

Featured Article

Early versus late MCI: Improved MCI staging using a neuropsychological approach

Emily C. Edmonds^{a,b,*}, Carrie R. McDonald^b, Anisa Marshall^b, Kelsey R. Thomas^{a,b}, Joel Eppig^c, Alexandra J. Weigand^c, Lisa Delano-Wood^{a,b}, Douglas R. Galasko^{a,b,d}, David P. Salmon^d, Mark W. Bondi^{a,b}, for the Alzheimer's Disease Neuroimaging Initiative¹

^aVeterans Affairs San Diego Healthcare System, San Diego, CA, USA

^bDepartment of Psychiatry, University of California San Diego, School of Medicine, La Jolla, CA, USA

^cSan Diego State University/University of California San Diego, Joint Doctoral Program in Clinical Psychology, San Diego, CA, USA

^dDepartment of Neurosciences, University of California San Diego, School of Medicine, La Jolla, CA, USA

Abstract

Introduction: The Alzheimer's Disease Neuroimaging Initiative (ADNI) separates “early” and “late” mild cognitive impairment (MCI) based on a single memory test. We compared ADNI's MCI classifications to our neuropsychological approach, which more broadly assesses cognitive abilities.

Methods: Three hundred thirty-six ADNI-2 participants were classified as “early” or “late” MCI. Cluster analysis was performed on neuropsychological test data, and participants were reclassified based on cluster results. These two staging approaches were compared on progression rates, cerebrospinal fluid biomarkers, and cortical thickness profiles.

Results: There was little correspondence between the two staging methods. ADNI's early MCI group included a large proportion of false-positive diagnostic errors. The reclassified neuropsychological MCI groups showed steeper survival curves and more abnormal biomarkers.

Conclusions: Our novel neuropsychological approach improved the staging of MCI by (1) capturing individuals at an early symptomatic stage, (2) minimizing false-positive cases, and (3) identifying a late MCI group further along the disease trajectory.

Published by Elsevier Inc. on behalf of the Alzheimer's Association.

Keywords:

Mild cognitive impairment; Early-stage MCI; Late-stage MCI; Alzheimer's disease; Dementia; Neuropsychology; Misdiagnosis; False positive; Cluster analysis

1. Introduction

Accurate identification of individuals in the early stages of dementia is vital to providing therapeutic interventions when they are likely to be most effective. Mild cognitive

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

*Corresponding author. Tel.: +858-552-8585 ext. 6369; Fax: +858-626-8537.

E-mail address: ecedmonds@ucsd.edu

impairment (MCI) is a transitional phase between normal cognitive aging and Alzheimer's disease (AD) in which individuals demonstrate objective cognitive impairment and report subjective complaints but have relatively intact functional abilities. In the Alzheimer's Disease Neuroimaging Initiative (ADNI), a multisite observational study of normal aging, MCI, and AD, participants are diagnosed with MCI based on their performance on a single episodic memory measure, subjective cognitive complaints, normal performance on a screening measure of global cognition, and a clinician's judgment of mild impairment based on a semistructured clinical interview [1].

This conventional diagnostic method, which is standard procedure for clinical trials and large-scale studies of MCI,

has been shown to be limited given its high susceptibility to false-positive diagnostic errors [2,3]. Within the ADNI MCI cohort, we have used statistical methods such as cluster analysis and latent profile analysis and found that a large proportion of participants diagnosed with MCI (up to one-third of ADNI's MCI sample) demonstrate intact performance on a more extensive neuropsychological test battery and show a low rate of progression to dementia [3,4]. Examination of AD biomarkers in this presumptive “false-positive” group has revealed normal cerebrospinal fluid (CSF) concentrations of β -amyloid and phosphorylated tau [3], normal levels of cortical amyloid burden [5], and normal cortical thickness profiles [6] as compared to a sample of robust normal controls. This false-positive group also tends to overreport subjective cognitive complaints, likely contributing to their MCI diagnosis [7], but they remain cognitively normal and functionally independent over time [8,9].

Our findings demonstrating a high rate of misdiagnosis via the conventional diagnostic criteria are consistent with previous research showing the unreliability of using a single test score [10–13] and subjective complaints [14,15] in the diagnosis of MCI. Several studies have explored alternative methods of MCI diagnosis by utilizing comprehensive neuropsychological data and requiring more than one impaired score for a diagnosis of MCI. Such neuropsychological criteria have been found to significantly improve diagnostic accuracy, produce MCI cohorts that show stronger associations with AD biomarkers, and improve prediction models for the development of AD [16–19].

In an attempt to identify participants as early as possible in the course of the disease, ADNI began subdividing MCI into “early” and “late” stages. Although individuals in both stages meet the conventional criteria for MCI, “early MCI” is thought to reflect those at an earlier point in the clinical spectrum [20]. ADNI assigns these stages based on different levels of impairment on the same single episodic memory measure that is used to diagnose MCI (i.e., one story from the Wechsler Memory Scale-Revised [WMS-R] Logical Memory II subtest).

Although it has been shown that the use of comprehensive neuropsychological data can improve the diagnosis of MCI [16–19], it is unknown how effectively these data can be applied to improve the staging of early and late MCI. Thus, we compared ADNI's early/late MCI classifications to our neuropsychological approach, which more broadly canvasses breadth and depth of cognitive functioning. We hypothesized that ADNI's early MCI group would include a substantial number of people with false-positive diagnoses due to the unreliability of MCI characterization based on a single memory measure. Furthermore, we expected that the neuropsychological approach would identify early and late MCI participants with higher rates of biomarker abnormalities and progression to AD relative to ADNI's conventional approach.

2. Methods

Data were obtained from the ADNI database (adni.loni.usc.edu). ADNI was launched in 2003 by the National Institute on Aging (NIA), National Institute of Biomedical Imaging and Bioengineering (NIBIB), Food and Drug Administration (FDA), private pharmaceutical companies, and nonprofit organizations. The primary goal of ADNI is to test whether neuroimaging, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. ADNI is the result of efforts of many coinvestigators from a range of academic institutions and private corporations, and participants have been recruited from over 50 sites across the U.S. and Canada. Participants in ADNI are between the ages of 55 and 90 years, completed at least 6 years of education, are fluent in English or Spanish, and are free of any significant neurological disease other than AD. The study was approved by the institutional review boards of all participating institutions. Informed written consent was obtained from all participants. For more information, see www.adni-info.org.

2.1. Participants

Participants were 336 individuals diagnosed with MCI and 294 cognitively normal (CN) individuals from the ADNI-2 MCI cohort. ADNI's MCI diagnoses were based on the following [1]: (1) subjective memory concern reported by the participant, study partner, or clinician; (2) abnormal memory function documented by scoring within education-adjusted ranges on delayed free recall of Story A from the WMS-R Logical Memory II subtest; (3) Mini-Mental State Examination (MMSE) score between 24 and 30; (4) global Clinical Dementia Rating (CDR) score of 0.5, with a Memory Box score of at least 0.5; and (5) general cognition and functional performance sufficiently preserved such that a diagnosis of AD could not be made. All MCI participants who met these criteria were further classified as “early MCI” ($n = 178$) or “late MCI” ($n = 158$) by ADNI based on the WMS-R Logical Memory II Story A score. The specific cutoff scores were as follows (out of a maximum score of 25): Early MCI (**ADNI-EMCI**) was assigned for a score of 9–11 for 16 or more years of education; a score of 5–9 for 8–15 years of education; or a score of 3–6 for 0–7 years of education. Late MCI (**ADNI-LMCI**) was assigned for a score of ≤ 8 for 16 or more years of education; a score of ≤ 4 for 8–15 years of education; or a score of ≤ 2 for 0–7 years of education.

CN participants in ADNI-2 were classified via the following criteria [1]: (1) no memory complaints, beyond what would be expected for age; (2) normal memory function documented by scoring above education-adjusted cutoffs on delayed free recall of Story A from WMS-R Logical Memory II (a score of ≥ 9 for 16 or more years of education; a score of ≥ 5 for 8–15 years of education; or a

score of ≥ 3 for 0–7 years of education); (3) MMSE score between 24 and 30; (4) global CDR of 0; and (5) absence of significant impairment in cognitive functions or activities of daily living.

2.2. Materials and procedure

2.2.1. Neuropsychological staging approach

Cluster analysis was performed on baseline neuropsychological test data from all MCI participants ($n = 336$). Neuropsychological scores examined included two measures of language (animal fluency, total score; 30-item Boston Naming Test, total score), two measures of attention/executive function (Trail Making Test, Parts A and B, time to completion), and two measures of memory (Rey Auditory Verbal Learning Test [RAVLT], 30-minute delayed free recall, number of words recalled; RAVLT corrected recognition score [hits minus false alarms]).

These measures were selected because they were administered to all ADNI participants and they were not used in ADNI's diagnostic criteria. Visuospatial/visuoconstructional measures were not included in the cluster analysis given that measures assessing this domain are limited in ADNI.

Raw scores for each of the six neuropsychological measures were transformed into standardized z-scores based on demographically corrected (age, sex, education) normative data, and z-scores were entered into a hierarchical cluster analysis. All MCI participants were *re-classified* as neuropsychological-early MCI (**NP-EMCI**), neuropsychological-late MCI (**NP-LMCI**), or “false positive” based on results of the cluster analysis.

2.2.2. CSF biomarkers

CSF data were available for 307 of the MCI participants and 255 of the CN participants. CSF was processed using Elecsys immunoassays. Biomarkers included β -amyloid ($A\beta_{1-42}$) concentration and the ratios of phosphorylated tau over $A\beta$ (pTau/ $A\beta_{1-42}$) and total tau over $A\beta$ (tTau/ $A\beta_{1-42}$). Published cutoff scores optimized for ADNI were used to determine biomarker positivity [21]: <977 pg/mL for $A\beta_{1-42}$, $>.025$ for pTau/ $A\beta_{1-42}$, and $>.27$ for tTau/ $A\beta_{1-42}$.

2.2.3. Cortical thickness biomarkers

A subset of the sample, including 172 of the MCI participants (95 ADNI-EMCI; 77 ADNI-LMCI) and 82 of the CN participants, had neuroimaging data (i.e., T1-weighted MRI) available, which passed our local quality control procedure (described in [6]). Images were processed using FreeSurfer software (v.5.3.0), and cortical thickness measurements were obtained using standard validated procedures [22,23]. Cortical thickness estimates for each participant were computed at each vertex (~ 1 mm spacing) across the cortical mantle and within 34 gyral-based regions of interest per hemisphere [24]. Mean thickness for each region of interest was calculated by averaging the cortical thickness

measurements across vertices within a given region based on unsmoothed data.

2.3. Statistical analyses

Z-scores for the six neuropsychological variables for each MCI participant were entered into a hierarchical cluster analysis using Ward's method, consistent with our previous work [2,3,16,25]. To examine how well the cluster solution fit the data, a discriminant function analysis was conducted using the six neuropsychological measures to predict group membership based on the number of clusters derived. The stability of the cluster solution was also examined using the leave-one-out cross-validation method.

Group differences in demographic variables, apolipoprotein E (*APOE*) $\epsilon 4$ status, and CDR Sum of Boxes (CDR-SOB) were examined using χ^2 and one-way analysis of variance with post hoc t-tests. Kaplan-Meier survival analysis was used to test for group differences in progression to AD, and curves were compared using a log-rank test. χ^2 analyses were used to compare rates of CSF biomarker positivity between diagnostic groups. To create cortical thickness surface maps, individual surfaces were resampled into a common spherical coordinate system that aligned cortical folding patterns across participants [26]. A general linear model was used to compare surface maps for each group relative to the CN group, and a false discovery rate correction was applied to account for multiple comparisons.

3. Results

3.1. Neuropsychological cluster groups

Cluster analysis of the six neuropsychological scores from the 336 MCI participants resulted in three groups: (1) **NP-EMCI** with impaired memory and below average naming ($n = 147$); (2) **NP-LMCI** with deficits across all three cognitive domains ($n = 49$); and (3) a “false positive” (**FP**) group that performed within normal limits on the more extensive neuropsychological testing despite their original MCI diagnosis ($n = 140$), consistent with our previous studies [3,4,16]; see Fig. 1. A discriminant function analysis using the six neuropsychological measures to predict group membership into these three clusters correctly classified 93.2% of the participants, and cross-validation using the leave-one-out method showed only a mild reduction in classification accuracy (91.4%). There were significant differences in age and rate of *APOE* $\epsilon 4$ carriers between groups, but no differences in education or sex (see Table 1).

3.2. Comparison of staging methods

Comparison of the neuropsychological staging method and ADNI's staging method is shown in Fig. 2. There was little correspondence between the two approaches, as participants were not consistently categorized as EMCI or LMCI. For participants in the ADNI-EMCI group, 56% fell into the

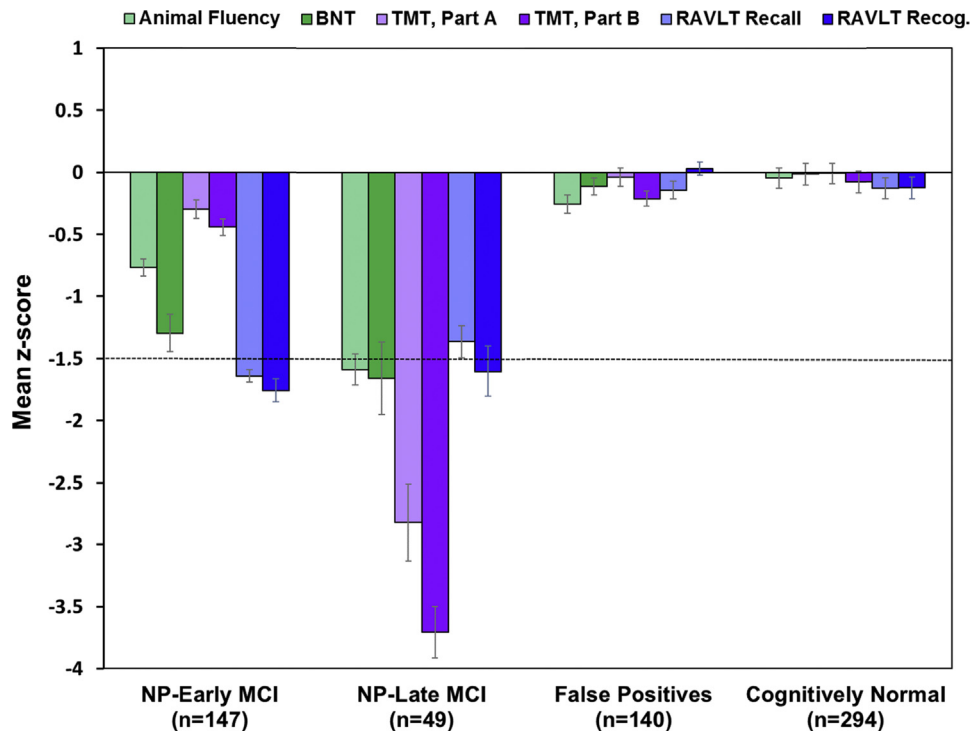


Fig. 1. Neuropsychological performance of cluster-derived groups and the CN group. Error bars denote standard error of the mean. The horizontal dotted line indicates the typical cutoff for impairment (-1.5 SDs). Abbreviations: BNT, Boston Naming Test; TMT, Trail Making Test; RAVLT, Rey Auditory Verbal Learning Test; NP, neuropsychological; MCI, mild cognitive impairment.

FP group, 34% were classified as NP-EMCI, and 10% were classified as NP-LMCI. For participants in the ADNI-LMCI group, 26% fell into the FP group, 54% were classified as NP-EMCI, and 20% were classified as NP-LMCI.

An analysis of variance showed significant group differences in the CDR (CDR-SOB: $F = 174.79$, $P < .001$, $\eta_p^2 = .48$), a measure used in ADNI's original diagnoses. Comparison of the two "early" groups showed that the NP-EMCI scored higher on the CDR-SOB (indicating greater impairment) than the ADNI-EMCI group ($P < .001$). Similarly, the NP-LMCI group scored higher than the ADNI-LMCI group ($P = .02$). The CDR-SOB did not differ between the NP-EMCI and the ADNI-LMCI groups ($P = .22$), which

is not surprising since these two groups included many of the same participants (see Fig. 2). In addition, the FP and ADNI-EMCI groups did not differ from one another, again due to these groups being comprised of many of the same participants.

3.3. Progression to AD

Follow-up data for an average of 30 months ($SD = 15.9$; range 0-60 months) showed that a subset of participants ($n = 185$) progressed to meet criteria for a diagnosis of probable AD based on NINCDS/ADRDA criteria. Kaplan-Meier survival curves showing rate of progression to AD in the

Table 1
Demographic characteristics and *APOE* $\epsilon 4$ status for each diagnostic group

Variable	ADNI-EMCI (n = 178)	ADNI-LMCI (n = 158)	NP-EMCI (n = 147)	NP-LMCI (n = 49)	FP (n = 140)	CN (n = 294)	F or χ^2	P value	Effect size
Age, years, mean (SD)	71.0 (7.3)	72.3 (7.6)	72.0 (7.0)	74.6 (7.4)	70.2 (7.7)	73.0 (6.0)	$F = 5.34^*$	$P < .001$	$\eta_p^2 = .03$
Education, years, mean (SD)	16.1 (2.7)	16.5 (2.6)	15.9 (2.7)	16.6 (2.8)	16.6 (2.6)	16.6 (2.5)	$F = 2.07$	$P = .07$	$\eta_p^2 = .01$
Sex (% male)	57.3%	51.9%	55.8%	61.2%	51.4%	45.6%	$\chi^2 = 9.39$	$P = .09$	$\phi_c = .10$
% <i>APOE</i> $\epsilon 4$ carriers	45.1%	56.7%	56.6%	59.2%	41.3%	30.1%	$\chi^2 = 47.26^\dagger$	$P < .002$	$\phi_c = .22$

Abbreviations: ADNI, Alzheimer's Disease Neuroimaging Initiative; *APOE*, apolipoprotein E; CN, cognitively normal; EMCI, early mild cognitive impairment; LMCI, late mild cognitive impairment; FP, false positive; NP, neuropsychological.

*ADNI-EMCI younger than NP-LMCI and CN; FP younger than NP-LMCI and CN.

† ADNI-EMCI higher rate of *APOE* $\epsilon 4$ carriers than CN; ADNI-LMCI higher than CN, FP, and ADNI-EMCI; NP-EMCI higher than CN, FP, and ADNI-EMCI; NP-LMCI higher than CN, FP, and ADNI-EMCI; FP higher than CN.

	ADNI-Early MCI	ADNI-Late MCI
False Positives	56% (n=99)	26% (n=41)
NP-Early MCI	34% (n=61)	54% (n=86)
NP-Late MCI	10% (n=18)	20% (n=31)
Total	100% (n=178)	100% (n=158)

Fig. 2. Percentage of participants in the early and late MCI groups by staging method. Abbreviations: ADNI, Alzheimer's Disease Neuroimaging Initiative; NP, neuropsychological; MCI, mild cognitive impairment.

groups using each staging method are shown in Fig. 3. A log-rank test revealed significant differences in survival curves [$\chi^2(5) = 232.53$; $P < .001$]. Comparison of the two “early” groups showed that the NP-EMCI curve was significantly steeper than the ADNI-EMCI curve ($P < .001$). Similarly, the NP-LMCI curve was significantly steeper than the ADNI-LMCI curve ($P = .01$). The survival curves did not differ between the NP-EMCI and the ADNI-LMCI groups ($P = .24$). Of note, there was no significant difference in the survival curves for the CN and FP groups, supporting the conclusion that many of these individuals likely represent false-positive cases and their original diagnosis of MCI should be revisited.

3.4. CSF biomarkers

χ^2 analyses showed significant group differences in CSF biomarker positivity for $A\beta_{1-42}$ ($\chi^2 = 92.13$,

$P < .001$, $\phi_c = .33$), $p\text{Tau}/A\beta_{1-42}$ ($\chi^2 = 106.11$, $P < .001$, $\phi_c = .35$), and $t\text{Tau}/A\beta_{1-42}$ ($\chi^2 = 100.29$, $P < .001$, $\phi_c = .34$). For each of the CSF variables, the proportion of biomarker-positive individuals did not differ between the CN and FP groups, nor between the FP and ADNI-EMCI groups. A higher rate of biomarker positivity was observed in the ADNI-LMCI, NP-EMCI, and NP-LMCI groups; however, these three groups did not significantly differ from one another; see Fig. 4.

3.5. Cortical thickness biomarkers

Differences in regional cortical thickness between each diagnostic group relative to the CN group are displayed at the vertex-wise level on the lateral and medial surfaces in Fig. 5. For the ADNI-EMCI and ADNI-LMCI groups, only very minimal thinning in the left medial temporal lobe (MTL) was observed. The NP-EMCI group showed thinning in bilateral MTLs and left lateral temporal lobe regions, whereas the NP-LMCI group showed widespread cortical thinning bilaterally. Surface maps for the FP group actually showed thicker cortex relative to the CN group in left occipitotemporal regions.

4. Discussion

We reclassified participants in ADNI's MCI cohort as either EMCI or LMCI based on a cluster analysis of performance on multiple neuropsychological measures. This

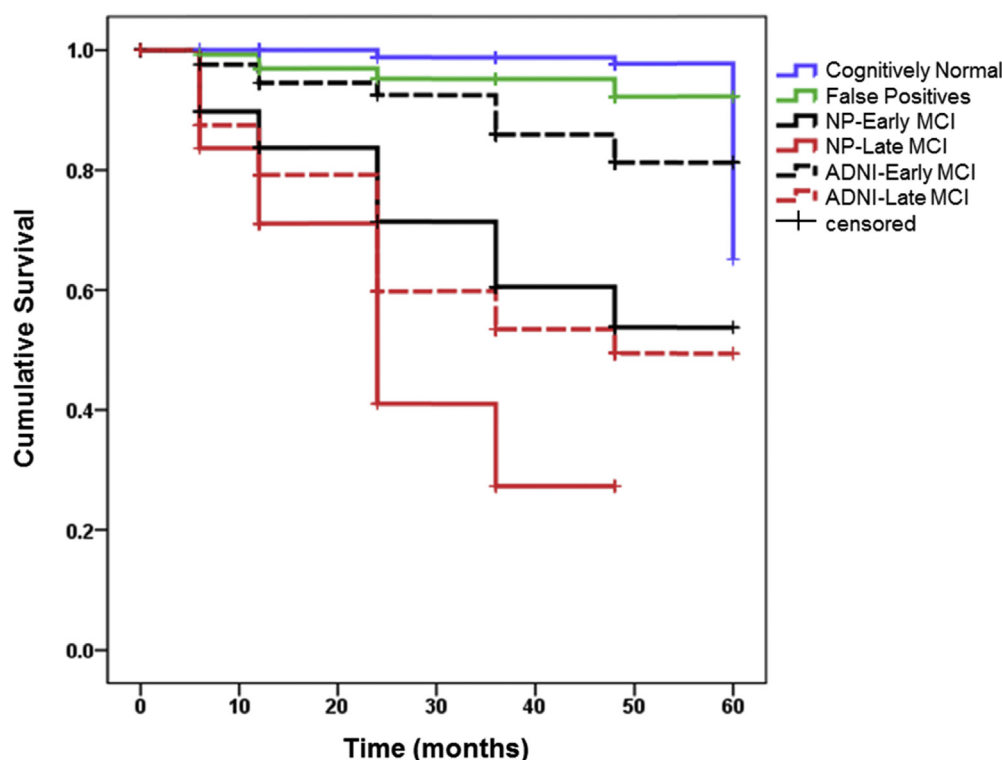


Fig. 3. Kaplan-Meier survival curves showing rate of progression to AD in the MCI groups by staging method. Abbreviations: ADNI, Alzheimer's Disease Neuroimaging Initiative; NP, neuropsychological; MCI, mild cognitive impairment.

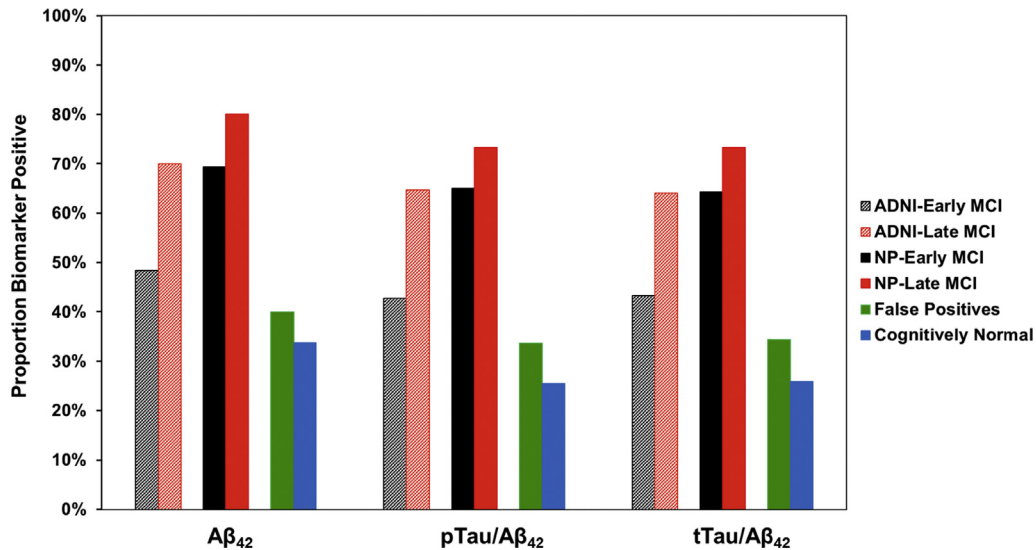


Fig. 4. Proportion of CSF AD biomarker-positive individuals in each group by staging method. Cutoffs used to determine biomarker positivity [21]: <977 pg/mL for $A\beta_{1-42}$, $>.025$ for $p\text{Tau}/A\beta_{1-42}$, and $>.27$ for $t\text{Tau}/A\beta_{1-42}$. Abbreviations: ADNI, Alzheimer's Disease Neuroimaging Initiative; NP, neuropsychological; MCI, mild cognitive impairment; CSF, cerebrospinal fluid.

neuropsychological method represents a novel approach to staging the early and late phases of MCI because staging has traditionally been based only on level of impairment on one memory measure. This conventional method has been used not only by ADNI [20] but also by studies published on early and late MCI in other samples (e.g., [27]). Results of our cluster analysis identified a large group of participants (42% of the MCI sample) who performed normally on neuropsychological testing despite having been diagnosed with MCI by ADNI based on subjective ratings (i.e., CDR; subjective cognitive complaints) and limited objective testing (i.e., MMSE; delayed memory for one story).

We have studied this FP group extensively in previous studies and found them to have normal CSF and imaging biomarkers [3–6], as well as intact cognitive performance and functional independence over time [8,9]. Consistent with this, the present study showed no differences in progression to AD, CSF biomarkers, or cortical thickness between the FP and CN participants. Prior work has also shown that FP individuals tend to overreport subjective cognitive complaints (relative to informant report) [7], which may have contributed to their CDR score of 0.5 and subsequent MCI diagnosis. These self-reported subjective complaints were found to be unrelated to objective cognitive functioning but associated with depressive symptoms [7]. Although the FP group is at higher genetic risk relative to the CN group based on *APOE* $\epsilon 4$ findings, a diagnosis of MCI is not warranted at this time given their intact performances on neuropsychological testing.

As hypothesized, a large proportion (56%) of the ADNI-EMCI group fell into the FP group. Thus, ADNI's attempt to identify participants as early as possible in the course of the

disease [20] appears to have resulted in an unreliable EMCI diagnosis that is overinclusive. Another concerning finding is that 26% of the ADNI-LMCI group appeared to represent false-positive diagnostic errors. This emphasizes the importance of considering multiple cognitive measures when assigning an MCI diagnosis, as performing below 1.5 SD on a single memory test is not atypical in healthy older adults (e.g., 39% of healthy older adults in the WMS-III standardization sample had at least one memory score 1.5 SD below the mean [11]). In addition, requiring two impaired scores within a cognitive domain has been shown to be more reliable than one impaired score when diagnosing MCI [16,19,28]. Support for this comes from a previous study which found that 20% of healthy older adults obtained one impaired score (-2.0 SD) in two different cognitive domains, but far fewer (approximately 5% or less) had two or more impaired scores within the same cognitive domain [29].

ADNI's use of education-adjusted cutoffs to classify EMCI and LMCI may have contributed to the false-positive error rate, particularly for older individuals. For example, a 75-year-old with 16 years of education who achieved a raw score of 11 on delayed recall of Story A from WMS-R Logical Memory II would be considered solidly average based on published age-adjusted normative data (scaled score = 10; [30]). However, this performance could lead to a classification of EMCI according to ADNI's education-adjusted cutoffs [20]. Similarly, a raw score of 8 would be considered intact for an 82-year-old (scaled score = 8; [30]) but would fall into ADNI's LMCI range if the individual had 16 years of education. ADNI's cutoffs alone, however, cannot entirely account for the high

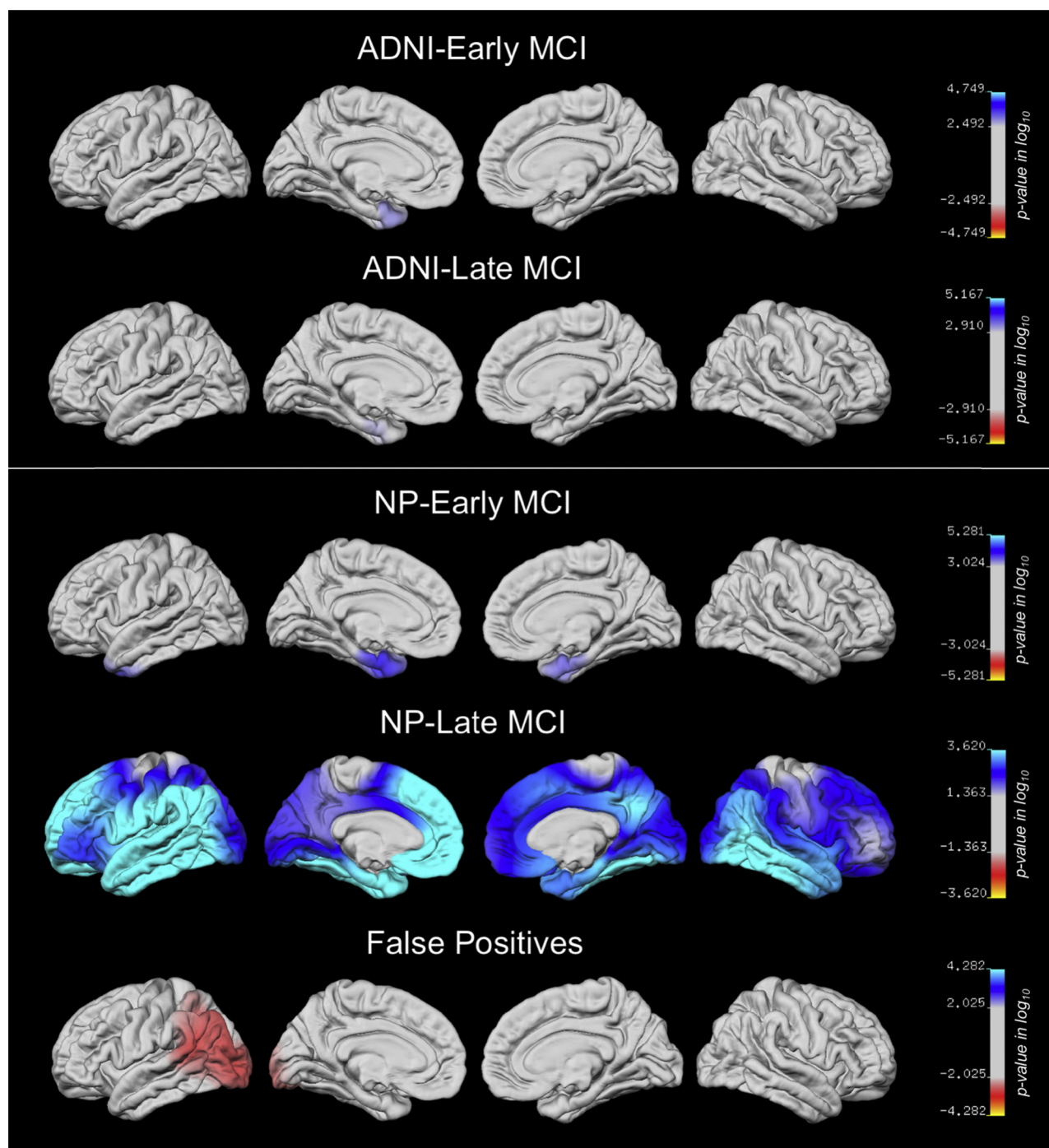


Fig. 5. T-value surface maps showing regional cortical thickness on the left and right lateral and medial pial surfaces for each group relative to the CN group with false discovery rate correction for multiple comparisons. The cyan/blue shades represent areas where the MCI subgroup has thinner cortex than the CN group, whereas the red regions represent areas where the subgroup has thicker cortex than the CN group. Abbreviations: ADNI, Alzheimer's Disease Neuroimaging Initiative; NP, neuropsychological; MCI, mild cognitive impairment; CN, cognitively normal.

frequency of false-positive cases, as the FP group was not older on average, and there were no differences in education between groups (see Table 1). Rather, it appears that the diagnostic approach itself is susceptible to errors. The

inclusion of false-positive cases in observational studies or clinical trials of MCI may lead to diluted results or a reduced ability to detect treatment effects [31]. In addition, false-positive diagnoses may have clinical implications,

including the potential for inappropriate medication use or undue anxiety that could result from an inaccurate diagnosis.

The notion that “level of impairment” (e.g., -1 SD vs. -1.5 SD) on a test corresponds to degree of disease (early/late or mild/severe) is a concept that is pervasive in the field. For instance, the DSM-5 [32] makes a distinction between cognitive test scores that are 1–2 SDs versus 2 or more SDs below the normative mean, with performance corresponding to a diagnosis of mild neurocognitive disorder or major neurocognitive disorder, respectively. Such an approach is limited by the psychometric properties of normative scores. For example, an 83-year-old participant with 12 years of education in the present study had a raw score of 0 on RAVLT delayed recall, which corresponded to a demographically adjusted z-score of -1.41 . This individual would be considered EMCI by the conventional staging approach but would never reach the -1.5 SD threshold on this test to be considered LMCI. Our neuropsychological method of staging MCI is less affected by these psychometric issues because it examines the number and pattern of impaired tests across multiple measures and several cognitive domains, rather than simply the degree of impairment on one measure.

Examination of rates of progression to AD in our study showed that our reclassified MCI groups had steeper survival curves than ADNI's MCI groups. Similarly, the reclassified groups had a higher proportion of individuals with abnormal CSF AD biomarkers, particularly the NP-EMCI group relative to the ADNI-EMCI group. Furthermore, analysis of a large subset of the sample revealed that ADNI's EMCI and LMCI groups showed whole-brain cortical thickness estimates that were similar to CN participants, with only minimal MTL thinning in both groups. Our reclassified groups, on the other hand, showed more robust MTL thinning in the NP-EMCI group and widespread bilateral cortical thinning in the NP-LMCI group. These findings in the reclassified groups correspond more closely to the known biological underpinnings of AD, with pathology (e.g., neurofibrillary tangles) initially restricted to MTL regions in early phases of the disease and progressing to cortical association areas in later phases [33].

Our cortical thickness findings are comparable to a previous study in a clinic-based sample, which found that EMCI participants had thinning in MTL and insular regions, whereas LMCI participants exhibited cortical thinning in multiple regions bilaterally [34]. This previous study examined more than one memory measure to classify their MCI groups (i.e., they used a verbal and a visual memory test, both of which were required to be impaired for an EMCI diagnosis) [34], which likely improved reliability relative to ADNI's use of delayed memory for one story. Taken together, progression rates and biomarker profiles indicate that the neuropsychological method for staging MCI improved the identification of at-risk individuals. This is

highlighted by the finding that, across analyses, participants classified as having *early* MCI based on our neuropsychological method were comparable to the *late* MCI group based on ADNI's criteria.

Results of the present study have implications for the interpretation of previous studies that have been conducted using ADNI's EMCI and LMCI designations. For example, one study examined biomarkers in ADNI's EMCI and cognitively healthy participants and found that cortical amyloid accumulation and lower CSF A β levels were associated with *APOE* $\epsilon 4$ status, but not with diagnostic group [35]. It can be speculated that this lack of difference between EMCI and control participants could be at least partially attributed to inclusion of false-positive diagnostic errors in the EMCI sample, thereby diminishing the utility of ADNI's EMCI group to differentiate between CN and at-risk individuals.

Other studies have shown differences between ADNI's EMCI and LMCI groups on variables such as amyloid burden and metabolism [36], hippocampal atrophy [37], and functional connectivity in the default mode network [38] and in the thalamus [39]. However, ADNI's MCI staging method limits the interpretation of these studies since the sample, particularly the EMCI group, is contaminated by the inclusion of a large number of false-positive diagnostic errors. This may have led to a somewhat artificial separation of ADNI's EMCI and LMCI groups and contributed to the differences found between these two stages in previous work.

A limitation of this study is that early and late MCI designations were assigned at a group level based on results of a cluster analysis (following the methodology of our previous work [2,3,16,25]), rather than based on an individual's specific test scores. The pattern of neuropsychological performance observed indicates that EMCI is consistent with a single-domain amnesic profile, whereas LMCI reflects multidomain cognitive impairment, but further research is needed to determine the optimal way of applying these stages at an individual level (e.g., [16]). Future studies should also focus on longitudinal staging of MCI to determine how a participant's MCI stage progresses over time, as well as examination of additional biomarkers that may distinguish EMCI and LMCI groups (e.g., neuroimaging of subcortical structures).

Strengths of our study include the examination of multiple neuropsychological tests across several cognitive domains to classify MCI stages, and examination of both CSF and neuroimaging biomarkers. Our novel neuropsychological method improved the staging of MCI by (1) capturing an early MCI group comprising memory-impaired individuals in an early symptomatic stage of MCI, (2) minimizing false-positive cases, and (3) identifying a late MCI group that is further along the disease trajectory given their multidomain cognitive impairment, higher rate of progression to AD, abnormal CSF

biomarkers, and extensive cortical thinning. Results have important implications for future studies attempting to examine early versus late stages of prodromal AD, as well as clinical trials testing medications designed to target AD pathology in order to modify the course of the disease and improve outcomes.

Acknowledgments

This work was supported by the Alzheimer's Association (AARG-17-500358 to E.C.E.; AARF-17-528918 to K.R.T.), the U.S. Department of Veterans Affairs Clinical Sciences Research and Development Service (Career Development Award-2 1K2 CX001415 to E.C.E.), and the NIH (R01 AG049810 to M.W.B.; K24 AG026431 to M.W.B.; P50 AG05131 to D.R.G.). 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.

Dr. Bondi is a consulting editor for the *Journal of the International Neuropsychological Society*; serves as a consultant for Eisai, Novartis, and Roche; and receives royalties from Oxford University Press. Dr. Galasko serves as editor for *Alzheimer's Research and Therapy* and as a paid consultant on Data Safety Monitoring Boards for Pfizer, Inc., Elan, Inc., and Balance Pharmaceuticals, Inc. Dr. Salmon serves as a paid consultant for Takeda Pharmaceuticals. The other authors report no disclosures.

RESEARCH IN CONTEXT

1. Systematic review: The authors searched PubMed for studies related to the classification of “early” and “late” mild cognitive impairment (MCI). Review of the literature revealed that the conventional method of staging MCI is based on level of impairment on a single memory test. Prior work has also shown high rates of false-positive errors based on conventional diagnostic criteria for MCI, which is a standard procedure for clinical trials and large-scale studies of MCI.
2. Interpretation: Our findings show that the conventional staging method results in an unreliable “early” MCI diagnosis that is overinclusive. Results emphasize the importance of considering multiple cognitive measures when diagnosing and staging MCI.
3. Future directions: Further research is needed to determine the optimal way of applying early and late MCI classification at an individual level and to determine whether our staging method has utility in improving the identification of individuals appropriate for clinical trials.

References

- [1] Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology* 2010;74:201–9.
- [2] Clark LR, Delano-Wood L, Libon DJ, McDonald CR, Nation DA, Bangen KJ, et al. Are empirically derived subtypes of mild cognitive impairment consistent with conventional subtypes? *J Int Neuropsychol Soc* 2013;19:1–11.
- [3] Edmonds EC, Delano-Wood L, Clark LR, Jak AJ, Nation DA, McDonald CR, et al. Susceptibility of the conventional criteria for mild cognitive impairment to false positive diagnostic errors. *Alzheimers Dement* 2015;11:415–24.
- [4] Eppig JS, Edmonds EC, Campbell L, Sanderson-Cimino M, Delano-Wood L, Bondi MW. Statistically derived subtypes and associations with cerebrospinal fluid and genetic biomarkers in mild cognitive impairment: a latent profile analysis. *J Int Neuropsychol Soc* 2017;23:564–76.
- [5] Bangen KJ, Clark AL, Werhane M, Edmonds EC, Nation DA, Evangelista N, et al. Cortical amyloid burden differences across empirically-derived mild cognitive impairment subtypes and interaction with APOE ε4 genotype. *J Alzheimers Dis* 2016;52:849–61.
- [6] Edmonds EC, Eppig J, Bondi MW, Leyden KM, Goodwin B, Delano-Wood L, et al. Heterogeneous cortical atrophy patterns in MCI not captured by conventional diagnostic criteria. *Neurology* 2016;87:2108–16.
- [7] Edmonds EC, Delano-Wood L, Galasko DR, Salmon DP, Bondi MW. Subjective cognitive complaints contribute to misdiagnosis of mild cognitive impairment. *J Int Neuropsychol Soc* 2014;20:836–47.
- [8] Edmonds EC, Weigand AJ, Thomas KR, Eppig J, Delano-Wood L, Galasko DR, et al. Increasing inaccuracy of self-reported subjective cognitive complaints over 24 months in empirically-derived subtypes

- of mild cognitive impairment. *J Int Neuropsychol Soc* 2018; 24:842–53.
- [9] Thomas KR, Edmonds EC, Delano-Wood L, Bondi MW. Longitudinal trajectories of informant-reported daily functioning in empirically-defined subtypes of mild cognitive impairment. *J Int Neuropsychol Soc* 2017;23:521–7.
 - [10] Brooks BL, Iverson GL, White T. Substantial risk of “accidental MCI” in healthy older adults: base rates of low memory scores in neuropsychological assessment. *J Int Neuropsychol Soc* 2007;13:490–500.
 - [11] Brooks BL, Iverson GL, Holdnack JA, Feldman HH. Potential for misclassification of mild cognitive impairment: a study of memory scores on the Wechsler Memory Scale-III in healthy older adults. *J Int Neuropsychol Soc* 2008;14:463–78.
 - [12] de Rotrou J, Wenisch E, Chausson C, Dray F, Fauconau V, Rigaud AS. Accidental MCI in healthy subjects: a prospective longitudinal study. *Eur J Neurol* 2005;12:879–85.
 - [13] Klekociuk SZ, Summers JJ, Vickers JC, Summers MJ. Reducing false positive diagnoses in mild cognitive impairment: the importance of comprehensive neuropsychological assessment. *Eur J Neurol* 2014; 21:1330–6.
 - [14] Lenehan ME, Klekociuk SZ, Summers MJ. Absence of a relationship between subjective memory complaint and objective memory impairment in mild cognitive impairment (MCI): is it time to abandon subjective memory complaint as an MCI diagnostic criterion? *Int Psychogeriatr* 2012;24:1505–14.
 - [15] Yates JA, Clare L, Woods RT. Subjective memory complaints, mood and MCI: a follow-up study. *Aging Ment Health* 2017;21:313–21.
 - [16] Bondi MW, Edmonds EC, Jak AJ, Clark LR, Delano-Wood L, McDonald CR, et al. Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and prediction of progression. *J Alzheimers Dis* 2014;42:275–89.
 - [17] Callahan BL, Ramirez J, Berezuk C, Duchesne S, Black SE. Predicting Alzheimer's disease development: a comparison of cognitive criteria and associated neuroimaging biomarkers. *Alzheimers Res Ther* 2015;7:68.
 - [18] Goerlich KS, Votinov M, Dicks E, Ellendt S, Csukly G, Habel U. Neuroanatomical and neuropsychological markers of amnesic MCI: a three-year longitudinal study in individuals unaware of cognitive decline. *Front Aging Neurosci* 2017;9:34.
 - [19] Jak AJ, Preis SR, Beiser AS, Seshadri S, Wolf PA, Bondi MW, et al. Neuropsychological criteria for mild cognitive impairment and dementia risk in the Framingham Heart Study. *J Int Neuropsychol Soc* 2016;22:937–43.
 - [20] Aisen PS, Petersen RC, Donohue MC, Gamst A, Raman R, Thomas RG, et al. Clinical core of the Alzheimer's Disease neuroimaging initiative: progress and plans. *Alzheimers Dement* 2010; 6:239–46.
 - [21] Hansson O, Seibyl J, Stomrud E, Zetterberg H, Trojanowski JQ, Bittner T, et al. CSF biomarkers of Alzheimer's disease concord with amyloid- β PET and predict clinical progression: a study of fully automated immunoassays in BioFINDER and ADNI cohorts. *Alzheimers Dement* 2018;14:1470–81.
 - [22] Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* 1999;9:179–94.
 - [23] Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A* 2000;97:11050–5.
 - [24] Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 2006;31:968–80.
 - [25] Delano-Wood L, Bondi MW, Sacco J, Abeles N, Jak AJ, Libon DJ, et al. Heterogeneity in mild cognitive impairment: differences in neuropsychological profile and associated white matter lesion pathology. *J Int Neuropsychol Soc* 2009;15:906–14.
 - [26] Fischl B, Sereno MI, Tootell RB, Dale AM. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp* 1999;8:272–84.
 - [27] Jessen F, Wolfgruber S, Wiese B, Bickel H, Mosch E, Kaduszkiewicz H, et al. AD dementia risk in late MCI, in early MCI, and in subjective memory impairment. *Alzheimers Dement* 2014;10:76–83.
 - [28] Jak AJ, Bondi MW, Delano-Wood L, Wierenga C, Corey-Bloom J, Salmon DP, et al. Quantification of five neuropsychological approaches to defining mild cognitive impairment. *Am J Geriatr Psychiatry* 2009;17:368–75.
 - [29] Palmer BW, Boone KB, Lesser IM, Wohl MA. Base rates of “impaired” neuropsychological test performance among healthy older adults. *Arch Clin Neuropsychol* 1998;13:503–11.
 - [30] Smith GE, Wong JS, Ivnik RJ, Malec JF. Mayo's Older American Normative Studies: separate norms for the WMS-R Logical Memory stories. *Psychol Assess* 1997;4:79–86.
 - [31] Edmonds EC, Ard MC, Edland SD, Galasko DR, Salmon DP, Bondi MW. Unmasking the benefits of donepezil via psychometrically precise identification of mild cognitive impairment: a secondary analysis of the ADCS vitamin E and donepezil in MCI study. *Alzheimers Dement (N Y)* 2018;4:11–8.
 - [32] American Psychiatric Association. Diagnostic and statistical manual of mental disorders: Fifth Edition (DSM-5). Arlington, VA: American Psychiatric Publishing; 2013.
 - [33] Braak H, Braak E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol* 1991;82:239–59.
 - [34] Ye BS, Seo SW, Yang JJ, Kim HJ, Kim YJ, Yoon CW, et al. Comparison of cortical thickness in patients with early-stage versus late-stage amnesic mild cognitive impairment. *Eur J Neurol* 2014;21:86–92.
 - [35] Risacher SL, Kim S, Shen L, Nho K, Foroud T, Green RC, et al. The role of apolipoprotein E (APOE) genotype in early mild cognitive impairment (E-MCI). *Front Aging Neurosci* 2013;5:11.
 - [36] Wu L, Rowley J, Mohades S, Leuzy A, Dauar MT, Shin M, et al. Dissociation between brain amyloid deposition and metabolism in early mild cognitive impairment. *PLoS One* 2012;7:e47905.
 - [37] Lee P, Ryoo H, Park J, Jeong Y. Morphological and microstructural changes of the hippocampus in early MCI: a study utilizing the Alzheimer's Disease Neuroimaging Initiative database. *J Clin Neurol* 2017;13:144–54.
 - [38] Lee ES, Yoo K, Lee YB, Chung J, Lim JE, Yoon B. Default mode network functional connectivity in early and late mild cognitive impairment. *Alzheimer Dis Assoc Disord* 2016; 30:289–96.
 - [39] Cai S, Huang L, Zou J, Jing L, Zhai B, Ji G, et al. Changes in thalamic connectivity in the early and late stages of amnesic mild cognitive impairment: a resting-state functional magnetic resonance study from ADNI. *PLoS One* 2015;10:e0115573.