Original Investigation

Comparison of Visual and Quantitative Florbetapir F 18 Positron Emission Tomography Analysis in Predicting Mild Cognitive Impairment Outcomes

Stefanie Schreiber, MD; Susan M. Landau, PhD; Allison Fero, BSc; Frank Schreiber, MSc; William J. Jagust, MD; for the Alzheimer's Disease Neuroimaging Initiative

IMPORTANCE The applicability of β-amyloid peptide (Aβ) positron emission tomography (PET) as a biomarker in clinical settings to aid in selection of individuals at preclinical and prodromal Alzheimer disease (AD) will depend on the practicality of PET image analysis. In this context, visual-based Aβ PET assessment seems to be the most feasible approach.

OBJECTIVES To determine the agreement between visual and quantitative Aβ PET analysis and to assess the ability of both techniques to predict conversion from mild cognitive impairment (MCI) to AD.

DESIGN, SETTING, AND PARTICIPANTS A longitudinal study was conducted among the Alzheimer's Disease Neuroimaging Initiative (ADNI) sites in the United States and Canada during a 1.6-year mean follow-up period. The study was performed from September 21, 2010, to August 11, 2014; data analysis was conducted from September 21, 2014, to May 26, 2015. Participants included 401 individuals with MCI receiving care at a specialty clinic (219 [54.6%] men; mean [SD] age, 71.6 [7.5] years; 16.2 [2.7] years of education). All participants were studied with florbetapir F 18 [^{18}F] PET. The standardized uptake value ratio (SUVR) positivity threshold was 1.11, and one reader rated all images, with a subset of 125 scans rated by a second reader.

MAIN OUTCOMES AND MEASURES Sensitivity and specificity of positive and negative [^{18}F] florbetapir PET categorization, which was estimated with cerebrospinal fluid Aβ1-42 as the reference standard. Risk for conversion to AD was assessed using Cox proportional hazards regression models.

RESULTS The frequency of Aβ positivity was 48.9% (196 patients; visual analysis), 55.1% (221 patients; SUVR), and 64.8% (166 patients; cerebrospinal fluid), yielding substantial agreement between visual and SUVR data (κ = 0.74) and between all methods (Fleiss κ = 0.71). For approximately 10% of the 401 participants in whom visual and SUVR data disagreed, interrater reliability was moderate (κ = 0.44), but it was very high if visual and quantitative results agreed (κ = 0.92). Visual analysis had a lower sensitivity (79% vs 85%) but higher specificity (96% vs 90%), respectively, compared with SUVR. The conversion rate was 15.2% within a mean of 1.6 years, and a positive [^{18}F] florbetapir baseline scan was associated with a 6.91-fold (SUVR) or 11.38-fold (visual) greater hazard for AD conversion, which changed only modestly after covariate adjustment for apolipoprotein ε4, concurrent fludeoxyglucose F 18 PET scan, and baseline cognitive status.

CONCLUSIONS AND RELEVANCE Visual and SUVR Aβ PET analysis may be equivalently used to determine Aβ status for individuals with MCI participating in clinical trials, and both approaches add significant value for clinical course prognostication.

JAMA Neurol. doi:10.1001/jamaneurol.2015.1633
Published online August 17, 2015.

Supplemental content at jamaneurology.com

Author Affiliations: Helen Wills Neuroscience Institute, University of California, Berkeley (S. Schreiber, Landau, Fero, Jagust); Department of Neurology, Otto-von-Guericke University, Magdeburg, Germany (S. Schreiber); German Center for Neurodegenerative Diseases, Magdeburg (S. Schreiber); Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California (Landau, Fero, Jagust); Institute of Control Engineering, Technische Universität Braunschweig, Braunschweig, Germany (F. Schreiber).

Group Information: Members of the Alzheimer's Disease Neuroimaging Initiative are listed in eAppendix 1 in the Supplement.

Corresponding Author: Stefanie Schreiber, MD, Helen Wills Neuroscience Institute, 132 Barker Hall, Mail Code 3190, University of California, Berkeley, CA 94720 (stefanie.schreiber@med.ovgu.de).

Copyright 2015 American Medical Association. All rights reserved.

Downloaded From: http://archneur.jamanetwork.com/ by a UCSF LIBRARY User on 08/24/2015

Increased brain β-amyloid peptide (Aβ) seen with positron emission tomography (PET) and decreased Aβ1-42 measured in cerebrospinal fluid (CSF) allow in vivo detection of Aβ, with substantial agreement.[1-4] These biomarkers have therefore been proposed as indicators of Alzheimer disease (AD) neuropathology, aiding the selection and monitoring of individuals with mild cognitive impairment (MCI) due to AD or prodromal AD in clinical trials.[5,6] In MCI, Aβ load assessment may be additionally useful for prognostication since Aβ PET positivity predicts a higher risk of cognitive decline and AD conversion.[7-15] With the exception of a few studies of relatively small samples,[7,13,16] most MCI studies evaluating the prognostic value of Aβ PET[8,9,11,14,17] and the agreement between Aβ PET with Aβ CSF markers[2-4] have used semiquantitative image assessments of standardized uptake value ratios (SUVRs). Although visual Aβ PET rating is relatively simple and is the standard in clinical practice, there is a lack of knowledge about its significance for prognostication in large MCI cohorts and its agreement with CSF Aβ1-42 data and more quantitative PET measures. In terms of participant selection for clinical trials, further research is needed to evaluate whether MCI due to AD could be equivalently identified by visual PET ratings.

The goals of our study were to investigate the concordance between visual and quantitative Aβ PET analysis and evaluate how each of those image assessments agrees with CSF Aβ1-42 data in MCI. We further aimed to examine the effect of visual and quantitative image categorization as Aβ PET negative or positive to predict longitudinal cognitive function and AD conversion risk. Our methodologic design intentionally corresponds to the setting of large clinical trials including a large MCI sample derived from various sites or centers using different PET scanners and performing Aβ imaging with florbetapir F 18 [^{18}F], a tracer approved for clinical use by the US Food and Drug Administration and European Medicines Agency.

Methods

Participants

Our analysis was performed on participants in the Alzheimer's Disease Neuroimaging Initiative (ADNI), a multisite study supported by the National Institutes of Health, private pharmaceutical companies, and nonprofit organizations, with a goal of using multimodal imaging, CSF, and cognitive measurements in elderly controls as well as patients with MCI or AD to standardize and validate biomarkers in AD clinical trials. Additional methodologic information on participants, image acquisition, CSF, and data analysis, is provided in eAppendix 2 in the Supplement. All participants provided written informed consent. The institutional review board of each participating institution approved this study. Provision of financial compensation depended on the local policies of the individual study sites. The study was performed from September 21, 2010, to August 11, 2014; data analysis was conducted from September 21, 2014, to May 26, 2015.

The study included 401 ADNI participants with one baseline [^{18}F] florbetapir and one concurrent fludeoxyglucose F 18 (FDG)-PET scan who were categorized at baseline PET into 2 groups: early MCI (EMCI) or late MCI (LMCI). Participants were monitored for at least 12 months after the baseline scan, with the final follow-up occurring on August 11, 2014. All MCI cases were single-domain or multidomain amnestic, had a subjective memory problem, had a Mini-Mental State Examination score between 24 and 30, and had a Clinical Dementia Rating of 0.5.[18] Assignment to the EMCI or LMCI group was based on the individuals' educational level-adjusted scores on the Logical Memory II subscale (Delayed Paragraph Recall, paragraph A only) from the Wechsler Memory Scale-Revised.[19] Conversion to probable AD according to the National Institute of Neurological and Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders Association criteria[20] was determined at each center. Participants with AD conversion were censored at the first visit during which dementia was diagnosed; the remaining MCI cases were censored at their most recent follow-up. Longitudinal cognitive function was assessed every 6 to 12 months from the baseline [^{18}F] florbetapir scan.

[^{18}F] Florbetapir Analysis

Preprocessed florbetapir scans and coregistered structural magnetic resonance images were analyzed as described previously.[4,11] [^{18}F] Florbetapir SUVRs were created from a volume-weighted average of the mean [^{18}F] florbetapir uptake from cortical gray matter (lateral and medial frontal, anterior and posterior cingulate, lateral parietal, and lateral temporal) normalized to the cerebellum (white and gray matter).

[^{18}F] florbetapir scans were also rated by 2 neurologists: reader 1 (S.S.), an inexperienced scan reader, and reader 2 (W.J.J.), an experienced scan reader; both were blinded to all clinical and other imaging characteristics of each participant. Reader 1 rated the [^{18}F] florbetapir scans of all 401 participants with MCI. Reader 2 subsequently rated the scans of a subsample (n = 125) including (1) the images of 72 randomly chosen participants with MCI whose [^{18}F] florbetapir scan assessments were concordant between visual analysis (reader 1) and SUVRs and (2) all 53 participants whose [^{18}F] florbetapir scan assessments were discordant between visual analysis (reader 1) and SUVRs. The ratings of reader 1 were used for statistical analysis.

Additional Biomarkers

For statistical analysis, CSF Aβ1-42 data for 256 of all 401 MCI participants (63.8%) were available. Dichotomized apolipoprotein E (APOE) ε4 carrier status and dichotomized FDG data of all 401 participants with MCI were included as covariates in Cox proportional hazards regression models.

The threshold values were 1.11 for [^{18}F] florbetapir SUVR,[4,11] 192 pg/mL for CSF Aβ1-42,[21] and less than or equal to 1.21 for an abnormal FDG-PET scan.[22]

Statistical Analysis

Dichotomous variables were dummy coded as 0 if they were negative ([^{18}F] florbetapir, APOE ε4) or normal (FDG) and as 1 if they were positive or abnormal. Intermethod agreement between any 2 Aβ biomarkers was determined using Cohen κ, and

Copyright 2015 American Medical Association. All rights reserved.

Downloaded From: http://archneur.jamanetwork.com/ by a UCSF LIBRARY User on 08/24/2015

Table 1. Descriptive Statistics of the Participants' Demographic Data, Cognitive Function, and Biomarkers[a]

Characteristic	All MCI	EMCI	LMCI
No. of participants[b]	401	256	145
Age at [^{18}F] florbetapir scan, y	71.6 (7.5)	71.4 (7.4)	72 (7.9)
Male, No. (%)	219 (54.6)	146 (57.0)	73 (50.3)
Educational level, y	16.2 (2.7)	16 (2.7)	16.6 (2.5)
APOE ε4 carriers, No. (%)	198 (49.4)	114 (44.5)	84 (57.9)
MMSE closest to [^{18}F] florbetapir scan	28.1 (1.7)	28.3 (1.6)	27.6 (1.8)
ADAS-cog closest to [^{18}F] florbetapir scan, baseline	9.4 (4.4)	8.1 (3.5)	11.6 (4.9)
AVLT closest to [^{18}F] florbetapir scan, baseline	36.8 (10.8)	39 (10.4)	33 (10.6)
[^{18}F] florbetapir, SUVR	1.21 (0.22)	1.18 (0.2)	1.28 (0.24)
[^{18}F] florbetapir positive, No. (%), visual reads	196 (48.9)	104 (40.6)	92 (63.4)
[^{18}F] florbetapir positive, No. (%), SUVR measurements	221 (55.1)	123 (48.0)	98 (67.6)
[^{18}F] florbetapir positive among converters, No. (%) [total No. of converters], visual reads	54 (88.5) [61]	11 (78.6) [14]	43 (91.5) [47]
[^{18}F] florbetapir positive among converters, No. (%) [total No. of converters], SUVR measurements	53 (86.9) [61]	10 (71.4) [14]	43 (91.5) [47]
Mean FDG[c]	1.27 (0.13)	1.28 (0.12)	1.23 (0.14)
CSF Aβ1-42 level, pg/mL	170.81 (51.42)	176.48 (52.17)	163.04 (49.58)
CSF Aβ1-42 positive, No. (% of all cases with available CSF data) [total No. of EMCI and LMCI cases with available CSF data]	166 (64.8) [256]	88 (59.5) [148]	78 (72.2) [108]
CSF Aβ1-42 positive among converters, No. (% of all converters with available CSF data) [total No. of converters]	33 (89.2) [61]	6 (85.7) [7]	27 (90.0) [30]
Interval between [^{18}F] florbetapir and FDG scan, d	8.3 (11.3)	8.5 (11.9)	7.9 (10.2)
Interval between [^{18}F] florbetapir scan and LP, d	25.6 (100.9)	36.7 (129.6)	9.4 (13.1)
Interval between [^{18}F] florbetapir scan and ADAS-cog baseline, d	11.3 (15.2)	12.6 (16.6)	9.0 (12)
Interval between LP and ADAS-cog baseline, d	25.2 (103.3)	36.2 (132.8)	9.3 (12.2)
Follow-up time from [^{18}F] florbetapir scan on for non-converters and converters, y	1.6 (0.7)	1.8 (0.8)	1.2 (0.5)
Follow-up time from [^{18}F] florbetapir scan for nonconverters, y	1.6 (0.7)	1.8 (0.8)	1.2 (0.5)
Follow-up time from [^{18}F] florbetapir scan for converters, y	1.5 (0.5)	1.6 (1.0)	1.3 (0.7)

Abbreviations: Aβ, β-amyloid peptide; AD, Alzheimer disease; ADAS-cog, Alzheimer's Disease Assessment Scale–Cognitive Subscale; AVLT, Auditory Verbal Learning Test; CSF, cerebrospinal fluid; EMCI, early mild cognitive impairment (MCI); FDG, fludeoxyglucose F 18; [^{18}F] florbetapir, amyloid positron emission tomographic scan; LMCI, late MCI; LP, lumbar puncture; MMSE, Mini-Mental State Examination; SUVR, [^{18}F] florbetapir standardized uptake values ratio.

[a] Data are reported as mean (SD) unless otherwise indicated. Cognitive tests included ADAS-cog ranging from 0 to 70 with higher scores indicating worse cognitive function, and AVLT score ranging from 0 to 60 with higher scores indicating better preserved cognitive function.[23,24]

[b] Missing data: ADAS-cog baseline, 1 EMCI; CSF Aβ1-42, 108 EMCI, 37 LMCI.

[c] Mean FDG represents a composite measure generated from the mean of predefined meta regions of interests (right and left inferior temporal and lateral parietal regions, bilateral posterior cingulate-precuneus region) relative to the mean of a pons and cerebellar vermis reference region.[25]

Fleiss κ was applied to indicate intermethod agreement between all 3 Aβ biomarkers (visual, quantitative PET, and CSF) simultaneously.

Age-, sex-, and educational level-adjusted linear regression models were used to examine the main effect of a visual or quantitative positive [^{18}F] florbetapir baseline scan (dichotomous variable) on longitudinal cognitive function. Age-, sex-, and educational level-adjusted Cox proportional hazards regression models were examined to calculate the MCI conversion hazard ratio for a positive [^{18}F] florbetapir compared with a negative [^{18}F] florbetapir scan at baseline (performed separately for visual and SUVR analysis). Analysis was related to time to censoring. Additional models included APOEε4, FDG-PET scan data (dichotomous variable), or cognitive function as the baseline Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-cog) score (continuous variable).[23] In the presence of these covariates, the positive [^{18}F] florbetapir to negative [^{18}F] florbetapir scan hazard ratio refers to APOEε4 positivity, abnormal mean FDG value, and 1-unit per score baseline ADAS-cog increase.

All analyses were performed using SPSS, version 22.0 (SAS Institute Inc). Statistical significance was defined as $P \leq .05$.

Results

Participants

Descriptive statistics of the participants are given in **Table 1**. The overall conversion rate was 15.2% over a mean of 1.6 years; the rates for EMCI and LMCI were 5.5% and 32.4%, respectively. Nonconverters and converters differed on baseline cognition and several biomarkers (eAppendix 2 in the Supplement).

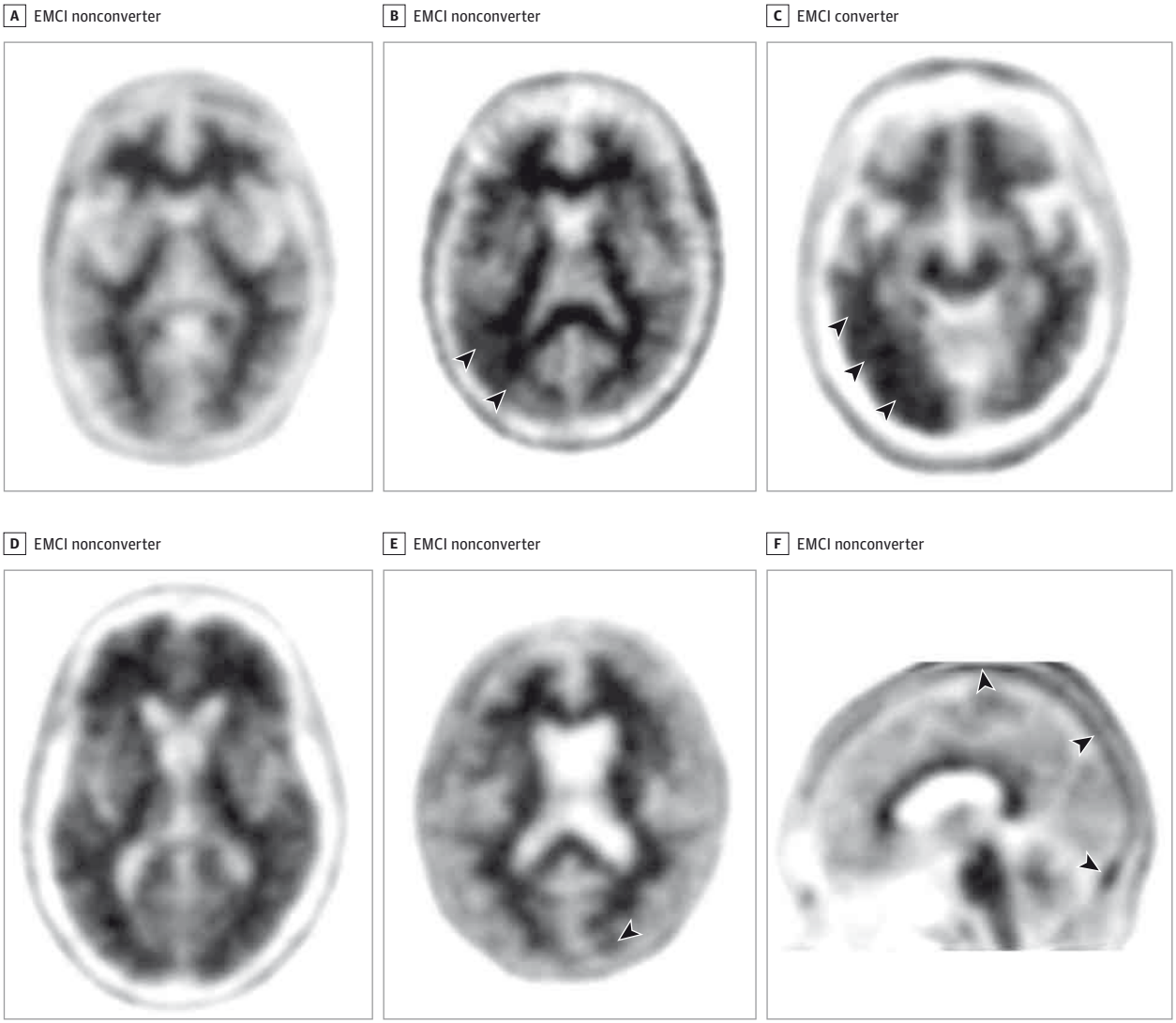
Agreement Between Aβ Biomarkers

At baseline, visual readings among all 401 participants were [^{18}F] florbetapir positive for 196 patients (48.9%; EMCI, 104 of 256 [40.6%]; and LMCI, 92 of 145 [63.4%]). The SUVR classifications were [^{18}F] florbetapir positive for 221 participants (55.1%; EMCI, 123 [48.0%]; and LMCI, 98 [67.6%]) (Table 1),

Copyright 2015 American Medical Association. All rights reserved.

Downloaded From: http://archneur.jamanetwork.com/ by a UCSF LIBRARY User on 08/24/2015

Figure 1. Grayscale Positron Emission Tomographic Images of Exemplary Participants With Mild Cognitive Impairment (MCI) With Discordant Results for Visual Florbetapir F 18 [^{18}F] Analysis and [^{18}F] Florbetapir Standardized Uptake Value Ratio (SUVR) Measurements



A, An early MCI (EMCI) nonconverter with a borderline SUVR value of 1.131 ([^{18}F] florbetapir positive) close to the threshold of 1.11 is demonstrated. Well-preserved contrast between white and gray matter with high nonspecific white matter florbetapir binding and absent cortical tracer uptake resulted in a visual [^{18}F] florbetapir-negative scan assessment among both readers; no cerebrospinal fluid (CSF) β-amyloid peptide (Aβ) values were available for this participant. B and C, Participants with EMCI who had high focal and asymmetric [^{18}F] florbetapir retention in the temporal cortex (B, arrowheads) and the temporooccipital cortex (C, arrowheads); both readers rated both scans as visual [^{18}F] florbetapir positive. The SUVR and CSF Aβ1-42 values for scan B were 1.097 ([^{18}F] florbetapir negative) and 164.5 pg/mL (CSF Aβ positive). For scan C, SUVR was 1.106 ([^{18}F] florbetapir negative) and CSF Aβ was 145.6 pg/mL (CSF Aβ positive). Participant B was a nonconverter; participant C converted to AD at age 72.7 years after a 12-month follow-up-period. D, An EMCI nonconverter's scan with slightly reduced contrast between white and gray matter; both readers rated the image as [^{18}F] florbetapir negative; the SUVR value of 1.242 indicated quantitative [^{18}F] florbetapir positivity. No CSF Aβ values were available for this participant. E and F, An EMCI nonconverter with noncortical [^{18}F] florbetapir uptake (arrowheads). Visual [^{18}F] florbetapir-negative scan assessment between both readers agreed with an Aβ-negative CSF value of 233.4 pg/mL; SUVR measurement of 1.117 was borderline [^{18}F] florbetapir positive.

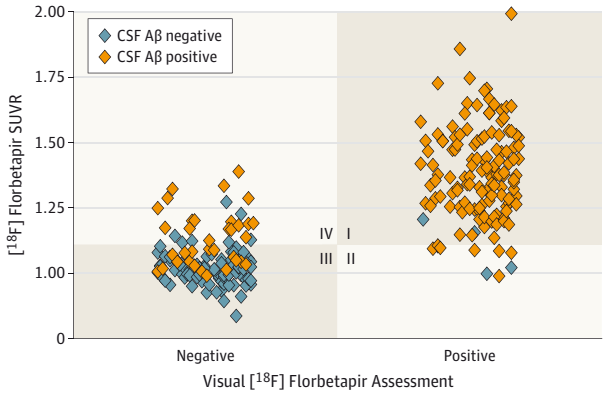
yielding substantial to very high intermethod agreement (κ = 0.74 for all patients: EMCI, κ = 0.66; LMCI, κ = 0.85). Discordance between visual and SUVR analysis occurred in 53 participants (13.2%; EMCI, 43 [16.8%]; and LMCI, 10 [6.9%]). [^{18}F] Florbetapir SUVR values of 23 of those 53 cases were within the ±5% CI of, and thus close to, the SUVR cutoff (1.11 [5% CI, 1.06-1.17][4]). [^{18}F] Florbetapir scans of exemplary discordant cases are demonstrated in **Figure 1**; the findings of a detailed visual inspection of all 53 discordant cases are given in eAppendix 2 in the Supplement. Demographic data, baseline cognition, and biomarkers did not differ significantly between participants with discordant and concordant visual and SUVR analysis (eAppendix 2 in the Supplement).

Compared with visual readings, SUVR tended to categorize more participants as [^{18}F] florbetapir positive and fewer cases as [^{18}F] florbetapir negative. Compared with visual read-

Copyright 2015 American Medical Association. All rights reserved.

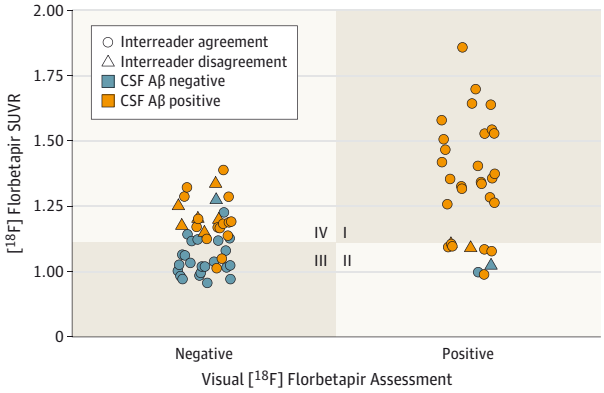
Downloaded From: http://archneur.jamanetwork.com/ by a UCSF LIBRARY User on 08/24/2015

Figure 2. Agreement Between Visual Interpretation, Quantitative Standardized Uptake Value Ratio (SUVR), and Cerebrospinal Fluid (CSF) β-Amyloid Peptide (Aβ) Measurement for 256 Participants With Early Mild Cognitive Impairment (EMCI) and Late MCI (LMCI)



Sector I and sector III demonstrate participants with concordant visual (reader 1) and quantitative florbetapir F 18 [^{18}F]–positive (sector I) and [^{18}F] florbetapir–negative (sector III) scan assessments (n = 221). Sector II and sector IV represent participants with discordant visual and quantitative florbetapir positron emission tomography (PET) data (n = 35). If visual and SUVR measures agreed, CSF Aβ data were highly concordant, particularly in case of scan assessment as [^{18}F] florbetapir positive (orange diamonds, sector I). If visual and quantitative PET analyses were discordant, agreement with CSF Aβ data was poor. When [^{18}F] florbetapir scans were assessed as positive, agreement between PET and CSF was very high (97% concordance for visual reads [orange diamonds in sectors I and II] and 94% concordance for SUVR measures [orange diamonds in sectors I and IV]). When scans were assessed as negative, concordance between PET and CSF was lower (71% concordance for visual reads [blue diamonds in sectors III and IV] and 76% concordance for SUVR measures [blue diamonds in sectors II and III]). The SUVR threshold of 1.11 (borders between sectors I and II and sectors III and IV) was derived from an independent sample.

Figure 3. Intermethod Agreement (Visual, Quantitative, Cerebrospinal Fluid [CSF]) and Interrater Agreement for a Subsample of Participants With Mild Cognitive Impairment (MCI) Visually Rated by Both Readers



The graph demonstrates all participants with available CSF β-amyloid peptide (Aβ) data (n = 77) of 125 cases rated by both readers. Sectors I and III illustrate individuals with MCI whose scan assessments were concordant between visual florbetapir F 18 [^{18}F] analysis of reader 1 and [^{18}F] florbetapir standardized uptake value ratio (SUVR) measurements. Sectors II and IV represent patients with MCI whose scan assessments were discordant between visual [^{18}F] florbetapir analysis of reader 1 and [^{18}F] florbetapir SUVR measurements. If visual and quantitative Aβ positron emission tomography evaluation agreed, interrater reliability was high (circles in sectors I and III) as was the concordance with CSF data (orange symbols in sector I, blue symbols in sector III). If visual and SUVR measures disagreed, concordance between both readers was only moderate (circles in sectors II and IV) and agreement with CSF data was very poor (orange symbols in sector II and blue symbols in sector IV for visual analysis, blue symbols in sector II and orange symbols in sector IV for SUVR analysis). The SUVR threshold of 1.11 (borders between sectors I and II and sectors III and IV) was derived from an independent sample.

ings, SUVR thus resulted in greater sensitivity (79% vs 85%) and less specificity (96% vs 90%), with CSF Aβ used here as the reference standard (eTable 1 in the Supplement).

Across all 256 participants with CSF Aβ, agreement between PET and CSF was substantial for visual [^{18}F] florbetapir readings (κ = 0.69 for all patients; EMCI, κ = 0.67; LMCI, κ = 0.71) (**Figure 2**; orange diamonds in sectors I and II, blue diamonds in sectors III and IV, with sectors divided at the SUVR threshold [1.11]) and SUVR measurements (κ = 0.72 for all MCI; EMCI, κ = 0.73; and LMCI, κ = 0.70) (Figure 2; orange diamonds in sectors I and IV, blue diamonds in sectors II and III). In 202 of those 256 cases (78.9%) all 3 methods agreed (Fleiss κ = 0.71 for all patients; EMCI, Fleiss κ = 0.68; LMCI, Fleiss κ = 0.75). Intermethod agreement between Aβ biomarkers remained substantial even if more conservative CSF and SUVR cutoffs were applied (eAppendix 2 in the Supplement).

When visual and SUVR analysis agreed, CSF was also highly concordant (κ = 0.82) (Figure 2; orange diamonds in sector I, blue diamonds in sector III); however, when visual and SUVR ratings disagreed (Figure 2; sectors II and IV), agreement with CSF was very poor (visual readings: κ = 0.05; SUVR, κ = 0.08). When [^{18}F] florbetapir scans were assessed as positive visually, agreement between PET and CSF was very high (Figure 2; orange diamonds in sectors I and II), and the same was true when [^{18}F] florbetapir scans were assessed as positive by SUVR (sectors I and IV). When the scans were assessed as negative, concordance between PET and CSF was lower (Figure 2; blue diamonds in sectors III and IV [visual readings] and sectors II and III [SUVR analysis]).

Figure 3 demonstrates relationships between visual readings, SUVRs, rater agreement, and CSF Aβ in the 77 cases that were read by 2 raters and in which CSF results were also available. When visual readings and SUVR agreed (42 [54.5%]), interrater agreement was very high (κ = 0.95) (Figure 3; circles in sectors I and III) as was the agreement between PET and CSF Aβ (κ = 0.90) (Figure 3; orange symbols in sector I, blue symbols in sector III). However, for the cases in which the visual reading and SUVR disagreed (35 [45.5%]), interrater agreement was only moderate (κ = 0.42) (Figure 3; circles in sectors II and IV), and agreement between CSF Aβ and visual analysis (κ = 0.05) (Figure 3; orange symbols in sector II, blue symbols in sector IV) and CSF Aβ and SUVRs (κ = −0.08) was very poor (Figure 3; orange symbols in sector IV, blue symbols in sector II).

Considering all 125 cases rated by both readers, interrater reliability was substantial (κ = 0.76). For those 53 participants with discordant visual (reader 1) and quantitative Aβ PET analysis, interreader agreement was only moderate (κ = 0.44), but it was very high for the remaining 72 individuals with concordant visual (reader 1) and SUVR florbetapir results (κ = 0.92).

Copyright 2015 American Medical Association. All rights reserved.

Downloaded From: http://archneur.jamanetwork.com/ by a UCSF LIBRARY User on 08/24/2015

Table 2. Cox Proportional Hazards Regression Models to Assess the Conversion HR for a [18F] Florbetapir-Positive vs a [18F] Florbetapir-Negative Baseline Scan

[18F] Florbetapir Scan Assessment (Dichotomous Predictor Variable)	Covariates	−2 Log Likelihood[a]	HR (95% CI)
Visual analysis	Age, sex, educational level	603.38	11.38 (5.10-25.39)[b]
	Age, sex, educational level, ADAS-cog baseline	549.14	5.67 (2.49-12.93)[b]
	Age, sex, educational level, APOEε4 status	634.31	8.71 (3.77-20.11)[b]
	Age, sex, educational level, FDG status[b]	578.09	8.02 (3.56-18.06)[b]
SUVR measurements	Age, sex, educational level	623.05	6.91 (3.26-14.67)[b]
	Age, sex, educational level, ADAS-cog baseline	559.11	3.63 (1.68-7.86)[b]
	Age, sex, educational level, APOEε4 status	614.75	4.88 (2.20-10.82)[b]
	Age, sex, educational level, FDG status[c]	593.43	4.87 (2.28-10.40)[b]

Abbreviations: ADAS-cog, Alzheimer's Disease Assessment Scale-Cognitive Subscale; APOE, apolipoprotein; FDG, fludeoxyglucose F 18; [18F] florbetapir, florbetapir F 18; HR, hazard ratio; SUVR, standardized uptake values ratio.

[a] The amount of unexplained variance in the model, with larger values indicating more unexplained data variance,

[b] $P < .001$.

[c] The FDG status was included as a dichotomous covariate and dummy-coded as 0 (normal status; mean FDG >1.21) and 1 (abnormal status; mean FDG ≤1.21).

Prediction of Longitudinal Cognitive Function and Conversion From MCI to AD

A positive [18F] florbetapir status assessed by visual and SUVR analysis was significantly associated with baseline and longitudinal cognitive function (eTable 2 in the Supplement). At baseline, 54 of all 61 converters (88.5%) were visual [18F] florbetapir positive, 53 (86.9%) were quantitative [18F] florbetapir positive, and 33 of all 37 converters (89.2%) with available CSF data were assessed as CSF Aβ positive (Table 1). Data on Aβ biomarker-negative converters are demonstrated in eAppendix 2 in the Supplement. Fifty percent of the visual and quantitative baseline [18F] florbetapir-positive cases converted to AD in approximately 2 years (eFigure in the Supplement). A visual [18F] florbetapir-positive baseline scan resulted in an 11.4-fold greater conversion hazard over a mean follow-up period of 1.6 years compared with a visual [18F] flor-betapir-negative baseline scan (**Table 2**). An approximately 6-fold to 9-fold greater conversion hazard for a visual [18F] flor-betapir-positive scan remained even after accounting for baseline ADAS-cog status and in the presence of an APOEε4 allele or an abnormal FDG-PET scan (Table 2). Although positive baseline SUVR data revealed a slightly lower conversion hazard than did the visual results, hazard ratios did not differ between quantitative and visual [18F] florbetapir PET analysis, as indicated by the ratios' 95% CI overlap (Table 2).

Discussion

We found substantial agreement between visual Aβ PET, quantitative Aβ PET, and CSF Aβ1-42 analysis in patients with MCI. In approximately 10% of all cases, agreement between visual and quantitative [18F] florbetapir analysis was poor, as was agreement between both readers and between PET and CSF data. Furthermore, both visual and quantitative Aβ PET assessment performed similarly in predicting longitudinal cognitive function and MCI to AD conversion.

Concordance between visual and quantitative image analysis was lower than that reported in a study[26] that included patients with AD. Similarly, acquisition of data in a single center, as opposed to this multicenter study, appears to result in a higher concordance between qualitative and SUVR measures (κ values up to 0.96).[13,26] In addition to its multicenter approach, our study differs from previous MCI investigations that compared visual and quantitative PET analysis by (1) the use of [18F] florbetapir images instead of Pittsburgh compound-B PET scans,[26] flutemetamol F 18 PET scans,[27] or florbetapen F 18 PET scans[13]; (2) the inclusion of individuals with EMCI; (3) the application of a less conservative SUVR threshold[13]; and (4) the examination of a much larger and likely more heterogeneous MCI sample.[7,13,26,27] All of those aspects may have contributed to the somewhat lower concordance between visual readings and SUVRs in the present study.

In contrast, our interrater reliability (κ = 0.92) was somewhat higher than that in another visual MCI PET study (κ = 0.46-0.86).[7] However, this very high interreader reliability was generated from a subsample of cases that already agreed between visual (reader 1) and SUVR analysis; therefore, it is likely to be biased toward higher concordance. When including all 125 cases rated by both readers (as described in the Methods section), interrater reliability was somewhat lower but still substantial (κ = 0.76). Because the subsample of 125 cases read by rater 2 specifically included the 53 participants with discordant qualitative (reader 1) and quantitative data, corresponding interrater agreement was biased toward lower concordance.

To have an objective, independent criterion, we considered CSF as a reference standard and found that visual readings resulted in lower sensitivity and higher specificity compared with SUVR analysis. The use of CSF as an external reference allowed us to compare visual and SUVR analysis with a third measure. Addition of this measure does not suggest that CSF Aβ should be considered as the reference standard or that our approach refers to general recommendations, especially since CSF and PET Aβ represent different aspects of cerebral Aβ pathology.[28,29] In fact, it has been proposed that CSF Aβ reduction indicates earlier stages of abnormality and increased Aβ PET retention reflects later stages of cerebral Aβ and AD pathology.[30]

In terms of its high specificity but slightly lower sensitivity, visual analysis may be useful for selection of individuals with MCI for participation in intervention trials aiming to avoid the treatment of true-negatives. Compared with SUVRs, visual readings may be less useful for clinical trials aiming to capture as many Aβ-positive cases as possible, especially those at

Copyright 2015 American Medical Association. All rights reserved.

Downloaded From: http://archneur.jamanetwork.com/ by a UCSF LIBRARY User on 08/24/2015

very early Aβ stages (eg, EMCI). Continuous SUVR measures may be superior to quantify treatment effects on Aβ, especially if therapeutic impact is moderate. Indeed, current and planned preclinical trials are designed to select and provide treatment for cognitively unimpaired individuals or patients with early cognitive impairment and Aβ biomarker evidence to show that therapy lessens Aβ burden and provides clinical benefit.[31,32] Thus, an approach combining visual and quantitative PET analysis may be best for selection of trial participants.[33] This idea is supported by the fact that visual and quantitative PET agreement was associated with very high overall intermethod concordance. The combined application of visual and quantitative PET analysis may thus be a valid approach to identify true Aβ-positive and true Aβ-negative cases. In a separate autopsy-validated study,[34] there was 100% agreement between visual and quantitative Aβ PET data for cases with concordance between visual [^{18}F] florbetapir and neuropathologic Aβ load classification, which supports this idea.

Classification of Aβ positivity or negativity was inconsistent for approximately 10% of all cases. Most (approximately 81%) of these cases were in the EMCI category. As has also been reported for elderly control participants,[33] focal and asymmetric [^{18}F] florbetapir retention explained some of the discrepancy between qualitative and quantitative analysis since it basically leads to visual Aβ positivity but quantitative Aβ negativity. Cerebrospinal fluid Aβ was positive in 85% of the visual [^{18}F] florbetapir-positive cases and quantitative [^{18}F] florbetapir-negative cases with focal Aβ burden, suggesting that focal cortical [^{18}F] florbetapir retention may account for some of the discordance between CSF and SUVR results.[3,4,30,35] Qualitative analysis was also superior in detecting [^{18}F] florbetapir retention in nonparenchymal brain structures, which contributed to its lower false-positive rate. Indeed, visual readings require significant efforts to maintain interrater reliability. Our data, however, show that interreader agreement can be substantial, even if visual analysis includes a fraction of challenging scans.

Our overall conversion rate was 15% within 1.6 years. With the exception of one study[7] reporting a similar conversion frequency (16% in 1.5 years), most MCI studies[8,10,13,36] comprising a mean observation period of approximately 2 years reported higher conversion rates (44%-59%). Higher conversion rates may be explained by the inclusion of MCI cohorts from specialized memory clinics[10,36] that are characterized by higher cortical Aβ retention,[10,13,36] lower baseline Mini-Mental State Examination scores,[8,36] and more APOE ε4 carriers.[8,10,36] Furthermore, the proportion of converters among our quantitative [^{18}F] florbetapir-positive cases (53 of 221 [24.0%]) was substantially lower than the proportions of 50% to 82% reported in other Aβ PET MCI studies,[2,8,9,12,13,17,36-38] which is explained by the inclusion of participants with EMCI and by our less conservative SUVR cutoff assessing comparably more individuals with MCI as Aβ positive. However, the proportion of converters among our qualitative [^{18}F] florbetapir-positive cases (54 of 196 [27.6%]) is in line with the proportions of 29% to 35% found in a previous MCI PET study[7,16] assessing [^{18}F] florbetapir scans visually. Our finding that approximately 90% of the converters were Aβ positive for both visual and SUVR analysis is in agreement with the commonly reported frequencies of Aβ positivity among converters[2,8,12-14,17,36,37] and supports the increased AD conversion risk in case of a [^{18}F] florbetapir-positive scan.

Amyloid-positive baseline PET predicted approximately 4-fold to 9-fold higher conversion risk even after adjustment for positive APOE ε4 carrier status, an abnormal concurrent FDG-PET scan, and 1-score worsening of a lower baseline ADAS-cog score. An Aβ-positive PET scan thus adds considerable predictive value even in the presence of genetic and cognitive status, as well as in the absence of additional biomarkers. Moreover, our data emphasize that the frequency of Aβ positivity and conversion rates increase with severity of cognitive symptoms.[11] Depending on the method used for Aβ load assessment, we found that 40.6% to 59.5% and 63.4% to 72.2% of EMCI and LMCI cases, respectively, were Aβ positive and that 5.5% and 32.4% of participants with EMCI and LMCI, respectively, converted to AD within 1.5 years. In general, all results were comparable between the qualitative and quantitative analyses.

Conclusions

Our findings suggest that visual readings and SUVRs are equivalent in their assignment of negative or positive Aβ status in MCI. Even including genetic, cognitive, and FDG status, qualitative and quantitative PET analysis adds significant value for clinical course prognostication. Our data thus support the applicability of a simpler case inclusion algorithm that may facilitate case selection in trials evaluating MCI due to AD.

ARTICLE INFORMATION

Accepted for Publication: June 3, 2015.

Published Online: August 17, 2015. doi:10.1001/jamaneurol.2015.1633.

Author Contributions: Dr S. Schreiber had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: S. Schreiber, Landau, Jagust.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: S. Schreiber, Jagust.
Critical revision of the manuscript for important intellectual content: Landau, Fero, F. Schreiber, Jagust.
Statistical analysis: S. Schreiber, Landau, F. Schreiber.
Obtained funding: Jagust.
Administrative, technical, or material support: Landau, Fero, F. Schreiber.
Study supervision: Jagust.

Conflict of Interest Disclosures: Dr Landau has served as a consultant to Biogen, Genentech, and Synarc, and Dr Jagust has served as a consultant to Banner Alzheimer Institute/Genentech, Novartis, and Synarc/Bioclinica. No other disclosures were reported.

Funding/Support: 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 Department of Defense ADNI (award number W81XWH -12-2-0012). ADNI is funded by the National Institute on Aging and the National Institute of Biomedical Imaging and Bioengineering and through generous contributions from the following organizations: Alzheimer's Association, Alzheimer's Drug Discovery Foundation, Araclon Biotech, BioClinica Inc, Biogen Idec Inc, Bristol-Myers Squibb Company, Eisai Inc, Elan Pharmaceuticals Inc, Eli Lilly and Company, EuroImmun, F. Hoffmann-La Roche Ltd and its affiliated company Genentech Inc, Fujirebio, GE Healthcare, IXICO Ltd, Janssen Alzheimer Immunotherapy Research & Development LLC, Johnson & Johnson Pharmaceutical Research & Development LLC, Medpace Inc, Merck & Co Inc, Meso Scale Diagnostics LLC, NeuroRx Research, Neurotrack

Copyright 2015 American Medical Association. All rights reserved.

Downloaded From: http://archneur.jamanetwork.com/ by a UCSF LIBRARY User on 08/24/2015

Technologies, Novartis Pharmaceuticals Corporation, Pfizer Inc, Piramal Imaging, Servier, Synarc Inc, and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (http://www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. This research was also supported by the German Research Foundation grant SCHR 1418/3-1.

Role of the Funder/Sponsor: The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Group Information: A complete listing of the Alzheimer's Disease Neuroimaging Initiative investigators can be found in eAppendix 1 in the Supplement.

Additional Information: Data used in preparation of this article were obtained from the ADNI database (http://adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report.

REFERENCES

1. Fagan AM, Mintun MA, Mach RH, et al. Inverse relation between in vivo amyloid imaging load and cerebrospinal fluid Aβ42 in humans. *Ann Neurol*. 2006;59(3):512-519.

2. Forsberg A, Engler H, Almkvist O, et al. PET imaging of amyloid deposition in patients with mild cognitive impairment. *Neurobiol Aging*. 2008;29(10):1456-1465.

3. Jagust WJ, Landau SM, Shaw LM, et al; Alzheimer's Disease Neuroimaging Initiative. Relationships between biomarkers in aging and dementia. *Neurology*. 2009;73(15):1193-1199.

4. Landau SM, Lu M, Joshi AD, et al; Alzheimer's Disease Neuroimaging Initiative. Comparing positron emission tomography imaging and cerebrospinal fluid measurements of β-amyloid. *Ann Neurol*. 2013;74(6):826-836.

5. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):270-279.

6. Dubois B, Feldman HH, Jacova C, et al. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol*. 2014;13(6):614-629.

7. Doraiswamy PM, Sperling RA, Coleman RE, et al; AV45-A11 Study Group. Amyloid-β assessed by florbetapir F 18 PET and 18-month cognitive decline: a multicenter study. *Neurology*. 2012;79(16):1636-1644.

8. Ewers M, Insel P, Jagust WJ, et al; Alzheimer's Disease Neuroimaging Initiative (ADNI). CSF biomarker and PIB-PET–derived β-amyloid signature predicts metabolic, gray matter, and cognitive changes in nondemented subjects. *Cereb Cortex*. 2012;22(9):1993-2004.

9. Jack CR Jr, Wiste HJ, Vemuri P, et al; Alzheimer's Disease Neuroimaging Initiative. Brain β-amyloid measures and magnetic resonance imaging atrophy both predict time-to-progression from mild cognitive impairment to Alzheimer's disease. *Brain*. 2010;133(11):3336-3348.

10. Koivunen J, Scheinin N, Virta JR, et al. Amyloid PET imaging in patients with mild cognitive impairment: a 2-year follow-up study. *Neurology*. 2011;76(12):1085-1090.

11. Landau SM, Mintun MA, Joshi AD, et al; Alzheimer's Disease Neuroimaging Initiative. Amyloid deposition, hypometabolism, and longitudinal cognitive decline. *Ann Neurol*. 2012;72(4):578-586.

12. Okello A, Koivunen J, Edison P, et al. Conversion of amyloid positive and negative MCI to AD over 3 years: an ^{11}C-PIB PET study. *Neurology*. 2009;73(10):754-760.

13. Ong KT, Villemagne VL, Bahar-Fuchs A, et al. αβ imaging with ^{18}F-florbetaben in prodromal Alzheimer's disease: a prospective outcome study. *J Neurol Neurosurg Psychiatry*. 2015;86(4):431-436.

14. Rowe CC, Ellis KA, Rimajova M, et al. Amyloid imaging results from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging. *Neurobiol Aging*. 2010;31(8):1275-1283.

15. Wolk DA, Price JC, Saxton JA, et al. Amyloid imaging in mild cognitive impairment subtypes. *Ann Neurol*. 2009;65(5):557-568.

16. Doraiswamy PM, Sperling RA, Johnson K, et al; AV45-A11 Study Group. Florbetapir F 18 amyloid PET and 36-month cognitive decline: a prospective multicenter study. *Mol Psychiatry*. 2014;19(9):1044-1051.

17. Nordberg A, Carter SF, Rinne J, et al. A European multicentre PET study of fibrillar amyloid in Alzheimer's disease. *Eur J Nucl Med Mol Imaging*. 2013;40(1):104-114.

18. Petersen R. Conceptual overview. In: Petersen RC, ed. *Mild Cognitive Impairment: Aging to Alzheimer's Disease*. New York, NY: Oxford University Press; 2003:1-14.

19. Wechsler D. *Wechsler Memory Scale—Revised Manual*. San Antonio, TX: Psychological Corp; 1987.

20. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984;34(7):939-944.

21. Shaw LM, Vanderstichele H, Knapik-Czajka M, et al; Alzheimer's Disease Neuroimaging Initiative. Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Ann Neurol*. 2009;65(4):403-413.

22. Landau SM, Harvey D, Madison CM, et al; Alzheimer's Disease Neuroimaging Initiative. Comparing predictors of conversion and decline in mild cognitive impairment. *Neurology*. 2010;75(3):230-238.

23. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry*. 1984;141(11):1356-1364.

24. Rey A. *L'examen Clinique en Psychologie*. Paris: Presses Universitaires de France; 1964.

25. Landau SM, Harvey D, Madison CM, et al; Alzheimer's Disease Neuroimaging Initiative. Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI. *Neurobiol Aging*. 2011;32(7):1207-1218.

26. Zwan MD, Ossenkoppele R, Tolboom N, et al. Comparison of simplified parametric methods for visual interpretation of ^{11}C-Pittsburgh compound-B PET images. *J Nucl Med*. 2014;55(8):1305-1307.

27. Thurfjell L, Lilja J, Lundqvist R, et al. Automated quantification of ^{18}F-flutemetamol PET activity for categorizing scans as negative or positive for brain amyloid: concordance with visual image reads. *J Nucl Med*. 2014;55(10):1623-1628.

28. Fagan AM, Mintun MA, Shah AR, et al. Cerebrospinal fluid tau and ptau$_{181}$ increase with cortical amyloid deposition in cognitively normal individuals: implications for future clinical trials of Alzheimer's disease. *EMBO Mol Med*. 2009;1(8-9):371-380.

29. Schöll M, Wall A, Thordardottir S, et al. Low PiB PET retention in presence of pathologic CSF biomarkers in Arctic APP mutation carriers. *Neurology*. 2012;79(3):229-236.

30. Mattsson N, Insel PS, Donohue M, et al; Alzheimer's Disease Neuroimaging Initiative. Independent information from cerebrospinal fluid amyloid-β and florbetapir imaging in Alzheimer's disease. *Brain*. 2015;138(pt 3):772-783.

31. Kozauer N, Katz R. Regulatory innovation and drug development for early-stage Alzheimer's disease. *N Engl J Med*. 2013;368(13):1169-1171.

32. Langbaum JB, Fleisher AS, Chen K, et al. Ushering in the study and treatment of preclinical Alzheimer disease. *Nat Rev Neurol*. 2013;9(7):371-381.

33. Cohen AD, Mowrey W, Weissfeld LA, et al. Classification of amyloid-positivity in controls: comparison of visual read and quantitative approaches. *Neuroimage*. 2013;71:207-215.

34. Clark CM, Pontecorvo MJ, Beach TG, et al; AV-45-A16 Study Group. Cerebral PET with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid-β plaques: a prospective cohort study. *Lancet Neurol*. 2012;11(8):669-678.

35. Zwan M, van Harten A, Ossenkoppele R, et al. Concordance between cerebrospinal fluid biomarkers and [^{11}C]PIB PET in a memory clinic cohort. *J Alzheimers Dis*. 2014;41(3):801-807.

36. Villemagne VL, Pike KE, Chételat G, et al. Longitudinal assessment of Aβ and cognition in aging and Alzheimer disease. *Ann Neurol*. 2011;69(1):181-192.

37. Grimmer T, Wutz C, Drzezga A, et al. The usefulness of amyloid imaging in predicting the clinical outcome after two years in subjects with mild cognitive impairment. *Curr Alzheimer Res*. 2013;10(1):82-85.

38. Rowe CC, Bourgeat P, Ellis KA, et al. Predicting Alzheimer disease with β-amyloid imaging: results from the Australian imaging, biomarkers, and lifestyle study of ageing. *Ann Neurol*. 2013;74(6):905-913.

Copyright 2015 American Medical Association. All rights reserved.

Downloaded From: http://archneur.jamanetwork.com/ by a UCSF LIBRARY User on 08/24/2015