Eur J Nucl Med Mol Imaging
DOI 10.1007/s00259-015-3170-y

ORIGINAL ARTICLE

Cross-validation of biomarkers for the early differential diagnosis and prognosis of dementia in a clinical setting

Daniela Perani[1,2,3] · Chiara Cerami[1,2,4] · Silvia Paola Caminiti[1,2] · Roberto Santangelo[5] · Elisabetta Coppi[5] · Laura Ferrari[5] · Patrizia Pinto[6] · Gabriella Passerini[7] · Andrea Falini[1,2,8] · Sandro Iannaccone[4] · Stefano Francesco Cappa[2,9] · Giancarlo Comi[1,5] · Luigi Gianolli[3] · Giuseppe Magnani[5]

Received: 6 April 2015 / Accepted: 10 August 2015

Abstract
Purpose The aim of this study was to evaluate the supportive role of molecular and structural biomarkers (CSF protein levels, FDG PET and MRI) in the early differential diagnosis of dementia in a large sample of patients with neurodegenerative dementia, and in determining the risk of disease progression in subjects with mild cognitive impairment (MCI).
Methods We evaluated the supportive role of CSF $A\beta_{42}$, t-Tau, p-Tau levels, conventional brain MRI and visual assessment of FDG PET SPM t-maps in the early diagnosis of dementia and the evaluation of MCI progression.
Results Diagnosis based on molecular biomarkers showed the best fit with the final diagnosis at a long follow-up. FDG PET SPM t-maps had the highest diagnostic accuracy in Alzheimer's disease and in the differential diagnosis of non-Alzheimer's disease dementias. The p-tau/$A\beta_{42}$ ratio was the only CSF biomarker providing a significant classification rate for Alzheimer's disease. An Alzheimer's disease-positive metabolic pattern as shown by FDG PET SPM in MCI was the best predictor of conversion to Alzheimer's disease.
Conclusion In this clinical setting, FDG PET SPM t-maps and the p-tau/$A\beta_{42}$ ratio improved clinical diagnostic accuracy, supporting the importance of these biomarkers in the emerging diagnostic criteria for Alzheimer's disease dementia. FDG PET using SPM t-maps had the highest predictive value by identifying hypometabolic patterns in different neurodegenerative dementias and normal brain metabolism in MCI, confirming its additional crucial exclusionary role.

Keywords Alzheimer's Disease · Mild Cognitive Impairment · Biomarkers · Dementia diagnosis · Clinical setting

Introduction

The World Alzheimer Report (2011) emphasized that dementia is recognized early in only 20 % to 50 % of patients [1]. The so-called "treatment gap" between the early stages of

✉ Daniela Perani
perani.daniela@hsr.it

Giuseppe Magnani
magnani.giuseppe@hsr.it

[1] Vita-Salute San Raffaele University, Via Olgettina, 58, 20132 Milan, Italy

[2] Division of Neuroscience, San Raffaele Scientific Institute, Via Olgettina, 58, 20132 Milan, Italy

[3] Nuclear Medicine Unit, San Raffaele Hospital, Via Olgettina, 60, 20132 Milan, Italy

[4] Clinical Neuroscience Department, San Raffaele Hospital, Via Olgettina, 60, 20132 Milan, Italy

[5] Department of Neurology, San Raffaele Hospital, Via Olgettina, 60, 20132 Milan, Italy

[6] Department of Neurology, Papa Giovanni XXIII Hospital, Piazza OMS, 1, 24127 Bergamo, Italy

[7] Servizio di Medicina di Laboratorio OSR, Via Olgettina, 60, 20132 Milan, Italy

[8] CERMAC – Department of Neuroradiology, San Raffaele Hospital, Via Olgettina, 60, 20132 Milan, Italy

[9] IUSS Pavia, Piazza della Vittoria, 15, 27100 Pavia, Italy

Published online: 04 September 2015

pathology and a correct diagnosis has led to an increase in costs and may be related to poor efficacy of treatments [2]. Moreover, different studies have found that between 12 % and 23 % of Alzheimer's disease (AD) diagnoses are not confirmed at autopsy ("misdiagnosed") [3]. These misdiagnoses may be due to the fact that different pathological conditions mimic the symptoms of AD. Although diagnostic criteria in dementia are constantly improving (for example [4–10]), several challenges in terms of differential and early diagnosis of dementias remain in the clinical setting. Thus, the validation of dementia biomarkers to support a more certain early and differential diagnosis has become urgent [11].

In AD, there is evidence of distinctive topographical and pathological markers for the symptomatic, prodromal and preclinical stages of AD [12–19]. Both the National Institute on Aging/Alzheimer's Association workgroup (NIA-AA) and the International Working Group (IWG) have proposed a significant revision of guidelines for preclinical/asymptomatic, prodromal (mild cognitive impairment, MCI, due to AD), and symptomatic AD conditions [4, 5, 10, 20], centring the diagnosis on the supportive use of biomarkers. Cerebrospinal fluid (CSF) levels of $A\beta_{42}$, total tau (t-Tau) and phosphorylated tau (p-Tau) proteins have been extensively investigated as potential in vivo markers of AD pathology, in typical and atypical AD conditions [21, 22], as well as in prodromal AD [15]. Low CSF $A\beta_{42}$ levels in combination with high t-Tau/p-Tau show high sensitivity and specificity in AD and high value for predicting progression of MCI to AD [23]. Since variations in the procedures may account for differences in accuracy [24], worldwide standard operating procedures have been suggested [24].

Structural MRI has been recognized as a useful measure of cerebral atrophy of both whole brain and, more specifically, the perirhinal and entorhinal cortices in AD [16]. Medial temporal lobe atrophy, however, is not specific for AD since it may also be present in other neurological conditions [25–27] and may be associated with advancing age [28]. For these reasons, it has limited value in the differential diagnosis of AD [16]. FDG PET provides in vivo information about the distribution of synaptic dysfunction in dementia and is the closest biomarker associated with the neuronal dysfunction in AD, being closely related to its clinical presentation [29, 30]. In the last 20 years, FDG PET imaging has achieved an increasing supportive role in the diagnostic algorithm of AD [31]. In addition, research studies in the memory clinic setting have shown that FDG PET has added value over the standard diagnostic work-up, influencing the final diagnosis [32, 33]. This is especially true when prior diagnostic confidence is low [34]. Since neurodegenerative dementias show selective neuronal vulnerabilities, according to which a specific neuronal population dies and others are resistant to neurodegeneration, specific topographical patterns of cerebral hypometabolism may also be detected at the single-subject level [14, 33, 35].

The simple use of biomarkers in the diagnostic algorithm is not sufficient alone to guarantee greater accuracy. Their accuracy depends on the type of imaging (e.g. amyloid PET imaging, FDG PET or MRI) as well as on the measurement method [36]. Semiquantitative or quantitative assessment and standard operating procedures are strongly recommended by scientific societies as key to the well-advised use of biomarkers in research, the clinical setting and in trials [37]. In addition, the combined use of different biomarkers and standardized procedures can be expected to facilitate the diagnostic process at the single-subject level [38–40].

Previous studies have provided evidence of the role of biomarkers in dementia diagnosis and in determining the risk of progression of MCI to dementia using the direct comparison of the three biomarkers (i.e. CSF protein levels, MRI and FDG PET) [41–46]. These studies have shown some discrepancy in the effectiveness of these biomarkers that might depend on different factors, such as the included population and possible clinical misdiagnoses, the differential role of the biomarkers during the disease course, and, crucially, the biomarker measurement methods [47]. The accuracy of imaging with these biomarkers can be indeed very different whether based on parametric approaches or on visual inspection depending greatly on the observer's experience and because of the lack of a clear cut-off between normal and pathological findings [36].

In this study, we evaluated the role of the biomarkers CSF protein levels, MRI and FDG PET as available in a routine clinical setting for the early differential diagnosis of AD and for prediction of progression of MCI to dementia. In particular, we evaluated the role of CSF $A\beta_{42}$, t-Tau and p-Tau proteins levels, of atrophy on visual assessment of conventional MRI images, and of visual assessment of FDG PET statistical parametrical mapping (SPM) t-maps [33, 48] in a large sample of subjects attending a memory clinic. We assessed the power of each biomarker and their combined use for the early differential diagnosis of dementia and their value in predicting progression of MCI to dementia. Given the previous evidence of very high diagnostic accuracy of our new validated procedure for the routine assessment of FDG PET imaging in the clinical setting [33, 48], we predicted that this biomarker would provide better performance than the other biomarkers.

Material and methods

Subjects

We retrospectively collected information about 86 patients with early dementia, including AD, frontotemporal lobar degeneration (FTLD) and dementia with Lewy bodies (DLB), and 35 subjects with MCI. All individuals were referred to the memory clinics of the San Raffaele Hospital (Milan, Italy)

with memory or other cognitive impairments during the period 2009 to 2012. At referral, they were evaluated by experienced behavioural neurologists and neuropsychologists. Clinical evaluation included a structured clinical interview, a full neurological examination, and a standard neuropsychological evaluation. The neuropsychological battery used in our clinical setting have been described in detail by Perani et al. [33]. All patients also underwent lumbar puncture and CSF analysis for $A\beta_{42}$, t-Tau and p-Tau levels as well as a FDG PET scan. Additionally, a conventional MRI scan was considered for visual analysis of atrophy in 60 individuals, 19 with AD, 8 with FTLD, 8 with DLB, and 19 with MCI.

All biomarker data were collected within 3 months of the baseline clinical visit. Each patient underwent a clinical and neuropsychological follow-up (at 27.48±10.43 months),that confirmed the diagnosis in the patients with dementia and in the subjects with MCI who converted. Accordingly, subjects with MCI were grouped into converters to AD and nonconverters.

Study design

Firstly, two neurologists who were experts in dementia and blinded to the previous clinical diagnosis reviewed the clinical data from all the included individuals in order to formulate a possible diagnosis based exclusively on the clinical information (i.e. case history, neurological examination and neuropsychological data), providing a "clinically-based classification". Of 121 patients, 16 were excluded from the initial sample (i.e. those without complete diagnostic agreement between the two experts and/or not fulfilling clinical criteria for neurodegenerative dementia or for MCI [7–10, 20, 49]). Thus, the final cohort included 75 patients with dementia (45 men and 30 women; mean age 66.71±7.22 years; Mini Mental State Examination (MMSE) score 18.41±5.26; disease duration 3.25±1.85 years) affected by AD, FTLD or DLB, and 30 subjects with MCI (15 men and 15 women; mean age 68.24± 6.8 years; MMSE score 25.96±2.25; Table 1). Dementia was defined according to the clinical criteria of each main neurodegenerative dementia subtype [6–10, 20]. MCI was defined as the presence of objective impairments in the neuropsychological evaluation in memory and/or other cognitive domains in the absence of functional impairment and dementia [49]. In addition, CSF $A\beta_{42}$, t-Tau and p-Tau levels as well as MRI and FDG PET scans were evaluated according to the procedures described below. In particular, CSF protein levels were considered as abnormal on the basis of established cut-off values [50]. Two expert neuroradiologists and two nuclear medicine physicians evaluated the MRI and FDG PET scans, respectively, blinded to the clinical/neuropsychological information, in order to provide a final classification of the biomarker results.

In a second step, in order to evaluate the additional value of the biomarkers, the same neurologists formulated a new diagnosis adding to the clinical data the information obtained with all the available biomarkers (i.e. CSF protein levels plus FDG PET, and MRI), thus providing a "clinically/biomarker-based classification". The two diagnostic classifications (i.e. clinically-based and clinically/biomarker-based) were then compared with the clinical diagnosis obtained with a follow-up of more than 2 years (follow-up diagnosis).

All subjects, or their informants/caregivers, gave informed consent to the each experimental procedure following detailed explanation, with the previous approval of the local ethics committee.

CSF acquisition and analysis

All patients underwent lumbar puncture in the L3–L4 or L4–L5 space after giving written informed consent, and following detailed explanation of the procedure. The procedure was always performed early in the morning. No serious adverse events were reported. CSF (8–10 ml) was collected in sterile polypropylene tubes. Routine chemical parameters were determined, and in the remaining CSF the levels of $A\beta_{42}$, t-Tau and p-Tau were determined. After centrifugation, the CSF samples were stored at −80 °C until analysis. The levels of

Table 1 Demographic and clinical characteristics of the included individuals

	AD (n=47)	FTLD (n=14)	DLB (n=14)	MCI converters (n=10)	MCI nonconverters (n=20)	p value
Age (years), mean (standard deviation)	66 (6.8)	65 (7.3)	72 (6)	69 (5.5)	68 (7.6)	<0.05[a]
Gender (m/f), n	26/21	8/6	11/3	7/4	8/11	NS
Disease duration (months), mean (standard deviation)	39 (24)	32 (19)	42 (22)	28 (10.9)	46.8 (38)	NS
MMSE score, mean (standard deviation)	18 (4.5)	20 (7.1)	16 (5.2)	26 (1.8)	26 (2.5)	<0.05[b]

AD Alzheimer's disease, *DLB* dementia with Lewy bodies, *FTLD* frontotemporal lobar degeneration, *MCI* mild cognitive impairment, *MMSE* Mini Mental State Examination

[a] DLB vs. the other groups

[b] MCI converters and MCI nonconverters vs. the other groups

$A\beta_{42}$, t-Tau and p-Tau were determined in the local laboratory (LABORAF, San Raffaele Hospital) by technicians blinded to the clinical diagnosis, using commercially available ELISA kits (Innogenetics, Gent, Belgium) according to the manufacturer's instructions. Cut-off values for AD reported in the literature [50] were adopted, i.e. $A\beta_{42}$ ≥500 ng/L, t-Tau ≤350 ng/L, p-Tau ≤61 ng/L.

MRI acquisition and analysis

Brain MRI scans were acquired using a 3-T Philips Achieva scanner (Philips Medical Systems, Best, The Netherlands) with an eight-channel head coil in a subgroup of 60 patients. The scanning session included a T1-weighted volumetric MR scan (220 slices, TR/TE 600/20 ms, voxel size 0.9 × 0.9 × 0.8 mm). Each scan was evaluated and independently classified by two expert raters. Brain atrophy was considered according to the validated MRI patterns of the three main neurodegenerative dementia conditions (i.e. AD, DLB and FTLD) [6–8, 10, 51]. Each scan was classified as not satisfying the criteria for brain atrophy (i.e. negative scan), as compatible with AD (AD-like pattern scan) or as compatible with a non-AD dementia (non-AD-like pattern scan).

FDG PET acquisition and analysis

FDG PET acquisitions were performed at the Nuclear Medicine Unit, San Raffaele Hospital following standardized procedures [13]. All images were acquired with a Discovery STE (GE Medical Systems, Milwaukee, WI) multiring PET tomography (PET/CT) system (with 45 min between injection and the start of the scan and scan duration 15 min). Images were reconstructed using an ordered subsets expectation maximization (OSEM) algorithm. Each PET phase was corrected for attenuation with the CT data of the corresponding phase. In each PET scan either 35 or 47 (depending to the different scanner characteristics) transaxial tomographic slices of 4.25 mm were acquired, and were reoriented into the coronal and sagittal planes. The emission images were then reconstructed using filtered back projection, using the software provided by the manufacturer.

Image preprocessing and statistical analysis were performed according to a new procedure developed in our laboratory [33, 48]. Each patient scan was then tested for relative "hypometabolism" by comparison with a large sample of FDG PET scans from a database of normal controls on a voxel-by-voxel basis [33]. Proportional scaling was used to remove intersubject global variation in PET intensities. Additionally, voxel-wise comparisons were made using a within-brain comparison-specific explicit FDG mask in order to remove emission counts outside the brain and to restrict subsequent analyses to within-brain voxels. The threshold was set at p=0.05, with correction for multiple comparisons using family-wise error correction (FWE) at the voxel level. Only clusters containing more than 100 voxels were deemed to be significant.

The FDG PET hypometabolic patterns were classified by two nuclear physicians, blinded to clinical/neuropsychological details, and the CSF and MRI results. Raters were asked to classify the SPM t-map as normal or abnormal on the basis that a normal SPM t map should not reveal any significant hypometabolic pattern at the FWE-corrected threshold, either at the voxel or the cluster level. In the event of abnormal findings, raters had to decide whether the hypometabolic pattern was suggestive of a specific neurodegenerative dementia subtype according to a well-established literature [6–8, 12, 33, 35, 52]. Thus each SPM t-map was classified as negative, AD-like, DLB-like or FTLD-like on the basis of criteria available elsewhere [14].

Statistical analysis

Data were analysed using the Statistical Package for Social Sciences software (SPSS v. 19.0). Sociodemographic differences between groups were assessed with the chi-squared statistic for dichotomous variables or Kruskal-Wallis one-way analysis of variance by ranks for continuous variables. First, we compared the clinically-based and the clinically/biomarker-based classification to establish whether they were discordant with respect to the reference follow-up diagnosis (Cohen's κ analysis). Cohen's κ coefficient was also used to evaluate the interrater agreement between the experts' classifications for the MRI and the FDG PET imaging, resulting in an "almost perfect agreement" (κ>0.85) in both cases. Thus, we selected a single set of classifications (i.e. the best) for the subsequent analyses.

We then performed separate logistic regression analyses to evaluate the accuracy of each biomarker in correctly differentiating dementia subtypes comparing, in particular, AD vs. non-AD (DLB + FTLD), FTLD vs. non-FTLD (DLB + AD) and DLB vs. non-DLB (FTLD + AD). This was crucial to establish the role of biomarkers in differentiating non-AD conditions. MRI was excluded from the last two of these analyses (FTLD vs. non-FTLD and DLB vs. non-DLB) because no specific DLB or FTLD classification was available for MRI scans. In a further step, the significant regressors derived from the logistic regression analyses were joined in a logistic regression multivariate model. The regressors resulting from the multivariate analysis were considered as the most informative for the specific type of dementia diagnosis. Age and gender variables were inserted as covariates both for separate and multivariate logistic regression analyses in order to evaluate their contribution to the model.

Subjects with MCI that progressed to AD during the follow-up period (27.48±10.43 months) were compared with those without progression to AD and who remained stable

using Cox proportional hazards regression analysis. CSF, MRI and FDG PET biomarkers were included in the multivariate analysis in order to find those biomarkers that were significantly associated with progression to AD. Each variable was entered in a stepwise forward manner. Thus, the final model was composed only of variables significantly predicting AD at the time of follow-up. Age and gender were inserted as covariates and follow-up time was used as a temporal variable of the model. The area under the curve (AUC), sensitivity, specificity and positive and negative likelihood ratios (LR+ and LR−) were calculated for each biomarker. All statistical results were considered as significant with threshold $p \leq 0.05$.

Results

Descriptive statistics

No significant age differences were found among the groups, except for DLB patients who were significantly older (72.36± 6.02 years; $p<0.001$) than all the other groups (Table 1). At the clinical follow-up, 10 of 30 MCI subjects had converted to dementia, 8 to AD and 2 to non-AD dementias (i.e. 1 FTLD and 1 DLB). The MCI converter and nonconverter groups did not show any significant demographic differences (Table 1).

Comparison between clinically-based and clinically/biomarker-based classifications

Cohen's κ analysis demonstrated a significant but not perfect agreement between the clinically-based and clinically/biomarker-based classifications (Cohen's $\kappa=0.75$, $p<0.001$), with a mismatch of 24 % between them. In relation to the follow-up diagnosis, the clinically/biomarker-based classification performed better than the clinically-based one.

Statistical analyses in dementia

AD vs. non-AD comparison

The best CSF measure for AD/non-AD classification was the ratio of p-Tau/Aβ_{42} (expβ=8.001, CI=2.55 – 25.15, $p<0.001$, AUC=0.81, sensitivity 83 %, specificity 64 %, LR+ 2.3; LR− 0.26). Although Aβ_{42} showed good sensitivity (85 %), its specificity was poor (46 %). Since AD-related pathology may also be present in patients with DLB, particularly in those with more severe cognitive impairment, we then excluded patient with DLB from the non-AD dementia group. Thus, performing this new comparison analysis comparing the AD and FTLD groups, we found that, among the CSF biomarkers, Aβ_{42} significantly influenced the final regression model (expβ=14.28; CI=3.49 – 58.54, $p<0.001$, AUC= 0.78, sensitivity 85 %, specificity 71 %). Compared to t-Tau, p-Tau showed a better performance although poorer than that of p-Tau/Aβ_{42} ratio (AUC=0.67, sensitivity 70 %, specificity 64 %).

Considering the two neuroimaging biomarkers, FDG PET was more accurate (expβ=88;,CI=18.17 – 426.13, $p<0.001$, AUC=0.90, sensitivity 94 %, specificity 86 %, LR+ 6.71, LR − 0.07), whereas MRI showed both low sensitivity (46 %) and specificity (50 %; Table 2). Notwithstanding the good performance of both FDG PET and p-Tau/Aβ_{42} ratio, when entered together in a multivariate logistic regression model, only FDG PET met the statistical significance threshold. Age and gender did not influence the results of the regression model.

FTLD vs. non-FTLD and DLB vs. non-DLB comparisons

In the binomial logistic regression FDG PET was the most accurate biomarker for discriminating FTLD from non-FTLD patients (expβ=251.33, CI=24.17 – 2,613.65, $p<0.001$, AUC 0.94, sensitivity 93 %, specificity 95 %, LR+ 18.6, LR− 0.07) and DLB from non-DLB patients (expβ=150, CI=15.17 – 1,483.15, $p<0.001$, AUC 0.94, sensitivity 71 %, specificity 98 %, LR+ 35.5, LR− 0.3). This result was confirmed also after entering age in the regression model as nuisance variable (expβ=187.013, CI=12.30 – 2845.60, $p<0.000$). Among CSF biomarkers, the Aβ_{42} value showed good accuracy in distinguishing FTLD from non-FTLD patients (expβ=0.07, CI=0.018 – 0.27, $p<0.001$). Notably, FDG PET and Aβ_{42} contributed differently to the correct discrimination of FTLD from non-FTLD patients (Figs. 1 and 2). In particular, while a positive FDG PET pattern for FTLD increased the likelihood of correctly classifying patients with FTLD, in contrast, positivity of Aβ_{42} showed good ability to classify non-FTLD patients (AD or DLB).

Statistical analyses in MCI

All biomarkers were considered in the multivariate Cox proportional hazards analysis comparing MCI converters and nonconverters to AD. FDG PET was the only predictor of conversion in the final step-wise model (expβ=8.62, CI= 1.02 – 72.74, $p<0.05$; Fig. 3). Further analysis of MCI conversion to other dementias (FTLD or DLB) was not possible because of the small number of patients. In the ROC analysis, both the t-Tau/Aβ_{42} and the p-Tau/Aβ_{42} ratios were informative in the prediction of conversion to AD. Though the p-Tau/Aβ_{42} ratio showed a higher specificity than FDG PET (96 % vs. 90 %), FDG PET had a higher sensitivity (86 % vs. 57 %). The lower FDG PET specificity was due to the number of MCI subjects with and AD-like FDG PET pattern who did not convert during the follow-up time, and thus were considered in the ROC analysis as false-positives.

Table 2 ROC analyses including sensitivity and specificity of each biomarker

	AD vs. non-AD Comparison			FTLD vs. non-FTLD Comparison			DLB vs. non-DLB Comparison		
	ROC AUC	Sensitivity (%)	Specificity (%)	ROC AUC	Sensitivity (%)	Specificity (%)	ROC AUC	Sensitivity (%)	Specificity (%)
$A\beta_{42}$	0.64	85	43	**0.22**	**29**	**15**	0.57	86	72
t-Tau	0.62	38	86	0.45	21	69	0.36	7.1	34
p-Tau	0.67	70	64	0.37	36	38	0.37	36	62
t-Tau/$A\beta_{42}$	0.80	79	68	0.20	36	23	0.31	43	77
p-Tau/$A\beta_{42}$	**0.81**	**83**	**64**	0.23	43	16	0.32	57	84
MRI	0.52	46	50	–	–	–	–	–	–
FDG PET	**0.90**	**94**	**86**	**0.94**	**93**	**95**	**0.94**	**71**	**98**

Values indicating significant discrimination in each comparison are indicated in bold

Discussion

In this study, we evaluated the supportive role of CSF $A\beta_{42}$, t-Tau and p-Tau levels, conventional brain MRI imaging and visual assessment of FDG PET SPM t-maps in the early diagnosis of dementia and in the evaluation of MCI progression. In particular, we investigated the accuracy of these biomarkers in recognizing AD and non-AD dementias (DLB and FTLD), and in predict progression of MCI to dementia.

The first key result of this study was a significant, but not perfect, agreement between the clinically-based and clinically/biomarker-based classifications. More specifically, only the clinically/biomarker-based classification resulted in a consistent diagnosis with reference to the follow-up diagnosis. This result indicates the significant supportive role of biomarkers in the clinical setting, improving the clinician's confidence in the correctness of the diagnosis. Diagnostic errors have a number of undesirable implications. Both patient and caregivers may experience significant emotional stress and psychological damage from misdiagnosis of an incurable neurodegenerative condition. In addition, a wrong AD diagnosis may result in the inappropriate long-term administration of pharmacological treatments (e.g. cholinesterase inhibitors) useless for non-AD conditions, and may also result in possibly reversible conditions such as vitamin deficiency and depression being missed.

As proved by the statistical analyses, the better performance of the clinically/biomarker-based classification was mainly due to the supportive role of the visual assessment of FDG PET SPM t-maps. The semiquantitative method used provided very high sensitivity (94 %) and specificity (86 %) in the discrimination of AD and non-AD dementias. FDG PET was the best biomarker among those investigated, also being able to identify non-AD dementias (i.e. sensitivity 93 % and specificity 95 %, and sensitivity

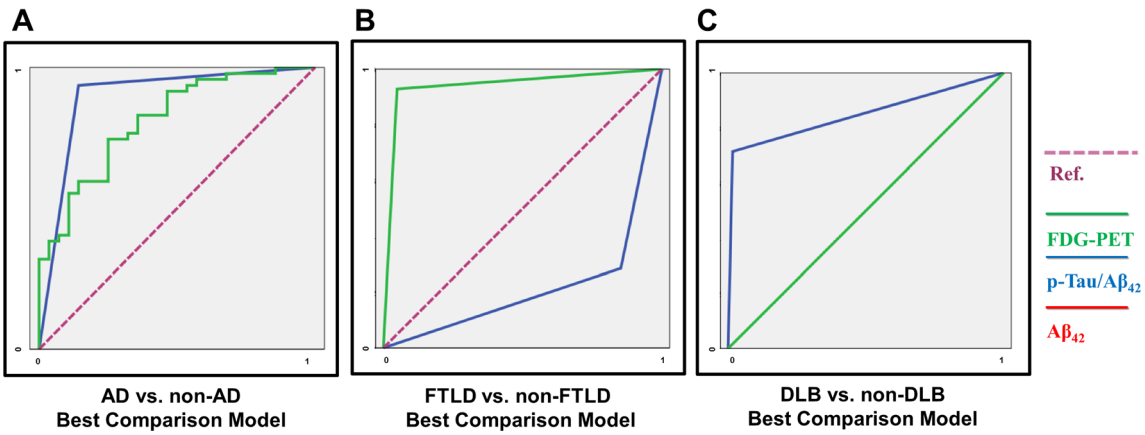

Fig. 1 ROC curve graphical representation of the best model for each dementia subgroup obtained from Logistic regression analyses for **a** AD vs. non-AD comparison, **b** FTLD vs. non-FTLD comparison, and **c** DLB vs. non-DLB comparison. Specificity values are shown on the horizontal axis and sensitivity values on the vertical axis (*Ref* reference line)

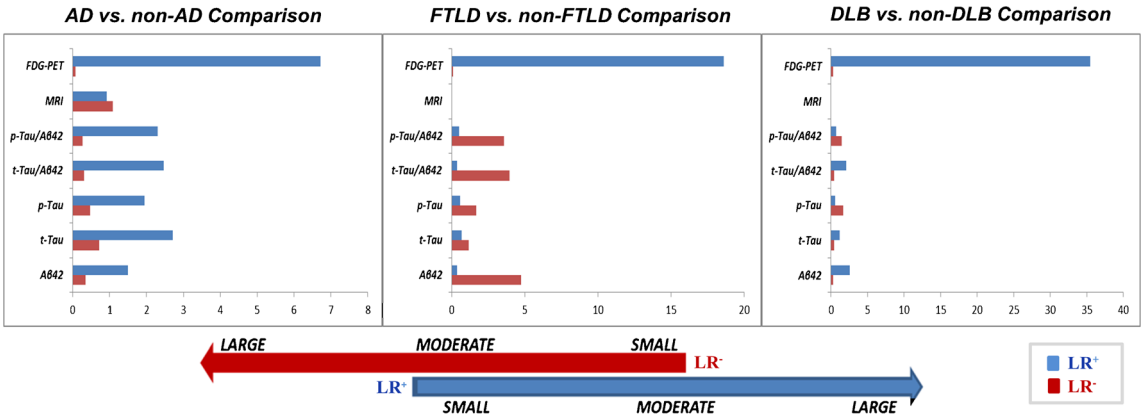

Fig. 2 Positive and negative likelihood ratio (LR+ and LR−) for correct classification of patients with Alzheimer's disease (AD), frontotemporal lobar degeneration (FTLD) and dementia with Lewy bodies (DLB). LR+ >5 indicates that the biomarker positive classification is more probably associated with the disease occurrence. LR− <0.2 indicates a relevant association between the negative biomarker classification and the absence of the dementia condition

71 % and specificity 98 % for the discrimination of FTLD from non-FTLD and DLB from non-DLB, respectively). This was a crucial part of our analysis. While the majority of FDG PET imaging studies in dementia have investigated AD [12, 39, 44], only a few have explored the discrimination of AD and non-AD dementias [33, 53]. Nevertheless, FDG PET has been included in the research diagnostic criteria of AD as well as non-AD dementias [4–7, 20, 54] .

In this study, we showed that the use of an optimized voxel-based method for FDG PET scan evaluation was crucial to achieving very high diagnostic accuracy. As previously validated by our group [33, 48], this tool had high sensitivity (96 %) and specificity (84 %) in recognizing dementias, as well as in predicting progression of to dementia or reversion of MCI to normal cognition [14]. Its use guarantees not only higher diagnostic accuracy compared with visual assessment of FDG uptake, but also a higher level of confidence even for moderately skilled FDG PET specialists [33].

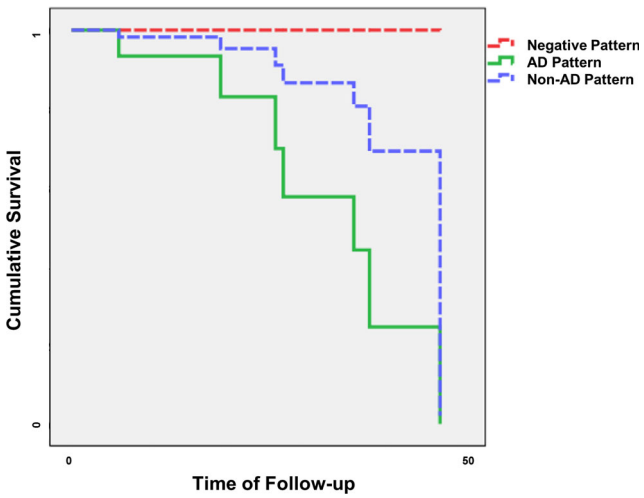

Fig. 3 Survival curves of MCI subjects according to their FDG PET pattern at baseline. None of the MCI subjects with a negative FDG PET pattern (n=8) converted to AD (*blue line*), one of the MCI subjects with an AD-positive FDG PET pattern converted to non-AD dementia (*grey line*), and six MCI subjects with an AD-positive FDG PET pattern converted to AD (*green line*) showing the higher risk in this group of patients (lower level of survival)

Among the CSF biomarkers, the p-Tau/Aβ_{42} ratio showed good accuracy in distinguishing AD from non-AD as well as in predicting MCI conversion to AD. These results agree with a recent meta-review [15] in which the p-Tau/Aβ_{42} ratio was found to be the most accurate CSF measure for differentiating patients with AD from controls and those with other dementias, and also for predicting MCI conversion to AD. In our study, the p-Tau/Aβ_{42} ratio, however, did not reach significance in the Cox regression analysis, indicating that CSF measures are less effective predictors of MCI conversion than imaging biomarkers [55, 56]. Moreover, the p-Tau/Aβ_{42} ratio had limited value for the differential diagnosis of dementia, and it showed a low specificity possibly because of the inclusion of patients with AD-related pathology in the non-AD group (i.e. patients with DLB). The CSF profile of DLB, characterized by very low Aβ_{42} with a normal A$\beta_{40/42}$ ratio and a moderate increase in t-Tau and p-Tau levels, is in fact comparable to the CSF profile of AD [57].

Qualitative evaluation of atrophy on MRI was the less informative biomarker both in supporting dementia classification and in predicting MCI conversion to AD. Its low accuracy is probably related to different factors. Above all, the assessment of brain focal atrophy on MRI that is a fully operator-dependent procedure could possibly be biased with respect to

semiquantitative and quantitative procedures. Semiquantitative and quantitative MRI approaches, however, are still unavailable for measurement at the single-subject level in the clinical setting because of the lack of validated cut-off values. In addition, the diagnostic value of the recognition of medial temporal lobe atrophy on MRI is very limited as it is a common neuroradiological finding in other neurodegenerative neurological conditions, including frontotemporal dementia [16, 25, 26, 58], as well as in normal ageing [28].

Statistical analyses in subjects with MCI confirmed that FDG PET imaging is the most useful biomarker for predicting further progression to dementia [14]. In particular, FDG PET SPM t-maps were able to identify at the single-subject level dysfunctional brain metabolic patterns typical of different dementia conditions. The high accuracy of FDG PET SPM t-maps in identifying heterogeneous hypometabolic patterns that are predictors of conversion to AD as well as to non-AD dementia, as also shown in a recent study [14], confirms their high diagnostic value in this clinical context and suggests their role in avoiding multiple examinations in subjects with MCI over months and years, which may lead to unnecessary delay in proper clinical management.

Conclusion

Considering the great overlap in clinical presentation among neurodegenerative disorders, diagnosis of dementia may be a difficult task in clinical practice, particularly in the early phase. Clinical and neuropsychological information alone may lead to a significant number of diagnostic errors and uncertainties, thus to unnecessary or inappropriate treatments. In this context, supportive findings from effective biomarkers have great value in helping to reach a more accurate early differential diagnosis. Here, in the routine clinical setting of our memory clinics, we have shown that the use of a validated and standardized method for semiquantitative assessment of FDG PET scans offers a solid support for diagnosis at the individual level with high accuracy and confidence.

Future multicentre studies, including larger samples of dementia patients and MCI subjects, will increase the generalizability of the present findings. The establishment of evidence-based standardized operation procedures is necessary for the correct use of biomarkers in clinical practice.

Compliance with ethical standards

Funding This research was funded by EU FP7 INMIND Project (FP7-HEALTH-2013, grant agreement no. 278850), the Italian Ministry of Health (Ricerca Finalizzata Progetto Reti Nazionale AD NET-2011-02346784), and Fondazione Eli-Lilly (Eli-Lilly grant 2011 "Imaging of neuroinflammation and neurodegeneration in prodromal and presymptomatic Alzheimer's disease phases").

Conflicts of interest None.

Ethical approval For this type of study formal consent is not required.

References

1. Prince M, Bryce R, Ferri C. World Alzheimer report 2011: the benefits of early diagnosis and intervention. London: Alzheimer's Disease International; 2011.
2. Geldmacher DS, Kirson NY, Birnbaum HG, Eapen S, Kantor E, Cummings AK, et al. Implications of early treatment among Medicaid patients with Alzheimer's disease. Alzheimers Dement. 2014;10:214–24.
3. Gaugler JE, Ascher-Svanum H, Roth DL, Fafowora T, Siderowf A, Beach TG. Characteristics of patients misdiagnosed with Alzheimer's disease and their medication use: an analysis of the NACC-UDS database. BMC Geriatr. 2013;13:137.
4. 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;7:270–9.
5. 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;7:280–92.
6. McKeith IG, Dickson DW, Lowe J, Emre M, O'Brien JT, Feldman H, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. Neurology. 2005;65:1863–72.
7. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain. 2011;134:2456–77.
8. Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, et al. Classification of primary progressive aphasia and its variants. Neurology. 2011;76:1006–14.
9. Armstrong MJ, Litvan I, Lang AE, Bak TH, Bhatia KP, Borroni B, et al. Criteria for the diagnosis of corticobasal degeneration. Neurology. 2013;80:496–503.
10. Dubois B, Feldman HH, Jacova C, Hampel H, Molinuevo JL, Blennow K, et al. Advancing research diagnostic criteria for Alzheimer ' s disease : the IWG-2 criteria. Lancet Neurol. 2014;13(6):614–29.
11. Shaw LM, Korecka M, Clark CM, Lee VM-Y, Trojanowski JQ. Biomarkers of neurodegeneration for diagnosis and monitoring therapeutics. Nat Rev Drug Discov. 2007;6:295–303.
12. 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;17:302–16.
13. Anchisi D, Borroni B, Franceschi M, Kerrouche N, Kalbe E, Beuthien-Beumann B, et al. Heterogeneity of brain glucose metabolism in mild cognitive impairment and clinical progression to Alzheimer disease. Arch Neurol. 2005;62:1728–33.
14. Cerami C, Della Rosa PA, Magnani G, Santangelo R, Marcone A, Cappa SF, et al. Brain metabolic maps in Mild Cognitive

Impairment predict heterogeneity of progression to dementia. Neuroimage Clin. 2014;7:187–94.
15. Ferreira D, Perestelo-Pérez L, Westman E, Wahlund LO, Sarría A, Serrano-Aguilar P. Meta-review of CSF core biomarkers in Alzheimer's disease: the state-of-the-art after the new revised diagnostic criteria. Front Aging Neurosci. 2014;6:47.
16. Frisoni GB, Fox NC, Jack CR, Scheltens P, Thompson PM. The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol. 2010;6:67–77.
17. Nordberg A, Carter SF, Rinne J, Drzezga A, Brooks DJ, Vandenberghe R, et al. A European multicentre PET study of fibrillar amyloid in Alzheimer's disease. Eur J Nucl Med Mol Imaging. 2013;40:104–14.
18. Lehmann M, Ghosh PM, Madison C, Laforce R, Corbetta-Rastelli C, Weiner MW, et al. Diverging patterns of amyloid deposition and hypometabolism in clinical variants of probable Alzheimer's disease. Brain. 2013;136:844–58.
19. Murray J, Tsui WH, Li Y, Mchugh P, Williams S, Pirraglia E, et al. FDG and amyloid PET in cognitively normal individuals at risk for late-onset Alzheimer's disease. Adv J Mol Imaging. 2014;4:15–26.
20. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack Jr CR, 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;7:263–9.
21. Coppi E, Ferrari L, Santangelo R, Caso F, Pinto P, Passerini G, et al. Further evidence about the crucial role of CSF biomarkers in diagnosis of posterior cortical atrophy. Neurol Sci. 2014;35:785–7.
22. Santangelo R, Coppi E, Ferrari L, Bernasconi MP, Pinto P, Passerini G, et al. Cerebrospinal fluid biomarkers can play a pivotal role in the diagnostic work up of primary progressive aphasia. J Alzheimers Dis. 2015;43:1429–40.
23. Ferreira D, Rivero-Santana A, Perestelo-Pérez L, Westman E, Wahlund L-O, Sarría A, et al. Improving CSF biomarkers' performance for predicting progression from mild cognitive impairment to Alzheimer's disease by considering different confounding factors: a meta-analysis. Front Aging Neurosci. 2014;6:287.
24. Blennow K, Dubois B, Fagan AM, Lewczuk P, de Leon MJ, Hampel H. Clinical utility of cerebrospinal fluid biomarkers in the diagnosis of early Alzheimer's disease. Alzheimers Dement. 2015;11:58–69.
25. Jack CR, Dickson DW, Parisi JE, Xu YC, Cha RH, O'Brien PC, et al. Antemortem MRI findings correlate with hippocampal neuropathology in typical aging and dementia. Neurology. 2002;58:750–7.
26. Barkhof F, Polvikoski TM, Van Straaten EC, Kalaria RN, Sulkava R, Aronen HJ, et al. The significance of medial temporal lobe atrophy: a postmortem MRI study in the very old. Neurology. 2007;69:1521–7.
27. Hodges JR. Alzheimer's disease and the frontotemporal dementias: contributions to clinico-pathological studies, diagnosis, and cognitive neuroscience. J Alzheimers Dis. 2013;33:S211–7.
28. Raji CA, Lopez OL, Kuller LH, Becker JT. Age, Alzheimer disease, and brain structure. Neurology. 2009;73(22):1899–905.
29. Perani D. Functional neuroimaging of cognition. Handb Clin Neurol. 2008;88:61–111.
30. Perani D. FDG-PET and amyloid-PET imaging: the diverging paths. Curr Opin Neurol. 2014;27:405–13.
31. Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Green RC, et al. The Alzheimer's Disease Neuroimaging Initiative: a review of papers published since its inception. Alzheimers Dement. 2013;9:e111–94.
32. Sánchez-Juan P, Ghosh PM, Hagen J, Gesierich B, Henry M, Grinberg LT, et al. Practical utility of amyloid and FDG-PET in an academic dementia center. Neurology. 2014;82:230–8.
33. Perani D, Della Rosa PA, Cerami C, Gallivanone F, Fallanca F, Vanoli EG, et al. Validation of an optimized SPM procedure for FDG-PET in dementia diagnosis in a clinical setting. Neuroimage. 2014;6:445–54.
34. Ossenkoppele R, Prins ND, Pijnenburg YAL, Lemstra AW, van der Flier WM, Adriaanse SF, et al. Impact of molecular imaging on the diagnostic process in a memory clinic. Alzheimers Dement. 2013;9:414–21.
35. Cerami C, Crespi C, Della Rosa PA, Dodich A, Marcone A, Magnani G, et al. Brain changes within the visuo-spatial attentional network in posterior cortical atrophy. J Alzheimers Dis. 2015;43:385–95.
36. Frisoni GB, Bocchetta M, Chételat G, Rabinovici GD, de Leon MJ, Kaye J, et al. Imaging markers for Alzheimer disease: which vs how. Neurology. 2013;81:487–500.
37. Frisoni GB, Perani D, Bastianello S, Bernardi G, Cappa SF, Trabucchi M. A roadmap to the use of biomarkers for the diagnosis of Alzheimer's disease in clinical practice: the Italian inter-societal consensus. Document based on a workshop held at the 3rd National Health Research Conference, Cernobbio, Como, 12 November 2012, organized by the Directorate for Research of the Italian Ministry of Health. http://www.centroalzheimer.org/iw/pdf/italian_roadmap.pdf. Accessed 14 August 2015.
38. Li Y, Rinne JO, Mosconi L, Pirraglia E, Rusinek H, DeSanti S, et al. Regional analysis of FDG and PIB-PET images in normal aging, mild cognitive impairment, and Alzheimer's disease. Eur J Nucl Med Mol Imaging. 2008;35:2169–81.
39. Edison P, Archer HA, Hinz R, Hammers A, Pavese N, Tai YF, et al. Amyloid, hypometabolism, and cognition in Alzheimer disease: an [11C]PIB and [18F]FDG PET study. Neurology. 2007;68:501–8.
40. Kasanuki K, Iseki E, Fujishiro H, Yamamoto R, Higashi S, Minegishi M, et al. Neuropathological investigation of the hypometabolic regions on positron emission tomography with [18F]fluorodeoxyglucose in patients with dementia with Lewy bodies. J Neurol Sci. 2012;314:111–9.
41. Landau SM, Harvey D, Madison CM, Reiman EM, Foster NL, Aisen PS, et al. Comparing predictors of conversion and decline in mild cognitive impairment. Neurology. 2010;75:230–8.
42. Prestia A, Caroli A, van der Flier WM, Ossenkoppele R, Van Berckel B, Barkhof F, et al. Prediction of dementia in MCI patients based on core diagnostic markers for Alzheimer disease. Neurology. 2013;80:1048–56.
43. Zhang D, Wang Y, Zhou L, Yuan H, Shen D. Multimodal classification of Alzheimer's disease and mild cognitive impairment. Neuroimage. 2011;55:856–67.
44. 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;30:285–92.
45. Shaffer JL, Petrella JR, Sheldon FC, Choudhury KR, Calhoun VD, Coleman RE, et al. Predicting cognitive decline in subjects at risk for Alzheimer disease by using combined cerebrospinal fluid, MR imaging, and PET biomarkers. Radiology. 2013;266:583–91.
46. Choo IH, Ni R, Schöll M, Wall A, Almkvist O, Nordberg A. Combination of (18)F-FDG PET and cerebrospinal fluid biomarkers as a better predictor of the progression to Alzheimer's disease in mild cognitive impairment patients. J Alzheimers Dis. 2013;33:929–39.
47. Alexopoulos P, Kriett L, Haller B, Klupp E, Gray K, Grimmer T, et al. Limited agreement between biomarkers of neuronal injury at different stages of Alzheimer's disease. Alzheimers Dement. 2014;10:684–9.
48. Della Rosa PA, Cerami C, Gallivanone F, Prestia A, Caroli A, Castiglioni I, et al. A standardized [18F]-FDG-PET template for

spatial normalization in statistical parametric mapping of dementia. Neuroinformatics. 2014;12:575–93.
49. Petersen RC, Roberts RO, Knopman DS, Boeve BF, Geda YE, Ivnik RJ, et al. Mild cognitive impairment: ten years later. Arch Neurol. 2009;66:1447–55.
50. Tapiola T, Alafuzoff I, Herukka SK, Parkkinen L, Hartikainen P, Soininen H, et al. Cerebrospinal fluid {beta}-amyloid 42 and tau proteins as biomarkers of Alzheimer-type pathologic changes in the brain. Arch Neurol. 2009;66:382–9.
51. Jack Jr CR, Albert MS, Knopman DS, McKhann GM, Sperling RA, Carrillo MC, et al. Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011;7:257–62.
52. Salmon E, Garraux G, Delbeuck X, Collette F, Kalbe E, Zuendorf G, et al. Predominant ventromedial frontopolar metabolic impairment in frontotemporal dementia. Neuroimage. 2003;20:435–40.
53. Foster NL, Heidebrink JL, Clark CM, Jagust WJ, Arnold SE, Barbas NR, et al. FDG-PET improves accuracy in distinguishing frontotemporal dementia and Alzheimer's disease. Brain. 2007;130:2616–35.
54. Gorno-Tempini ML, Dronkers NF, Rankin KP, Ogar JM, Phengrasamy L, Rosen HJ, et al. Cognition and anatomy in three variants of primary progressive aphasia. Ann Neurol. 2004;55:335–46.
55. Fellgiebel A, Scheurich A, Bartenstein P, Müller MJ. FDG-PET and CSF phospho-tau for prediction of cognitive decline in mild cognitive impairment. Psychiatry Res Neuroimaging. 2007;155:167–71.
56. Walhovd KB, Fjell AM, Brewer J, McEvoy LK, Fennema-Notestine C, Hagler DJ, et al. Combining MR imaging, positron-emission tomography, and CSF biomarkers in the diagnosis and prognosis of Alzheimer disease. AJNR Am J Neuroradiol. 2010;31:347–54.
57. Kaerst L, Kuhlmann A, Wedekind D, Stoeck K, Lange P, Zerr I. Using cerebrospinal fluid marker profiles in clinical diagnosis of dementia with Lewy bodies, Parkinson's disease, and Alzheimer's disease. J Alzheimers Dis. 2014;38:63–73.
58. Van de Pol LA, Hensel A, Barkhof F, Gertz HJ, Scheltens P, van der Flier WM. Hippocampal atrophy in Alzheimer disease: age matters. Neurology. 2006;66:236–8.