Early identification of MCI converting to AD: a FDG PET study

Marco Pagani[1,2] · Flavio Nobili[3] · Silvia Morbelli[4] · Dario Arnaldi[3] · Alessandro Giuliani[5] · Johanna Öberg[6] · Nicola Girtler[3,7] · Andrea Brugnolo[3] · Agnese Picco[3] · Matteo Bauckneht[4] · Roberta Piva[4] · Andrea Chincarini[8] · Gianmario Sambuceti[4] · Cathrine Jonsson[9] · Fabrizio De Carli[10]

Abstract

Purpose Mild cognitive impairment (MCI) is a transitional pathological stage between normal ageing (NA) and Alzheimer's disease (AD). Although subjects with MCI show a decline at different rates, some individuals remain stable or even show an improvement in their cognitive level after some years. We assessed the accuracy of FDG PET in discriminating MCI patients who converted to AD from those who did not.

Methods FDG PET was performed in 42 NA subjects, 27 MCI patients who had not converted to AD at 5 years (nc-MCI; mean follow-up time 7.5 ± 1.5 years), and 95 MCI patients who converted to AD within 5 years (MCI-AD; mean conversion time 1.8 ± 1.1 years). Relative FDG uptake values in 26 meta-volumes of interest were submitted to ANCOVA and support vector machine analyses to evaluate regional differences and discrimination accuracy.

Results The MCI-AD group showed significantly lower FDG uptake values in the temporoparietal cortex than the other two groups. FDG uptake values in the nc-MCI group were similar to those in the NA group. Support vector machine analysis discriminated nc-MCI from MCI-AD patients with an accuracy of 89% (AUC 0.91), correctly detecting 93% of the nc-MCI patients.

Conclusion In MCI patients not converting to AD within a minimum follow-up time of 5 years and MCI patients converting within 5 years, baseline FDG PET and volume-based analysis identified those who converted with an accuracy of 89%. However, further analysis is needed in patients with amnestic MCI who convert to a dementia other than AD.

Keywords Alzheimer's disease · Mild cognitive impairment · Conversion to AD · Positron emission tomography · Support vector machine · Volume of interest analysis

✉ Marco Pagani
marco.pagani@istc.cnr.it

[1] Institute of Cognitive Sciences and Technologies, CNR, Via Palestro 32, 00185 Rome, Italy

[2] Department of Nuclear Medicine, Karolinska Hospital Stockholm, Stockholm, Sweden

[3] Clinical Neurology, Department of Neuroscience (DINOGMI), University of Genoa and IRCCS AOU San Martino-IST, Genoa, Italy

[4] Department of Nuclear Medicine, Department of Health Science (DISSAL), University of Genoa and IRCCS AOU San Martino-IST, Genoa, Italy

[5] Environment and Health Department, Istituto Superiore di Sanità, Rome, Italy

[6] Department of Hospital Physics, Karolinska Hospital, Stockholm, Sweden

[7] Clinical Psychology, IRCCS AOU San Martino-IST, Genoa, Italy

[8] National Institute of Nuclear Physics (INFN), Genoa section, Genoa, Italy

[9] Medical Radiation Physics and Nuclear Medicine, Imaging and Physiology, Karolinska University Hospital, Stockholm, Sweden

[10] Institute of Molecular Bioimaging and Physiology, CNR - Genoa Unit, AOU San Martino-IST, Genoa, Italy

Introduction

One of the main clinical issues in a memory clinic is to predict whether or not a subject presenting with objective mild cognitive impairment (MCI) will convert to dementia. MCI patients convert at an average rate of 10–17% per year [1–3], yet a substantial proportion of them remain stable or improve after some years [4, 5]. These 'nonconverter' MCI patients (nc-

MCI) represent 20% to 40% of the MCI cohort [6] and are affected by a variety of conditions that may mimic a neurodegenerative disease [7]. In clinical practice, neuropsychological assessment tools might differentiate both normal subjects and those with only subjective cognitive impairment from MCI patients.

Structural and functional neuroimaging have been used to discriminate normal subjects from MCI patients who convert to dementia, mainly Alzheimer's disease (MCI-AD), with relatively high accuracy [8–10]. On the contrary, distinguishing MCI-AD from nc-MCI patients remains a major challenge. The identification of MCI patients at imminent risk of conversion may improve patient management by providing the opportunity to benefit from newly developed drugs, as well as to avoid the implementation of inappropriate therapies. To answer this crucial question, both unimodal [8, 11–15] and multimodal [8, 15–20] approaches have been applied, and both approaches have shown accuracies in discriminating the two conditions of up to 83%. The analyses performed included interviewing patients and care-givers, neuropsychological and genetic tests, cerebrospinal protein analysis, and pathological, anatomical and functional neuroimaging. Among these tools, ^{18}F FDG PET may ideally assist in assessing synapse dysfunction which is correlated with cognitive impairment [21]. Moreover, FDG PET has been found to be the only examination that significantly improves the value of demographic covariates in predicting the development of AD [15], and to be a better predictor of conversion than MRI [22]. Specifically, FDG PET alone has shown accuracies in predicting progression of MCI to AD ranging between 70% and 83% [2, 13, 15, 23, 24], which, although statistically significant, is not considered sufficient to provide adequate assistance in clinical diagnosis or as a target for drug trials.

There are several reasons that partly explain such lack of consistency. These tend to negatively affect the value of multicentre studies, and include differences in image processing methods, patient heterogeneity, study design and unsatisfactory camera performance. However, the most important reason may be limited patient follow-up that may prevent timely assessment of AD conversion and minimization of the risk of late conversion. This last effect could hamper the selection of appropriate conversion-related regions, since nc-MCI may also show hypometabolism in critical areas [25].

Applying multivariate analysis techniques based on principal component analysis [26] and independent component analysis [27], our group has recently found progressive disgregation of brain networks in the transition from normal ageing (NA) to MCI and AD. Our studies have also shown the potential of FDG PET for discriminating patients with nc-MCI from those with MCI-AD and full-blown AD. Accuracies close to 90% have been reported, and even higher values have been found for the theoretical condition of repeated PET examinations. These results support the development of new statistical tools for assisting diagnosis in clinical practice, aiding visual inspection and providing high accuracy in the prognosis of MCI evolution, based on evaluation of baseline regional FDG uptake.

The aims of the present study were to: (1) assess the accuracy of an analysis based on volumes of interest (VOIs) in discriminating between nc-MCI and AD-MCI patients, and (2) identify the regions that mostly contribute to the differentiation of the two groups. For reliable discrimination, the classification was based on a follow-up of at least 5 years after the baseline PET examination.

Materials and methods

Participants

The study population consisted of 164 subjects divided into the following three groups: 42 normal elderly subjects (NA), 27 MCI patients who had not converted to AD after a follow-up of at least 5 years after the first FDG PET scan (nc-MCI; mean follow-up 7.5 ± 1.5 years, range 5.0–9.8 years), and 95 MCI patients who converted to AD within 5 years of the baseline PET scan (MCI-AD; mean time to conversion 1.8 ± 1.1 years, range 0.4–5.0 years; Table 1). This retrospective cohort study involved patients who had had amnestic MCI (a-MCI) since 2007. Since our aim was to evaluate the fate of those developing AD dementia in relation to those who remained stable or normalized during follow-up, we chose not to consider a-MCI patients who converted to dementias other than AD. The reasons of this choice were twofold. First, although a sufficient number of a-MCI patients (95) converted to AD dementia for statistical analysis, in the same period only seven patients converted to dementia with Lewy bodies (DLB) and five to frontotemporal dementia (FTD) making these groups too small for statistical analysis. Second, the study was specifically designed to investigate AD-related metabolic changes.

All patients had given their consent to undergo FDG PET within the framework of a longitudinal, long-term observational study. The exclusion criteria have previously been described [26]. Formal permission to conduct the study was given by the Institutional Ethics Committee, in accordance with the principles of the Declaration of Helsinki.

MCI patients

The MCI patients were subjects referred to our memory clinic for a first diagnostic assessment of a memory complaint. The baseline evaluation included standard blood and urine tests, and morphological (MRI) and functional (FDG PET) neuroimaging. The patients underwent extensive neuropsychological examination including tests of: (1) categorical and

Table 1 Demographic data

Group	Education (years)	Age at PET (years)	MMSE score[a]	Gender (male/female)
NA	10.0 ± 4.1	68.8 ± 9.7	29.1 ± 0.9	11/31
nc-MCI	8.9 ± 3.7	71.9 ± 6.4	26.8 ± 1.5	15/12
MCI-AD	10.1 ± 2.1	75.2 ± 5.4	26.0 ± 1.0	31/64

MMSE Mini-Mental State Examination

The data are presented as means ± SD, except gender as numbers of males and females

[a] Normalized to level of education

phonological verbal fluency, (2) episodic verbal memory, (3) executive functions, (4) visuospatial ability, and (5) attention and working memory. According to the MCI criteria of Petersen et al. patients who showed impairment in a memory test but were not demented [30] were included in the MCI group, either with or without involvement of other cognitive domains (multidomain and single-domain a-MCI, respectively) [28]. A 15-item Geriatric Depression Scale (GDS) score of ≤10 was required for inclusion, yet a depressed trait was not considered an exclusion criterion. Patients with MRI evidence of a major stroke or brain lesions and those meeting the criteria for vascular cognitive impairment were excluded [29]. By contrast, a Wahlund score of <3 in all regions [30], and the presence of white matter hyperintensities, leukoaraiosis and lacunae were not exclusion criteria. All patients had a modified Hachinski ischaemic score of <3 [31].

The 27 nc-MCI patients who were followed up for a minimum of 5 years received a different diagnosis at the last follow-up. Four patients underwent an amyloid PET scan, and three of these patients were positive for amyloidosis. Two more patients showed progressive deterioration, although this did not meet the criteria for dementia; amyloid PET imaging was not performed, as both patients were older than 80 years. These two patients, as well as the three with brain amyloidosis (i.e. predementia AD according to the International Working Group-2 research criteria [32]), were considered to be still at risk for the development of AD (very late converters). Two patients had very low levels of school education and were socially and environmentally deprived, and these factors most likely contributed to their poor neuropsychological performance. In the remaining 20 patients, a definite diagnosis was not reached. In these patients, a combination of conditions, including moderate vascular lesions, medical comorbidities, drug therapies and psychosocial factors, or the presence of suspected non-AD pathology may have accounted for their ongoing MCI.

Controls

The control subjects were healthy volunteers carefully evaluated by clinical examination and by their general medical history. Only subjects with a normal Mini-Mental State Examination (MMSE) score (i.e. >26) and with a Clinical Dementia Rating score of 0 were considered. They were chosen so that their age, gender and educational level distributions matched those of the patients. They underwent both FDG PET and MRI scans and were subjected to extended neuropsychological testing. In order to exclude the presence of impairment in specific cognitive domains, a z-score threshold of −1.5 adjusted for age and education level computed using Italian normative values for each test was applied [33]. Of the control subjects, 25 (60%) were followed up with both clinical and neuropsychological examinations after 12 to 107 months (mean 41.8 ± 31 months), which confirmed their healthy status. No further follow-up data were available in 17 control subjects.

Data acquisition and preprocessing

^{18}F-FDG PET protocol and SPM preprocessing.

PET Images were acquired using a Siemens Biograph 16 PET/CT scanner. Scans were acquired in 3-D mode with an acquisition time of 15 min. Images were reconstructed using an ordered subsets expectation maximization algorithm, with 16 subsets and six iterations, and a reconstructed voxel size of 1.33 × 1.33 × 2.00 mm. The CT scan was used for attenuation correction. Brain FDG PET DICOM files were exported and converted into Analyze files. FDG PET images were preprocessed using Statistical Parametric Mapping (SPM8) stand-alone version for spatial normalization into MNI space (Wellcome Department of Cognitive Neurology, London, UK). The spatially normalized set of images were then smoothed with a 10-mm isotropic Gaussian filter and all the default choices of SPM8 were followed, except spatial normalization, for which a customized brain FDG PET template optimized for dementia patients was chosen [34].

Region of interest identification

FDG uptake values were calculated in 45 anatomical VOIs in each hemisphere, as defined by the Automated Anatomical Labeling Atlas [35], and analysed using a Matlab-based script created in-house that automatically processed the mean FDG

uptake value from each of the VOIs bilaterally. The extracted values were then normalized in each subject to the average intensity of the cerebellar VOIs on the basis that the cerebellum is poorly affected by the AD pathological process. In order to decrease the number of variables for statistical analysis, the number of VOIs was reduced by merging regions with similar anatomofunctional characteristics into the following 13 meta-VOIs in each hemisphere:

1. Occipital cortex (calcarine/lingual/inferior occipital/middle occipital/superior occipital gyri)
2. Putamen/pallidum/caudate
3. Parahippocampal gyrus/amygdala/hippocampus/insula
4. Orbitofrontal cortex (inferior frontal/medial frontal/middle frontal gyri)
5. Frontal cortex (middle frontal/superior frontal/superior-medial frontal/superior-orbital frontal/inferior frontal gyri)
6. Cuneus/fusiform gyrus/precuneus
7. Postcentral gyrus/precentral gyrus/supplementary motor area
8. Parietal lobe (inferior parietal/superior parietal gyri)
9. Anterior cingulate gyrus
10. Posterior cingulate gyrus
11. Temporal lobe (inferior temporal/middle temporal/superior temporal gyri)
12. Temporal pole (middle temporal pole/superior temporal pole gyri)
13. Thalamus

Statistical analysis

Neuropsychological tests

Due to the lack of data in some patients, we were able to reliably compare only some of the administered neuropsychological tests between the two MCI groups, assessing verbal memory and executive functions. One-way analysis of variance was used to evaluate the significance of differences between the MMSE scores, Rey Auditory Verbal Learning Test immediate total recall, Rey Auditory Verbal Learning Test delayed recall, and Trail Making Test Parts A and B.

Analysis of variance

Considering the small age difference between the groups, a preliminary analysis of FDG uptake in relation to patient age was performed in controls using a repeated measures linear model applied to the meta-VOI values. The analysis was applied to controls assuming that the effect of age could be better estimated in this group in the absence of pathological hypometabolism. In order to remove the effect of age, the regression coefficients estimated for controls were used to correct individual data in all groups as a function of the subject's age. Differences between mean regional values in the three groups of patients were evaluated by repeated measures analysis of variance considering one between-subject factor (group), one within-subject factor (region) and their interaction. Post-hoc between-group pair-wise comparisons were performed using the Tukey-Kramer method for multiple comparisons. FDG uptake patterns in the two MCI groups were compared with the pattern in the control subjects to reproduce the clinical routine in which visual or visual plus semiquantitative evaluations are carried out by comparison with the normal metabolic pattern. Deviations in the distribution of FDG in each group from a normal distribution is ordinarily used to identify differences between nc-MCI and AD-MCI patients and provide reference patterns for an immediate diagnostic evaluation.

Support vector machine

To determine the value of meta-VOI-based FDG uptake in predicting the possible onset of AD, we applied a support vector machine (SVM) model using age-corrected baseline data to discriminate between nc-MCI and MCI-AD patients. The SVM model was implemented with a radial basis function (Gaussian) kernel [36] with a constant scale factor derived from a previous study [10]. A stepwise selection procedure was applied to identify the sets of regions that provided the highest discrimination. The performance of the model was evaluated by a cross-validation procedure. Considering the relatively small sample size, the subjects were divided into 21 subgroups (21-fold cross-validation), each including a proportional share of subjects from the two groups (about 5%). The training and testing procedure was repeated 21 times. For each subgroup, the SVM model was fitted to all remaining subjects and then applied to each subject in the subgroup to produce a score measuring group membership, and a consequent classification. Thus an individual score was produced for each investigated subject based on a virtually independent training set, and this enabled a receiver operating characteristic (ROC) curve to be built.

To evaluate the accuracy, sensitivity and specificity of the method, a cut-off value was chosen as the minimum the distance from the upper left corner of the ROC curve (where specificity = sensitivity = 1). Accuracy, sensitivity, specificity and the area under the ROC curve (AUC) were determined along with their confidence intervals (CIs). The Wald interval, with exact binomial probabilities, was used for sensitivity, specificity and accuracy [37], and the CIs for the ROC AUCs were estimated using the bootstrap method described by Qin and Hotilovac [38]. Statistical analyses were performed using the Statistics Toolbox of Matlab R2015b (MathWorks, Natick, MA).

Results

Rey Auditory Verbal Learning Test immediate total recall and Rey Auditory Verbal Learning Test delayed recall showed significant differences between the two MCI groups ($p < 0.05$ and $p < 0.001$, respectively). On the other hand, no significant changes were found in the MMSE scores and Trail Making Test Parts A and B.

Repeated measures analysis of variance showed highly significant differences among the mean values associated with all the effects: group ($F_{2,161}$ 13.65, $p < 0.0001$), region ($F_{25,4025}$ 801.84, $p < 0.0001$) and their interaction ($F_{50,4025}$ 3.54, $p < 0.0001$). Post hoc analysis showed significantly lower mean values in the MCI-AD group than in the two other groups. Table 2 shows the regional mean values in the three groups with associated probabilities of pair-wise differences, as evaluated using the Tukey-Kramer method but uncorrected for multiple regional comparisons. The pattern of regional mean values in the nc-MCI group was close to that in the NA group in most regions, and was intermediate between the patterns in the latter group and the MCI-AD group (Table 2). Most regions, including the parahippocampal gyrus/amygdala/hippocampus/insula, the cuneus/fusiform gyrus/precuneus, the posterior cingulate gyrus, and the parietal and temporal lobes, showed significantly lower values in the MCI-AD group than in the NA group. In some of these regions, a similar decreasing trend was also found with respect to the nc-MCI group and the difference survived Bonferroni correction in the left cuneus/fusiform gyrus/precuneus, the left posterior cingulate gyrus and the right temporal lobe (Table 2).

Application of the SVM model resulted in a fair partition between MCI-AD and nc-MCI patients with a set of only three meta-VOIs (putamen/pallidum/caudate, posterior cingulate gyrus and temporal lobe, all in the left hemisphere; Fig. 1) that, after cross-validation, was able to discriminate the two groups with 82.0% accuracy (CI 75.2–88.8%), 81.1% sensitivity (detection of MCI-AD; CI 73.2–88.9%) and 85.2% specificity (CI 71.8–98.6%), and an AUC of 0.834 (CI 0.742–0.900; Fig. 2). Increasing the number of selected regions from 12 to 20 increased the accuracy of the SVM in discriminating between MCI-AD and nc-MCI patients after cross-validation to about 88%. Including more regions resulted in a gradual decrease in discrimination power, probably due to overfitting and/or to a confounding effect of uninfluential regions.

The best solution was found with the following set of 13 meta-VOIs: occipital cortex and cuneus/fusiform gyrus/precuneus bilaterally, thalamus, putamen/pallidum/caudate and orbitofrontal and frontal cortices, anterior cingulate gyrus, and the postcentral gyrus/precentral gyrus/supplementary motor area on the left side, and the right temporal pole, and parietal and temporal lobes (Fig. 3). This model was able to discriminate between AD-MCI and nc-MCI patients with 88.5% accuracy (CI 82.9–94.2%), 87.4% sensitivity (detection of AD-MCI; CI 80.7–94.1%) and 92.6% specificity (detection of nc-MCI; CI 82.7–100.0%), and an AUC of 0.911 (CI 0.835–0.957; Fig. 2).

Among these meta-VOIs, the greatest decreases in mean FDG uptake values in the MCI-AD patients with respect to the nc-MCI patients were found in the right temporal and parietal lobes, and in the cuneus/fusiform gyrus/precuneus of both hemispheres. The difference in the AUCs between the two models (three meta-VOIs and 13 meta-VOIs) was of borderline significance in the Hanley and McNeil test ($z = 1.7$, $p = 0.045$; Fig. 2).

Discussion

Both in clinical routine and in research setting, it is of paramount importance to identify the early hallmarks of conversion of MCI to AD, possibly before the onset of more severe AD-related neurobiological changes and dementia symptoms. In this respect, the use of FDG PET as a tool to predict the progression of neuronal injury is based on its stronger association with cognitive deterioration in comparison with markers of cerebral amyloidosis, the latter being already elevated in the asymptomatic stage and showing very little variation when overt symptoms are already present. Moreover, amyloidosis biomarkers can also be increased from the eighth decade in asymptomatic subjects [39].

The most relevant result of the present study is the identification of large regions (meta-VOIs) in which metabolism was found to be specifically impaired in patients with MCI converting to AD, namely the bilateral cuneus/fusiform gyrus/precuneus, the posterior cingulate gyrus, and the parietal and temporal lobes (regions that are known to be implicated in AD pathology). FDG uptake in these areas was clearly reduced in MCI-AD patients, not only compared with uptake in NA patients, but also with that in nc-MCI patients, whose values were found to be similar to those in NA patients (Table 2). Metabolic data from brain regions of 122 MCI patients were segmented using a freely available anatomical atlas and extracted by a Matlab-based script developed in-house that pooled them into 13 meta-VOIs, bilaterally, which were then analysed using the SVM. In clinical routine, this would enable results to be interpreted in a very short time at the level of the individual patient, since the processing time is shorter than with similar analyses using predefined VOIs where automation procedures may be more complex [40].

The proposed methodology was able to discriminate the two MCI groups, that showed very similar neuropsychological profiles as early as at the baseline examination, with an overall accuracy of 89% and a well-balanced

Table 2 Mean uptake values in 13 merged regions in each hemisphere (normalized to the cerebellar region) in each patient group and evaluation of the differences between the values in each group by repeated measures analysis of variance and Tukey-Kramer-based post-hoc analyses

Region	Patient group			Probability values		
	NA	nc-MCI	MCI-AD	NA vs. nc-MCI	NA vs. MCI-AD	Nc-MCI vs. MCI-AD
Amygdala/hippocampus/insula						
Left	0.852	0.834	0.816	0.200	**0.0000**[a]	0.134
Right	0.900	0.875	0.857	0.104	**0.0000**[a]	0.246
Cuneus/fusiform gyrus/precuneus						
Left	1.175	1.173	1.117	0.993	**0.0000**[a]	**0.0008**[a]
Right	1.141	1.132	1.078	0.878	**0.0000**[a]	**0.0028**
Anterior cingulate gyrus						
Left	1.055	1.029	0.992	0.462	**0.0003**[a]	0.124
Right	0.949	0.923	0.913	0.224	**0.0069**	0.760
Posterior cingulate gyrus						
Left	1.159	1.137	1.048	0.546	**0.0000**[a]	**0.0000**[a]
Right	0.828	0.833	0.784	0.962	**0.0026**	**0.0051**
Frontal cortex						
Left	1.093	1.084	1.043	0.913	**0.0038**	0.061
Right	1.039	1.024	0.986	0.713	**0.0009**[a]	0.075
Occipital cortex						
Left	1.193	1.196	1.163	0.988	0.091	0.126
Right	1.147	1.143	1.094	0.983	**0.0010**[a]	**0.0121**
Orbitofrontal cortex						
Left	1.132	1.112	1.081	0.470	**0.0002**[a]	0.087
Right	0.994	0.974	0.947	0.391	**0.0002**[a]	0.127
Putamen/pallidum/caudate						
Left	0.943	0.907	0.906	**0.0382**	**0.0018**[a]	0.990
Right	1.092	1.058	1.056	0.116	**0.0143**	0.990
Parietal lobe						
Left	1.125	1.122	1.055	0.990	**0.0004**[a]	**0.0056**
Right	1.051	1.041	0.978	0.909	**0.0001**[a]	**0.0067**
Postcentral/precentral gyri						
Left	1.133	1.147	1.110	0.804	0.411	0.172
Right	1.083	1.088	1.065	0.971	0.476	0.424
Temporal lobe						
Left	1.078	1.058	1.005	0.509	**0.0000**[a]	**0.0030**
Right	1.028	1.010	0.954	0.501	**0.0000**[a]	**0.0004**[a]
Temporal pole						
Left	0.824	0.807	0.771	0.500	**0.0000**[a]	**0.0277**
Right	0.720	0.706	0.672	0.488	**0.0000**[a]	**0.0035**
Thalamus						
Left	1.018	0.994	0.964	0.383	**0.0002**[a]	0.147
Right	1.061	1.045	1.022	0.578	**0.0044**	0.263

Values indicating remarkable differences uncorrected for multiple comparisons are highlighted in bold

[a] Values lower than the Bonferroni-corrected threshold

sensitivity and specificity, and correctly classified 87% of patients converting to AD and 93% of patients not converting to AD. These values are better than those obtained in similar studies in which FDG PET was the only

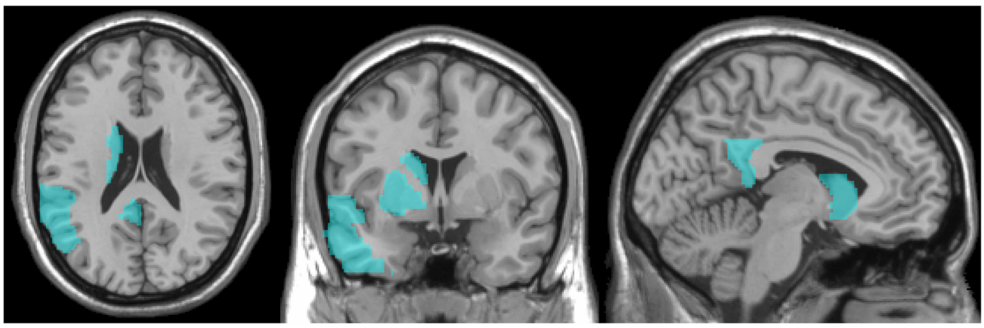

Fig. 1 Topographic representation of the three meta-VOIs set in the left hemisphere with the best discrimination ability superimposed on the Montreal Neurological Institute template in the transverse (*left*), coronal (*centre*) and sagittal (*right*) views

feature under investigation and predictions of conversion were not higher than 83% [2, 13, 15, 24].

Identification of MCI subjects who will develop AD is not a trivial task, since demographic [15, 41], neuropsychological [16, 42, 43], structural and longitudinal [11, 12], and functional [13] variables are involved, and automated multimodal classification systems, as well as algorithms, have often been used to integrate data [44]. However, to the best of our knowledge, the reported accuracies obtained with computer-assisted support tools are around 80% or lower [16–18, 20, 43, 45], including when investigating the conversion from no cognitive impairment to MCI [46]. Such reported accuracies although promising, yet do not put an end to the question of their prognostic use in routine clinical setting .

Structural [47, 48] and functional [2, 13, 15] investigation have shown significant differences between nc-MCI and AD-MCI patients in various regions, mainly the limbic system and the temporoparietal cortex. Structural and metabolic changes in these regions have been found to be correlated with age of onset, symptom severity and time of conversion [41]. In a comprehensive meta-analysis of seven neuroimaging studies involving the use of different modalities, Schroeter et al. [49] identified the inferior parietal lobule (Brodmann areas 7, 39 and 40) and the precuneus (Brodmann areas 7 and 31) as the regions most useful for discriminating converting and nonconverting MCI patients. Structural and functional differences were found in these two regions and in the hippocampal/parahippocampal and orbitofrontal cortices between controls and both converting and nonconverting MCI patients.

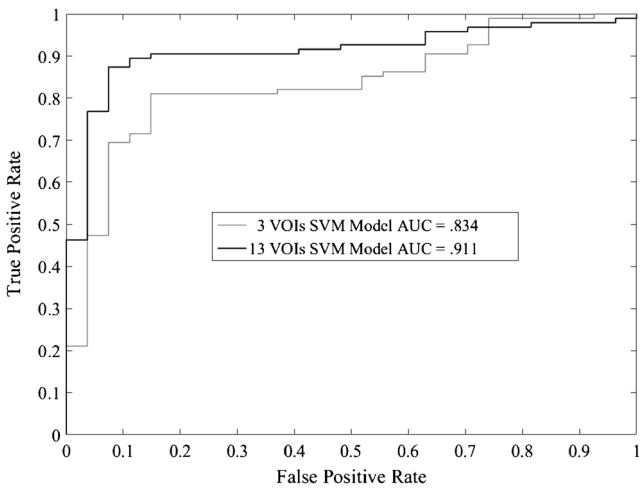

Fig. 2 Receiver operating characteristic curves showing the accuracy of the SVM model in discriminating between non-converter MCI patients and MCI patients who subsequently converted to dementia. The *grey curve* was obtained using three selected regions; the *black curve* was obtained using 13 selected regions (see text)

More recently, Teipel et al. [50], analysing pathological, structural and functional data using a logistic regression model, found an accuracy after cross-validation of 72% in the prediction of conversion. The authors found that the bilateral parietotemporal and posterior cingulate cortices, precuneus and medial temporal lobe had lower metabolic rates in MCI-AD patients than in nc-MCI patients. Furthermore, as with FDG PET, regions mostly predicting time to conversion were the hippocampus, anterior cingulate cortex, and the posterior temporal and occipital lobes. All these findings from large multicentre databases of MRI, SPECT and PET data confirm the differences between MCI-AD and nc-MCI patients, yet overlap only partially with our results. We found a remarkable decrease in uptake levels only in the left basal ganglia in nc-MCI patients in relation to NA subjects, but the probability level did not reach the Bonferroni-corrected threshold, while only a decreasing trend was observed in other regions.

What is especially intriguing is that we found significantly different metabolic rates in most of the association cortices between MCI-AD patients and NA subjects, while the differences between MCI-AD and nc-MCI patients were focused on a smaller set of regions, where values in the nc-MCI patients were not significantly different from those in NA subjects. Regions surviving subtraction of the differences between the nc-MCI patients and NA subjects were the most robust predictors of conversion; these regions included the parietal lobe, the posterior cingulate gyrus, the temporal pole and lobe, and the cuneus/fusiform gyrus/precuneus (Table 2). Differences in metabolic values in other regions, including the medial temporal lobe (parahippocampal gyrus/amygdala/hippocampus/insula) and the anterior cingulate gyrus, were not sufficient to discriminate between MCI-AD and nc-MCI patients. This finding is especially useful in the clinical context, since other conditions leading to MCI, including depression [51] and

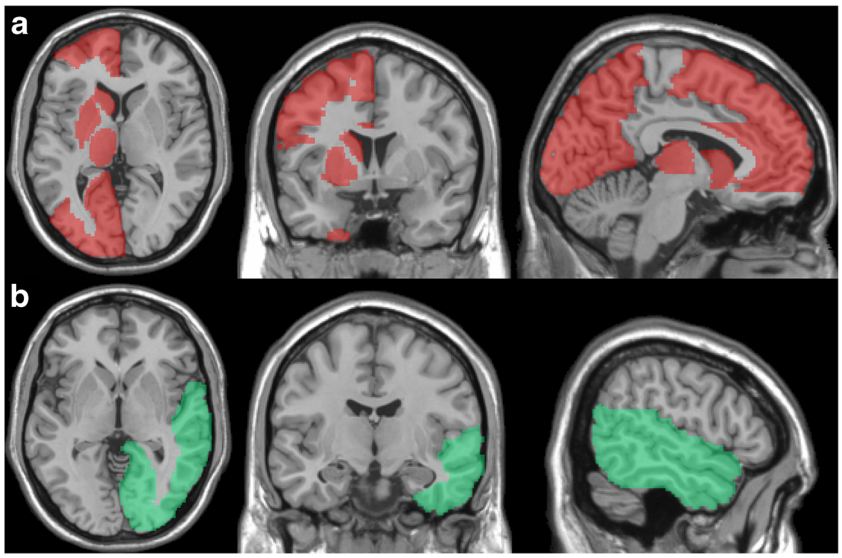

Fig. 3 Topographic representation of the 13 meta-VOI sets with the best discrimination ability superimposed on the Montreal Neurological Institute template. **a** Left hemisphere set (*red*. *Left:* transverse view, *centre:* coronal view, *right:* sagittal view). **b** Right hemisphere set (*green*. *Left*: transverse view, *centre:* coronal view, *right:* sagittal view). See the text for the composition of each meta-VOI

vascular cognitive impairment [52], might affect brain metabolism in these areas.

It is worth noting that the putamen/pallidum/caudate was among the regions selected for SVM classification. This region is generally considered less affected by the pathological AD process, and in this study did not show significant differences between nc-MCI and MCI-AD patients (Table 2). The presence of relatively spared regions has been previously observed [53] and the identification of such regions probably contributes to discrimination in that the differences between affected and spared regions are highlighted, and are apparent in the early and intermediate stages of the pathological process.

One of the strengths of this study was the very homogeneous method of investigation performed by the same clinical group, in which the same camera was used, minimizing the likelihood of data variability and a lack of robustness in the results that would have occurred if inhomogeneous subject samples had been included and acquisitions had been obtained using different cameras. One of the drawbacks of multicentre studies (for example the ADNI, AIBL and BioFINDER studies) is that differences in diagnostic procedures, clinical factors, selection of patients and healthy control cohorts, demographic characteristics, and acquisition protocols may result in potentially confounding effects of uncontrolled variables, hence weakening statistical power. Another strength of our study was the extended follow-up (minimum 5 years) in nc-MCI patients, which, to the best of our knowledge, exceeds that in previous studies. This relatively long follow-up minimized the likelihood of not detecting very late converters (possibly 5 of 122 MCI patients), reinforcing the metabolic pattern in nc-MCI and allowing better differentiation from the other groups, potentially leading to the identification of a robust gold-standard for single-patient discrimination. Considering the high accuracy obtained by the present meta-VOI-based SVM analysis of a highly homogeneous dataset, a further step for its application as an automatic tool in the clinical setting should be its fitting and validation in a larger and multicentre dataset.

This study was specifically designed to investigate the metabolic abnormalities in a-MCI patients converting to AD dementia, while not investigating the metabolic pattern of neurodegenerative dementing disorders other than AD. Indeed, both DLB and FTD can be recognized in the prodromal stage, but prodromal DLB more often manifests as non-amnestic MCI (na-MCI), and prodromal FTD as either behavioural impairment or na-MCI, or both. Moreover, comparative investigations of metabolic patterns specific to neurodegenerative diseases other than AD would require an ad-hoc designed study and other selection criteria, including a-MCI, na-MCI, mild behavioural impairment, and the presence of mild dementia, to achieve suitable numbers for analysis, since both FTD and DLB patients are less likely to be diagnosed in the predementia stage. However, excluding from the analysis a-MCI patients who developed dementia other than AD remains a selection bias that restricts the diagnostic value of our results since these patients might have been missed by FDG PET, or wrongly included in one of the three groups. Therefore, our conclusions are limited to the differential diagnosis between converters and non-converters to AD. Further analysis with an adequate number of patients in the other neurodegenerative dementia classes is needed.

While the combination of different diagnostic tools, ranging from cognitive tests to different neuroimaging biomarkers, can help diagnostic assessment in the clinic, it is worth keeping the different domains separate, instead of trying to pursue a 'globally large automated procedure' that provides a single evaluation score from heterogeneous information sources. This opinion stems from both practical and theoretical considerations. The practical side involves the possible lack of

information needed for the application of the complete model. The theoretical consideration takes into account the need for clinicians to be able to critically evaluate any contrasting evidence without the use of a common index. In this respect, our work attempts to generate an easy to use 'pure PET' index for diagnosis (e.g. by a previous standardization of the individual systems).

Another limitation of the study was the lack of a systematic search for a marker of amyloidosis in nc-MCI patients due to further ageing of several already old participants during follow-up, and because amyloidosis biomarkers were still unavailable in our centre. We also tried to minimize the risk of including in the control group subjects with a preclinical neurodegenerative disease by follow-up clinical and neuropsychological examinations which were available in 60% of subjects after an average time of 41.8 ± 31 months; these examinations confirmed the healthy status of all the control subjects.

In conclusion, our results confirm the prognostic power of FDG PET in patients with MCI assessed at their first visit to a memory clinic with the clinical suspicion of prodromal AD. The optimized semiquantitative procedure correctly identified 88% of patients as either converting to AD dementia or not. In this context, 93% of nc-MCI patients were identified, and this is particularly relevant as these patients represent the natural control group for MCI-AD. Their metabolic distribution pattern was not identical to that in the normal control group, and they thus acted as 'confounders' in the early identification of converters. Finally, there is an urgent need to harmonize the interpretation and reporting of FDG PET scans, as the method adopted for analysis can be as important as the choice of the biomarker itself [54], providing it is also tested in a-MCI patients later developing dementia other than AD. Although several valuable techniques have been proposed to assist the nuclear medicine physician in reporting scans, head-to-head comparisons have led to somewhat disappointing results [40], thus emphasizing the need for a greater effort towards harmonization [55].

Acknowledgments The authors thank Ms. Katja Gasperini for helping with the English editing.

Compliance with ethical standards

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest None.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the principles of the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards

References

1. Davatzikos C, Bhatt P, Shaw LM, Batmanghelich KN, Trojanowski JQ. Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. Neurobiol Aging. 2011;32:2322.e19–e27. doi:10.1016/j.neurobiolaging.2010.05.023.
2. 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. doi:10.1212/WNL.0b013e3181e8e8b8.
3. 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. doi:10.1001/archneurol.2009.266.
4. Pagani M, Dessi B, Morbelli S, Brugnolo A, Salmaso D, Piccini A, et al. MCI patients declining and not-declining at mid-term follow-up: FDG-PET findings. Curr Alzheimer Res. 2010;7:287–94.
5. Hamel R, Kohler S, Sistermans N, Koene T, Pijnenburg Y, van der Flier W, et al. The trajectory of cognitive decline in the pre-dementia phase in memory clinic visitors: findings from the 4C-MCI study. Psychol Med. 2015;45:1509–19. doi:10.1017/s0033291714002645.
6. Mitchell AJ, Shiri-Feshki M. Rate of progression of mild cognitive impairment to dementia – meta-analysis of 41 robust inception cohort studies. Acta Psychiatr Scand. 2009;119:252–65. doi:10.1111/j.1600-0447.2008.01326.x.
7. Barnes DE, Alexopoulos GS, Lopez OL, Williamson JD, Yaffe K. Depressive symptoms, vascular disease, and mild cognitive impairment: findings from the Cardiovascular Health Study. Arch Gen Psychiatry. 2006;63:273–9. doi:10.1001/archpsyc.63.3.273.
8. Arbizu J, Prieto E, Martinez-Lage P, Marti-Climent JM, Garcia-Granero M, Lamet I, et al. Automated analysis of FDG PET as a tool for single-subject probabilistic prediction and detection of Alzheimer's disease dementia. Eur J Nucl Med Mol Imaging. 2013;40:1394–405. doi:10.1007/s00259-013-2458-z.
9. Cuingnet R, Gerardin E, Tessieras J, Auzias G, Lehericy S, Habert MO, et al. Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. Neuroimage. 2011;56:766–81. doi:10.1016/j.neuroimage.2010.06.013.
10. Pagani M, De Carli F, Morbelli S, Oberg J, Chincarini A, Frisoni GB, et al. Volume of interest-based [18F]fluorodeoxyglucose PET discriminates MCI converting to Alzheimer's disease from healthy controls. A European Alzheimer's Disease Consortium (EADC) study. Neuroimage Clin. 2015;7:34–42. doi:10.1016/j.nicl.2014.11.007.
11. Runtti H, Mattila J, van Gils M, Koikkalainen J, Soininen H, Lotjonen J, et al. Quantitative evaluation of disease progression in a longitudinal mild cognitive impairment cohort. J Alzheimers Dis. 2014;39:49–61. doi:10.3233/JAD-130359.
12. Ardekani BA, Bermudez E, Mubeen AM, Bachman AH; Alzheimer's Disease Neuroimaging Initiative. Prediction of incipient Alzheimer's disease dementia in patients with mild cognitive impairment. J Alzheimers Dis. 2017;55:269–81. doi:10.3233/JAD-160594.
13. Lange C, Suppa P, Frings L, Brenner W, Spies L, Buchert R. Optimization of statistical single subject analysis of brain FDG PET for the prognosis of mild cognitive impairment-to-Alzheimer's disease conversion. J Alzheimers Dis. 2016;49:945–59. doi:10.3233/JAD-150814.
14. Lee SH, Bachman AH, Yu D, Lim J, Ardekani BA. Predicting progression from mild cognitive impairment to Alzheimer's disease using longitudinal callosal atrophy. Alzheimers Dement (Amst). 2016;2:68–74. doi:10.1016/j.dadm.2016.01.003.
15. 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. doi:10.1148/radiol.12120010.
16. Liu Y, Mattila J, Ruiz MA, Paajanen T, Koikkalainen J, van Gils M, et al. Predicting AD conversion: comparison between prodromal AD guidelines and computer assisted PredictAD tool. PLoS One. 2013;8:e55246. doi:10.1371/journal.pone.0055246.
17. Young J, Modat M, Cardoso MJ, Mendelson A, Cash D, Ourselin S, et al. Accurate multimodal probabilistic prediction of conversion to Alzheimer's disease in patients with mild cognitive impairment. Neuroimage Clin. 2013;2:735–45. doi:10.1016/j.nicl.2013.05.004.
18. Yu P, Dean RA, Hall SD, Qi Y, Sethuraman G, Willis BA, et al. Enriching amnestic mild cognitive impairment populations for clinical trials: optimal combination of biomarkers to predict conversion to dementia. J Alzheimers Dis. 2012;32:373–85. doi:10.3233/JAD-2012-120832.
19. Ewers M, Brendel M, Rizk-Jackson A, Rominger A, Bartenstein P, Schuff N, et al. Reduced FDG-PET brain metabolism and executive function predict clinical progression in elderly healthy subjects. Neuroimage Clin. 2014;4:45–52. doi:10.1016/j.nicl.2013.10.018.
20. Moradi E, Pepe A, Gaser C, Huttunen H, Tohka J. Alzheimer's disease neuroimaging I. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. Neuroimage. 2015;104:398–412. doi:10.1016/j.neuroimage.2014.10.002.
21. Haense C, Buerger K, Kalbe E, Drzezga A, Teipel SJ, Markiewicz P, et al. CSF total and phosphorylated tau protein, regional glucose metabolism and dementia severity in Alzheimer's disease. Eur J Neurol. 2008;15:1155–62. doi:10.1111/j.1468-1331.2008.02274.x.
22. Walhovd KB, Fjell AM, Dale AM, McEvoy LK, Brewer J, Karow DS, et al. Multi-modal imaging predicts memory performance in normal aging and cognitive decline. Neurobiol Aging. 2010;31:1107–21. doi:10.1016/j.neurobiolaging.2008.08.013.
23. 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. doi:10.1016/j.neuroimage.2011.01.049.
24. Ito K, Fukuyama H, Senda M, Ishii K, Maeda K, Yamamoto Y, et al. Prediction of outcomes in mild cognitive impairment by using 18F-FDG-PET: a multicenter study. J Alzheimers Dis. 2015;45:543–52. doi:10.3233/JAD-141338.
25. Fouquet M, Desgranges B, Landeau B, Duchesnay E, Mezenge F, de la Sayette V, et al. Longitudinal brain metabolic changes from amnestic mild cognitive impairment to Alzheimer's disease. Brain. 2009;132:2058–67. doi:10.1093/brain/awp132.
26. Pagani M, Giuliani A, Oberg J, Chincarini A, Morbelli S, Brugnolo A, et al. Predicting the transition from normal aging to Alzheimer's disease: a statistical mechanistic evaluation of FDG-PET data. Neuroimage. 2016;141:282–90. doi:10.1016/j.neuroimage.2016.07.043.
27. Pagani M, Giuliani A, berg J, De Carli F, Morbelli S, Girtler N, et al. Progressive disintegration of brain networking from normal aging to Alzheimer's Disease. Analysis of Independent Components of 18F-FDG PET Data. J Nucl Med. 2017;58:1132–1139. doi:10.2967/jnumed.116.184309.
28. Petersen RC, Negash S. Mild cognitive impairment: an overview. CNS Spectr. 2008;13:45–53.
29. Gorelick PB, Scuteri A, Black SE, Decarli C, Greenberg SM, Iadecola C, et al. Vascular contributions to cognitive impairment and dementia: a statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2011;42:2672–713. doi:10.1161/STR.0b013e3182299496.
30. Wahlund LO, Barkhof F, Fazekas F, Bronge L, Augustin M, Sjogren M, et al. A new rating scale for age-related white matter changes applicable to MRI and CT. Stroke. 2001;32:1318–22.
31. Loeb C, Gandolfo C. Diagnostic evaluation of degenerative and vascular dementia. Stroke. 1983;14:399–401.
32. 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;13:614–29. doi:10.1016/s1474-4422(14)70090-0.
33. Picco A, Polidori MC, Ferrara M, Cecchetti R, Arnaldi D, Baglioni M, et al. Plasma antioxidants and brain glucose metabolism in elderly subjects with cognitive complaints. Eur J Nucl Med Mol Imaging. 2014;41:764–75. doi:10.1007/s00259-013-2638-x.
34. 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. doi:10.1007/s12021-014-9235-4.
35. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage. 2002;15:273–89. doi:10.1006/nimg.2001.0978.
36. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20:273–97. doi:10.1007/bf00994018.
37. Brown LD, Cai TT, DasGupta A. Interval estimation for a binomial proportion. Statist Sci. 2001;16:101–33. doi:10.1214/ss/1009213286.
38. Qin G, Hotilovac L. Comparison of non-parametric confidence intervals for the area under the ROC curve of a continuous-scale diagnostic test. Stat Methods Med Res. 2008;17:207–21. doi:10.1177/0962280207087173.
39. Ossenkoppele R, Jansen WJ, Rabinovici GD, Knol DL, van der Flier WM, van Berckel BN, et al. Prevalence of amyloid PET positivity in dementia syndromes: a meta-analysis. JAMA. 2015;313:1939–49. doi:10.1001/jama.2015.4669.
40. Caroli A, Prestia A, Chen K, Ayutyanont N, Landau SM, Madison CM, et al. Summary metrics to assess Alzheimer disease-related hypometabolic pattern with 18F-FDG PET: head-to-head comparison. J Nucl Med. 2012;53:592–600. doi:10.2967/jnumed.111.094946.
41. Dukart J, Mueller K, Villringer A, Kherif F, Draganski B, Frackowiak R, et al. Relationship between imaging biomarkers, age, progression and symptom severity in Alzheimer's disease. Neuroimage Clin. 2013;3:84–94. doi:10.1016/j.nicl.2013.07.005.
42. Casanova R, Hsu FC, Sink KM, Rapp SR, Williamson JD, Resnick SM, et al. Alzheimer's disease risk assessment using large-scale machine learning methods. PLoS One. 2013;8:e77949. doi:10.1371/journal.pone.0077949.
43. Hinrichs C, Singh V, Xu G, Johnson SC. Alzheimers disease neuroimaging I. Predictive markers for AD in a multi-modality framework: an analysis of MCI progression in the ADNI population. Neuroimage. 2011;55:574–89. doi:10.1016/j.neuroimage.2010.10.081.
44. Ritter K, Schumacher J, Weygandt M, Buchert R, Allefeld C, Haynes JD. Multimodal prediction of conversion to Alzheimer's disease based on incomplete biomarkers. Alzheimers Dement (Amst). 2015;1:206–15. doi:10.1016/j.dadm.2015.01.006.
45. Sanchez-Catasus CA, Stormezand GN, van Laar PJ, De Deyn PP, Sanchez MA, Dierckx RA. FDG-PET for prediction of AD dementia in mild cognitive impairment. A review of the state of the art with particular emphasis on the comparison with other neuroimaging modalities (MRI and perfusion SPECT). Curr Alzheimer Res. 2017;14:127–42.
46. Rizk-Jackson A, Insel P, Petersen R, Aisen P, Jack C, Weiner M. Early indications of future cognitive decline: stable versus declining controls. PLoS One. 2013;8:e74062. doi:10.1371/journal.pone.0074062.

47. Scola E, Bozzali M, Agosta F, Magnani G, Franceschi M, Sormani MP, et al. A diffusion tensor MRI study of patients with MCI and AD with a 2-year clinical follow-up. J Neurol Neurosurg Psychiatry. 2010;81:798–805. doi:10.1136/jnnp.2009.189639.
48. Wei R, Li C, Fogelson N, Li L. Prediction of conversion from mild cognitive impairment to Alzheimer's disease using MRI and structural network features. Front Aging Neurosci. 2016;8:76. doi:10.3389/fnagi.2016.00076.
49. Schroeter ML, Stein T, Maslowski N, Neumann J. Neural correlates of Alzheimer's disease and mild cognitive impairment: a systematic and quantitative meta-analysis involving 1351 patients. Neuroimage. 2009;47:1196–206. doi:10.1016/j.neuroimage.2009.05.037.
50. Teipel SJ, Kurth J, Krause B, Grothe MJ; Alzheimer's Disease Neuroimaging Initiative. The relative importance of imaging markers for the prediction of Alzheimer's disease dementia in mild cognitive impairment – beyond classical regression. Neuroimage Clin. 2015;8:583–93. doi:10.1016/j.nicl.2015.05.006.
51. Brendel M, Reinisch V, Kalinowski E, Levin J, Delker A, Darr S, et al. Hypometabolism in brain of cognitively normal patients with depressive symptoms is accompanied by atrophy-related partial volume effects. Curr Alzheimer Res. 2016;13:475–86.
52. Seo SW, Cho SS, Park A, Chin J, Na DL. Subcortical vascular versus amnestic mild cognitive impairment: comparison of cerebral glucose metabolism. J Neuroimaging. 2009;19:213–9. doi:10.1111/j.1552-6569.2008.00292.x.
53. Chen K, Langbaum JB, Fleisher AS, Ayutyanont N, Reschke C, Lee W, et al. Twelve-month metabolic declines in probable Alzheimer's disease and amnestic mild cognitive impairment assessed using an empirically pre-defined statistical region-of-interest: findings from the Alzheimer's Disease Neuroimaging Initiative. Neuroimage. 2010;51:654–64. doi:10.1016/j.neuroimage.2010.02.064.
54. Frisoni GB, Bocchetta M, Chetelat G, Rabinovici GD, de Leon MJ, Kaye J, et al. Imaging markers for Alzheimer disease: which vs how. Neurology. 2013;81:487–500. doi:10.1212/WNL.0b013e31829d86e8.
55. Garibotto V, Herholz K, Boccardi M, Picco A, Varrone A, Nordberg A, et al. Clinical validity of brain fluorodeoxyglucose positron emission tomography as a biomarker for Alzheimer's disease in the context of a structured 5-phase development framework. Neurobiol Aging. 2017;52:183–195. doi:10.1016/j.neurobiolaging.2016.03.033