NeuroImage: Clinical 4 (2014) 45–52

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Reduced FDG-PET brain metabolism and executive function predict clinical progression in elderly healthy subjects☆,☆☆

Michael Ewers [a,b,c,*], Matthias Brendel [d], Angela Rizk-Jackson [e], Axel Rominger [e], Peter Bartenstein [e], Norbert Schuff [b,c], Michael W. Weiner [b,c], for the Alzheimer's Disease Neuroimaging Initiative (ADNI)

[a] Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig-Maximilian-University, Munich, Germany
[b] Department of Radiology, University of California, San Francisco, CA, USA
[c] VA Medical Center, Center for Imaging of Neurodegenerative Diseases, San Francisco, CA, USA
[d] Department of Nuclear Medicine, Klinikum der Universität München, Ludwig-Maximilian-University, Munich, Germany
[e] Clinical & Translational Science Institute, University of California, San Francisco, CA, USA

ARTICLE INFO

Article history:
Received 21 August 2013
Received in revised form 12 October 2013
Accepted 27 October 2013
Available online 4 November 2013

Keywords:
Preclinical AD
Conversion
Diagnosis
FDG-PET
Gray matter volume
Executive function

ABSTRACT

Brain changes reminiscent of Alzheimer disease (AD) have been previously reported in a substantial portion of elderly cognitive healthy (HC) subjects. The major aim was to evaluate the accuracy of MRI assessed regional gray matter (GM) volume, 18F-fluorodeoxyglucose positron emission tomography (FDG-PET), and neuropsychological test scores to identify those HC subjects who subsequently convert to mild cognitive impairment (MCI) or AD dementia. We obtained in 54 healthy control (HC) subjects *a priori* defined region of interest (ROI) values of medial temporal and parietal FDG-PET and medial temporal GM volume. In logistic regression analyses, these ROI values were tested together with neuropsychological test scores (free recall, trail making test B (TMT-B)) as predictors of HC conversion during a clinical follow-up between 3 and 4 years. In voxel-based analyses, FDG-PET and MRI GM maps were compared between HC converters and HC non-converters. Out of the 54 HC subjects, 11 subjects converted to MCI or AD dementia. Lower FDG-PET ROI values were associated with higher likelihood of conversion ($p = 0.004$), with the area under the curve (AUC) yielding 82.0% (95% CI = (95.5%, 68.5%)). The GM volume ROI was not a significant predictor ($p = 0.07$). TMT-B but not the free recall tests were a significant predictor (AUC = 71% (95% CI = 50.4%, 91.7%)). For the combination of FDG-PET and TMT-B, the AUC was 93.4% (sensitivity = 82%, specificity = 93%). Voxel-based group comparison showed reduced FDG-PET metabolism within the temporo-parietal and prefrontal cortex in HC converters. In conclusion, medial temporal and-parietal FDG-PET and executive function show a clinically acceptable accuracy for predicting clinical progression in elderly HC subjects.

1. Introduction

The recently proposed research diagnosis of preclinical Alzheimer's disease is primarily based on brain abnormalities detected by amyloid PET, FDG-PET, or volumetric MRI in elderly HC subjects (Sperling et al., 2011). From a clinical point of view, a crucial question is the accuracy of such neuroimaging measures for the prediction of clinical progression to AD dementia in HC subjects. Previous studies showed that medial temporal lobe FDG-PET hypometabolism was associated with subsequent conversion to MCI or AD dementia (de Leon et al., 2001; Mosconi et al., 2008). However, the accuracy of medial temporal lobe FDG-PET remained below a clinically significant level for the prediction of progression to MCI (Mosconi et al., 2008). Results from studies assessing volumetric MRI in elderly HC subjects suggested that regional GM volume decrease was associated with increased risk to develop AD dementia (den Heijer et al., 2006; Dickerson et al., 2011). However, the accuracy of FDG-PET and GM volume for predicting AD dementia in HC is still unclear, and to our best knowledge, no studies have compared the utility of FDG-PET and MRI for the detection of preclinical AD so far. Importantly, although slight cognitive changes have been proposed to be part of the preclinical stage of AD (Sperling et al., 2011), the diagnostic benefit of combining neuropsychological test and neuroimaging modalities has not been tested.

Therefore, the major aim of the current study was to assess the predictive accuracy of regional FDG-PET metabolism, MRI assessed

☆☆ Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this report. Complete listing of ADNI investigators is available at http://www.loni.ucla.edu/ADNI/Collaboration/ADNI_Manuscript_Citations.pdf

* Corresponding author at: Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig Maximilian University, Max-Lebsche-Platz 30, 81377 Munich, Germany. Tel.: +49 89 7095 8367; fax: +49 89 7095 8369.
E-mail address: michael.ewers@med.uni-muenchen.de (M. Ewers).

2213-1582/$ – see front matter © 2013 The Authors. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.nicl.2013.10.018

gray matter volume and neuropsychological test performance for predicting clinical progression to MCI or AD dementia in elderly HC subjects. For neuropsychological test scores, we used tests of free recall and the trail making test B (TMT-B) — a measure of executive function — since these tests were previously shown to be predictive of the clinical progression from MCI to AD dementia (Ewers et al., 2012).

2. Data and methods

2.1. Subjects

All subjects were recruited within the multicenter Alzheimer's Disease Neuroimaging Initiative (ADNI, http://www.adni-info.org/) and all data and scans were obtained from ADNI's publicly available databank (http://adni.loni.ucla.edu/). In ADNI, the general inclusion criteria were an age between 55 and 90 years, a modified Hachinski score ≤ 4, education of at least 6 grade level, and stable treatment of at least 4 weeks in case of treatment with permitted medication (for full list see http://www.adni-info.org, Procedures Manual).

A total of 54 HC subjects were included for testing the prediction of the development of MCI or AD dementia. To be included in the current study, HC subjects had to have 1) HC diagnosis at baseline, 2) baseline MRI and FDG-PET scan, 3) conversion to MCI or AD during clinical follow-up or a clinical follow-up interval of at least 36 months, and 4) a clinical dementia rating (CDR) score = 0 at each follow-up time point in HC non-converters (see for flow chart Fig. 1). The CDR requirement in the HC non-converters was done in order to have a clean control group to be contrasted against HC subjects who show clinical progression. Such a study design was considered useful to boost power for establishing a prediction model that clearly distinguishes between HC converters and those HC subjects who remain stable. The requirement of at least 36 months of follow-up time for all HC subjects was implemented in order to reduce censoring effects.

The ROIs of FDG-PET and GM volume as predictors of clinical progression were *a priori* established in a training sample including 78 amnestic MCI with abnormal CSF $A\beta_{1-42}$ levels and 30 HC subjects with normal CSF $A\beta_{1-42}$ levels. The inclusion criteria were 1) diagnosis of "MCI of the AD type" (Albert et al., 2011) or "prodromal AD" (Dubois et al., 2007) as recently proposed by the National Institute on Aging and Alzheimer's Association workgroup (Albert et al., 2011; McKhann et al., 2011), *i.e.* amnestic MCI subjects with abnormal CSF $A\beta_{1-42}$ levels (<192 pg/ml, henceforth called MCI (Aβ+)). HC subjects were included if they had CSF $A\beta_{1-42}$ levels > 192 pg/ml, henceforth called HC (Aβ−) (Shaw et al., 2009). Due to sample size restriction, the sample of 30 HC (Aβ−) subjects in the training sample included also 15 HC subjects that were part of the 54 HC subjects (test sample) in whom the prediction of subsequent conversion to MCI or AD was tested. Statistical analysis ensured that the overlap between both data sets did not bias the results (see Statistics section below).

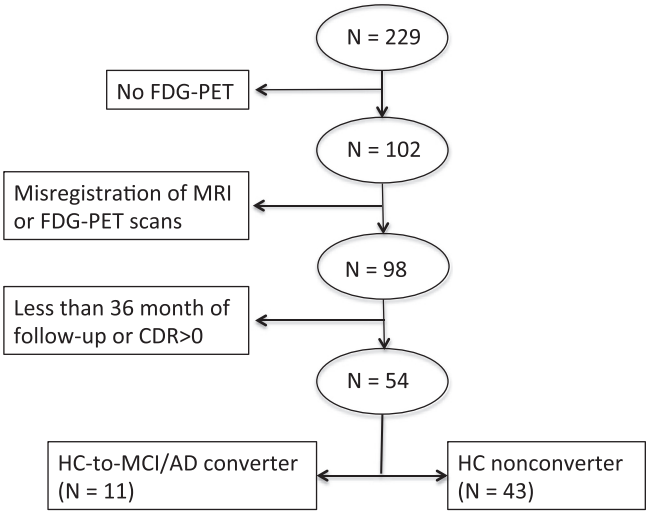

Fig. 1. Flow chart of the selection of HC subjects for the test set.

2.2. The ADNI study

ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public–private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). The initial goal of ADNI was to recruit 800 adults, ages 55 to 90, to participate in the research — approximately 200 cognitively normal older individuals to be followed for 3 years, 400 people with MCI to be followed for 3 years, and 200 people with early AD to be followed for 2 years.

2.3. MRI and FDG-PET acquisition

All MRI data were acquired on 1.5 T MRI scanners using volumetric T1-weighted sequences to map brain structures, optimized for the different scanners as indicated at http://www.loni.ucla.edu/ADNI/Research/Cores/index (Jack et al., 2008).

FDG-PET scans were acquired on multiple scanners of varying resolution. FDG scans were collected as 6 × 5-min frames beginning 30 min after injection of approximately 5 mCi of tracer (details at: http://adni.loni.ucla.edu/wp-content/uploads/2010/09/PET-Tech_Procedures_Manual_v9.5.pdf). Prior to download, FDG-PET images had been preprocessed at the University of Berkeley according to a standard ADNI procedure (frames were co-registered, averaged, oriented along the anterior–posterior commission line and resliced to a 1.5 mm isotopic voxel space). The ADNI preprocessing protocol can be found at: http://adni.loni.ucla.edu/methods/pet-analysis/pre-processing/. In order to reduce the multicenter variability of the FDG-PET images, correction factors for image reconstruction were obtained on the basis of scanning a 3D Hoffman phantom object at all participating ADNI sites. This allowed to derive scanner-specific correction factors. Image reconstruction was done centrally at the ADNI PET imaging core laboratory (for details see (Joshi et al., 2009)).

2.4. Spatial normalization of MRI and FDG-PET scans

The T1 MRI scans were segmented and the resulting GM partitions were spatially normalized using high-dimensional large-deformation diffeomorphic registration as implemented in the DARTEL tool of SPM8 (http://www.fil.ion.ucl.ac.uk/spm/software/) (Ashburner, 2007). In the first step, the MRI scans were segmented into GM, white matter, and CSF partitions (Ashburner and Friston, 2005). Rigid body transformation was applied in order to bring GM segments into approximate alignment. The estimation of the non-linear spatial normalization parameters was calculated in iterative steps of the template generation. Within each of 6 successive cycles, an average of the GM tissue class images was generated and all GM maps were spatially normalized using a linear set of basis functions, where the optimization of the parameter estimates followed the Levenberg–Marquardt (LM) algorithm as the minimum energy optimization procedure (Ashburner, 2007). The normalization parameters were applied to the GM images. In the next cycle, the GM images that had been normalized during the first step were averaged and once more subjected to the spatial normalization procedure. This cycle was iterated 6 times, successively relaxing

the regularization parameters as the image registration improves and the spatial resolution of the template became crisper (see also manual of SPM 8 at http://www.fil.ion.ucl.ac.uk/spm/doc/manual.pdf). In the final step, the custom GM template was registered onto the ICBM GM template in MNI space included in the SPM8 package and combined with the non-linear flow fields estimated during the diffeomorphic normalization procedure described above. GM segments were modulated (*i.e.* multiplied by the Jacobian determinants of the spatial transformation) in order to yield a GM concentration in proportion to the dilation or contraction a voxel underwent during spatial normalization. The total intracranial volume (ICV) was estimated as the sum of the GM, WM and CSF tissue class images in native space obtained during the spatial normalization procedure (described above). For FDG-PET, each FDG-PET scan was coregistered to the individual's corresponding GM segment in native space. The spatial normalization parameters that had been estimated on the basis of the MRI images *via* DARTEL were subsequently applied to the coregistered FDG-PET images so that both the GM maps and the FDG-PET images were in MNI space.

For intensity normalization of the FDG-PET scans to the pons, a mask was manually drawn onto the GM template in MNI space, using a 3-D display in the program MIPAV (http://mipav.cit.nih.gov/). The binarized mask was multiplied with each spatially normalized FDG-PET image to calculate the average FDG-PET within the pons region. Each voxel of a subject's FDG-PET image was subsequently divided by the individual average FDG-PET value of the pons.

Correction for partial volume effects (PVEs) was done using the geometric transfer matrix (GTM) method (Rousset et al., 1998) as implemented in the software program PMOD (PMOD Technologies Ltd, Zurich, Switzerland). Briefly, the spill-over effects between the different brain compartments are estimated by convolving binarized voxels of interest (GM *vs.* background) with the point spread function to derive the weights of the spill-over effects. Based on these weights, the PVE corrected FDG-PET images were estimated (http://www.pmod.com/technologies/pdf/doc/PNEURO.pdf).

All GM and FDG-PET images were finally spatially smoothed with a Gaussian kernel at 8 mm FWHM.

2.5. Construction of FDG-PET and GM ROIs

Voxel-based ANCOVAs were used to compare group differences between HC (Aβ−) and MCI (Aβ+) subjects (training sample), using *group* as the independent variable, and in the case of GM maps, ICV as an additional covariate. Potentially confounding variables including age, gender, and education did not differ between the HC (Aβ−) and MCI (Aβ+) groups (Supplementary Table 1), and were therefore not included as covariates in the voxel based regression analyses. The MMSE score was lower ($t = -4.9$, $df = 106$, $p < 0.001$) and ApoE ε4 genotype was more frequent in the MCI (Aβ+) compared to the HC (Aβ−) groups ($\chi^2 = 23.3$, $df = 1$, $p < 0.001$). In addition to CSF $A\beta_{1-42}$ (group selection variable), CSF total tau (Welch two-sample t-test: $t = -5.9$, $df = 105$, $p < 0.001$) and CSF p-tau_{181} (Welch two-sample t-test: $t = -7.6$, $df = 96$, $p < 0.001$) differed between groups (see also Supplementary Table 1).

For voxel-based comparison of the FDG-PET and GM maps, the statistical parametric maps of the t-scores were thresholded at the voxel-level at $p = 0.001$ (uncorrected), with the cluster extent threshold set to the number of expected voxel per cluster (k), including $k = 624$ (for FDG-PET) and $k = 254$ (for GM maps).

Results showed significantly lower FDG-PET metabolism primarily within the medial temporal and medial–parietal brain regions in MCI (Aβ+) compared to HC (Aβ−) (Supplementary Table 2, Supplementary Fig. 1A). For GM differences, voxels with significantly reduced GM volume (MCI (Aβ+) < HC (Aβ−)) clustered within the inferior and medial temporal lobe (Supplementary Table 3, Supplementary Fig. 1B). No brain region showed significantly larger FDG-PET metabolism or GM volume in MCI (Aβ+) compared to HC (Aβ−). Clusters of significant group differences of FDG-PET or GM volume were binarized to create a single ROI mask for each modality. FDG-PET and GM volume within the ROI masks were subsequently obtained from the baseline FDG-PET and GM maps of each of the 54 HC subjects.

2.6. Neuropsychological tests

Episodic memory was assessed by the Rey Auditory Verbal Learning Test (RAVLT), which includes a list of 15 words to be recalled immediately after each of the 5 verbal presentations and after a 30-min delay (Rey, 1964). A test of frontal lobe functions included the trail making test B (TMT-B, score: total number of seconds to complete the test) (Reitan and Wolfson, 1985).

2.7. Statistics

"Conversion" was defined as a change in diagnosis from HC to MCI or AD dementia during the follow-up in the 54 HC subjects (test set). Logistic regression analyses were used to predict conversion status (non-converter *vs.* converters). The predictors included baseline FDG-PET ROI (both uncorrected or corrected for PVE), GM volume ROI, RAVLT immediate free recall, RAVLT delayed free recall, and the TMT-B test. Separate logistic regression models were constructed for each predictor and significant predictors were subsequently combined in a final model to test additive effects. Each model was corrected for age, gender, and ApoE genotype (ε4 allele carriers *vs.* ε4 allele non-carriers). Normal distribution of each predictor was tested by Shapiro–Wilks tests, and — in case of non-normal distribution — the variables were rank transformed. All predictors were transformed into z-scores before entering them into the logistic regression model to yield better comparison of the regression weights between predictors. ROC analyses for significant predictors were conducted, using bootstrapping with $n = 1000$ resampling iterations to estimate the area under the curve (AUC) and the specificity at a fixed sensitivity of 80%. The bootstrapped 95% confidence intervals (95% CI) of the AUC and specificity are reported. In order to examine the robustness of the results, the logistic regression analyses for the prediction of conversion in the 54 HC subjects were repeated, now excluding those 15 HC subjects who had also been part of the training sample for the establishment of the ROIs (see above "Construction of FDG-PET and GM ROIs"). All of these analyses were done with the program R (version 2.12.1, www.r-project.org).

Voxel-based comparisons of FDG-PET and GM maps between HC converters vs. non-converters were conducted through ANCOVAs, controlled for age (that differed significantly between converters and non-converters). For the voxel-based group comparison of GM volume, the ICV was an additional covariate in the regression analyses. The statistical parametric maps were thresholded at $p = 0.001$ (uncorrected at voxel level), where only clusters exceeding the expected number of voxels per cluster (cluster extent threshold) were reported. Voxel-wise maps of the effect size *d* (Cohen, 1992) were computed to display group differences unbiased by sample size.

3. Results

3.1. Conversion from HC to MCI or AD dementia

The descriptive statistics of the HC converters and HC non-converters are displayed in Table 1. Over a time interval of 54 months, 11 out of 54 HC subjects progressed to MCI or AD (MCI: $n = 9$, AD dementia: $n = 2$). All of the 11 subjects had converted after 3 years, except for one subject (time to conversion = 4 years). For the non-converters, all of the subjects had been assessed at least for 3 years of clinical follow-up duration. Nineteen out of 43 non-converters (44%) had also been assessed after at least 4 years. HC converters were significantly older than HC non-converters ($t = -3.5$, $df = 16.6$, $p = 0.002$).

Table 1
Demographic, clinical and biomarker statistics in HC converters and HC non-converters. The mean and standard deviation (SD) are indicated where appropriate.

Group	N	Age (yrs)	Sex (f/m)	Edu (yrs)	ApoE ε4 (+/−)	MMSE	TMT-B (in sec.)	Recall immed.	Recall del.
HC converter	11	78.9 (3.7)	5/6	16.3 (2.9)	4/7	29.0 (0.8)	141.2 (45.1)	40 (10.4)	5.6 (4.4)
HC non-converter	43	74.3 (4.6)	15/28	16.6 (3.0)	12/31	28.4 (1.2)	111.2 (32.3)	44.1 (10.1)	7.7 (4.1)

ApoE (+) = ApoE ε4 carriers, ApoE (−) = ApoE ε4 non-carriers, Edu = education, Recall immed. = immediate free recall, Recall del. = delayed free recall, yrs = years, f = female, m = male, sec. = seconds.

3.2. Prediction of conversion from HC to MCI or AD dementia

Logistic regression analyses showed that the FDG-PET ROI (including medial temporal and parietal brain regions) was a significant predictor of HC conversion (B = −1.4 z-core units, SE = .5, p = 0.006), independent of age, gender, and ApoE ε4 status. The results of the ROC analysis are shown in Table 2 and Fig. 2. Specifically, the ROC analysis for the age-adjusted FDG-PET ROI yielded an AUC of 77.8% (95% CI = (62.2%, 93.6%)). The bootstrapped median specificity was 51.2% (95% CI = (27.9%, 93%)) at a fixed sensitivity of 80%. The FDG-PET ROI after PVE correction remained a significant predictor of conversion (B = 1.4 z-score units, SE = .5, p = 0.004), with the AUC remaining virtually the same when compared to the FDG-PET model without PVE correction: AUC = 80.3% (95% CI = (66.2%, 94.5%)).

The ROI of GM volume (medial temporal lobe regions) was not a significant predictor of conversion (p = 0.07). Because the significance threshold was missed by a small margin, the ROC analysis was computed nevertheless. At a sensitivity of 80%, the mean specificity of the GM volume ROI was below 50%.

For neuropsychological tests, the logistic regression analyses showed that TMT-B test scores discriminated HC converters from HC non-converters (B = 0.88 z-score units, SE = 0.4, p = 0.03), adjusted for age, gender, and ApoE ε4 genotype. The tests of free recall were not significant predictors. ROC analysis showed for age-adjusted TMT-B scores an AUC of 71% (95% CI = 50.4%, 91.7%). At a fixed sensitivity of 80%, the median specificity was 34.9% (95% CI = 9.3%, 95.3%). In a combined model, TMT-B test scores significantly added to the predictive accuracy of the FDG-PET ROI (TMT-B: B = 1.8 z-score units, SE = 0.7, p = 0.01, FDG-PET ROI: B = −2.3 z-score units, SE = 0.8, p = 0.005). The AUC of the combined model of FDG-PET plus TMT-B increased to 89.4% (95% CI = 79.9%, 98.9%). The sensitivity was 81.8% and the specificity was 79.1% for the combined model.

When repeating the regression analyses excluding those 15 HC subjects who had been also part of the training sample, the result pattern remained the same for all analyses (data not shown), indicating that the results were not influenced by the partial overlap between training and test samples.

CSF biomarker measurements were available in a subset of the 54 HC subjects including 7 HC converters and 22 HC non-converters. Abnormal CSF $A\beta_{1-42}$ levels (<192 pg/ml) (Shaw et al., 2009) were present in 5 (71%) HC converters and 9 (41%) HC non-converters. Abnormal CSF p-tau_{181} levels (>23 pg/ml) (Shaw et al., 2009) were present in 3 (75%) HC converters and 6 (27%) of HC non-converters.

Table 2
Specificity at a fixed sensitivity of 80% and AUC for the classification of HC converter *vs.* HC non-converter.

Predictor	% AUC (95% CI)	% specificity (at 80% sensitivity)
FDG-PET ROI	77.8 (62.2, 93.6)	51.2
PVE corrected FDG-PET ROI	80.3 (66.2, 94.5)	67.4
TMT-B	71 (50.4, 91.7)	34.9
Combined FDG-PET & TMT-B	89.4 (79.9, 98.9)	79.1[a]
Combined PVE corrected FDG-PET & TMT-B	92 (84.2, 99.8)	80.7[a]

[a] Exact sensitivity was 81.8%.

3.3. Voxel based comparison of FDG-PET and MRI between HC converters and HC stable

In the voxel-based comparison of FDG-PET, HC converters showed hypometabolism within the areas of the medial frontal gyrus and the temporo-parietal brain areas (Fig. 3, for cluster coordinates see Supplementary Table 4). No brain areas showed increased FDG-PET in HC converters when compared to HC stable subjects. The comparison of GM volume between HC converters and HC non-converters produced no significant clusters of group differences.

Lastly, we mapped the effect size *d* of FDG-PET and GM volume for the comparison between HC converters and HC non-converters without thresholding the maps on the basis of p-values. Visual inspection of the effect size maps shows that the effect sizes of group differences (*i.e.* lower FDG-PET in HC converters) are highest within the temporal, frontal, and parietal brain areas (Fig. 4A), and are higher in most brain areas when compared to the group difference in GM volume (Fig. 4B). For GM volume, the highest effect sizes of decreased GM volume within the HC converters group are primarily centered within the medial temporal lobe.

4. Discussion

Our findings of the prediction of the conversion from HC to MCI and AD dementia based on the *a priori* determined FDG-PET ROI (located in the medial temporal and parietal lobes) is consistent with previous studies showing that medial temporal and parietal FDG-PET was

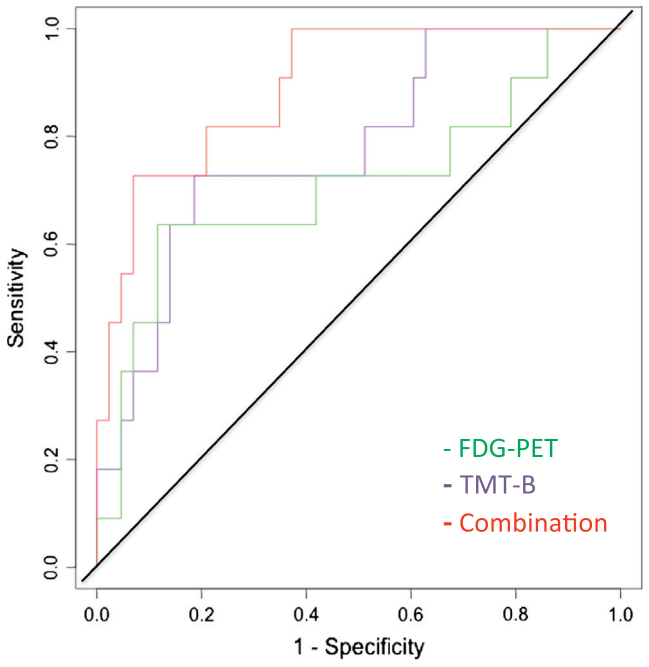

Fig. 2. ROC curves for the classification of HC converters *vs.* HC non-converters based on the FDG-PET ROI (green), TMT-B test (blue) and the combination of both FDG-PET ROI and TMT-B (red). The AUC values are indicated in Table 2.

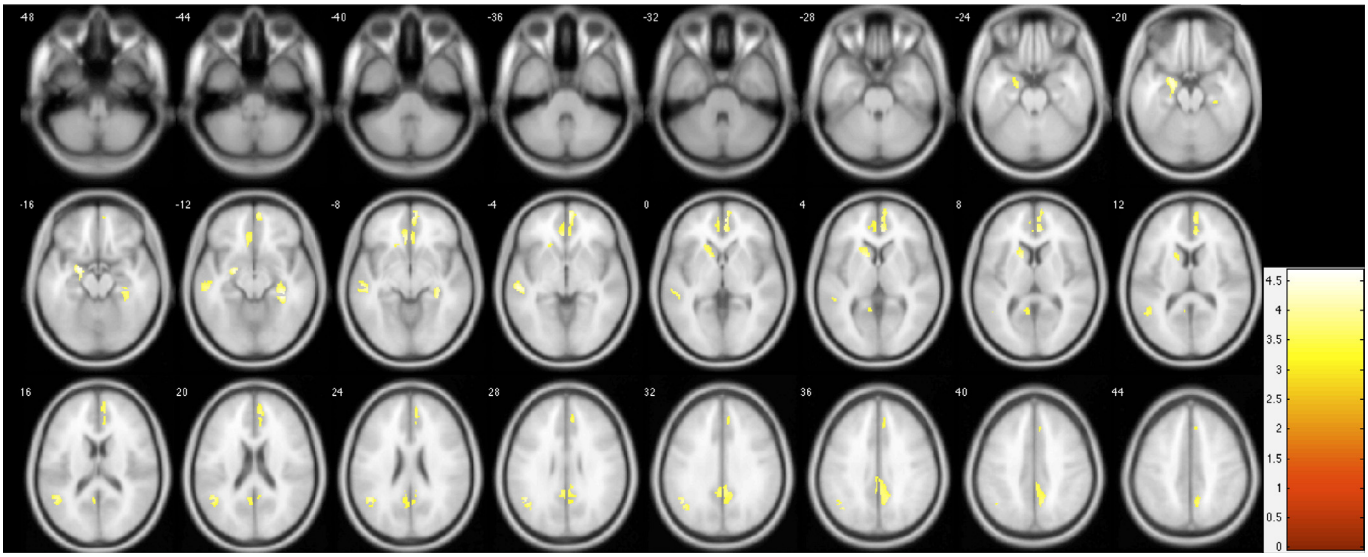

Fig. 3. Projection of voxel-based t-statistic map onto axial brain slices. Yellow designates t-values for significantly higher FDG-PET in HC non-converters compared to HC converters. The opposite contrast (HC converters > HC stable) was not significant.

associated with faster cognitive decline (Jagust et al., 2006) and hippocampal/entorhinal FDG-PET was sensitive to preclinical AD (de Leon et al., 2001; Mosconi et al., 2008). These regions of reduced FDG-PET in medial temporal and medial parietal brain areas can be considered a subset of those temporo-parietal areas (*i.e.* medial and lateral parietal lobes & medial temporal lobes) that are typically affected in AD dementia (Mosconi, 2005). Whereas we found a predictive effect of the FDG-PET ROI on subsequent clinical progression, we did not detect a significant predictive value of the medial temporal GM volume ROI. The findings are in line with MRI studies reporting that baseline total GM or medial temporal GM in HC subjects did not predict global cognitive decline (Jagust et al., 2006) or progression to MCI or AD (Rusinek et al., 2003). However, the results are in disagreement with other studies suggesting that medial temporal GM volume is affected in preclinical AD (Chiang et al., 2011; Dickerson et al., 2012). In particular, recent studies showed that a summary measure of cortical thickness ROIs including several temporal, parietal and prefrontal brain regions showed predictive value for the classification of subjects who subsequently converted to AD dementia (Dickerson et al., 2011). Inconsistencies between the current and previous findings might be explained by methodological differences including differences in cortical thickness *vs.* morphometric measures (Hutton et al., 2009; Winkler et al., 2010), or a more advanced state of disease severity of HC subjects included in those studies which detected gray matter atrophy at the preclinical stage (Dickerson et al., 2011). The most parsimonious explanation is, however, that the subtle gray matter differences were already present but remained below the significance threshold in the current study. The examination of the effect size maps of gray matter volume differences confirmed that reductions in gray matter volume were mostly confined to the medial temporal lobe, which is in general agreement with previous reports of voxel-based analysis showing that GM volume decreases exclusively within the medial temporal and angular gyrus in HC subjects who later convert to MCI (Smith et al., 2007). Importantly, the effect size map confirmed that FDG-PET hypometabolism was widely spread out within the gray matter brain areas, supporting the notion that FDG-PET was more affected compared to GM volume in subjects who showed clinical worsening that may correspond to preclinical AD.

For neuropsychological tests, the TMT-B test but not measures of episodic memory were predictive of clinical progression. We previously reported that the TMT-B test outperformed other neuropsychological tests including tests of free recall, category fluency, digit span, and the Boston naming test for predicting conversion from MCI to AD dementia (Ewers et al., 2012). Together with the previous findings, the current results suggest that TMT-B test might reflect cognitive function that is affected early in the course of AD. The TMT-B test is thought to be a measure of visual attention and cognitive flexibility associated with frontal and parietal lobe function (Zakzanis et al., 2005). It is not known exactly which functional and structural brain changes in AD underlie the cognitive performance changes on the TMT-B test early in the course of AD. However, from a clinical point of view, the combination of the TMT-B and regional FDG-PET provides high sensitivity and specificity for the detection of clinical worsening in elderly HC subjects — possibly at an early stage of AD dementia. It should be noted that no systematic screening of a large range of neuropsychological tests or biomarkers candidates was done in the current study. Thus, neuropsychological tests or biomarkers other than those tested here may be relevant as well (Chary et al., 2013; Condret-Santi et al., 2013; Sperling et al., 2011; Tondelli et al., 2012). In the current study, some of the *a priori* selected predictors that were previously shown to be sensitive to prodromal AD (Ewers et al., 2012) were confirmed to be successful for the prediction of clinical progression in HC subjects. It is for larger future studies to validate the current results and define a potentially extended test set of best predictors of preclinical AD. Apart from the current objective measures of neuropsychological ability, subjective memory complaints have previously been shown to be a potential predictor of subsequent memory decline (Chamberlain et al., 2011; Scheef et al., 2012) in elderly cognitively normal subjects and may be tested in future studies as well.

For the interpretation of the current results, it is important to note that not all HC subjects who converted to MCI may eventually progress to AD dementia. Compared to the HC non-converters, a higher proportion of HC converters showed abnormal measures on primary AD pathology as measured by CSF Aβ and p-tau$_{181}$ levels that were available in a subset of the patients, supporting the notion that the HC converters were afflicted by preclinical AD. The development to AD dementia need to be confirmed, however, in future follow-ups within the ADNI project. It should also be noted that the subset of 54 HC subjects represents a subsample of the overall total of 229 HC subjects assessed in ADNI I. The sample size reduction was majorly due to the ADNI I study design, where a subset of the 229 HC subjects received FDG-PET at baseline. For the current study further requirements included follow-up duration (>3 years) and a CDR of 0 for the HC non-converters. We considered

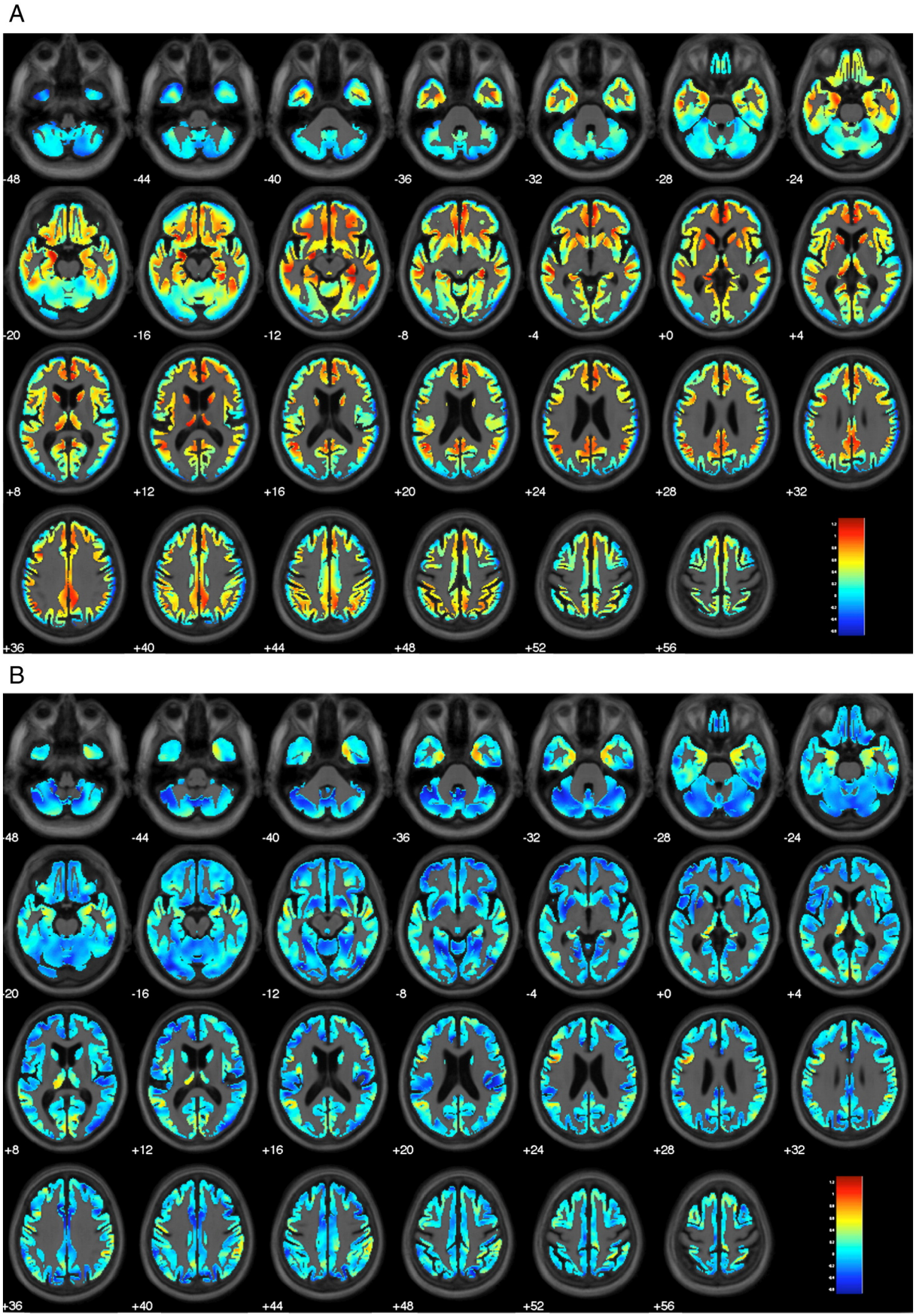

Fig. 4. Maps of effect size *d* for the group comparison of HC converter *vs.* HC non-converter are superimposed onto axial brain slices for FDG-PET (A) and MRI gray matter (B). Red colors indicate a positive effect size, *i.e.* HC non-converter > HC converter, the bluish colors indicate the opposite effect. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

these requirements useful in order to be able to clearly define a control group to establish a sensitive prediction model of preclinical AD.

Increased brain Aβ PET binding in the brain has previously been shown to be predictive for the conversion to AD dementia in MCI subjects and may also predict clinical progression at the preclinical stage of AD (Nordberg et al., 2013; Okello et al., 2009). Thus the question needs to be addressed to what extent the current markers add to measures of primary pathology (tau or Aβ) to the detection of preclinical AD (Desikan et al., 2012; Prestia et al., 2013).

Acknowledgments

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc., F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., as well as from the non-profit partners, the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the U.S. Food and Drug Administration. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation.

Dr. Weiner has been on advisory boards for Lilly, Araclon and Institut Catala de Neurociencies Aplicades, Gulf War Veterans Illnesses Advisory Committee, VACO, Biogen Idec, Elan/Wyeth Alzheimer's Immunotherapy Program North American Advisory Board, Novartis Misfolded Protein Scientific Advisory Board Meeting, Banner Alzheimer's Institute Alzheimer's Prevention Initiative Advisory Board Meeting, and the Research Advisory Committee on Gulf War Veterans' Illnesses. He has been a consultant for Elan/Wyeth, Novartis, Forest, Ipsen, Daiichi Sankyo, Inc., Astra Zeneca, Araclon, Medivation/Pfizer, TauRx Therapeutics LTD, Bayer Healthcare, Biogen Idec, Exonhit Therapeutics, SA, Servier, and Synarc. Dr. Weiner has received research support from Merck, Avid, the NIH, the DOD, and the VA, and holds stock options with Synarc and Elan. Dr. Weiner has received funding for Travel from: Elan/Wyeth Alzheimer's Immunotherapy Program North American Advisory Board, Alzheimer's Association, Forest, University of California, Davis, Tel-Aviv University Medical School, Colloquium Paris, Ipsen, Wenner-Gren Foundations, Social Security Administration, Korean Neurological Association, National Institutes of Health, Washington University at St. Louis, Banner Alzheimer's Institute, Clinical Trials on Alzheimer's Disease, Veterans Affairs Central Office, Beijing Institute of Geriatrics, Innogenetics, New York University, NeuroVigil, Inc., CHRU-Hopital Roger Salengro, Siemens, AstraZeneca, Geneva University Hospitals, Lilly, University of California, San Diego — ADNI, Paris University, Institut Catala de Neurociencies Aplicades, University of New Mexico School of Medicine, Ipsen, CTAD (Clinical Trials on Alzheimer's Disease), Pfizer, AD PD meeting, Paul Sabatier University, and Novartis. Dr. Weiner has been on the Editorial Advisory Boards of the Journals Alzheimer's and Dementia and MRI. He has received honoraria from the American Academy of Neurology, Ipsen, NeuroVigil, Inc. and Insitut Catala de Neurociencies Aplicades.

Organizations contributing to the Foundation for NIH and thus to the NIA funded Alzheimer's Disease Neuroimaging Initiative are Abbott, Alzheimer's Association, Alzheimer's Drug Discovery Foundation, Anonymous Foundation, AstraZeneca, Bayer Healthcare, BioClinica, Inc. (ADNI 2), Bristol-Myers Squibb, Cure Alzheimer's Fund, Eisai, Elan, Gene Network Sciences, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson & Johnson, Eli Lilly & Company, Medpace, Merck, Novartis, Pfizer Inc., Roche, Schering Plough, Synarc, and Wyeth. Dr. Weiner holds stocks of the companies Synarc and Elan. Mr. Brendel and Drs. Ewers, Rominger, Bartenstein, Schuff, and Rizk-Jackson report no disclosures.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.nicl.2013.10.018.

References

Albert, M.S., DeKosky, S.T., Dickson, D., Dubois, B., Feldman, H.H., Fox, N.C., Gamst, A., Holtzman, D.M., Jagust, W.J., Petersen, R.C., Snyder, P.J., Carrillo, M.C., Thies, B., Phelps, C.H., 2011. 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. 7, 270–279.

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. NeuroImage 38, 95–113.

Ashburner, J., Friston, K.J., 2005. Unified segmentation. NeuroImage 26, 839–851.

Chamberlain, S.R., Blackwell, A.D., Nathan, P.J., Hammond, G., Robbins, T.W., Hodges, J.R., Michael, A., Semple, J.M., Bullmore, E.T., Sahakian, B.J., 2011. Differential cognitive deterioration in dementia: a two year longitudinal study. J. Alzheimers Dis. 24, 125–136.

Chary, E., Amieva, H., Pérès, K., Orgogozo, J.-M., Dartigues, J.-F., Jacqmin-Gadda, H., 2013. Short- *versus* long-term prediction of dementia among subjects with low and high educational levels. Alzheimers Dement. 9, 562–571.

Chiang, G.C., Insel, P.S., Tosun, D., Schuff, N., Truran-Sacrey, D., Raptentsetsang, S., Jack Jr., C.R., Weiner, M.W., 2011. Identifying cognitively healthy elderly individuals with subsequent memory decline by using automated MR temporoparietal volumes. Radiology 259, 844–851.

Cohen, J., 1992. A power primer. Psychol. Bull. 112, 155–159.

Condret-Santi, V., Barbeau, E.J., Matharan, F., Le Goff, M., Dartigues, J.F., Amieva, H., 2013. Prevalence of word retrieval complaint and prediction of dementia in a population-based study of elderly subjects. Dement. Geriatr. Cogn. Disord. 35, 313–324.

de Leon, M.J., Convit, A., Wolf, O.T., Tarshish, C.Y., DeSanti, S., Rusinek, H., Tsui, W., Kandil, E., Scherer, A.J., Roche, A., Imossi, A., Thorn, E., Bobinski, M., Caraos, C., Lesbre, P., Schlyer, D., Poirier, J., Reisberg, B., Fowler, J., 2001. Prediction of cognitive decline in normal elderly subjects with 2-[(18)F]fluoro-2-deoxy-D-glucose/positron-emission tomography (FDG/PET). Proc. Natl. Acad. Sci. U. S. A. 98, 10966–10971.

den Heijer, T., Geerlings, M.I., Hoebeek, F.E., Hofman, A., Koudstaal, P.J., Breteler, M.M., 2006. Use of hippocampal and amygdalar volumes on magnetic resonance imaging to predict dementia in cognitively intact elderly people. Arch. Gen. Psychiatry 63, 57–62.

Desikan, R.S., McEvoy, L.K., Thompson, W.K., Holland, D., Brewer, J.B., Aisen, P.S., Sperling, R.A., Dale, A.M., 2012. Amyloid-beta-associated clinical decline occurs only in the presence of elevated P-tau. Arch. Neurol. 69, 709–713.

Dickerson, B.C., Stoub, T.R., Shah, R.C., Sperling, R.A., Killiany, R.J., Albert, M.S., Hyman, B.T., Blacker, D., Detoledo-Morrell, L., 2011. Alzheimer-signature MRI biomarker predicts AD dementia in cognitively normal adults. Neurology 76, 1395–1402.

Dickerson, B.C., Wolk, D.A., 2012. MRI cortical thickness biomarker predicts AD-like CSF and cognitive decline in normal adults. Neurology 78, 84–90.

Dubois, B., Feldman, H.H., Jacova, C., Dekosky, S.T., Barberger-Gateau, P., Cummings, J., Delacourte, A., Galasko, D., Gauthier, S., Jicha, G., Meguro, K., O'Brien, J., Pasquier, F., Robert, P., Rossor, M., Salloway, S., Stern, Y., Visser, P.J., Scheltens, P., 2007. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurol. 6, 734–746.

Ewers, M., Walsh, C., Trojanowski, J.Q., Shaw, L.M., Petersen, R.C., Jack Jr., C.R., Feldman, H.H., Bokde, A.L., Alexander, G.E., Scheltens, P., Vellas, B., Dubois, B., Weiner, M., Hampel, H., 2012. Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. Neurobiol. Aging 33, 1203–1214.

Hutton, C., Draganski, B., Ashburner, J., Weiskopf, N., 2009. A comparison between voxel-based cortical thickness and voxel-based morphometry in normal aging. NeuroImage 48, 371–380.

Jack Jr., C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., Whitwell, J.L., Ward, C., Dale, A.M., Felmlee, J.P., Gunter, J.L., Hill, D.L., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C.S., Krueger, G., Ward, H.A., Metzger, G.J., Scott, K.T., Mallozzi, R., Blezek, D., Levy, J., Debbins, J.P., Fleisher, A.S., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., Weiner, M.W., 2008. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J. Magn. Reson. Imaging 27, 685–691.

Jagust, W., Gitcho, A., Sun, F., Kuczynski, B., Mungas, D., Haan, M., 2006. Brain imaging evidence of preclinical Alzheimer's disease in normal aging. Ann. Neurol. 59, 673–681.

Joshi, A., Koeppe, R.A., Fessler, J.A., 2009. Reducing between scanner differences in multi-center PET studies. NeuroImage 46, 154–159.

McKhann, G.M., Knopman, D.S., Chertkow, H., Hyman, B.T., Jack Jr., C.R., Kawas, C.H., Klunk, W.E., Koroshetz, W.J., Manly, J.J., Mayeux, R., Mohs, R.C., Morris, J.C.,

Rossor, M.N., Scheltens, P., Carrillo, M.C., Thies, B., Weintraub, S., Phelps, C.H., 2011. 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. 7, 263–269.

Mosconi, L., 2005. Brain glucose metabolism in the early and specific diagnosis of Alzheimer's disease. FDG-PET studies in MCI and AD. Eur. J. Nucl. Med. Mol. Imaging 32, 486–510.

Mosconi, L., De Santi, S., Li, J., Tsui, W.H., Li, Y., Boppana, M., Laska, E., Rusinek, H., de Leon, M.J., 2008. Hippocampal hypometabolism predicts cognitive decline from normal aging. Neurobiol. Aging 29, 676–692.

Nordberg, A., Carter, S.F., Rinne, J., Drzezga, A., Brooks, D.J., Vandenberghe, R., Perani, D., Forsberg, A., Langstrom, B., Scheinin, N., Karrasch, M., Nagren, K., Grimmer, T., Miederer, I., Edison, P., Okello, A., Van Laere, K., Nelissen, N., Vandenbulcke, M., Garibotto, V., Almkvist, O., Kalbe, E., Hinz, R., Herholz, K., 2013. A European multicentre PET study of fibrillar amyloid in Alzheimer's disease. Eur. J. Nucl. Med. Mol. Imaging 40, 104–114.

Okello, A., Koivunen, J., Edison, P., Archer, H.A., Turkheimer, F.E., Nagren, K., Bullock, R., Walker, Z., Kennedy, A., Fox, N.C., Rossor, M.N., Rinne, J.O., Brooks, D.J., 2009. Conversion of amyloid positive and negative MCI to AD over 3 years: an 11C-PIB PET study. Neurology 73, 754–760.

Prestia, A., Caroli, A., van der Flier, W.M., Ossenkoppele, R., Van Berckel, B., Barkhof, F., Teunissen, C.E., Wall, A.E., Carter, S.F., Scholl, M., Choo, I.H., Nordberg, A., Scheltens, P., Frisoni, G.B., 2013. Prediction of dementia in MCI patients based on core diagnostic markers for Alzheimer disease. Neurology 80, 1048–1056.

Reitan, R., Wolfson, D., 1985. The Halstead-Reitan Neuropsychological Test Battery. Neuropsychological Press, Tucson.

Rey, A., 1964. L'examen clinique en psychologie. Press Universitaires de France, Paris.

Rousset, O.G., Ma, Y., Evans, A.C., 1998. Correction for partial volume effects in PET: principle and validation. J. Nucl. Med. 39, 904–911.

Rusinek, H., De Santi, S., Frid, D., Tsui, W.H., Tarshish, C.Y., Convit, A., de Leon, M.J., 2003. Regional brain atrophy rate predicts future cognitive decline: 6-year longitudinal MR imaging study of normal aging. Radiology 229, 691–696.

Scheef, L., Spottke, A., Daerr, M., Joe, A., Striepens, N., Kolsch, H., Popp, J., Daamen, M., Gorris, D., Heneka, M.T., Boecker, H., Biersack, H.J., Maier, W., Schild, H.H., Wagner, M., Jessen, F., 2012. Glucose metabolism, gray matter structure, and memory decline in subjective memory impairment. Neurology 79, 1332–1339.

Shaw, L.M., Vanderstichele, H., Knapik-Czajka, M., Clark, C.M., Aisen, P.S., Petersen, R.C., Blennow, K., Soares, H., Simon, A., Lewczuk, P., Dean, R., Siemers, E., Potter, W., Lee, V.M., Trojanowski, J.Q., 2009. Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. Ann. Neurol. 65, 403–413.

Smith, C.D., Chebrolu, H., Wekstein, D.R., Schmitt, F.A., Jicha, G.A., Cooper, G., Markesbery, W.R., 2007. Brain structural alterations before mild cognitive impairment. Neurology 68, 1268–1273.

Sperling, R.A., Aisen, P.S., Beckett, L.A., Bennett, D.A., Craft, S., Fagan, A.M., Iwatsubo, T., Jack Jr., C.R., Kaye, J., Montine, T.J., Park, D.C., Reiman, E.M., Rowe, C.C., Siemers, E., Stern, Y., Yaffe, K., Carrillo, M.C., Thies, B., Morrison-Bogorad, M., Wagster, M.V., Phelps, C.H., 2011. 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. 7, 280–292.

Tondelli, M., Wilcock, G.K., Nichelli, P., De Jager, C.A., Jenkinson, M., Zamboni, G., 2012. Structural MRI changes detectable up to ten years before clinical Alzheimer's disease. Neurobiol. Aging 33, 825.e825–825.e836.

Winkler, A.M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P.T., Duggirala, R., Glahn, D.C., 2010. Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. NeuroImage 53, 1135–1146.

Zakzanis, K.K., Mraz, R., Graham, S.J., 2005. An fMRI study of the Trail Making Test. Neuropsychologia 43, 1878–1886.