Published in final edited form as:
Alzheimers Dement. 2013 November ; 9(6): 677–686. doi:10.1016/j.jalz.2012.09.016.

Diagnostic accuracy of markers for prodromal Alzheimer's disease in independent clinical series

A Prestia, PsyD[(a)], **A Caroli, PhD**[(a),(b)], **K Herholz, MD**[(c)], **E Reiman, MD**[(d)], **K Chen, PhD**[(d)], **WJ Jagust, MD**[(e)], **GB Frisoni, MD**[(a),*], and **Translational Outpatient Memory Clinic – TOMC - Working Group**[#] **for the Alzheimer's Disease Neuroimaging Initiative**[##]

[(a)]LENITEM – Laboratory of Epidemiology Neuroimaging and Telemedicine, and Unit for the Clinical Translation of Research, IRCCS Centro San Giovanni di Dio FBF, Brescia, Italy

[(b)]Medical Imaging Unit, Biomedical Engineering Department, Mario Negri Institute for pharmacological research, Bergamo, Italy

[(c)]University of Manchester, UK

[(d)]Banner Alzheimer's Institute, Phoenix (AZ), USA

[(e)]Helen Wills Neuroscience Institute, University of California, Berkeley (CA), USA

Abstract

Background—Different biomarkers could help to diagnose Alzheimer's dementia (AD) since the earliest stages. We capitalized on data from different clinical series to compare sensitivity and specificity of individual biomarkers for predicting mild cognitive impairment (MCI) progression to AD

Methods—Medial temporal atrophy, cortical hypometabolism, and cerebrospinal fluid (CSF) biomarkers were assessed in 18 MCI patients with prodromal AD (pAD, conversion time = 26±12 months) and 18 stable MCI patients (sMCI) from TOMC cohort, as well as in 24 pAD patients (conversion time = 36±12 months) and 33 sMCI patients from ADNI cohort. Medial temporal atrophy was measured by manual, semi automated and automated hippocampal volumetry, cortical hypometabolism was measured using several indices of AD-related hypometabolism pattern, and

[*]**Corresponding Author:** Giovanni B Frisoni, Centro San Giovanni di Dio FBF, The National Centre for Research and Care of Alzheimer's and Mental Diseases, via Pilastroni 4, 25125 Brescia, Italy - Telephone +39 030 3501361, gfrisoni@fatebenefratelli.it.

[#]Data used in this article were collected by the Traslational Outpatient Memory Clinic – TOMC – working group. G Amicucci, S Archetti, L Benussi, G Binetti, L Bocchio-Chiavetto, M Bonetti, E Canu, F Caobelli, E Cavedo, E Chittò, D Costardi, M Cotelli, M Gennarelli, S Galluzzi, C Geroldi, R Ghidoni, R Giubbini, UP Guerra, G Kuffenschin, G Lussignoli, D Moretti, A Orlandini, B Paghera, M Parapini, D Paternicò, C Porteri, M Romano, S Rosini, C Scarpazza, I Villa, R Zanardini, O Zanetti.

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

Disclosure: All authors report no conflict of interests.

CSF markers were Aβ42 and total tau protein concentrations. For each biomarker, sensitivity for prodromal AD, specificity for sMCI, and diagnostic accuracy were computed.

Results—Sensitivity to predict MCI conversion to AD in ADNI and TOMC cohorts was 79% and 94% based on Aβ42, 46% and 28% based on hippocampal volumes, 33 to 66% and 56 to 78% based on different hypometabolism indices, and 46% and 61% based on total tau levels, respectively. Specificity to exclude sMCI was 27% and 50% based on Aβ42, 76% and 94% based on hippocampal volumes, 58 to 67% and 55 to 83% based on different hypometabolism indices, and 61% and 83% based on total tau levels.

Conclusion—Current findings suggest that Aβ42 concentrations and hippocampal volumes may be used in combination to best identify prodromal AD.

Keywords

Alzheimer's disease; MCI; MRI; PET; Diagnostic test assessment; Diagnostic accuracy

1. Background

A deeper understanding of the pathophysiology of Alzheimer's disease (AD) and the availability of its *in vivo* biomarkers has led to a substantial revision of diagnostic criteria for AD (1), in order to capture the full spectrum of the disease since its earliest stages (2). The revised NIA-AA criteria (3-5) go beyond the view of AD as a clinical-pathological entity (6), suggesting that confidence in diagnosing AD could be improved by structural and biological evidence of Alzheimer's pathology (5). At the dementia stage, biomarkers might improve the differential diagnosis of AD and other neurodegenerative diseases; at the mild cognitive impairment (MCI) stage, they might allow to differentiate MCI due to AD from MCI due to other forms of memory impairment (5); and at the asymptomatic stage they might allow to capture healthy persons at risk of developing AD.

New biomarkers of AD are indicative of: i) brain amyloidosis, namely abnormal tracer retention on amyloid PET imaging and low concentration of Aβ42 protein in the cerebrospinal fluid (CSF), and ii) neuronal degeneration, namely elevated CSF concentration of tau protein, decreased cortical fluorodeoxyglucose uptake on PET and medial temporal atrophy on structural magnetic resonance (MRI). In this paper we focus on biomarkers available in our memory clinic (i.e. all but amyloid imaging). Moreover, since the way a biomarker is measured can make a difference in diagnostic accuracy (6-7), we assessed cortical glucose metabolism using three different metrics, and hippocampal volume using manual, semi-automated, and fully automated methods.

Aim of this study is to compare the sensitivity and specificity of individual AD biomarkers to predict progression from MCI to AD in 2 different cohorts.

2. Methods

2.1 Subjects

Patients come from two independent datasets: the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Translational Outpatient Memory Clinic (TOMC). ADNI was

launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public-private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials. The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California – San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 adults, aged 55 to 90, to participate in the research, approximately 200 cognitively normal older individuals to be followed for 3 years, 400 people with MCI to be followed for 3 years and 200 people with early AD to be followed for 2 years. For up-to-date information, see www.adni-info.org. At baseline, all subjects receive a comprehensive neuropsychological evaluation, blood drawing (for APOE genotyping) and structural MR. Subsets of subjects undergo lumbar puncture (for CSF sampling) or PIB-PET, and half of the subjects undergo FDG-PET. All subjects undergo yearly follow-up visits. Moreover, MCI patients are visited every 6-month (follow-up time = 36±12 months [12 to 48 months]) in order to assess conversion to dementia.

The Translational Outpatient Memory Clinic (TOMC) of the Scientific Institute for the Research and Care of Alzheimer's Disease, IRCCS Centro San Giovanni di Dio Fatebenefratelli in Brescia - Italy, instituted in 2006, is a multidisciplinary team of neurologists, geriatricians, neuropsychologists, neuroscientists expert in image analysis, biologists, neurophysiologists, and geneticists. All patients are self-referred or referred by general practitioners or specialists, complaining memory or other cognitive disturbance unaccountable by focal cerebral, physical, psychiatric, or metabolic diseases. Routinely, they undergo blood drawing (for APOE genotyping), clinical and cognitive assessment as well as high resolution MR scan, 18 F-FDG PET, and lumbar puncture whenever accepted by the patients and caregivers or not contraindicated. Diagnoses are based on clinical criteria (NINCDS-ADRDA criteria (8) for AD and Petersen criteria (9) for MCI), without taking potential positivity to biomarkers into account. MCI patients undergo follow-up visits every 12 months (follow-up time = 26±12 months [12 to 36 months]) until development of probable AD (10). Though baseline biomarker results are at clinician's disposal, progression to AD is ascertained based on clinical criteria (worsening in cognitive and behavioural function).

2.1.1 Patients—Patients included in the study are all MCI patients with prodromal AD taken from ADNI and TOMC databases, with available baseline structural MR, FDG-PET and CSF sampling: ADNI prodromal AD patients (pAD, mean age 75±8 years) (n = 34); TOMC prodromal AD patients (n = 18, mean age 71±7 years). Prodromal AD was defined

as a diagnosis of MCI at baseline followed by conversion to AD during follow-up period; MCI patients who converted to non-AD dementia were excluded from the study.

2.1.2 Control group—The control group, used in statistical analyses to assess specificity and diagnostic accuracy of each biomarker, included 33 stable MCI (sMCI, mean age 75±8 years) from ADNI and 18 sMCI (mean age 72±8 years) from TOMC database.

2.2 Biomarkers of AD neuropathology

2.2.1 FDG-PET metrics—18F-FDG PET imaging was performed at the Nuclear Medicine Service, Spedali Civili of Brescia, Brescia, Italy (TOMC dataset) and at any of the ADNI PET sites in North America (ADNI dataset). TOMC FDG-PET images were acquired with a 24 rings General Electric 3D PET/CT device (Discovery ST PET with Light Speed CT), isotropic resolution of 5.99 mm; 15.7 cm axial FOV; 70 cm transaxial FOV. Attenuation correction was performed using the CTAC method. Images were reconstructed using the FORE-Iterative algorithm (48 subsets, 5 iterations) with xy and z filter (cut off of 4 mm), yielding 128×128 matrix with a pixel size of 1.95 mm (11). ADNI FDG-PET images were acquired using 17 different scanner models from three vendors (Philips, Siemens and General Electric). PET sites were initially approved for PET scanning by performing a pair of phantom scans on the 3-D Hoffman brain phantom following a protocol that matched the acquisition and reconstruction parameters to be used for human phase of the ADNI project. Quality control of phantom scans helped develop a method for assessing and correcting for differences in PET images across sites (12). In both TOMC and ADNI dataset, each FDG-PET image underwent a stringent quality control procedure to assess image quality.

FDG-PET indices of AD-related hypometabolism were PALZ (13), hypometabolic convergence index (HCI) (14), and metaROI average (15) (7). All metrics are based on voxel-by-voxel analysis of FDG-PET images and provide a single measure of AD-related hypometabolism, however, they are computed using different processing procedures.

PALZ score (13) combines voxel-based parametric mapping with the diagnostic information on brain regions typically affected in AD. Each FDG-PET image is compared to a fixed database of normal elderly scans through a voxel-wise t-test, including age as confounding variable, and the PALZ score is computed as voxel-by-voxel sum of t-scores in a pre-defined AD-pattern mask.

The AD-related hypometabolic convergence index (HCI) (14) is generated using fully automated algorithms based on SPM: each individual PET is compared to those of healthy subjects from a pre-defined normative database through a voxel-wise t-test, and HCI is calculated as the inner-product of the resulting individual t-score map (converted to a z-score map) and a pre-defined AD Z-score map.

MetaROI average is computed on spatially and intensity normalized PET images as average of the mean counts in 5 metaROI volumes, originally computed based on a metanalysis of studies carrying out direct whole-brain contrasts of FDG-PET data and reporting Z-scores or T-values in voxels showing significantly different FDG uptake between patients (AD or MCI) and controls (15). As metaROI does not account for age, pertinent age-corrected

scores, hereafter named as W-scores, were computed, following a previously published procedure (16).

2.2.2 Hippocampal volume—TOMC subjects underwent brain T1- weighted MRI in the Neuroradiology Department of the Città di Brescia Hospital, Brescia, Italy, while ADNI subjects underwent it in any of the ADNI MRI sites in North America, as previously described ((17) and (18)). Hippocampal volumes were manually traced (TOMC dataset) or semi-automatically computed (ADNI dataset); in addition, for all subjects, hippocampal volumes were automatically computed.

TOMC MR images were first corrected for magnetic field non-uniformities, intensity normalized and brain-to-brain linearly registered to a standard template in the stereotaxic space (ICBM152) based on the Talairach atlas (19). In case of failure of automated registration, manual registration was performed, based on eleven anatomical landmarks. Left and right hippocampal volumes were manually segmented on the reoriented and normalized TOMC images by a single tracer. The hippocampal boundaries were outlined using Display (http://packages.bic.mni.mcgill.ca/tgz/, Montreal Neurological Institute, McGill University, Montreal). Manual tracing was performed on approximately 30-40 contiguous coronal slices, following a standardized and validated protocol (19). TOMC hippocampal volumes were finally normalized to the manually traced total intracranial volume (TIV), and resulting volumes were retained for statistical analyses.

On all ADNI MR images, left and right hippocampal volumes were semi automatically computed at University of California, San Francisco, using a previously validated (20) commercially available brain mapping tool (Medtronic Surgical Navigation Technologies (SNT), Louisville, CO) based on fluid image transformation (21). The software requires to manually place 2 global landmarks on AC and PC location, and 44 local landmarks surrounding the left and right hippocampus; once scans are fully landmarked, they are processed by algorithms which produce hippocampal boundaries and volumes; boundaries are checked by qualified reviewers and in case of failure can be manually edited. ADNI hippocampal volumes were finally normalized to the estimated intracranial volume (eTIV, available in the ADNI dataset and generated by automatically segmenting the baseline MR image), and resulting volumes were retained for statistical analyses.

Automatically segmented left and right hippocampal volumes were further computed on both ADNI and TOMC MR images using a previously well described, validated, commercially available software (Freesurfer http://surfer.nmr.mgh.harvard.edu/) (22). Obtained hippocampal volumes were normalized to the estimated intracranial volume (eTIV), and resulting volumes were retained for statistical analyses.

As no hippocampal volume processing procedures account for age, pertinent age-corrected scores, hereafter named as W-scores, were computed, following a previously published method (16). Individual left and right hippocampal W-scores were finally averaged to have a single hippocampal volume measure.

2.2.3 CSF biomarkers—CSF biomarkers were Aβ42 and total tau protein concentrations. CSF was obtained by lumbar puncture performed with a 20- or 24-gauge spinal needle between L4 and L5 or L3 and L4 and collected in polypropylene tubes. TOMC CSF samples were maintained at +4°C and centrifuged at 2000g for 5 minutes, and then aliquoted and stored at −80°C until assay (10); ADNI CSF samples were thawed for 1 hour, gentle mixed, aliquoted and frozen on dry ice at −80°C. Aβ42 and total tau protein concentrations were determined by a commercially available enzyme-linked immunosorbent assay (ELISA, Innogenetics, Ghent, Belgium) (10) (TOMC samples) and by xMAP Luminex platform (Luminex Corp, Austin TX) with Innogenetics (INNOBIA AlzBio3, Ghent, Belgium) immunoassay kit–based reagents (23).

2.3 Biomarker cut-offs

For each biomarker, a threshold of abnormality was considered; in case previously published cut-offs were not available, they were computed for the purpose of this study, and defined as the threshold associated with 95% specificity in an independent normal control dataset.

PET biomarkers cut-offs were specifically computed based on their performance in correctly identifying 148 normal elders (all subjects included in a previously described reference FDG-PET normative dataset (7)), imposing 95% level of specificity. Despite a threshold for abnormality already being available for PALZ, its cut-off was newly determined on the same normative dataset, in order to maximize the reliability of the comparison among different PET indices performance.

Hippocampal volume cut-offs were specifically computed based on hippocampal volume performance in correctly identifying a group of 287 cognitively healthy elders taken from a reference normative database (24) (manual segmentation) or a group of 66 ADNI cognitively healthy elders (both semi-automated and automated procedures).

For both CSF total tau and Aβ42 concentrations, previously published cut-offs were adopted; ADNI-related cut-offs were based on CSF biomarkers performance in correctly identifying a normative group of 52 autopsy-defined normal controls (23), while TOMC-related cut-offs were based on CSF biomarkers performance in correctly identifying a group of 231 normal controls using a rank-based method (25). Biomarkers cut-offs are summarized in table 1.

The study was reviewed and approved by the local ethics committee (CEIOC).

2.4 Statistical analysis

Pairwise comparison between groups in any sociodemographic, cognitive and biomarker feature was performed through independent sample t-test (continuous variables) and paired chi-squared test (dichotomous variables). To investigate the potential effect of differences in sociodemographic and clinical features between datasets on automated hippocampal volumes and on markers of cortical hypometabolism, a multivariate general linear model was used, setting dataset (TOMC vs ADNI) as independent variable and hippocampal volumes and PET metrics as dependent variables, and adding age and MMSE score as covariates. To assess diagnostic performance of each biomarker, sensitivity for prodromal

AD (in comparison with stable MCI), specificity for stable MCI, and diagnostic accuracy (with proper 95% confidence interval) were computed in each dataset separately. Biomarkers were ranked by diagnostic performance, and biomarkers showing best sensitivity and/or best specificity were identified. Positive likelihood ratios with 95% confidence intervals were finally computed to investigate whether positivity to both biomarkers increased the likelihood to progress to AD in the pooled ADNI and TOMC sample.

3. Results

33 sMCI and 24 pAD from ADNI and 28 sMCI and 18 pAD from TOMC dataset were enrolled. Table 1 summarizes cut-offs used for each diagnostic biological marker as thresholds of abnormality. Different cut-offs are provided for ADNI and TOMC datasets, as most of the biomarkers were collected in the two datasets with different and not directly comparable assaying procedures. In both clinical series, PALZ score, HCI and tau concentration are defined as abnormal (i.e. indicative of AD pathology) when equal or higher than the pertinent cut-off. On the contrary, metaROI average, hippocampal volume (manually, semi automatically and automatically computed) and Aβ42 concentration are defined as abnormal when equal or lower than the pertinent cut-off.

Table 2 shows main sociodemographic and clinical features, and hypometabolism, medial temporal atrophy and CSF markers of all subjects. As expected, in both ADNI and TOMC datasets, sMCI and pAD showed similar baseline MMSE scores, but significant differences in markers of AD pathology: pAD showed higher hypometabolism ($p < .05$ according to HCI score) in ADNI and performed worse ($p < .05$) on all biomarkers except for manually segmented (TOMC only) and automatically segmented hippocampal volume (both TOMC and ADNI datasets). The comparison between ADNI and TOMC biomarkers revealed significant differences in PALZ ($p=.004$) and wMetaROI scores ($p=.017$) and no significant differences in HCI and automatically computed hippocampal volumes. A multivariate analysis showed that such differences could be explained by differences in socio-demographic and clinical features (the inclusion of age and MMSE score as covariates into a multivariate general linear model resulted in no significant differences in PET biomarkers between datasets, $p=.492$).

Table 3 shows sensitivity for pAD, specificity for sMCI, and diagnostic accuracy of each biomarker, providing biomarker ranking. In both clinical series, the highest sensitivity for the detection of pAD was achieved by low CSF Aβ42 (79% in ADNI and 94% in TOMC), and the best specificity for the exclusion of sMCI was achieved by manually or semi automatically computed high hippocampal volume (76% in ADNI and 94% in TOMC). The diagnostic accuracy of the 3 indices of hypometabolism was intermediate and variable (between 52% and 61% in ADNI and 64% and 75% in TOMC dataset), sensitivity ranging between 33% and 66% in ADNI and 56% and 78% in TOMC, and specificity between 58% and 67% and 55% and 83%. The accuracy pattern of tau was similar to that of metrics of hypometabolism. Different hippocampal volume computation procedures showed consistent diagnostic performance; manual segmentation and semiautomatic computation achieved best

specificity and diagnostic accuracy, while automatic computation achieved best sensitivity in both datasets.

The positive likelihood ratio (LR+) of positivity to Aβ42, manual/semi-automated hippocampal volume, or both biomarkers was 1.325 (0.895-1.960 95%C.I.), 1.754 (1.097-2.804 95%C.I.), and 2.082 (1.275-3.400 95% C.I.) respectively (pooled ADNI and TOMC sample).

4. Discussion

The present study provides indications on the hierarchy of biomarkers for AD.

As expected, patients with pAD showed greater severity in markers of AD pathology as compared to sMCI, but no difference in global cognitive function as measured by MMSE. These results are in agreement with the early appearance of AD pathology, which anticipates overt cognitive deficit (26), providing evidence that biomarkers can indeed help in identifying AD from its earliest stage.

Abnormal Aβ42 concentration provided highest sensitivity to predict AD progression, in agreement with previous studies supporting Aβ42 as the most effective AD biomarker (27) (28) and showing that markers of amyloidosis become abnormal already at the earliest and preclinical AD stage (29). On the other hand, poor Aβ42 specificity suggests the need to combine it with other markers, such as tau or phospho-tau, in line with recent studies showing the power of combining CSF biomarkers (30), or CSF with MRI or FDG-PET biomarkers (31).

Manual or semi automated hippocampal volumes (previously shown to have comparable performance (32)) achieved the best specificity to correctly recognize sMCI patients, closely followed by automated volumes. Our findings are in line with studies showing that hippocampal volume closely correlates with disease progression (33), and that structural and amyloid markers have complementary performance in correctly identifying normal from MCI subjects (34). On the other hand, high hippocampal volume specificity could be more related to clinical progression to dementia than to progression toward AD itself, as documented by a recent study finding hippocampal sclerosis in elderly subjects clinically ascertained as AD and later, at autopsy, categorized as having non-AD pathologies (35).

Despite variable between datasets, the three metrics of hypometabolism showed intermediate sensitivity and specificity patterns. These results are in line with recent studies finding that parietotemporal and posterior cingulate hypometabolism on FDG-PET is less sensitive than CSF Aβ42 at milder AD stages (36), and FDG-PET closely follows MRI biomarkers performance in predicting clinical changes (31).

The combination of CSF Aβ42 and hippocampal atrophy increased likelihood of developing AD, in line with recent findings showing the power of combining MRI and CSF biomarkers (31), and suggesting that assessing CSF Aβ42 concentration and hippocampal atrophy at a time could be the best choice to identify prodromal AD cases.

Diagnostic performance and ranking of the biomarkers were not completely consistent across datasets. All biomarkers but hippocampal volume showed higher performance in the TOMC cohort than in the ADNI sample, probably due to several factors that may have played a role in terms of cognitive reserve, such as educational level, premorbid intelligence or ApoE genotype; further studies taking into account these additional relevant factors are needed to clarify this issue. Furthermore, despite main findings being similar in both TOMC and ADNI samples, biomarkers ranking quite differed between the two datasets, with largest differences in accuracy.

The choice of biomarker cut-offs is not trivial and should be made with caution. In this study, cut-offs were defined as thresholds leading to 95% specificity in a proper normal control dataset (independent of any group enrolled in the study), following a procedure used in most routine laboratory analyses. Cut-offs used for CSF and structural biomarkers were dataset-dependent, due to different procedures adopted, but such differences did not affect the comparison across datasets. Ideally, to this study purpose, it would have been best to use MCI-specific cut-offs maximizing the classification accuracy. However, not to define misleading cut-offs, final data on progression to dementia should have been available and this was not the case. Future studies are needed to determine biomarker cut-offs best predicting MCI conversion to AD.

This study has several strengths. To our knowledge, this is the first study providing indications on the hierarchy among biomarkers and ranking them based on sensitivity and specificity for AD in its early stages. Previous studies (see (37) for a review) were either focused on a subset of biomarkers or limited to the investigation of overt disease stages. In our study, follow-up period for clinical assessment was longer than in most previous studies, and potential differences in findings need to be interpreted taking this into account. As there are many available measures of cortical hypometabolism on FDG-PET and several segmentation procedures to assess hippocampal volume on MRI but no evidence on which one performs best (7), (38), (39), in this study we included the most widely used; the consistent performance across different measures supports generalizability of the findings. Last, specificity and diagnostic accuracy in pAD were computed based on biomarkers performance in sMCI patients, in order to answer the pertinent clinical question (i.e. separate pAD from sMCI patients).

Some limitations should be taken into account. First, pAD and sMCI sample size, especially in TOMC dataset, is quite small. Secondly, as each biomarker could potentially be better refined to improve its performance, current findings should be considered as preliminary. In addition, biomarker specificity in sMCI and diagnostic accuracy in pAD could be biased by the high heterogeneity of the sMCI group, including patients who will indeed remain stable and who will revert to cognitively normal status but potentially including also patients who will progress to dementia later on. Larger groups of MCI patients clinically followed for longer follow-up periods are needed to confirm current findings. Our results overall rely on univariate prediction models, thus providing evidence of diagnostic performance of each biomarker, separately. Future studies based on multivariate predictor models, aimed at assessing diagnostic accuracy of different biomarker combinations, are needed to confirm our preliminary findings. Finally, in this study we had to exclude MCI individuals who

progressed to non-AD dementia due to their limited number, which prevented us for having any reliable result. Moreover, we did not consider phosphorylated tau, previously shown to have high specificity for AD (40), due to the lack of available data in TOMC sample.

In conclusion, these preliminary findings provide further evidence to the revised NIA-AA criteria for AD and suggest that CSF Aβ42 and hippocampal measurements may be used in combination to best identify MCI patients “at risk”, optimizing costs/benefits ratio in clinical trials.

Acknowledgments

Data collection and sharing for this project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott; Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Amorfix Life Sciences Ltd.; AstraZeneca; Bayer HealthCare; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Rev August 16 2011 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 California, Los Angeles. This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation.

This study was funded in part by the National Institute of Mental Health (R01MH57899), the National Institute on Aging (R01AG031581 and P30AG19610) and the state of Arizona.

This work was also supported by the grants: sottoprogetto finalizzato Strategico 2006: “Strumenti e procedure diagnostiche per le demenze utilizzabili nella clinica ai fini della diagnosi precoce e differenziale, della individuazione delle forme a rapida o lenta progressione e delle forme con risposta ottimale alle attuali terapie”; Programma Strategico 2006, Convenzione 71; Programma Strategico 2007, Convenzione PS39, Ricerca Corrente Italian Ministry of Health, and was also supported in part by a grant from the Associazione Fatebenefratelli per la Ricerca (AFaR): “Proteomica Clinica nelle Malattie Neurodegenerative ad esordio tardivo”. Some of the costs related to patient assessment and imaging and biomarker detection were paid by an ad hoc grant from the Fitness e Solidarieta‘ 2006 and 2007 campaigns.

References

1. Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J, et al. Research criteria for the diagnosis of Alzheimer’s disease: revising the NINCDS-ADRDA criteria. Lancet Neurol. 2007; 6:734–746. [PubMed: 17616482]
2. Dubois B, Feldman HH, Jacova C, Cummings JL, Dekosky ST, Barberger-Gateau P, et al. Revising the definition of Alzheimer’s disease: a new lexicon. Lancet Neurol. 2010; 9:1118–1127. [PubMed: 20934914]
3. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, 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; 7:270–279. [PubMed: 21514249]
4. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer’s disease: Recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease. Alzheimers Dement. 2011; 7:280–292. [PubMed: 21514248]
5. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, et al. The diagnosis of dementia 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; 7:263–269. [PubMed: 21514250]

6. Geuze E, Vermetten E, Bremner JD. MR-based in vivo hippocampal volumetrics: 1. Review of methodologies currently employed. Mol Psychiatry. 2005; 10:147–159. [PubMed: 15340353]
7. Caroli A, Prestia A, Chen K, Ayutyanont N, Landau SM, Madison CM, et al. Summary metrics to assess Alzheimer's disease-related hypometabolic pattern with FDGPET:head-to-head comparison. J Nucl Med. Feb 17.2012
8. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDAWork Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. Neurology. 1984; 34:939–944. [PubMed: 6610841]
9. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. Arch Neurol. 1999; 56:303–308. [PubMed: 10190820]
10. Frisoni GB, Prestia A, Zanetti O, Galluzzi S, Romano M, Cotelli M, et al. Markers of Alzheimer's disease in a population attending a memory clinic. Alzheimers Dement. 2009; 5:307–317. [PubMed: 19560101]
11. Caroli A, Lorenzi M, Geroldi C, Nobili F, Paghera B, Bonetti M, et al. Metabolic compensation and depression in Alzheimer's disease. Dement Geriatr Cogn Disord. 2010; 29:37–45. [PubMed: 20110699]
12. Jagust WJ, Bandy D, Chen K, Foster NL, Landau SM, Mathis CA, et al. The Alzheimer's Disease Neuroimaging Initiative positron emission tomography core. Alzheimers Dement. 2010; 6:221–229. [PubMed: 20451870]
13. Herholz K, Salmon E, Perani D, Baron JC, Holthoff V, Frölich L, et al. Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. Neuroimage. 2002; 17:302–316. [PubMed: 12482085]
14. Chen K, Ayutyanont N, Langbaum JB, Fleisher AS, Reschke C, Lee W, et al. Characterizing Alzheimer's disease using a hypometabolic convergence index. Neuroimage. 2011; 56:52–60. [PubMed: 21276856]
15. Landau SM, Harvey D, Madison CM, Koeppe RA, Reiman EM, Foster NL, et al. Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI. Neurobiol Aging. 2011; 32:1207–1218. [PubMed: 19660834]
16. Jack CR Jr, Petersen RC, Xu YC, Waring SC, O'Brien PC, Tangalos EG, et al. Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. Neurology. 1997; 49:786–794. [PubMed: 9305341]
17. Caroli A, Testa C, Geroldi C, Nobili F, Guerra UP, Bonetti M, et al. Brain perfusion correlates of medial temporal lobe atrophy and white matter hyperintensities in mild cognitive impairment. J Neurol. 2007; 254:1000–1008. [PubMed: 17375260]
18. Jack CR Jr, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J Magn Reson Imaging. 2008; 27:685–691. [PubMed: 18302232]
19. Pruessner JC, Li LM, Serles W, Pruessner M, Collins DL, Kabani N, et al. Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: minimizing the discrepancies between laboratories. Cerebral Cortex. 2000; 10:433–442. [PubMed: 10769253]
20. Hsu YY, Schuff N, Du AT, Mark K, Zhu X, Hardin D, et al. Comparison of automated and manual MRI volumetry of hippocampus in normal aging and dementia. J Magn Reson Imaging. 2002; 16:305–310. [PubMed: 12205587]
21. Christensen GE, Joshi SC, Miller MI. Volumetric transformation of brain anatomy. IEEE Trans Med Imaging. 1997; 16:864–877. [PubMed: 9533586]
22. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron. 2002; 33(3):341–355. [PubMed: 11832223]

23. Shaw LM, Vanderstichele H, Knapik-Czajka M, Clark CM, Aisen PS, Petersen RC, et al. Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. Ann Neurol. 2009; 65:403–413. [PubMed: 19296504]
24. Galluzzi S, Testa C, Boccardi M, Bresciani L, Benussi L, Ghidoni R, et al. The Italian Brain Normative Archive of structural MR scans: norms for medial temporal atrophy and white matter lesions. Aging Clin Exp Res. Aug-Oct;2009 ;21(4-5):266–76.
25. Sjögren M, Vanderstichele H, Agren H, Zachrisson O, Edsbagge M, Wikkelsø C, et al. Tau and Abeta42 in cerebrospinal fluid from healthy adults 21-93 years of age: establishment of reference values. Clin Chem. 2001; 47:1776–1781. [PubMed: 11568086]
26. Vemuri P, Wiste HJ, Weigand SD, Shaw LM, Trojanowski JQ, Weiner MW, et al. MRI and CSF biomarkers in normal, MCI, and AD subjects: predicting future clinical change. Neurology. 2009; 73:294–301. [PubMed: 19636049]
27. Mattsson N, Zetterberg H, Hansson O, Andreasen N, Parnetti L, Jonsson M, et al. CSF biomarkers and incipient Alzheimer disease in patients with mild cognitive impairment. JAMA. 2009; 302:385–393. [PubMed: 19622817]
28. Hampel H, Bürger K, Teipel SJ, Bokde AL, Zetterberg H, Blennow K. Core candidate neurochemical and imaging biomarkers of Alzheimer's disease. Alzheimers Dement. 2008; 4:38–48. [PubMed: 18631949]
29. Jack CR Jr, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. Lancet Neurol. 2010; 9:119–128. [PubMed: 20083042]
30. Johansson P, Mattsson N, Hansson O, Wallin A, Johansson JO, Andreasson U, et al. Cerebrospinal fluid biomarkers for Alzheimer's disease: diagnostic performance in a homogeneous mono-center population. J Alzheimers Dis. 2011; 24:537–546. [PubMed: 21297262]
31. Walhovd KB, Fjell AM, Brewer J, McEvoy LK, Fennema-Notestine C, Hagler DJ Jr, et al. Combining MR imaging, positron-emission tomography, and CSF biomarkers in the diagnosis and prognosis of Alzheimer disease. AJNR Am J Neuroradiol. 2010; 31:347–354. [PubMed: 20075088]
32. Hsu YY, Schuff N, Du AT, Mark K, Zhu X, Hardin D, Weiner MW. Comparison of automated and manual MRI volumetry of hippocampus in normal aging and dementia. J Magn Reson Imaging. 2002; 16:305–10. [PubMed: 12205587]
33. Frisoni GB, Fox N, Jack CR Jr, Scheltens P, Thompson PM. The clinical use of structural MRI in Alzheimer's disease. Nat Rev Neurol. 2010; 6:67–77. [PubMed: 20139996]
34. Sluimer JD, Bouwman FH, Vrenken H, Blankenstein MA, Barkhof F, van der Flier WM, et al. Whole-brain atrophy rate and CSF biomarker levels in MCI and AD: a longitudinal study. Neurobiol Aging. 2010; 31:758–764. [PubMed: 18692273]
35. Pao WC, Dickson DW, Crook JE, Finch NA, Rademakers R, Graff-Radford NR. Hippocampal sclerosis in the elderly: genetic and pathologic findings, some mimicking Alzheimer disease clinically. Alzheimer Dis Assoc Disord. 2011; 25:364–368. [PubMed: 21346515]
36. Morinaga A, Ono K, Ikeda T, Ikeda Y, Shima K, Noguchi-Shinohara M, et al. A comparison of the diagnostic sensitivity of MRI, CBF-SPECT, FDG-PET and cerebrospinal fluid biomarkers for detecting Alzheimer's disease in a memory clinic. Dement Geriatr Cogn Disord. 2010; 30:285–292. [PubMed: 20861634]
37. Mayeux R, Reitz C, Brickman AM, Haan MN, Manly JJ, Glymour MM, et al. Operationalizing diagnostic criteria for Alzheimer's disease and other age-related cognitive impairment-Part 1. Alzheimers Dement. 2011; 7:15–34. [PubMed: 21255741]
38. Cherbuin N, Anstey KJ, Réglade-Meslin C, Sachdev PS. In vivo hippocampal measurement and memory: a comparison of manual tracing and automated segmentation in a large community-based sample. PLoS One. 2009; 4:e5265. [PubMed: 19370155]
39. Morey RA, Petty CM, Xu Y, Hayes JP, Wagner HR 2nd, Lewis DV, LaBar KS, Styner M, McCarthy G. A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. Neuroimage. 2009; 45:855–866. [PubMed: 19162198]
40. Prvulovic D, Hampel H. Amyloid β (Aβ) and phospho-tau (p-tau) as diagnostic biomarkers in Alzheimer's disease. Clin Chem Lab Med. 2011; 49:367–374. [PubMed: 21342022]

Research in Context

Systematic review. The revised NIA-AA criteria suggest that confidence in diagnosing AD could be improved by structural and biological evidence of Alzheimer's pathology and that the new identified biomarkers of AD are indicative of brain amyloidosis and neuronal degeneration, but indications on the hierarchy of biomarkers for AD are still missing.

Interpretation. Our findings suggest that CSF Aβ42 and hippocampal measurements may be used in combination to best identify MCI patients "at risk", optimizing costs/benefits ratio in clinical trials.

Future directions. Larger groups of MCI patients clinically followed for longer follow-up periods are needed to confirm current findings. Moreover, future studies based on multivariate predictor models, aimed at assessing diagnostic accuracy of different biomarker combinations, and that consider also MCI individuals who progress to non-AD dementia are strongly requested.

Table 1

Cut-offs for diagnostic biological markers.

	Biomarker assaying procedure	Cut-off	Reference
ADNI			
PALZ	sum of t-scores within a pre-defined AD mask	t= 13,48	on 148 normal controls, imposing 95% specificity
MetaROI average	average of mean counts in 5 metaROI volumes	w=−2.60	same
HCI*	inner-product of individual and pre-defined AD Z-score maps	1,055	same
Hippocampal volume, semi automatic	Medtronic Surgical Navigation Technologies	w=−2.95	On 66 ADNI normal controls, imposing 95% specificity
Hippocampal volume, automatic	Freesurfer	w=−2.14	same
Tau	xMAP Luminex platform with Innogenetics immunoassay kit–based reagents	93 pg/ml	on 52 normal controls at autopsy (23)
Aβ42	same	192 pg/ml	same
TOMC			
PALZ	sum of t-scores within a pre-defined AD mask	t=13,481	on 148 normal controls, imposing 95% specificity
MetaROI average	average of mean counts in 5 metaROI volumes	w=−2.60	same
HCI*	inner-product of individual and pre-defined AD Z-score maps	1,055	same
Hippocampal volume, manual	manual segmentation	w=−2.76	on 287 TOMC normal controls, imposing 95% specificity
Hippocampal volume, automatic	Freesurfer	w=−2.14	On 66 ADNI normal controls, imposing 95% specificity
Tau	enzyme-linked immunosorbent assay (ELISA)	age 51-70: 450 pg/ml age 71-93: 500 pg/ml	on 231 normal controls using a rank-based method (25)
Aβ42	same	500 pg/ml	same

* HCI score is expressed in arbitrary units. MetaROI average and hippocampal volumes are expressed in terms of age-corrected W-scores.

Table 2

Sociodemographic, clinical features and diagnostic biological markers in the TOMC and ADNI groups.

	ADNI		TOMC	
	Stable MCI	**Prodromal AD**	**Stable MCI**	**Prodromal AD**
	N=33	**N=24**	**N=18**	**N=18**
General variables				
Age [Range]	75±8 [55÷89]	75±8 [58÷89]	72±8 [51÷85]	71±8 [54÷83]
Gender (females)	13 (39%)	10 (42%)	9 (50%)	12 (67%)
MMSE [Range]	27±2 [24÷30]	28±2 [24÷29]	26±2 [24÷29]	26±2 [24÷29]
Markers of cortical hypometabolism				
PALZ score [Range]	13,930±14,570 [1,534÷70,049]	16,484±17,490 [1,819÷67,757]	15,288±15,764 [1,721÷65,292]	34,206±30,261 [6,837÷116,026]
Abnormal (%)	11 (33%)	8 (33%)	5 (28%)	14 (78%)
MetaROI average [Range]	−1.7±2.4 [−7.1÷3.9]	−2.6±1.4 [−5.6÷0.13]	−1.3±2.0 [−5.6÷3.7]	−3.3±2.3 [−7.4÷1.2]
Abnormal (%)	14 (42%)	11 (46%)	3 (17%)	10 (56%)
HCI [Range]	859±931 [−1,111÷3,134]	1,334±648 [202÷2,895]	746±1,031 [−2,030÷2,739]	1,656±1,283 [−588÷4,000]
Abnormal (%)	13 (39%)	15 (63%)	8 (44%)	13 (72%)
Markers of medial temporal atrophy				
Hippoc. Volume Automatic* [Range]	−1.4±1.6 [−3.9÷1.4]	−2.2±1.5 [−4.5÷1.9]	−0.6±1.7 [−3.9÷2.2]	−1.9±2.1 [−4.6÷2.2]
Abnormal (%)	13 (39%)	12 (50%)	4 (27%)	8 (44%)
Hippoc. Volume Manual [Range]	---	---	−1.4±0.7 [−2.8÷0.4]	−2.1±1.8 [−6.7÷1.8]
Abnormal (%)			1 (6%)	5 (28%)
Hippoc. Volume* Semi Automatic [Range]	−1.5±2.1 [−5.9÷2.5]	−2.5±1.9 [−5.2÷1.6]	---	---
Abnormal (%)	8 (24%)	11 (46%)		
CSF markers				
Tau* [Range]	94±50 [28÷226]	104±53 [47÷274]	333±164 [137÷713]	622±273 [268÷1,200]
Abnormal (%)	13 (39%)	11 (46%)	3 (17%)	11 (61%)
Aβ42* [Range]	161±60 [53÷276]	164±50 [86÷266]	576±252 [257÷1,074]	370±81 [230÷530]
Abnormal (%)	24 (73%)	19 (79%)	9 (50%)	17 (94%)

Values are mean ± standard deviations for continuous variables or frequency (percentage) for gender.

*Variables not directly comparable between ADNI and TOMC due to different assaying procedures. MetaROI average and hippocampal volume (both manually or semi automatically and automatically computed) are expressed in terms of w-scores (age-corrected scores).

[†] Missing data for 3 sMCI and 1 AD TOMC patients due to failure of automated processing.

Table 3

Sensitivity for prodromal AD (pAD), specificity for stable MCI (sMCI), and accuracy of each biomarker in the TOMC and ADNI datasets.

		Sensitivity		Specificity		Accuracy	
		Rank	pAD (n=24)	Rank	sMCI (n=33)	Rank	pAD+sMCI (n=57)
ADNI	Aβ42*	**1**	79% [59-91]	**5**	27% [15-44]	**6**	49% [37-62]
	HCI	2	63% [43-79]	3	61% [44-75]	2	61% [48-73]
	PALZ	**5**	33% [18-53]	2	67% [50-80]	5	52% [40-65]
	MetaROI	4	46% [28-65]	4	58% [41-73]	5	52% [40-65]
	Tau*	4	46% [28-65]	3	61% [44-75]	4	54% [42-67]
	Hippo volume† Semi Auto	4	46% [28-65]	**1**	76% [59-87]	**1**	63% [50-74]
	Hippo volume†† Auto	3	50% [31-68]	3	61% [44-75]	3	56% [43-68]
		Sensitivity		Specificity		Accuracy	
		Rank	pAD (n=18)	Rank	sMCI (n=18)	Rank	pAD+sMCI (n=36)
TOMC	Aβ42**	**1**	94% [74-99]	**6**	50% [29-71]	2	72% [56-84]
	HCI	3	72% [49-87]	5	55% [34-75]	4	64% [47-77]
	PALZ	2	78% [55-91]	4	72% [49-87]	**1**	75% [59-86]
	MetaROI	5	56% [34-75]	2	83% [61-94]	3	69% [53-82]
	Tau**	4	61% [39-80]	2	83% [61-94]	2	72% [56-84]
	Hippo volume††† Manual	**7**	28% [12-51]	**1**	94% [74-99]	5	61% [45-75]
	Hippo volume†† Auto‡	6	44% [25-66]	3	73% [48-89]	**6**	57% [41-73]

For each statistical measure of performance, values [95% CI] and biomarker ranking are reported. MetaROI average and hippocampal volume (both manually or semi automatically and automatically computed) are expressed in terms of w-scores (age-corrected scores). Aβ42 and total tau protein concentrations were determined by * xMAP Luminex platform with Innogenetics immunoassay kit– based reagents, or ** a commercially available enzyme-linked immunosorbent (ELISA) assay. Hippocampal volume was † semiautomatically computed using Medtronic Surgical Navigation Technologies, †† automatically computed using Freesurfer, or ††† manually segmented. ‡Missing data for 3 sMCI and 1 AD TOMC patients due to failure of automated processing.