brain regions as well as smaller brain ventricles relative to MCI_C subjects.

As for CSF factors, scores of factor 1 (amyloid-tau) were significantly higher for MCI_C subjects than for MCI_{NC} subjects. Here, higher scores of factor 1 indicate higher CSF levels of t-tau and p-tau as well as lower CSF levels of $A\beta_{42}$, which probably reflects higher intensity of intracellular neurofibrillary tangles and higher levels of deposition of $A\beta_{42}$ in plaques (Blennow, 2004; Welge et al., 2009). Scores of the other two factors did not differ significantly between MCI_{NC} and MCI_{NC} subjects.

CFAs

As mentioned, we used the same 5-factor solution for NP variables and 3-factor solutions for MRI and CSF variables in CFAs as in EFAs, and we made necessary modifications based on modification indices. The final factor solutions of the three CFAs for NP, MRI, and CSF variables (Figures S1, S2, and S3) and the CFA factor loadings (Table S1) are presented in the Supplementary Materials. We relied on comparative fit index (CFI), root mean square of approximation (RMSEA), and standardized root mean squared residual (SRMR) for evaluating the CFA model fits, with CFI ≥ .90, RMSEA ≤

Table 2. Exploratory factor analysis loadings of neuropsychological predictors for 160 ADNI 1 subjects with mild cognitive impairment at baseline.

NP Predictors	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
ADAS Delayed Recall	**.74**	−.04	−.07	−.05	.02
ADAS Recognition	**.53**	.01	−.02	−.09	−.07
RAVLT Trial 1–5 Recall	**.67**	.14	−.03	.17	.01
RAVLT Forgetting	**−.72**	.04	−.13	.03	−.05
CDT Command	.07	.01	−.04	.00	**.69**
CDT Copy	−.07	.09	−.05	−.02	**.54**
Digit Span Forward	−.04	.07	.06	**.65**	−.03
Digit Span Backward	.03	−.05	−.09	**.72**	.03
CFT Animal	.11	**.59**	−.19	−.03	−.08
CFT Vegetable	.24	**.47**	−.15	−.03	−.02
Boston Naming Test	−.04	**.78**	.02	.01	.08
ADAS Naming	−.05	**.54**	.14	.10	.13
TMT Part A	.06	.03	**.72**	.01	−.08
TMT Part B	−.07	−.03	**.63**	−.14	−.06
ADAS Cancellation	−.02	.10	**−.58**	−.05	−.02
SS loadings	1.96	1.70	1.50	1.05	.92

Note. ADAS = Alzheimer's Disease Assessment Scale; RAVLT = Rey Auditory Verbal Learning test; CDT = Clock Drawing Test; CFT = Categorical Fluency Test; TMT = Trail Making Test.

RAVLT Trial 1–5 Recall was the average recall across trials 1 to 5. RAVLT Forgetting was measured as the percentage of recall decline between the 5th and the delayed trial. Scores derived from the ADAS battery were all reversed-coded so higher scores indicate better performance. Loadings ≥ .40 were highlighted in bold font. SS loadings = sum of squared loadings.

Table 3. Exploratory factor analysis loadings of magnetic resonance imaging predictors for 160 ADNI 1 subjects with mild cognitive impairment at baseline.

MRI Predictors	Factor 1	Factor 2	Factor 3
Entorhinal Cortex CV	−.08	**.82**	−.02
Entorhinal Cortex TA	.08	**.86**	−.09
Hippocampus SV	.00	**.65**	.29
Amygdala SV	−.09	**.78**	.04
Temporal Pole TA	.20	**.66**	−.02
Middle Temporal Gyrus CV	**.44**	.25	.26
Middle Temporal Gyrus TA	**.90**	.05	.08
Inferior Temporal Gyrus TA	**.67**	.22	.12
Superior Temporal Gyrus TA	**.65**	.23	.12
Inferior Parietal Gyrus TA	**.96**	−.04	.00
Superior Parietal Gyrus TA	**.87**	−.02	−.25
Cerebral Cortex SV	**.45**	.15	.32
Precuneus TA	**.93**	−.04	−.13
Supramarginal Gyrus TA	**.90**	−.09	.18
Inferior Lateral Ventricle SV	−.06	−.27	**−.60**
Lateral Ventricle SV	.01	.07	**−.96**
Nucleus Accumbens SV	.06	.17	**.48**
Cerebral Cortex White Matter SV	.14	.04	**.59**
Ventral Dorsal Column SV	−.08	.16	**.53**
SS loading	5.78	3.64	2.89

Note. MRI = magnetic resonance imaging. SV = subcortical volume. CV = cortical volume. TA = cortical thickness average. Subcortical and cortical volumes were normalized by intracranial volume. The unit of measurement for cortical thickness is mm.
Loadings ≥ .40 were highlighted in bold font. SS loadings = sum of squared loadings.

Table 4. Exploratory factor analysis loadings of cerebrospinal fluid predictors for 160 ADNI 1 subjects with mild cognitive impairment at baseline.

CSF Predictors	Factor 1	Factor 2	Factor 3
t-tau	**.82**	.28	−.05
p-tau	**.89**	.38	−.03
$A\beta_{42}$	**−.74**	.06	−.03
t-tau/$A\beta_{42}$	**.92**	.10	.01
p-tau/$A\beta_{42}$	**.98**	−.10	.02
Neurogranin	**.57**	.31	−.16
Interleukin 6	.15	−.30	**.41**
Factor H	−.06	.21	**.76**
Complement C3	−.08	.06	**.90**
Progranulin	.18	−.21	**.66**
Clusterin	−.06	**.70**	.31
Chromogranin-A	.09	**.63**	−.10
Cystatin C	.15	**.79**	.14
Fatty Acid-Binding Protein 3	**.40**	**.58**	.03
sTREM2	.14	.38	.24
SS loading	4.58	2.73	2.30

Note. CSF = cerebrospinal fluid; t-tau = total tau; p-tau = phosphorylated tau; sTREM2 = soluble triggering receptor expressed on myeloid cell 2.
Loadings ≥ .40 were highlighted in bold font. SS loading = sum of squared loadings.

.08, and RMSEA ≤ .08 indicating adequate model fits (Browne & Cudek, 1993; Hu & Bentler, 1999). The fit of the CFA model for NP variables was excellent (CFI = .97, RMSEA = .03, SRMR = .04), but the fits of the CFA models for MRI variables (CFI = .90, RMSEA = .11, SRMR = .10) and for CSF variables (CFI = .91, RMSEA = .13, SRMR = .08) were slightly worse. Here, we noted that although all three fit indices delivered a good fit for the NP model, there were disagreements between CFI, RMSEA, and SRMR for the MRI and CSF models. However, because CFIs indicated adequate fits for the latter two CFA models and disagreements between these fit statistics are typically not diagnostic of model specification problems (Lai & Green, 2016), we continued to generate factor scores using the current factor structures.

Table 5. Exploratory factor analysis factor scores for 160 ADNI 1 subjects with mild cognitive impairment at baseline.

	Group			
Factors	MCI_C (N = 68)	MCI_{NC} (N = 92)	Cohen's d	p value
NP-Factor 1	−.60 (.84)	.33 (1.20)	.85	**< .001**
NP-Factor 2	−.09 (1.08)	.13 (1.14)	.21	.259
NP-Factor 3	.45 (1.24)	−.28 (.84)	.57	**< .001**
NP-Factor 4	.08 (1.17)	.09 (1.27)	.09	.605
NP-Factor 5	−.002 (1.18)	.23 (1.32)	.37	**.038**
MRI-Factor 1	−.39 (1.01)	.31 (.82)	.81	**< .001**
MRI-Factor 2	−.35 (.88)	.18 (1.01)	.60	**< .001**
MRI-Factor 3	−.18 (.96)	.26 (.92)	.47	**.008**
CSF-Factor 1	.36 (.91)	−.21 (1.01)	.50	**.005**
CSF-Factor 2	−.19 (1.11)	.12 (1.10)	.29	.114
CSF-Factor 3	.06 (1.45)	−.04 (.79)	.15	.417

Note. Standard deviations are included in parentheses.
MCI_C = MCI subjects who converted to AD within 24 months; MCI_{NC} = MCI subjects who did not convert to AD within 24 months; NP = neuropsychological; MRI = magnetic resonance imaging; CSF = cerebrospinal fluid.
The factors with suffixes indicate the separate factors extracted for the NP, MRI, and CSFpredictors.
The p values of two-sample t tests between MCI_C and MCI_{NC} were corrected for multiple comparisons using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995), with p values < .05 highlighted in bold font.

Logistic regressions

Using EFA factor scores as predictors

Next, we turn to the logistic regression models in which factor scores from the EFAs were used as predictors of the MCI to AD conversion. As shown in rows 1–3 of Table 6, when factor scores of a single modality are analyzed, NP factor scores produced an AUC of .80 (sensitivity = .66 and specificity = .77 when c = .5).[10] Similarly, MRI factor scores produced an AUC of .80 (sensitivity = .66, specificity = .78), and CSF factor scores produced an AUC of .70 (sensitivity = .56, specificity = .79). A visual inspection of Figure 2 reveals that the AUC of CSF factor scores is smaller than that of NP or MRI factor scores. We conducted one-sided DeLong' tests (DeLong et al., 1988; Sun & Xu, 2014) to determine whether such differences were statistically significant.[11] The tests showed that the AUC for NP factors was larger than that of CSF factors, z = 1.98, p = .033, and the AUC for MRI factors was larger than that of CSF factors, z = 2.10, p = .027. There was no significant difference between the AUCs of NP and of MRI factors.

Furthermore, DeLong' tests showed that the combination of NP and MRI factors increased AUC compared to NP factors alone, z = 2.26, p = .027, or MRI factors

Table 6. Summary of logistic regression analyses with exploratory factor analysis factor score(s) as predictor(s) for 160 ADNI 1 subjects with mild cognitive impairment at baseline.

		MCI_C vs. MCI_{NC} (for 160 MCI subjects)			
Predictor type	Predictor	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy (95% CI)
Single modality					
	NP	.80 (.73, .86)	.66 (.55, .77)	.77 (.69, .86)	.73 (.66, .79)
	MRI	.80 (.74, .87)	.66 (.55, .77)	.78 (.70, .87)	.73 (.66, .80)
	CSF	.70 (.62, .79)	.56 (.44, .68)	.79 (.71, .88)	.69 (.62, .77)
Two modalities					
	NP + MRI	.85 (.80, .91)	.69 (.58, .80)	.83 (.75, .90)	.77 (.70, .83)
	NP + CSF	.80 (.74, .87)	.62 (.50, .73)	.82 (.74, .89)	.73 (.66, .80)
	MRI + CSF	.81 (.75, .88)	.60 (.49, .72)	.79 (.71, .88)	.71 (.64, .78)
Three modalities					
	NP + MRI + CSF	.86 (.80, .91)	.69 (.58, .80)	.82 (.74, .89)	.76 (.70, .83)
		MCI_C vs. MCI_{NC} (for leave-one-out cross-validation)			
		AUC	Sensitivity	Specificity	Accuracy
Single modality					
	NP	.74	.56	.73	.66
	MRI	.76	.57	.77	.69
	CSF	.62	.47	.73	.62
Two modalities					
	NP + MRI	.79	.66	.74	.71
	NP + CSF	.72	.56	.76	.68
	MRI + CSF	.75	.56	.76	.68
Three modalities					
	NP + MRI + CSF	.78	.63	.71	.68

Note. MCI_C = MCI patients who converted to AD within 24 months; MCI_{NC} = MCI patients who did not convert to AD within 24 months; NP = neuropsychological; MRI = magnetic resonance imaging; CSF = cerebrospinal fluid.
NP, MRI, and CSF without suffixes means that all separate factors within the respective modality are used as predictors in one logistic regression model. Sensitivity and specificity are estimated with threshold set to .5

alone, $z = 2.37$, $p = .022$. In addition, the NP + CSF combination increased the AUC compared to CSF factors alone ($z = 2.76$, $p = .015$), but it did not increase the AUC relative to NP factors alone. Meanwhile, the MRI + CSF combination increased the AUC relative to CSF factors alone ($z = 3.04$, $p = .009$), but it did not increase AUC relative to MRI factors alone. Thus, it seems that CSF factor scores did not offer additive predictive power relative to NP or MRI factor scores.

Finally, when we used factors from all three modalities as predictors, we found that the AUC for the NP + MRI + CSF combination was comparable to that of the NP + MRI combination. This provides converging support for the view that CSF factor scores did not add independent predictive power to NP and MRI factors. Nevertheless, the NP + MRI + CSF combination produced a larger AUC than the NP + CSF combination, $z = 2.27$, $p = .022$, and a larger AUC than the MRI + CSF combination, $z = 2.10$, $p = .027$. This again shows that combining NP or MRI factors with factors from other data modalities improves the prediction. Additionally, the AUC for the NP + MRI + CSF combination was larger than that for NP factors alone ($z = 2.28$, $p = .022$), for MRI factors alone ($z = 2.58$, $p = .019$), or for CSF factors alone ($z = 3.78$, $p = .001$).

To evaluate the robustness of our findings, we conducted a leave-one-out cross-validation. The results for the cross-validation analyses are presented in the lower half of Table 6. As can be seen there, the AUCs were slightly reduced compared to the full data, with the shrinkage ranging from .04 to .08. Importantly, the cross-validation analyses showed very similar patterns that NP and MRI factor scores delivered similar AUCs, which were larger than that of CSF factor scores, that adding NP or MRI factors to CSF factors increased AUCs, and that combining CSF factors with either NP or MRI factors produced no improvement in AUCs. All of this points to the conclusion that the model results for the full sample of MCI subjects are quite robust.

Using CFA factor scores as predictors

Similarly, we used factor scores from the CFAs as predictors of MCI to AD conversion in a series of logistic regressions. A summary of the logistic regression and leave-one-out cross-validation results is presented in Table 7, and the ROC curves for the logistic regression

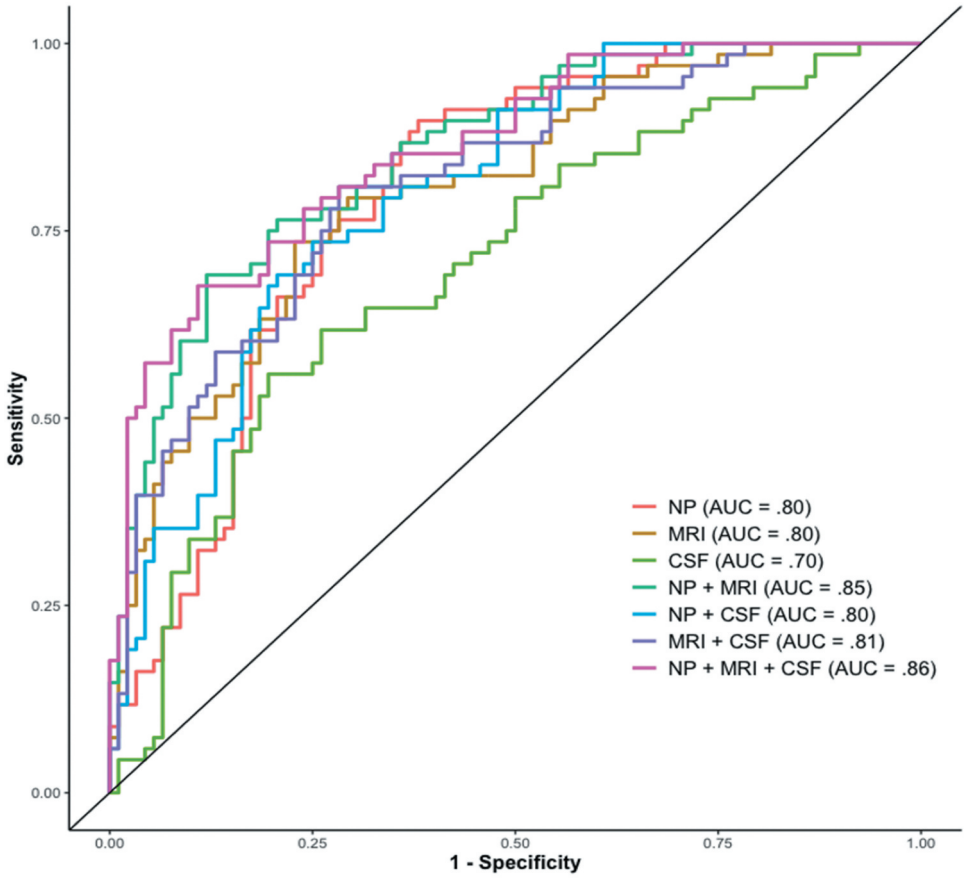

Figure 2. Receiver operating characteristics curves for all logistic regression models that used exploratory factor analysis factor scores to predict conversions from mild cognitive impairment to Alzheimer's disease for the 160 ADNI 1 subjects at baseline. NP = neuropsychological; MRI = magnetic resonance imaging; CSF = cerebrospinal fluid.

models are displayed in Figure 3. A visual inspection of Table 7 and Figure 3 reveals that the patterns were very similar to the results when using EFA factor scores as predictors. We again conducted Delong's tests to compare the AUCs. The tests showed that the AUC of MRI factors was larger than that of CSF factors, $z = 2.33$, $p = .021$, and there was no difference in the AUCs between NP factors and MRI factors or between NP factors and CSF factors. The NP + MRI combination produced an increase in AUC relative to NP factors ($z = 2.37$, $p = .021$) but not relative to MRI factors. The NP + CSF combination produced an increase in AUC relative to CSF factors ($z = 2.42$, $p = .021$) but not relative to NP factors. Similarly, the MRI + CSF combination produced an increase in AUC relative to CSF factors ($z = 3.12$, $p = .007$) but not relative to MRI factors.

When factors of all three data modalities are used as predictors, the AUC was larger than that of NP factors alone ($z = 2.50$, $p = .021$) and of CSF factors alone ($z = 3.57$, $p = .003$). Last, the NP + MRI + CSF combination produced larger AUC than the NP + CSF combination, $z = 2.44$, $p = .021$, but it produced no increase in AUC compared to the NP + MRI combination or the MRI + CSF combination. In summary, these results supported the previous findings that NP and MRI factor scores delivered similar AUCs, that adding NP or MRI factors to CSF factors increased AUCs for the latter, and that combining CSF factors with either NP or MRI factors produced no improvement in AUCs.

Additional analyses[12]

As follow-up analyses, we examined whether our results were robust if MCI subjects were diagnosed with the Jak/Bondi criteria (Bondi et al., 2014). The methods (Appendix A), logistic regression results (Table S2), and the ROC curves (Figure S4) for the additional analyses are presented in the Supplementary Materials. In brief, the AUCs were numerically reduced relative to the ADNI criteria, and the leave-one-out cross-validation analyses showed that the model results were not

Table 7. Summary of logistic regression analyses with confirmatory factor analysis factor score(s) as predictor(s) for 160 ADNI 1 subjects with mild cognitive impairment at baseline.

		MCI_C vs. MCI_{NC} (for 160 MCI subjects)			
Predictor type	Predictor	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy (95% CI)
Single modality					
	NP	.76 (.68, .83)	.60 (.49, .72)	.77 (.69, .86)	.70 (.63, .77)
	MRI	.80 (.73, .86)	. 60 (.49, .72)	.80 (.72, .89)	.72 (.65, .79)
	CSF	.69 (.61, .77)	.54 (.43, .66)	.78 (.70, .87)	.68 (.61, .75)
Two modalities					
	NP + MRI	.82 (.76, .89)	.66 (.55, .77)	.78 (.70, .87)	.73 (.66, .80)
	NP + CSF	.77 (.70, .84)	.60 (.49, .72)	.79(.71, .88)	.71 (.64, .78)
	MRI + CSF	.81 (.74, .87)	.65 (.53, .76)	.77 (.69, .86)	.72 (.65, .79)
Three modalities					
	NP + MRI + CSF	.83 (.77, .89)	.66 (.55, .77)	.80 (.72, .89)	.74 (.68, .81)
		MCI_C vs. MCI_{NC} (for leave-one-out cross-validation)			
		AUC	Sensitivity	Specificity	Accuracy
Single modality					
	NP	.69	.51	.75	.65
	MRI	.75	.60	.75	.69
	CSF	.62	.50	.76	.65
Two modalities					
	NP + MRI	.75	.60	.73	.68
	NP + CSF	.68	.56	.74	.66
	MRI + CSF	.74	.59	.74	.68
Three modalities					
	NP + MRI + CSF	.74	.62	.73	.68

Note. MCI_C = MCI patients who converted to AD within 24 months; MCI_{NC} = MCI patients who did not convert to AD within 24 months; NP = neuropsychological; MRI = magnetic resonance imaging; CSF = cerebrospinal fluid.
NP, MRI, and CSF without suffixes means that all separate factors within the respective modality are used as predictors in one logistic regression model.
Sensitivity and specificity are estimated with threshold set to .5

very robust. However, the basic qualitative pattern was preserved because the NP and MRI factors seemed more predictive than the CSF factors, and multimodal factors produced larger AUCs than single-modal factors. Delong's tests revealed that only two pairwise comparisons of AUCs yielded significant differences: NP + MRI + CSF produced a larger AUC than NP factors alone ($z = 2.61$, $p = .034$) or than CSF factors alone ($z = 2.68$, $p = .034$). Here, it is worth mentioning that the reduced sample size (from 160 to 99) leads to a higher type II error rate, which may explain why some previously reliable differences in AUCs were no longer significant.

Discussion

The current research was aimed at evaluating the ability of NP, MRI, and CSF factor scores to predict future transitions from MCI to AD over a 2-year interval, and to determine whether combinations of multimodal factor scores can further improve predictive power. We conducted separate EFAs and CFAs for NP, MRI and CSF measures to isolate their respective factor structures for ADNI 1 MCI subjects. This yielded a five-factor structure for NP (memory, language, executive function, attention, and visuospatial), along with three-factor structures for MRI (temporal-parietal, entorhinal-amygdala, and ventricle) and for CSF (tau-amyloid, neuroprotection, and neuroinflammation). Then, we conducted a series of logistic regressions for predicting MCI to AD conversion, using either single-modal factor scores or multimodal factor scores as predictors.

Overall, we saw that NP and MRI factor scores delivered quite good predictive power over a 2-year interval. Our results echoed prior studies that demonstrated that factor scores are potentially powerful predictors of MCI to AD conversion (Chapman et al., 2011; Eckerström et al., 2013; Giraldo et al., 2017; Wilhalme et al., 2017). As noted in the Introduction, there are several reasons for using factor scores instead of individual test scores as predictors, including that factor scores have the advantage of improving measurement precision by reducing inherent measurement biases in the single predictors (Crane et al., 2012; Wilhalme et al., 2017) and by combining predictors that are differentially sensitive to AD conversion in different sub-populations (Gibbons et al., 2012). Additionally, the use of factor scores also controls multicollinearity through reductions in the number of predictors while simultaneously preserving information about the individual variables.

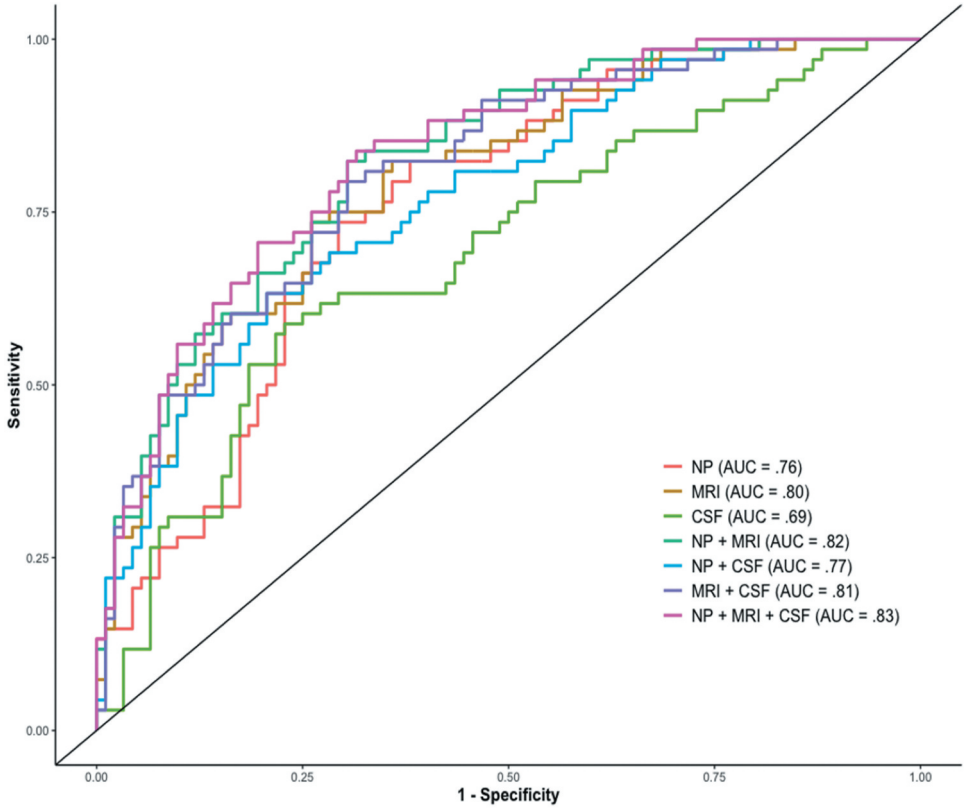

Figure 3. Receiver operating characteristics curves for all logistic regression models that used confirmatory factor analysis factor scores to predict conversions from mild cognitive impairment to Alzheimer's disease for the 160 ADNI 1 subjects at baseline. NP = neuropsychological; MRI = magnetic resonance imaging; CSF = cerebrospinal fluid.

Moreover, this approach also has important theoretical and empirical advantages. On the theoretical side, compared to pure data-driven approaches such as machine learning algorithms of feature selection (e.g., Cui et al., 2011), the latent-variable approach has the key advantage of interpretability. Factor analysis is a data-reduction method that is aimed at identifying theoretical latent variables from a set of observed variables. As a result, individual variables that are entered into factor analysis are grouped under theoretically meaningful latent constructs, which is essential for testing theoretical assumptions. Empirically, the latent variable approach is easy to implement, low-cost, and efficient. Further, factor scores can be "harmonized" across different AD studies (Chapman et al., 2011, 2010; Hampton et al., 2020). That means it would be possible to compare factor scores directly across different batteries, as long as their tests tap the same or similar latent constructs. Therefore, the latent-variable approach can potentially improve the standardization of diagnosis and prediction methods across different AD studies. After the predictive ability of factor scores is verified and the factor scores are standardized across different studies, it is possible that future research can develop cutoff criteria for the factor scores based on norming data and implement the latent-variable approach in clinical practice, much as how the Wechsler Adult Intelligence Scale (Cohen, 1957) is implemented for tests of intelligence.

In Table 6, Table 7, Figures 2 and 3, we can see that adding NP or MRI factors to CSF factors significantly improved predictive power, but it was not the other way around: CSF factor scores did not add any additional predictive power to either NP or MRI factors alone or to their combination. This finding is conceptually consistent with prior findings that NP and MRI variables usually generate a better prediction for MCI to AD conversion than CSF predictors with the ADNI subjects (Cui et al., 2011; Devanand et al., 2012; Eckerström et al., 2013; Gomar et al., 2011, 2014; Li et al., 2017). Here, Jack et al.'s (2010) hypothetic model of the Alzheimer's pathological cascade provides one possible explanation for CSF variables' weaker predictive power for MCI to AD conversion. According to this model, different biomarkers display abnormalities in an ordered manner, and thus different biomarkers may dominate different stages of pathological changes. In that connection, certain CSF markers (e.g., CSF

concentration of $A\beta_{42}$) reach a plateau at an earlier stage than other biomarkers. Therefore, it is possible that around the given two-year window of MCI to AD conversion, some CSF variables have reached a plateau. Therefore, at this stage, their predictive power wanes while other biomarkers predominate the pathological processes.

On the one hand, there is general consensus in prior studies that combining NP and MRI predictors produces more accurate predictions (Barnes et al., 2014; Cui et al., 2011; Moradi et al., 2015; Ye et al., 2012; but see, Richard et al., 2013). Our results are in line with those findings in that the combination of NP and MRI factor scores produced improvement in prediction. It is also worth mentioning with respect to AUC, our NP + MRI model outperformed the results from multiple past ADNI studies that used multimodal data (Cheng et al., 2015; Cui et al., 2011; Davatzikos et al., 2011; Gomar et al., 2011; Ye et al., 2012; Zhang et al., 2012).

On the other hand, past findings about the incremental benefits of CSF predictors have been mixed: Some studies have found that combining CSF predictors with predictors from other modalities slightly improved predictive power (Cui et al., 2011; Davatzikos et al., 2011; Eckerström et al., 2013; Shaffer et al., 2013), whereas other studies found no increase in predictive power (Gomar et al., 2011, 2014; Richard et al., 2013; Ye et al., 2012). Clearly, our results aligned with the latter. There are multiple possible explanations for the discrepant findings, such as variations in sample size and predictor selections. In the former connection, the studies that found positive results for CSF predictors were usually conducted with smaller samples than those that found negative results. In the latter connection, the aforementioned studies used different methods for selecting predictor variables, such as the minimum redundancy and maximum relevance filter method (e.g., Cui et al., 2011) or sparse logistic regression with stability selection (e.g., Ye et al., 2012). As a result, the final sets of variables that are entered into the prediction model are quite different, leading to different covariance structures among variables. Thus, it is possible that CSF variables can provide additional predictive power with certain variables from other modalities, but not with others. Here, we stress that considering the invasive nature of CSF sampling methods, uncertainty about the incremental predictive value of CSF markers should be weighed when evaluating their continued use in clinical diagnosis.

Because CSF data were only collected from a subset of the ADNI subjects, the results discussed so far were restricted to a relatively small sample ($N = 160$). In that connection, as CSF factor scores provided no added value relative to the combination of NP and MRI factor scores, we conducted a follow-up analysis without the CSF data. Because NP and MRI data are available for many more ADNI 1 subjects than CSF data, by only considering the former two types of data, we were able to repeat the logistic regression analyses with a larger sample ($N = 319$, including 134 MCI_C subjects and 185 MCI_{NC} subjects). This increased the statistical power and reliability of our analyses. The logistic regression results with EFA factor scores of NP and MRI data (Table S3) and the ROC curves (Figure S5) are presented in the Supplementary Materials. It can be seen there that the AUCs for NP factor scores, MRI factor scores, and the NP + MRI combination were very similar to those in Table 6. Specifically, the AUCs for NP, MRI, and NP + MRI factor scores were .80, .80, and .85 with the 160-subject sample and .82, .80, and .86 with the 319-subject sample. Meanwhile, there was no difference in AUC between NP and MRI factors, while the AUC for NP + MRI factors was larger than that for NP factors, $z = 2.59$, $p = .007$, and for MRI factors, $z = 3.05$, $p = .003$. Further, the AUCs in the leave-one-out cross-validation analyses were only marginally lower, with the shrinkage in AUCs ranging from .02 to .03. On a related note, when 192 MCI subjects were reclassified based on the Jak/Bondi criteria, the logistic regression results with only NP and MRI EFA factors showed a similar pattern (see Appendix B, Table S4, and Figure S6 in the Supplementary Materials).

Moreover, the logistic regression results with CFA factor scores for the same 319 MCI subjects displayed a very similar pattern as those with EFA factor scores (see Table S5 and Figure S7 in the Supplementary Materials). First, the results in Table S5 resemble those in Table 7: the AUCs for NP, MRI, and NP + MRI factor scores were .76, .80, and .82 with the 160-subject sample and .79, .80, and .84 with the 319-subject sample. Second, the AUC for NP + MRI factors was larger than that for NP factors alone, $z = 3.01$, $p = .004$, and for MRI factors alone, $z = 2.29$, $p = .016$, with no difference in AUC between NP and MRI factors. Third, the AUCs only shrank slightly in the leave-one-out cross-validation analyses, indicating that all the observed patterns were robust. In summary, these follow-up analyses illustrate that two observed patterns were robust in the larger baseline sample of MCI subjects: NP and MRI factor scores are highly predictive of MCI to AD conversion, and the combination of NP and MRI factor scores outperformed either NP or MRI factor scores alone.

Last, we note some limitations of the current study. First, as shown in Table 1, the ADNI MCI subjects

overall have a high educational level: Mean years of education are 15.35 for MCI_C subjects and 16.28 for MCI_{NC} subjects. Thus, the average ADNI subject completed or nearly completed an undergraduate degree, which is not nationally representative. The high educational level is potentially concerning as more years of education are associated with more cognitive reserve and lower risk of AD (Sando et al., 2008; Stern, 2012). A broader range of variables might be reliable predictors of AD conversion in a nationally representative sample. To provide some control of the possible effects of education level, we included education level as a covariate in all our logistic regression models.

Second, we only analyzed ADNI data that were collected during the first 24 months. It is conceivable that some MCI_C subjects in our current sample eventually convert to AD later. In that case, it is possible that the latent factor structures and the predictive power of latent factors can be different for MCI subjects with slower conversion to AD. Therefore, it is recommended that the current findings should be validated with a longer follow-up interval.

Third, it remains an open question whether our findings hold when an alternative definition of AD is applied. The ADNI diagnostic criteria, as described in the Method section, focus on clinical symptoms. Alternatively, a biological definition of AD was recently proposed in the updated National Institute on Aging and Alzheimer's Association (NIA-AA) research framework (Jack et al., 2018). In this framework, the definition of AD is based on biomarker (PET or CSF) evidence of both Aβ deposition and pathological tau, which was divorced from clinical symptoms of AD. However, due to the limited PET and CSF data for ADNI 1 subjects between baseline and the 24-month follow-up session, we were unable to implement the NIA-AA framework in the current study. Thus, further research is recommended to verify our findings with the biological definition of AD, which will be possible when more PET or CSF data become available in later ADNI phases.

Conclusion

In the current study, we found that factor scores of NP and MRI data were strong predictors of MCI to AD conversion over a 2-year interval, demonstrating that the latent-variable approach is a useful method for predicting MCI to AD conversion. Further, the combination of factor scores from different modalities increases predictive power. More explicitly, incorporating both NP and MRI factor scores into a single model increased predictive power, relative to either NP or MRI factor scores alone. In contrast, adding CSF factor scores did not increase predictive power, regardless of whether they were combined with factor scores for NP, MRI, or both. Cross-validation and follow-up analyses demonstrated that our model results were robust and stable. Last, we recommend future research to validate our results with a longer follow-up interval and alternative diagnostic criteria for AD. To advance the clinical utility of the latent-variable approach, we also encourage future studies to establish clinically appropriate cutoff values for the latent factor scores based on large-scale norming data.

Notes

1. PCA and EFA are often confused with each other. The produced index variables are called components in PCA and factors in EFA. The two methods differ in that PCA is aimed at accounting for most variance of the manifest variables without considering the latent structure of these variables, but EFA is meant to identify the number of latent variables and the latent structure that can explain the correlations between the manifest variables. Thus, PCA does not distinguish between the shared and unique variance of a manifest variable, but EFA does (Costello & Osborne, 2005).
2. This means that the MCI sample in ADNI is all amnestic-MCI (a-MCI). Note that a-MCI is the subtype of MCI that is at increased risk of converting to AD (Petersen, 2011).
3. One subject, who was diagnosed as AD at 12 months but converted back to MCI at 24 months, was removed from the final analyses.
4. Among the 92 MCI_{NC} subjects, only 6 reverted back to a not-impaired diagnosis within 24 months. Thus, the classification of MCI is very reliable in the ADNI 1 dataset.
5. The MRI data were acquired during a screening session, which occurred within 28 days of the baseline session. Although MRI data were also gathered during the baseline session, the screening session data were richer, and hence, those data figured in our analyses.
6. We thank an anonymous reviewer for suggesting CFA analyses in addition to EFA analyses.
7. Modification indices are the changes in chi-squared values if a certain path was added or a certain constraint was removed.
8. Notably, all variables loaded on this factor (entorhinal cortex, amygdala, hippocampus, and temporal pole) are involved in memory consolidation (Izquierdo & Medina, 1993; Landi et al., 2021).
9. Neuroinflammation refers to inflammatory responses within the central neural system (CNS), which are triggered by CNS insults, such as protein misfolding and aggregation. Recent evidence has suggested that excessive neuroinflammation can cause neuron damage and contribute to deterioration in brain diseases (Calsolaro & Edison, 2016; Heneka et al., 2015).

10. We followed the convention of setting the threshold at .5 in ROC analyses throughout the paper. Please note that the threshold can be adjusted to improve sensitivity or specificity at a relative cost to each other.
11. All *p* values of the DeLong' tests were corrected for multiple comparisons using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995).
12. We thank an anonymous reviewer for suggesting additional analyses based on the Jak/Bondi criteria (Bondi et al., 2014).
13. Again, one subject, who was diagnosed as AD at 12 months but converted back to MCI at 24 months, was removed from the final analyses.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported by National Institutes of Health Grant 1RC1AG036915-02 to the second author.Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Data availability statement

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 with the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this report. A complete listing of ADNI investigators can be found at: https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

References

Alzheimer's Association. (2021). Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, *17*(3), 327–406. https://doi.org/10.1002/alz.12328

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

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Blennow, K. (2004). Cerebrospinal fluid protein biomarkers for Alzheimer's disease. *NeuroRX*, *1*(2), 213–225. https://doi.org/10.1602/neurorx.1.2.213

Boggs, L. N., Fuson, K. S., Baez, M., Churgay, L., McClure, D., Becker, G., & May, P. C. (1996). Clusterin (Apo J) protects against in vitro amyloid-β(1–40) neurotoxicity. *Journal of Neurochemistry*, *67*(3), 1324–1327. https://doi.org/10.1046/j.1471-4159.1996.67031324.x

Bondi, M. W., Edmonds, E. C., Jak, A. J., Clark, L. R., Delano-Wood, L., McDonald, C. R., Nation, D. A., Libon, D. J., Au, R., Galasko, D., & Salmon, D. P., for the Alzheimer's Disease Neuroimaging Initiative. (2014). Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates, *Journal of Alzheimer's Disease*, *42*(1), 275–289. https://doi.org/10.3233/JAD-140276.

Browne, M., & Cudek, R. (1993). Alternative ways of assessing model fit. In K. Bollen & J. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.

Calsolaro, V., & Edison, P. (2016). Neuroinflammation in Alzheimer's disease: Current evidence and future directions. *Alzheimer's & Dementia*, *12*(6), 719–732. https://doi.org/10.1016/j.jalz.2016.02.010

Cattell, R. B., & Vogelmann, S. (1977). A comprehensive trial of the scree and KG criteria for determining the number of factors. *Multivariate Behavioral Research*, *12*(3), 289–325. https://doi.org/10.1207/s15327906mbr1203_2

Chang, M., & Brainerd, C. J. (2021). Factor analyses of the ADNI neuropsychological battery: An examination of diagnostic and longitudinal invariance. *Neuropsychology*, *35*(4), 434–450. https://doi.org/10.1037/neu0000736

Chapman, R. M., Mapstone, M., McCrary, J. W., Gardner, M. N., Porsteinsson, A., Sandoval, T. C., Guillily, M. D., DeGrush, E., & Reilly, L. A. (2011). Predicting conversion from mild cognitive impairment to Alzheimer's disease using neuropsychological tests and multivariate methods. *Journal of Clinical and Experimental Neuropsychology*, *33*(2), 187–199. https://doi.org/10.1080/13803395.2010.499356

Chapman, R. M., Mapstone, M., Porsteinsson, A. P., Gardner, M. N., McCrary, J. W., DeGrush, E., Reilly, L. A., Sandoval, T. C., & Guillily, M. D. (2010).

Diagnosis of Alzheimer's disease using neuropsychological testing improved by multivariate analyses. *Journal of Clinical and Experimental Neuropsychology*, *32*(8), 793–808. https://doi.org/10.1080/13803390903540315

Chen, T., Zeng, D., & Wang, Y. (2015). Multiple kernel learning with random effects for predicting longitudinal outcomes and data integration. *Biometrics*, *71*(4), 918–928. https://doi.org/10.1111/biom.12343

Cheng, B., Liu, M., Zhang, D., Munsell, B. C., & Shen, D. (2015). Domain transfer learning for MCI conversion prediction. *IEEE Transactions on Biomedical Engineering*, *62*(7), 1805–1817. https://doi.org/10.1109/TBME.2015.2404809

Cohen, J. (1957). A factor-analytically based rationale for the wechsler adult intelligence scale. *Journal of Consulting Psychology*, *21*(6), 451–457. https://doi.org/10.1037/h0044203

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, *10*(7), 1–9. https://doi.org/10.7275/jyj1-4868

Crane, P. K., Carle, A., Gibbons, L. E., Insel, P., Mackin, R. S., Gross, A., Jones, R. N., Mukherjee, S., Curtis, S. M., Harvey, D., Weiner, M., & Mungas, D. (2012). Development and assessment of a composite score for memory in the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Brain Imaging and Behavior*, *6*(4), 502–516. https://doi.org/10.1007/s11682-012-9186-z

Cui, Y., Liu, B., Luo, S., Zhen, X., Fan, M., Liu, T., Zhu, W., Park, M., Jiang, T., & Jin, J. S., & the Alzheimer's Disease Neuroimaging Initiative. (2011). Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors. *PLOS ONE*, *6* (7), e21896. https://doi.org/10.1371/journal.pone.0021896

Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M.-O., Chupin, M., Benali, H., & Colliot, O. (2011). Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database. *NeuroImage*, *56*(2), 766–781. https://doi.org/10.1016/j.neuroimage.2010.06.013

Davatzikos, C., Bhatt, P., Shaw, L. M., Batmanghelich, K. N., & Trojanowski, J. Q. (2011). Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiology of Aging*, *32*(12), 2322.e19–2322.e27. https://doi.org/10.1016/j.neurobiolaging.2010.05.023

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, *44*(3), 837–845. https://doi.org/10.2307/2531595

Deming, Y., Xia, J., Cai, Y., Lord, J., Holmans, P., Bertelsen, S., Holtzman, D., Morris, J. C., Bales, K., Pickering, E. H., Kauwe, J., Goate, A., & Cruchaga, C. (2016). A potential endophenotype for Alzheimer's disease: Cerebrospinal fluid clusterin. *Neurobiology of Aging*, *37*(208), e1–208.e9. https://doi.org/10.1016/j.neurobiolaging.2015.09.009

Devanand, D. P., Liu, X., Brown, P. J., Huey, E. D., Stern, Y., & Pelton, G. H. (2012). A two-study comparison of clinical and MRI markers of transition from mild cognitive impairment to Alzheimer's disease. *International Journal of Alzheimer's Disease*, *2012*, e483469. https://doi.org/10.1155/2012/483469

Eckerström, C., Olsson, E., Bjerke, M., Malmgren, H., Edman, Å., Wallin, A., & Nordlund, A. (2013). A combination of neuropsychological, neuroimaging, and cerebrospinal Fluid markers predicts conversion from mild cognitive iImpairment to dementia. *Journal of Alzheimer's Disease*, *36*(3), 421–431. https://doi.org/10.3233/JAD-122440

Ewers, M., Walsh, C., Trojanowski, J. Q., Shaw, L. M., Petersen, R. C., Jack, C. R., Feldman, H. H., Bokde, A. L. W., Alexander, G. E., Scheltens, P., Vellas, B., Dubois, B., Weiner, M., & Hampel, H., & North American Alzheimer's Disease Neuroimaging Initiative. (2012). Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. *Neurobiology of Aging*, *33*(7), 1203–1214.e2. https://doi.org/10.1016/j.neurobiolaging.2010.10.019

Fjell, A. M., Walhovd, K. B., Fennema-Notestine, C., McEvoy, L. K., Hagler, D. J., Holland, D., Brewer, J. B., & Dale, A. M., for the Alzheimer's Disease Neuroimaging Initiative. (2010). CSF biomarkers in prediction of cerebral and clinical change in mild cognitive impairment and Alzheimer's disease. *Journal of Neuroscience*, *30*(6), 2088–2101. https://doi.org/10.1523/JNEUROSCI.3785-09.2010

Gibbons, L. E., Carle, A. C., Mackin, R. S., Harvey, D., Mukherjee, S., Insel, P., Curtis, S. M., Mungas, D., & Crane, P. K. (2012). A composite score for executive functioning, validated in Alzheimer's Disease Neuroimaging Initiative (ADNI) participants with baseline mild cognitive impairment. *Brain Imaging and Behavior*, *6*(4), 517–527. https://doi.org/10.1007/s11682-012-9176-1

Giraldo, D. L., Sijbers, J., & Romero, E. (2017). Quantifying cognition and behavior in normal aging, mild cognitive impairment, and Alzheimer's disease. *13th International Conference on Medical Information Processing and Analysis*, *10572*, 105720H. https://doi.org/10.1117/12.2287036

Gomar, J. J., Bobes-Bascaran, M. T., Conejero-Goldberg, C., Davies, P., Goldberg, T. E., & Initiative, A. D. N. (2011). Utility of combinations of biomarkers, cognitive markers, and risk factors to predict conversion from mild cognitive impairment to *Alzheimer* disease in patients in the *Alzheimer*'s disease neuroimaging initiative. *Archives of General Psychiatry*, *68*(9), 961–969. https://doi.org/10.1001/archgenpsychiatry.2011.96

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

Goodglass, H., Kaplan, E., & Weintraub, S. (1983). *Boston naming test*. Lea & Febiger.

Hampel, H., Teipel, S. J., Padberg, F., Haslinger, A., Riemenschneider, M., Schwarz, M. J., Kötter, H. U., Scheloske, M., Buch, K., Stübner, S., Dukoff, R., Lasser, R., Müller, N., Sunderland, T., Rapoport, S. I., & Möller, H.-J. (1999). Discriminant power of combined cerebrospinal fluid τ protein and of the soluble interleukin-6 receptor

complex in the diagnosis of Alzheimer's disease. *Brain Research*, *823*(1), 104–112. https://doi.org/10.1016/S0006-8993(99)01146-4

Hampton, O. L., Mukherjee, S., Properzi, M. J., Schultz, A. P., Crane, P. K., Gibbons, L. E., Hohman, T. J., Maruff, P. T., Lim, Y. Y., Amariglio, R., Papp, K. V., Johnson, K. A., Rentz, D., Sperling, R. A., & Buckley, R. F. (2020). Harmonizing the preclinical Alzheimer cognitive composite for multi-cohort studies. *Alzheimer's & Dementia*, *16* (S9), e047423. https://doi.org/10.1002/alz.047423

Harrison, J. E., Buxton, P., Husain, M., & Wise, R. (2000). Short test of semantic and phonological fluency: Normal performance, validity and test-retest reliability. *British Journal of Clinical Psychology*, *39*(2), 181–191. https://doi.org/10.1348/014466500163202

Heneka, M. T., Carson, M. J., Khoury, J. E., Landreth, G. E., Brosseron, F., Feinstein, D. L., Jacobs, A. H., Wyss-Coray, T., Vitorica, J., Ransohoff, R. M., Herrup, K., Frautschy, S. A., Finsen, B., Brown, G. C., Verkhratsky, A., Yamanaka, K., Koistinaho, J., Latz, E., Halle, A., & Kummer, M. P. (2015). Neuroinflammation in Alzheimer's disease. *The Lancet Neurology*, *14*(4), 388–405. https://doi.org/10.1016/S1474-4422(15)70016-5

Hershberger, S. L. (2005). Factor score estimation. In B. S. Everitt & D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (pp. 726). John Wiley & Sons, Ltd. https://doi.org/10.1002/0470013192.bsa726

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/BF02289447

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hu, W. T., Watts, K. D., Tailor, P., Nguyen, T. P., Howell, J. C., Lee, R. C., Seyfried, N. T., Gearing, M., Hales, C. M., Levey, A. I., Lah, J. J., & Lee, E. K., for the Alzheimer's Disease Neuro-Imaging Initiative. (2016). CSF complement 3 and factor H are staging biomarkers in Alzheimer's disease, *Acta Neuropathologica Communications*, *4*(1), 14. https://doi.org/10.1186/s40478-016-0277-8

Izquierdo, I., & Medina, J. H. (1993). Role of the amygdala, hippocampus and entorhinal cortex in memory consolidation and expression. *Brazilian Journal of Medical and Biological Research*, *26*(6), 573–589.

Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., Holtzman, D. M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J. L., Montine, T., Phelps, C., Rankin, K. P., Rowe, C. C., Scheltens, P., Siemers, E., Snyder, H. M., … Silverberg, N. (2018). NIA-AA research framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, *14*(4), 535–562. https://doi.org/10.1016/j.jalz.2018.02.018

Jack, C. R., Knopman, D. S., Jagust, W. J., Shaw, L. M., Aisen, P. S., Weiner, M. W., Petersen, R. C., & Trojanowski, J. Q. (2010). Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet Neurology*, *9*(1), 119–128. https://doi.org/10.1016/S1474-4422(09)70299-6

Jutten, R. J., Sikkes, S. A. M., Amariglio, R. E., Buckley, R. F., Properzi, M. J., Marshall, G. A., Rentz, D. M., Johnson, K. A., Teunissen, C. E., Van Berckel, B. N. M., Van der Flier, W. M., Scheltens, P., Sperling, R. A., & Papp, K. V., for the Alzheimer Disease Neuroimaging Initiative, & National Alzheimer's Coordinating Center, the Harvard Aging Brain Study, and the Alzheimer Dementia Cohort. (2021). Identifying sensitive measures of cognitive decline at different clinical stages of Alzheimer's disease, *Journal of the International Neuropsychological Society*, *27*(5), 426–438. https://doi.org/10.1017/S1355617720000934

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*(3), 187–200. https://doi.org/10.1007/BF02289233

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, *35*(4), 401–415. https://doi.org/10.1007/BF02291817

Kaiser, H. F., & Rice, J. (1974). Little jiffy, mark IV. *Educational and Psychological Measurement*, *34*(1), 111–117. https://doi.org/10.1177/001316447403400115

Kaur, G., & Levy, E. (2012). Cystatin C in Alzheimer's disease. *Frontiers in Molecular Neuroscience*, *5*, 79. https://doi.org/10.3389/fnmol.2012.00079

Khan, W., Aguilar, C., Kiddle, S. J., Doyle, O., Thambisetty, M., Muehlboeck, S., Sattlecker, M., Newhouse, S., Lovestone, S., Dobson, R., Giampietro, V., Westman, E., Simmons, A., & Initiative, A. D. N. (2015). A subset of cerebrospinal fluid proteins from a multi-analyte panel associated with brain atrophy, disease classification and prediction in Alzheimer's disease. *PLOS ONE*, *10*(8), e0134368. https://doi.org/10.1371/journal.pone.0134368

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, *28*(5), 1–26. https://doi.org/10.18637/jss.v028.i05

Kumar, R. G., Diamond, M. L., Boles, J. A., Berger, R. P., Tisherman, S. A., Kochanek, P. M., & Wagner, A. K. (2015). Acute CSF interleukin-6 trajectories after TBI: Associations with neuroinflammation, polytrauma, and outcome. *Brain, Behavior, and Immunity*, *45*, 253–262. https://doi.org/10.1016/j.bbi.2014.12.021

Lai, K., & Green, S. B. (2016). The problem with having two watches: Assessment of fit when RMSEA and CFI disagree. *Multivariate Behavioral Research*, *51*(2–3), 220–239. https://doi.org/10.1080/00273171.2015.1134306

Landi, S. M., Viswanathan, P., Serene, S., & Freiwald, W. A. (2021). A fast link between face perception and memory in the temporal pole. *Science*, *373*(6554), 581–585. https://doi.org/10.1126/science.abi6671

Li, K., Chan, W., Doody, R. S., Quinn, J., & Luo, S. (2017). Prediction of conversion to Alzheimer's disease with longitudinal measures and time-to-event data. *Journal of Alzheimer's Disease*, *58*(2), 361–371. https://doi.org/10.3233/JAD-161201

Martens, L. H., Zhang, J., Barmada, S. J., Zhou, P., Kamiya, S., Sun, B., Min, S.-W., Gan, L., Finkbeiner, S., Huang, E. J., & Farese, R. V. (2012). Progranulin deficiency promotes neuroinflammation and neuron loss following toxin-induced injury. *The Journal of Clinical Investigation*, *122*(11), 3955–3959. https://doi.org/10.1172/JCI63113

Mattsson, N., Insel, P., Nosheny, R., Trojanowski, J. Q., Shaw, L. M., Jack, C. R., Tosun, D., & Weiner, M. (2014).

Effects of cerebrospinal fluid proteins on brain atrophy rates in cognitively healthy older adults. *Neurobiology of Aging*, *35*(3), 614–622. https://doi.org/10.1016/j.neurobiolaging.2013.08.027

Miners, J. S., Clarke, P., & Love, S. (2017). Clusterin levels are increased in A lzheimer's disease and influence the regional distribution of Aβ. *Brain Pathology*, *27*(3), 305–313. https://doi.org/10.1111/bpa.12392

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

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

Morenas-Rodríguez, E., Cervera-Carles, L., Vilaplana, E., Alcolea, D., Carmona-Iragui, M., Dols-Icardo, O., Ribosa-Nogué, R., Muñoz-Llahuna, L., Sala, I., Belén Sánchez-Saudinós, M., Blesa, R., Clarimón, J., Fortea, J., & Lleó, A. (2016). Progranulin protein levels in cerebrospinal fluid in primary neurodegenerative dementias. *Journal of Alzheimer's Disease*, *50*(2), 539–546. https://doi.org/10.3233/JAD-150746

Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*, *43*(11), 2412–2414. https://doi.org/10.1212/WNL.43.11.2412-a

Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C. R., Jagust, W., Trojanowski, J. Q., Toga, A. W., & Beckett, L. (2005). Ways toward an early diagnosis in *Alzheimer*'s disease: The *Alzheimer*'s Disease Neuroimaging Initiative (ADNI). *Alzheimer's & Dementia*, *1*(1), 55–66. https://doi.org/10.1016/j.jalz.2005.06.003

Park, L. Q., Gross, A. L., McLaren, D. G., Pa, J., Johnson, J. K., Mitchell, M., & Manly, J. J., Alzheimer's Disease Neuroimaging Initiative. (2012). Confirmatory factor analysis of the ADNI neuropsychological battery. *Brain Imaging and Behavior*, *6*(4), 528–539. https://doi.org/10.1007/s11682-012-9190-3

Paterson, R. W., Bartlett, J. W., Blennow, K., Fox, N. C., Shaw, L. M., Trojanowski, J. Q., Zetterberg, H., & Schott, J. M. (2014). Cerebrospinal fluid markers including trefoil factor 3 are associated with neurodegeneration in amyloid-positive individuals. *Translational Psychiatry*, *4*(7), e419–e419. https://doi.org/10.1038/tp.2014.58

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, *256*(3), 183–194. https://doi.org/10.1111/j.1365-2796.2004.01388.x

Petersen, R. C. (2011). Mild cognitive impairment. *The New England Journal of Medicine*, *364*(23), 2227–2234. https://doi.org/10.1056/NEJMcp0910237

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

Pogue, A. I., Li, Y. Y., Cui, J.-G., Zhao, Y., Kruck, T. P. A., Percy, M. E., Tarr, M. A., & Lukiw, W. J. (2009). Characterization of an NF-κB-regulated, miRNA-146a-mediated down-regulation of complement factor H (CFH) in metal-sulfate-stressed human brain cells. *Journal of Inorganic Biochemistry*, *103*(11), 1591–1595. https://doi.org/10.1016/j.jinorgbio.2009.05.012

Portelius, E., Zetterberg, H., Skillbäck, T., Törnqvist, U., Andreasson, U., Trojanowski, J. Q., Weiner, M. W., Shaw, L. M., & Mattsson, N., for the Alzheimer's Disease Neuroimaging Initiative. (2015). Cerebrospinal fluid neurogranin: Relation to cognition and neurodegeneration in Alzheimer's disease, *Brain*, *138*(11), 3373–3385. https://doi.org/10.1093/brain/awv267

R Core Team. (2019). *R: A language and environment for statistical computing*. https://www.R-project.org/

Reitan, R. M., & Wolfson, D. (1985). *The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation* (Vol. 4). Reitan Neuropsychology.

Revelle, W. (2016). Psych: Procedures for personality and psychological research (Version 1.6. 12)[Computer software]. *Evanston, IL: Northwestern University*, *30*, 2017.

Rey, A. L. (1964). *Presses Universitaires de France*. Presses Universitaires de France.

Richard, E., Schmand, B. A., Eikelenboom, P., & Van Gool, W. A. (2013). MRI and cerebrospinal fluid biomarkers for predicting progression to Alzheimer's disease in patients with mild cognitive impairment: A diagnostic accuracy study. *BMJ Open*, *3*(6), e002541. https://doi.org/10.1136/bmjopen-2012-002541

Risacher, S. L., Saykin, A. J., West, J. D., Shen, L., Firpi, H. A., & McDonald, B. C. (2009). Baseline MRI predictors of conversion from MCI to probable AD in the ADNI cohort. *Current Alzheimer Research*, *6*(4), 347–361. https://doi.org/10.2174/156720509788929273

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*(1), 77. https://doi.org/10.1186/1471-2105-12-77

Rosen, W. G., Mohs, R. C., & Davis, K. L. (1984). A new rating scale for Alzheimer's disease. *The American Journal of Psychiatry*, *141*(11), 1356–1364. https://doi.org/10.1176/ajp.141.11.1356

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin*, *94*(1), 18–38. https://doi.org/10.1037/0033-2909.94.1.18

Sando, S. B., Melquist, S., Cannon, A., Hutton, M., Sletvold, O., Saltvedt, I., White, L. R., Lydersen, S., & Aasly, J. (2008). Risk-reducing effect of education in Alzheimer's disease. *International Journal of Geriatric Psychiatry*, *23*(11), 1156–1162. https://doi.org/10.1002/gps.2043

Shaffer, J. L., Petrella, J. R., Sheldon, F. C., Choudhury, K. R., Calhoun, V. D., Coleman, R. E., & Doraiswamy, P. M. (2013). Predicting cognitive decline in subjects at risk for Alzheimer disease by using combined cerebrospinal fluid, MR Imaging, and PET biomarkers. *Radiology*, *266*(2), 583–591. https://doi.org/10.1148/radiol.12120010

Shaw, L. M., Vanderstichele, H., Knapik-Czajka, M., Clark, C. M., Aisen, P. S., Petersen, R. C., Blennow, K., Soares, H., Simon, A., Lewczuk, P., Dean, R., Siemers, E., Potter, W., Lee, V. M.-Y., & Trojanowski, J. Q. (2009). Cerebrospinal fluid biomarker

signature in Alzheimer's disease neuroimaging initiative subjects. *Annals of Neurology*, *65*(4), 403–413. https://doi.org/10.1002/ana.21610

Shi, F., Liu, B., Zhou, Y., Yu, C., & Jiang, T. (2009). Hippocampal volume and asymmetry in mild cognitive impairment and Alzheimer's disease: Meta-analyses of MRI studies. *Hippocampus*, *19*(11), 1055–1064. https://doi.org/10.1002/hipo.20573

Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *Lancet Neurology*, *11*(11), 1006–1012. https://doi.org/10.1016/S1474-4422(12)70191-6

Suárez-Calvet, M., Kleinberger, G., Araque Caballero, M. Á., Brendel, M., Rominger, A., Alcolea, D., Fortea, J., Lleó, A., Blesa, R., Gispert, J. D., Sánchez-Valle, R., Antonell, A., Rami, L., Molinuevo, J. L., Brosseron, F., Traschütz, A., Heneka, M. T., Struyfs, H., Engelborghs, S., … Haass, C. (2016). STREM 2 cerebrospinal fluid levels are a potential biomarker for microglia activity in early-stage Alzheimer's disease and associate with neuronal injury markers. *EMBO Molecular Medicine*, *8*(5), 466–476. https://doi.org/10.15252/emmm.201506123

Sun, X., & Xu, W. (2014). Fast implementation of DeLong's algorithm for comparing the areas under correlated receiver operating characteristic curves. *IEEE Signal Processing Letters*, *21*(11), 1389–1393. https://doi.org/10.1109/LSP.2014.2337313

Toledo, J. B., Korff, A., Shaw, L. M., Trojanowski, J. Q., & Zhang, J., & the Alzheimer's Disease Neuroimaging Initiative. (2014). Low levels of cerebrospinal fluid complement 3 and factor H predict faster cognitive decline in mild cognitive impairment. *Alzheimer's Research & Therapy*, *6* (3), 36. https://doi.org/10.1186/alzrt266

Wechsler, D. (1987). *Wechsler memory scale-revised*. Psychological Corporation.

Weiner, M. W., Veitch, D. P., Aisen, P. S., Beckett, L. A., Cairns, N. J., Green, R. C., Harvey, D., Jack, C. R., Jagust, W., Morris, J. C., Petersen, R. C., Saykin, A. J., Shaw, L. M., Toga, A. W., & Trojanowski, J. Q. (2017). Recent publications from the Alzheimer's Disease Neuroimaging Initiative: Reviewing progress toward improved AD clinical trials. *Alzheimer's & Dementia*, *13* (4), e1–e85. https://doi.org/10.1016/j.jalz.2016.11.007

Welge, V., Fiege, O., Lewczuk, P., Mollenhauer, B., Esselmann, H., Klafki, H.-W., Wolf, S., Trenkwalder, C., Otto, M., Kornhuber, J., Wiltfang, J., & Bibl, M. (2009). Combined CSF tau, p-tau181 and amyloid-β 38/40/42 for diagnosing Alzheimer's disease. *Journal of Neural Transmission*, *116*(2), 203–212. https://doi.org/10.1007/s00702-008-0177-6

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

Wilhalme, H., Goukasian, N., De Leon, F., He, A., Hwang, K. S., Woo, E., Elashoff, D., Zhou, Y., Ringman, J. M., & Apostolova, L. G. (2017). A comparison of theoretical and statistically derived indices for predicting cognitive decline. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *6*, 171–181. https://doi.org/10.1016/j.dadm.2016.10.002

Williams, B., Onsman, A., & Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Australasian Journal of Paramedicine*, *8*(3), 1–13. https://doi.org/10.33151/ajp.8.3.93

Ye, J., Farnum, M., Yang, E., Verbeeck, R., Lobanov, V., Raghavan, N., Novak, G., DiBernardo, A., & Narayan, V. A., for the Alzheimer's Disease Neuroimaging Initiative. (2012). Sparse learning and stability selection for predicting MCI to AD conversion using baseline ADNI data. *BMC Neurology*, *12*(1), 46. https://doi.org/10.1186/1471-2377-12-46

Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, *9*(2), 79–94. https://doi.org/10.20982/tqmp.09.2.p079

Zhang, D., Shen, D., & Initiative, A. D. N. (2012). Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers. *PLOS ONE*, *7*(3), e33182. https://doi.org/10.1371/journal.pone.0033182

Appendix A

Methods for Additional Analyses with the 99 ADNI 1 Subjects Reclassified Based on the Jak/Bondi Criteria

According to Bondi et al. (2014), the Jak/Bondi criteria were based on six measures in the ADNI NP battery, separately for three cognitive domains: Categorical Fluency and Boston Naming Test for the language domain, Trial Making Tests Parts A and B for the speed/executive function domain, and RAVLT delayed recall and recognition for the memory domain. Subjects were classified as MCI if they had (1) scores > 1 SD below the age-corrected normative mean on both tests in at least one cognitive domain; (2) scores > 1 SD below the age-corrected normative mean on one test in each of the three cognitive domains; or (3) scores = 9 on the Functional Assessment Questionnaire (FAQ; Pfeffer et al., 1982).

To run the additional analyses with subjects classified based on the Jak/Bondi criteria, we referred to Bondi et al.'s (2014) Supplementary Materials, in which ADNI normal and MCI subjects were already reclassified. Here, 234 ADNI 1 subjects were classified as MCI based on the Jak/Bondi criteria at baseline. Because EFA and CFA factor scores produced very similar results for the 160-subject sample classified based on the ADNI criteria, we only conducted EFAs in the follow-up analyses. We included all the 234 subjects in the EFAs of NP predictors and of MRI predictors. However, because CSF data were only available for 113 out of 234 MCI subjects, the EFA of CSF predictors was restricted to those subjects. The factor solutions for the three EFAs were the same as for the 160 MCI subjects classified based on the ADNI criteria. The logistic regressions using EFA factor scores as predictors were restricted to the baseline MCI subjects for whom NP, MRI, and CSF data were available for baseline and conversion statuses were identifiable for the following 24 months (N = 99).[13] These subjects were split into two subgroups: MCI_C (N = 46) and MCI_{NC} (N = 53).

Appendix B

Results for Follow-Up Analyses with the 192 ADNI 1 Subjects Reclassified Based on the Jak/Bondi Criteria

We reran the logistic regressions with only NP and MRI EFA factors for the 192 MCI subjects (including 108 MCI_C subjects and 84 MCI_{NC} subjects) who were diagnosed with the Jak/Bondi criteria. The regression results (see Table S4 and Figure S6 in the Supplementary Materials) were similar to the results reported for the 319-subject sample classified based on the ADNI criteria. The AUCs for NP, MRI, and NP + MRI were .78, .78 and .83, respectively. Most important, the Delong's tests revealed that NP and MRI factors did not differ in AUCs, but NP + MRI produced larger AUC than NP factors alone (z = 2.33, p = .020) and MRI factors alone (z = 2.21, p = .020).