ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 40 (2008) 1503–1515

Multivariate and univariate neuroimaging biomarkers of Alzheimer's disease☆

Christian Habeck,[a,*] Norman L. Foster,[b] Robert Perneczky,[c] Alexander Kurz,[c] Panagiotis Alexopoulos,[c,d] Robert A. Koeppe,[e] Alexander Drzezga,[f] and Yaakov Stern[a]

[a]*Cognitive Neuroscience Division, Taub Institute, Columbia University Medical Center, PH18, 622 West 168th Street, New York, NY 10032, USA*
[b]*Center for Alzheimer's Care, Imaging and Research, Department of Neurology, University of Utah, 650 Komas #106A, Salt Lake City, UT 84108, USA*
[c]*Department of Psychiatry and Psychotherapy, Technical University Munich, Ismaninger Str. 22, 81675 München, Germany*
[d]*Department of Psychiatry and Psychotherapy, University Erlangen, Schwabachanlage 6, 91054 Erlangen, Germany*
[e]*Department of Radiology, University of Michigan, 3480 Kresge III/Box 0552, Ann Arbor, MI 48109, USA*
[f]*Department of Nuclear Medicine, Technical University Munich, Ismaninger Str. 22, 81675 München, Germany*

Received 2 August 2007; revised 14 December 2007; accepted 22 January 2008
Available online 14 February 2008

We performed univariate and multivariate discriminant analysis of FDG–PET scans to evaluate their ability to identify Alzheimer's disease (AD). FDG–PET scans came from two sources: 17 AD patients and 33 healthy elderly controls were scanned at the University of Michigan; 102 early AD patients and 20 healthy elderly controls were scanned at the Technical University of Munich, Germany. We selected a derivation sample of 20 AD patients and 20 healthy controls matched on age with the remainder divided into 5 replication samples. The sensitivity and specificity of diagnostic AD-markers and threshold criteria from the derivation sample were determined in the replication samples. Although both univariate and multivariate analyses produced markers with high classification accuracy in the derivation sample, the multivariate marker's diagnostic performance in the replication samples was superior. Further, supplementary analysis showed its performance to be unaffected by the loss of key regions. Multivariate measures of AD utilize the covariance structure of imaging data and provide complementary, clinically relevant information that may be superior to univariate measures.
Published by Elsevier Inc.

Introduction

The prevalence of Alzheimer's disease (AD) is increasing rapidly, and a large-scale research effort is devoted to understanding, characterizing and curing the disease to avoid its financial and social burden as the population ages. One theme of this research has been to identify diagnostic biomarkers through neuroimaging that enable early detection of the disease, even before the full clinical symptoms of AD have become manifest. Neuroimaging has identified a wide array of biomarkers that can differentiate AD patients from healthy control subjects such as volume loss measured with morphometry (e.g., Bozzali et al., 2006; Chetelat et al., 2005; Teipel et al., 2005; Thomann et al., 2006), cerebral blood flow (e.g., Kasama et al., 2005; Nakano et al., 2006; Trollor et al., 2005), glucose metabolism (Burdette et al., 1996; Chetelat et al., 2003b; Foster et al., 2007; Herholz et al., 2002; Higdon et al., 2004) and beta-amyloid deposition (e.g., Engler et al., 2006; Nichols et al., 2006). Often these analyses target specific locations (entorhinal cortex, hippocampus, prefrontal cortex, amygdala, etc.) that are known to be implicated in the natural history of the disease. If proceeding on a brain-wide level, researchers usually employ univariate analysis that voxel-by-voxel contrasts the particular imaging signal of choice between healthy controls and AD patients.

In this article, we report on the derivation and prospective application of multivariate biomarkers obtained from two different clinical populations, scanned with ^{18}F-fluorodeoxyglucose positron emission tomography (FDG–PET). Multivariate analytic techniques aimed at identifying diagnostic neural networks have been applied less frequently than univariate ones, particularly in Alzheimer's disease. There have been recent successful applications of neural network techniques using back propagation (Warkentin et al., 2004) and Bayesian non-linear analysis of structural MRI data from MCI subjects (Davatzikos et al., 2006). Further, a default-mode resting network, identified with independent components analysis from normal elders, has been shown to be abnormal in AD subjects (Greicius et al., 2004) and could therefore serve as a sensitive biomarker. Apart from these analyses, there recently has been work from our (Devanand et al., 2006; Scarmeas et al., 2004) and other laboratories (Huang et al., 2007; Kerrouche et al., 2006) utilizing

☆ Grant support: National Institute of Biomedical Imaging and Bioengineering (R01EB006204) and National Institute on Aging (R01AG22394 and R01AG08671).
* Corresponding author.
E-mail address: ch629@columbia.edu (C. Habeck).

1053-8119/$ - see front matter. Published by Elsevier Inc.
doi:10.1016/j.neuroimage.2008.01.056

principal components analysis (PCA) for the detection of dementia. Our aim was to build on our and other groups' successful usage of PCA for the detection of neurodegenerative diseases. PCA has a long history with successful replications and prospective applications of diagnostic PET patterns in Parkinson's disease (e.g., Carbon et al., 2004; Eidelberg et al., 1991; Eidelberg et al., 1997; Trost et al., 2002; Trost et al., 2006), and even in normal aging (e.g., Alexander et al., 2006; Brickman et al., 2007; Eidelberg et al., 1997; Moeller et al., 1996).

We extended the PCA -approach to the FDG–PET data from two clinical populations and analyzed the spatially correlated metabolism as a function of disease status. Despite the absence of precise mechanistic knowledge about spatially correlated changes brought about by Alzheimer's disease, we reason that such correlation of the imaging signal is plausible when considering spatially spreading disease pathology that induces regionally correlated changes in blood flow or metabolism. For instance, deposition of neurofibrillary plaques and tangles spread in a pattern that is the reverse of cortical myelination (Braak et al., 1998; Delieu and Keady, 1996; Pearson, 1996). Thus, any accompanying reduction in glucose level can be expected to manifest a spreading spatial pattern as well. This means that, even when there is no spatial correlation induced by any form of long-range connectivity that 'communicates' pathologic changes between distant regions, spatial covariance might be able to identify covariance patterns that can signal the disease onset. Further, even a relatively isolated pathologic change can induce a pattern of metabolic change across a larger area.

We previously performed a cross-sectional comparison of normal healthy controls and subjects with early AD (Scarmeas et al., 2004) and using single photon emission computed tomography (SPECT) identified correlated changes in blood flow as a function of disease status, while univariate analysis and visual interpretation both failed to detect any differences as a function of disease status. Furthermore, the AD-related covariance pattern that we obtained from contrasting early AD subjects and healthy controls passed several independent plausibility tests: (1) when prospectively applied to an independent sample of MCI subjects, the magnitude of the pattern expression correlated with the correct directionality with several neuropsychological variables, i.e. the more the MCI subjects expressed the AD-related patterns, the worse they performed. (2) Further, subject expression of the AD-pattern in the MCI subjects could also predict future decline in subjects' neuropsychological scores (Devanand et al., 2006): the more subjects expressed the AD-related pattern at baseline, the faster they declined. These findings give a brief glimpse of the potential of multivariate analytic approaches.

In addition to conceptual and neuroscientific plausibility, multivariate analyses also have enhanced statistical power to detect non-focal changes in imaging signal. This is intuitively obvious when keeping in mind that a bran scan consists of the order of 10^4 to 10^5 image elements (voxels). Performing a multivariate analysis is akin to finding the major source of variance of the data in this high-dimensional space along which there is a significant group- or condition-related difference in expression of the brain scans. If such major source of variance is not identified with a PCA first, a voxel-by-voxel comparison of the group- or condition-related difference might risk losing significance, since now many independent comparisons have to be performed leading to a multiplication of the *p*-level by the number of voxels, which is a very stringent threshold that might result in 'correcting away' true effects of interest in the data.

Methods

Subject and image acquisition

FDG–PET scans were obtained from two scanning sites: the Technical University of Munich and the University of Michigan. Details of subject recruitment and image preprocessing analysis will be given for both centers separately.

Munich sample

Patients with AD and healthy controls without memory complaints (=spouses of AD patients), who where all examined between 2001 and 2003 at the centre for cognitive disorders of the technical university of Munich were identified in an electronic database. All study participants were administered with an IV bolus of 370 mBq ^{18}F-FDG at rest 30 min prior to PET scanning. Scans were performed under standard resting conditions with the patient's eyes closed in dimmed ambient light. Exactly the same scanning protocol was applied to every subject. Imaging was performed on a Siemens ECAT/EXACT HR + PET scanner (CTI, Knoxville, TN). A sequence of three frames (10 min, 5 min, and 5 min) was started [3-dimensional (3D) mode, total axial field of view of 15.52 cm] and later combined into a single frame. Patients were positioned with the canthomeatal line parallel to the detector rings to obtain transaxial images parallel to the intercommissural line. Attenuation correction was performed using a transmission scan. Data were corrected for random, dead time and scatter, and images were reconstructed by filtered back-projection with a Hamm filter (cut-off frequency 0.5 cycles/projection element) resulting in 63 slices in a 128×128 pixel matrix (pixel size 2.06 mm) and interplane separation of 2.425 mm. The clinical diagnosis was established by an experienced psychiatrist according to National Institute of Neurological and Communicative Diseases and Stroke/Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria for AD (McKhann et al., 1984) and revised Mayo criteria for MCI (Winblad et al., 2004). The neuropsychological evaluation was based on the Consortium to Establish a Registry for Alzheimer's Disease neuropsychological assessment battery (CERAD-NAB) (Berg et al., 1988), which incorporates the Mini-Mental-State Examination (Folstein et al., 1975). Overall severity of dementia was rated with the Clinical Dementia Rating scale (CDR) (Rubin et al., 1993). Only patients with questionable (CDR=0.5) or mild dementia (CDR=1) were included. The healthy controls underwent functional brain imaging according to the same routine protocol. Patients were excluded who fulfilled diagnostic criteria of neurodegenerative causes for dementia other than AD, such as Lewy body disease (McKeith et al., 1996), frontotemporal lobar degeneration (Neary et al., 1998), or Parkinson's disease (Hughes et al., 1992). Patients were also excluded if they had significant cerebrovascular lesions on their structural brain scans, relevant functional psychiatric disorders such as major depression, or a history of traumatic brain injury, stroke, cerebral tumors, epilepsy, or alcohol abuse. All consecutive individuals who fulfilled the inclusion and exclusion criteria were entered into the study. Written informed consent forms were available for all patients. The ethics committee of the technical university of Munich approved of the scanning of the healthy control group.

All study participants were administered with an IV bolus of 370 mBq ^{18}F-FDG at rest 30 min prior to PET scanning. Scans were performed under standard resting conditions, with the participants' eyes closed in dimmed ambient light on a Siemens 951 R/31 scanner (CTI). Exactly the same scanning protocol was applied in

every case. A sequence of three 10-min frames was acquired and later combined into a single frame. The acquisitions were performed in 2D mode with a total axial field of view of 10.5 cm and no interplane dead space. Patients were positioned with the canthomeatal line parallel to the detector rings to obtain transaxial images parallel to the intercommissural line. Attenuation correction was performed using a transmission scan from the end of the session. Corrections for random, dead time, and scatter were performed after data acquirement, and images were reconstructed by filtered backprojection with a Hamm filter (cut-off frequency 0.5 cycles/projection element), resulting in 47 slices in a 128 × 128 pixel matrix (pixel size 2.0 mm) and interplane separation of 3.447 mm (Drzezga et al., 2005).

Michigan sample

Patients received an FDG–PET study on the ECAT EXACT scanner at the University of Michigan between December 1993 and February 2001 and subsequently received a postmortem examination documenting a histopathological diagnosis of AD, uncomplicated by other pathology such as stroke or significant numbers of cortical Lewy bodies. Only individuals with retrievable parametric PET images that included most of the brain in the field of view were considered. In distinction to other similar series from this center (Foster et al., 2007; Higdon et al., 2004), only studies from a single scanner comparable to that used in Munich were included. A total of 17 individuals (10 men and 7 women; mean ± SD age at scan, 65.5 ± 13.9 years; age range, 33–82 years) were identified who had been evaluated by dementia-specialist neurologists at the Michigan Alzheimer's Disease Research Center. During their clinical evaluation, NINCDS-ADRDA criteria were met in 16 for probable AD in one possible AD. MMSE score at the time of scan is available for 13 of these individuals (mean ± SD, 16.0 ± 6.3; range, 2–24). (These individuals with MMSE scores were 8 men and 5 women; mean ± age at scan 67.3 ± 13.7; range, 33–82.) All met NIA-Reagan neuropathological criteria for either high (12 cases) or intermediate (5 cases) likelihood of AD (NIA and Reagan Institute Working Group, 1997). Cortical Lewy bodies meeting neuropathological criteria for dementia with Lewy bodies also were present in 6 of the AD subjects (McKeith et al., 1996). Although not evident on structural brain imaging or clinical examination when they were scanned, 5 patients had vascular pathology at autopsy. A large, acute cerebral infarct caused the death of two subjects and a massive acute intracerebral hemorrhage in another. Two additional patients had small lacunar infarcts of indeterminate age. For comparison, we used data from 33 normal individuals (19 men and 14 women; mean ± age at scan, 68.5 ± 8.2 years; range, 58–91 years) who were scanned contemporaneously and had served as control subjects for a several other research studies.

Table 1
Scanner parameters for Munich (left column) and Michigan (right column)

Characteristic	Siemens/CTI ECAT/Exact HR+	Siemens/CTI ECAT/Exact 47
Detector	Bismuth germanium oxide	
Mode/image planes	3D/63	2D or 3D/47
Acquisition matrix	128 × 128	
XY pixel dimension (mm)	2.0	1.91
In-plane × axial resolution (mm)	4.4 × 4.1	8.0 × 5.0
Slice thickness (mm)	2.46	3.375
File format	ECAT 7	CTI-6 or 7

Table 2
Descriptive statistics of all subject groups

	Michigan AD	Michigan HE	Munich AD	Munich HE
N	17	33	102	25
Age, years	65.5 ± 13.9	67.5 ± 7.8	69.3 ± 9.9	58.2 ± 10.8
MMSE	16.0 ± 6.3	N/A	22.8 ± 3.2	N/A

MMSE was not available in all subjects.

Subjects were scanned using an ECAT EXACT 47-image plane PET scanner with bismuth germanium oxide detectors (CTI). A transmission scan with a ^{68}Ga/^{68}Ge pin source was obtained for attenuation correction. Emission scans were acquired in 2D or 3D mode. The scanner has a total axial field of view of 16.2 cm and no interplane dead space. PET studies were performed with the subjects at rest with eyes closed and ears unplugged, comfortably lying in a darkened and quiet room. All subjects fasted for at least 4 h before PET scanning and had a normal blood glucose level at the time of the scan. Thirty minutes after injection of approximately 370 MBq ^{18}F-FDG, a sequence of three 10-min frames was acquired and later summed into a single frame. Images were reconstructed by filtered back projection with a Hanning filter (cutoff frequency at 0.5 cycles/projection element). The dimensions of the reconstructed PET images were 128 × 128 with *XY* pixel dimensions of 1.91 mm. The in-plane resolution was 8.0 mm and axial resolution was 5.0 mm. There was an interslice distance of 3.375 mm without any gap between slices (Table 1).

In the Munich sample, age and MMSE was available for all MCI and AD patients, but only age was available for healthy elderly controls. In the Michigan sample, age and MMSE was only available for the AD subjects, with 4 people out of 17 AD patients missing their MMSE information. We summarized all the information available for all sub samples of subjects in Table 2.

Image processing

All PET images were spatially normalized and smoothed at FWHM = 12 mm, and masked according the canonical SPM-supplied probabilistic gray matter mask. Voxels with a gray matter probability bigger than 50% were retained in the analysis. This resulted in 125 regional resolution elements (RESELs), each one containing 1475 voxels. Further, since no absolute quantification of FDG–PET signal was undertaken with the image acquisition, all images were normalized (by division through their individual whole-brain means) to yield an average unit value. For this averaging calculation, only the voxels within the probabilistic gray-matter mask were included. Thus, only statements about *relative* metabolic level can be made. In particular, increased *relative* metabolism in Alzheimer's patients compared to control subjects is now possible and unsurprising, even if *absolute* metabolism might be lower.

PCA approach

We give a brief mathematical sketch for those unfamiliar with covariance analysis and the idea of dimensionality reduction. We denote the data array to be analyzed as $Y(s, \mathbf{x})$, with the index s symbolizing subjects, while $\mathbf{x}$ stands for the 3D voxel location in the brain.

The goal of any multivariate decomposition is to separate subject from voxel indices in the following series:

$$Y(s,\mathbf{x}) = SSF_1(s)v_1(\mathbf{x}) + SSF_2(s)v_2(\mathbf{x}) + SSF_3(s)v_3(\mathbf{x}) + \cdots$$

The data array has been decomposed into a series of terms of diminishing variance contribution. The terms are made up of subject-dependent subject scaling factors, SSF_i, and voxel-dependent principal component topographies, i.e. brain images, v_i.[1] Usually, only the first few (typically less than a third of the number of subjects) terms are retained to represent the data. For instance, if we decide (by some hitherto unspecified criterion) that we want to retain the first two terms, we have two principal component images, $v_1(\mathbf{x})$ and $v_2(\mathbf{x})$, which are fixed across subjects and tasks, and two sets of subject scaling factors, $SSF_1(s)$ and $SSF_2(s)$, that are varying across subjects. These latter two sets of numbers can be used for further analysis and correlation with clinical and experimental variables. [In the current study, we used clinical status (1 = AD/0 = healthy) as the dependent variable in this linear regression.] The goal of dimensionality reduction thus becomes clear: rather than keeping track of subject- and task-dependent numbers at every voxel location, we only have two subject- and task-dependent variables to deal with.

Derivation of AD-related pattern

To identify network-correlates of early dementia and age, we employed the Scaled Subprofile Model (SSM), a covariance-analysis method (Alexander et al., 1999; Moeller et al., 1987) that has been used previously in resting imaging studies of normal aging and a variety of neurological diseases (Alexander et al., 1999; Hutchinson et al., 2000; Moeller et al., 1996; Nakamura et al., 2001). This analysis was applied to the FDG–PET images acquired for AD and control subjects. SSM captures the major sources of between- and within-group variation in these images and produces a series of principal components (PCs). Using Akaike's information criterion for the goodness-of-fit estimation (Burnham and Anderson, 2002), the optimal number of PCs were identified that should be included as predictors in a linear regression using group membership (AD vs. control) as the outcome measure. The best-fitting linear combination of these PCs will define an AD-related covariance pattern, while accounting for most of the CBF variance in all subjects. We summarize the steps of this algorithm in detail; no room was left for arbitrary decisions on the part of the analyst. Because of the age-discrepancy documented above, age was carefully matched by picking a subset of 20 subjects of both AD and controls groups who did not differ in age.

1. The realigned, spatially transformed and smoothed FDG–PET images from both the AD patients and healthy elderly subjects were simultaneously included in a Principal Components Analysis, which capture the major sources of between- and within-groups variation, producing a series of principal components. Voxels participating in each PC may have either a positive or a negative loading. Voxels with positive loadings can be conceived as exhibiting concomitant increased signal and those with negative loadings can be conceived as exhibiting concomitant decreased signal. These loadings are fixed and the same for all subjects.
2. The expression of each PC for each subject was quantified by a subject-scaling factor (SSF). A higher SSF value indicates more prominent concomitantly increased signal of the voxels with positive loadings and more prominent concomitantly decreased signal of the voxels with negative loadings. Therefore, the SSFs express the degree of subjects' expression of the fixed PC.
3. To identify a covariance pattern that best discriminates AD patients from controls, each subject's expression of the specified PCs derived from Step 1 will be entered into a linear regression model as the independent variable. Group membership (AD versus controls) will be the dependent variable. This regression results in a linear combination of the PCs that best discriminated the 2 groups. This linear combination of the PCs can itself be conceived of as signifying a 'pattern' or 'network'. We used Akaike's information criterion (Burnham and Anderson, 2002) to determine how many PCs should be included in the regression in order to achieve optimal bias–variance trade-off. The set of PCs that yields the lowest value in Akaike's information criterion will be selected as predictors in the regression model. We restricted the set of PCs to a cumulative one out of the first few PCs that explain at least 75% of the variance. This means, for instance, if 6 PCs account for 75% of the variance, we would check the following 6 sets for their value of the Akaike criterion (AIC) in the group discrimination: PC1, PC1–2, PC1–3, PC1–4, PC1–5, PC1–6. The set that yields the lowest value of AIC will be selected for construction of the discriminant pattern.

Bootstrap estimation of stability of regional weights in covariance pattern

Patterns resulting from any multivariate analysis assign different weights to all voxels included in the analysis, depending on the salience of their covariance contribution. Positive voxel weights indicate a positive correlation between the subject expression value and the associated regional glucose metabolism, whereas negative weights indicate a negative correlation. This means that as the expression of a pattern *increases*, activation in the positively weighted regions *increases* as well, whereas activation in the negatively weighted regions *decreases*. Whether a voxel weight is reliably different from zero is assessed by a bootstrap estimation procedure (Efron and Tibshirani, 1994). This procedure entails performing the complete computation listed above (Steps 1–3) several hundred times (~500) on data that were resampled with replacement from the pool of AD patients and controls. Denoting the point estimate of a voxel weight as $\mathbf{w}$ and the standard deviation around $\mathbf{w}$ resulting from the bootstrap resampling procedure as $s_{\mathbf{w}}$, we can compute a Z-score according to $z = \mathbf{w}/s_{\mathbf{w}}$. Sufficiently small variability of a voxel weight around its point estimate value in the resampling processes results in Z-value of large magnitude, and indicates a reliable contribution to the covariance pattern. As the threshold criterion, we chose $|z| > 3.36$; under the assumptions of a standard-normal distribution, i.e. $z \sim N(0,1)$, this corresponds to a one-tailed probability of 0.05 with a multiple-comparisons correction for 125 RESELs (resulting from our smoothing and masking steps explained above). To aid in clarification, we mention here that the bootstrap estimation procedure and the ensuing Z-map are used for *visualization* purposes only. For any supplementary analysis discussed in this paper, the point estimate of the covariance pattern itself was used for any prospective application.

[1] We emphasize that the subject-dependent scaling factors are not to be confused with any other scaling operation that is performed as a corrective measure *prior* to any group-level data analysis.

Univariate analysis

Similar to the multivariate analysis for the production of the AD-related covariance pattern, we also conducted a univariate contrast between healthy elderly controls and AD subjects. 20 scans from both AD subjects and healthy elderly controls were subjected to a voxel-wise *T*-test. The threshold for significance was adjusted by the number RESELs (125) to a yield a corrected two-tailed *p*-level of 0.05 and came out as $T=3.87$.

Prospective application to replication samples

To test how well univariate and multivariate group-differences between AD patients and controls generalize to a more representative pool of subjects, we split our pool of subjects into a *derivation sample*, employed in both our univariate and multivariate analysis to identify AD-related group differences first, and several *replication samples*. The AD-related differences identified from the derivation sample were then tested in all replication samples as well: expression of the AD-related covariance pattern was tested for its diagnostic ability as well as relative FDG level at any voxel locations that showed significant AD-related effects in the derivation sample. The subject expression of the AD-related pattern in the replication sample was obtained as follows: for every subject in a replication sample, the subject's scan was multiplied with the loadings of the covariance pattern in a voxel-wise manner; the resulting product image was then summed over all voxel locations to yield the level of subject expression as a single number.

For the derivation sample, we picked up 20 Munich AD-patients (out of a total of 102) and 20 Munich controls that were age-matched. The remaining 82 Munich AD-patients were split into 3 samples of 20 subjects, and 1 sample of 22 subjects. No additional control subjects were available, so that these 4 replication samples could only test the sensitivity, but not the specificity, of the derived AD-markers independently. A fifth replication sample was provided by the Michigan subjects (17 ADs/33 controls), which provided a test of both specificity and sensitivity of the AD-markers.

To simulate the realistic context of a clinical application of both AD-markers on a subject-by-subject basis, we used the threshold criteria in the replication samples for which the AD-markers gave the optimal sensitivity and specificity values in the derivation sample. Just relying on the area under the ROC-curve is not enough, since this statistic summarizes the discriminability across the range of all possible values of the threshold criterion and cannot be computed for just one human subject in isolation, without any reference to other human subjects. Further, it is unaffected by uniform shifts in the value of the AD-marker, regardless of the magnitude of the shift. This means in particular that in spite of impressive ROC-curves that result from the prospective application of the AD-markers in the replication samples, the values for sensitivity and specificity *at the thresholds chosen in the derivation sample* might not be as impressive. This is the crucial test whether the AD-marker is useful in a clinical context where a diagnosis has to be offered based on its absolute value in single subjects.

Results

Clinical evaluation

Subject information other than disease status was only available in form of years of age and Mini-Mental Status Examination (MMSE), but not for every person in the sample. In the Munich sample, age and MMSE was available for all MCI and AD patients, but only age was available for healthy elderly controls. In the Michigan sample, MMSE was only available for the AD subjects, with 4 people missing their MMSE information. We therefore restricted our comparisons to the AD portions of both samples. Two-sample *t*-tests revealed no differences in age ($p=0.18$), but a significant difference in MMSE (Michigan: 16.0 ± 6.3, Munich: 22.8 ± 3.2, $p<0.0001$), indicating that the AD patients from the Michigan sample are more advanced in their disease course than the Munich patients.

Derivation of AD-related pattern

As mentioned in the previous section, for the derivation of the AD-related pattern, we matched 20 subjects of both patient and control groups on age on the mean, before employing our covariance analysis. Age was now *not* significantly different as a function of disease-status any longer (AD derivation group: 62.0 ± 9.3; healthy elderly derivation group: 62.0 ± 9.0). The first 2 principal components accounted for 29% of the variance in the data. Using these principal components, we constructed a single pattern whose subject expression distinguished between subjects and controls successfully. We then performed the bootstrap resampling procedure described before to arrive at a *Z*-map, sampling with replacement from the all subjects in the derivation sample. Robust regional weights, i.e. *Z*-values that surpass the threshold $|Z|>3.36$, in the covariance pattern were listed in Table 3; robust *positive* weights indicating an increase in PET signal in the AD relative to the control subjects were mainly found in frontal lobe areas. *Negative* weights, indicating a *decrease* in signal in the AD groups vs. controls, were found in parietal, temporal and prefrontal areas. We emphasize again that these are *relative* changes; no absolute quantification was possible.

Table 4 summarizes the performance characteristics of the AD-related pattern in all derivation and replication samples. The threshold for optimal sensitivity and specificity of the group discrimination in the derivation was deemed to be -10.0. For the prospective application in the replication samples, subjects were therefore diagnosed as AD when they expressed the covariance pattern to a degree >-10.0.

One can appreciate from Table 4 that the successful group discrimination achieved by the AD-related pattern in the derivation sample generalizes to all replication samples as well. The specificity values are fixed for all Munich samples at 0.95 since the same 20 control subjects were used in all samples with the same decision threshold (-10), but the sensitivity fluctuated in the range [0.85,1]. Reassuringly, sensitivity and ROC-area values were *not* highest for the derivation sample itself, underscoring the generalizability of the results beyond the derivation sample. Even in the Michigan sample, whose data were acquired on a different scanner with a different protocol, the Munich-derived AD-pattern has good diagnostic ability.

We also tested the association between expression of AD-related pattern and age in all replication samples; since there were clear age differences between control subjects and AD patients in all 4 Munich replication samples, we restricted our correlational analysis to the AD patients only in the Munich replication samples. A significant *negative* correlation between age and AD-pattern expression was found in replication sample 1 ($R^2=0.40$, $p<0.005$), but otherwise no correlations reached significance ($p>0.3$ in replication samples 2, 3, and 4). For the Michigan sample, age was comparable in both patients and controls, so the correlation between AD-pattern expression and age was computed for all 50 subjects; again, no significant correlation was obtained ($R^2=0.05$, $p=0.16$). Thus, apart from a

Table 3
MNI coordinates and listings of local *Z*-maxima farther than 8 mm apart, used to construct the covariance pattern indicating their weighting for either increased and decreased glucose uptake relative to the mean in the Alzheimer's groups as compared to the control subjects

X	*Y*	*Z*	Laterality/lobe	Structure	Brodmann label	*Z*
Positive weights = relatively increased FDG signal in AD						
18	−4	−8	R Sub-lobar	Lentiform nucleus	Medial globus pallidus	9.8539
58	2	10	R Frontal lobe	Precentral gyrus	Brodmann area 6	9.8216
34	−4	−22	R Limbic lobe	Parahippocampal gyrus	Brodmann area 36	7.4967
−64	−12	16	L Parietal lobe	Postcentral gyrus	Brodmann area 43	5.7304
−44	−4	14	L Sub-lobar	Insula	Brodmann area 13	5.3955
−36	−2	12	L Sub-lobar	Insula	Brodmann area 13	5.1361
2	−40	68	R Frontal lobe	Paracentral lobule	Brodmann area 4	5.6647
−6	−30	58	R Frontal lobe	Medial frontal gyrus	Brodmann area 6	4.1707
40	−28	44	R Parietal lobe	Postcentral gyrus	Brodmann area 2	3.9627
−14	18	−14	L Frontal lobe	Subcallosal gyrus	Brodmann area 47	3.8058
−48	−20	36	L Parietal lobe	Postcentral gyrus	Brodmann area 3	3.571
−28	−10	10	L Sub-lobar	Lentiform nucleus	Putamen	3.5399
−50	−24	0	L Temporal lobe	Superior temporal gyrus	Brodmann area 21	3.5134
Negative weights = relatively decreased FDG signal in AD						
50	−62	26	R Temporal lobe	Superior temporal gyrus	Brodmann area 39	14.1039
40	−68	48	R Parietal lobe	Inferior parietal lobule	Brodmann area 7	8.1061
−4	−38	42	L Limbic lobe	Cingulate gyrus	Brodmann area 31	10.6563
10	−54	36	L Parietal lobe	Precuneus	Brodmann area 7	8.5122
−50	−58	38	L Parietal lobe	Inferior parietal lobule	Brodmann area 40	10.4684
−58	−48	36	L Parietal lobe	Supramarginal gyrus	Brodmann area 40	9.9976
−44	−60	16	L Temporal lobe	Superior temporal gyrus	Brodmann area 39	9.5121
−44	10	46	L Frontal lobe	Middle frontal gyrus	Brodmann area 8	7.9338
42	10	32	R Frontal lobe	Inferior frontal gyrus	Brodmann area 9	7.881
44	8	46	R Frontal lobe	Middle frontal gyrus	Brodmann area 8	7.565
−54	−48	−16	L Temporal lobe	Inferior temporal gyrus	Brodmann area 20	4.8012
−64	−42	−4	L Temporal lobe	Middle temporal gyrus	Brodmann area 21	4.3843
44	22	−6	R Frontal lobe	Inferior frontal gyrus	Brodmann area 47	3.7552
−10	4	10	L Sub-lobar	Caudate	Caudate Body	3.6025
−28	54	12	L Frontal lobe	Middle frontal gyrus	Brodmann area 10	3.4568
−34	46	14	L Frontal lobe	Middle frontal gyrus	Brodmann area 10	3.4463

The areas listed have a *Z*-value $|Z|>3.36$ corresponding to a one-tailed *p*-value of 0.05, corrected for the number of regional resolution elements (125). Only nearest gray matter areas were listed. The location of the most positive loading [−28 −50 −32], was not in any gray matter region and is not listed below.

negative correlation between pattern expression and age in Munich replication sample 1, which defies a straightforward explanation on account of the sign of the correlation, we have no indication that our AD-related pattern was influenced by age.

Univariate analysis

We also conducted a univariate contrast using values in a single region of interest in the derivation sample, and then mirrored the forward application of the AD-related pattern to all replication samples, by checking the group separation at the location of the largest *T*-value in the derivation sample. The right parietotemporal area (MNI: [50 −58 30]) showed the largest separation between the groups, with a decreased relative cerebral glucose level in the AD subjects. There were no areas of *increased* relative metabolic signal in the AD subjects. We tabulated all areas showing a significant difference between groups in Table 5.

We then interrogated the ability of the parietotemporal area's FDG–PET signal to discriminate between patients and control in the derivation sample and all replication samples. We chose an optimal

Table 4
Descriptive statistics of the subject strengths of AD subjects ("AD") and healthy elderly controls ("HE") and expression of the AD-related covariance pattern derived from the Munich derivation sample in all Munich samples and in the Michigan sample

	N (AD)/*N* (HE)	Age (AD)/Age (HE)	Sensitivity	Specificity	ROC area
Munich derivation sample	20/20	62.0±9.3/62.0±9.0	0.90	0.95	0.96
1st replication Munich sample	20/20	69.2±10.4/62.0±9.0	1	0.95	0.97
2nd replication Munich sample	20/20	70.7±8.9/62.0±9.0	0.85	0.95	0.92
3rd replication Munich sample	20/20	70.6±9.2/62.0±9.0	1	0.95	0.97
4th replication Munich sample	22/20	73.5±8.5/62.0±9.0	0.91	0.95	0.87
Michigan replication sample	17/33	65.5±13.9/68.5±7.6	0.80	1	0.90

All but one Munich samples contain the same healthy elderly control subjects, while differing in the AD subjects. The specificity is thus fixed at the same value for all Munich samples. Sensitivity for all samples was calculated using the optimal threshold (−10) from the Munich derivation sample to simulate genuine prospective application, rather than optimizing the threshold for each sample anew.

Table 5
MNI coordinates and listing of brain areas showing significant differences in FDG uptake relative to mean values between the groups in the derivation sample as identified by the univariate comparisons

X	X	Z	Laterality/Lobe	Structure	BA label	T
50	−58	30	R Temporal Lobe	Superior Temporal Gyrus	Brodmann area 39	8.02
38	−70	50	R Parietal Lobe	Superior Parietal Lobule	Brodmann area 7	6.45
−52	−56	32	L Parietal Lobe	Supramarginal Gyrus	Brodmann area 40	7.81
−50	−58	42	L Parietal Lobe	Inferior Parietal Lobule	Brodmann area 40	7.62
−46	−62	20	L Temporal Lobe	Superior Temporal Gyrus	Brodmann area 39	7.56
−4	−46	38	L Limbic Lobe	Cingulate Gyrus	Brodmann area 31	7.56
6	−50	36	R Parietal Lobe	Precuneus	Brodmann area 31	7.29
4	−70	40	R Parietal Lobe	Precuneus	Brodmann area 7	5.99
−44	14	46	L Frontal Lobe	Middle Frontal Gyrus	Brodmann area 8	6.40
−28	20	52	L Frontal Lobe	Middle Frontal Gyrus	Brodmann area 8	5.85
−44	12	32	L Frontal Lobe	Middle Frontal Gyrus	Brodmann area 9	5.66
−62	−38	−8	L Temporal Lobe	Middle Temporal Gyrus	Brodmann area 21	5.35
−56	−50	−14	L Temporal Lobe	Middle Temporal Gyrus	Brodmann area 37	5.10
−52	−34	−24	L Temporal Lobe	Fusiform Gyrus	Brodmann area 20	4.16
44	16	48	R Frontal Lobe	Middle Frontal Gyrus	Brodmann area 8	4.92
30	20	52	R Frontal Lobe	Middle Frontal Gyrus	Brodmann area 8	4.83
62	−38	−18	R Temporal Lobe	Middle Temporal Gyrus	Brodmann area 21	4.81

Displayed are local T-maxima farther than 8 mm apart with at $p<0.05$, after multiple-comparison correction for the number of resolution elements (125) and the *T*-threshold chosen was $|T|=3.87$. There were no areas of *increased* relative FDG level in the AD subjects. Only nearest gray matter areas were listed.

threshold of 1.05 as the criterion with the best sensitivity and specificity value, yielding a value of 0.95 for both. For the application in the replication samples, the diagnostic algorithm therefore was as follows: inspect the signal value at the parietotemporal location; if it is less than 1.05 the average whole-brain signal value the subject is diagnosed as AD, otherwise deemed normal. Table 6 lists sensitivity, specificity, *T*-value of the healthy–AD contrast, and ROC area associated with the parietotemporal location. Specificity for all Munich replication samples was again fixed at 0.95, since the same control subjects were used.

Fig. 1 graphically summarizes the information displayed in Tables 2 and 4, and enables a visual comparison of multivariate and univariate markers. Fig. 2 shows the discrimination performance of both biomarkers in the Munich derivation sample. One can appreciate that both univariate and multivariate biomarkers achieve good discriminability across all samples. This manifests itself in superior *T*-contrasts and ROC-areas. Notably, the displayed ROC-curves for the derivation sample look almost identical. However, of real clinical relevance are sensitivity and specificity values *at the decision threshold chosen from the derivation sample*, since *T*-contrasts and ROC-areas are summary statistics whose computation requires a multitude of subjects in the replication sample. This obviates their use for a more realistic subject-by-subject application. Further, similarly to a simple correlation coefficient, they are invariant with respect to arbitrary, but uniform, shifts in the AD-marker value, which is also unsatisfactory.

When comparing the univariate and multivariate AD-marker on the more rigorous performance criterion of sensitivity and specificity of diagnosis at the fixed threshold level obtained from the derivation sample, we can see that the multivariate marker does substantially better than the univariate one. One can see this in particular for replication samples 3 and 4. For these samples, the discriminability between patients and controls was somewhat diminished, but the sensitivity at a threshold of 1.05 suffered more and only yielded poor values of 0.55 and 0.59, respectively. In these samples, the multivariate marker performed much better with sensitivities of 1.0 and 0.91, respectively. Expression of the AD-related covariance pattern is hardly affected by overly high signal values at location [50 −58 30].

For the Michigan sample, on the other hand, we can see that, despite very good T-contrast, ROC-area and sensitivity at threshold 1.05, the univariate marker has poor specificity (0.67). Obviously, AD patients have lower signal values at location [50 −58 30] than controls, but *everybody's* value seems to be depressed — and many

Table 6
Descriptive statistics of subject strengths of AD subjects ("AD") and healthy elderly controls ("HE") and PET signal at the parietal location [50 −58 30], the area of maximum contrast in the Healthy – AD contrast for the Munich derivation sample

	N(AD)/N(HE)	Sensitivity	Specificity	T-value (Healthy – AD)	ROC area
Munich Derivation sample	20/20	0.95	0.95	8.02	0.97
1st replication Munich sample	20/20	0.90	0.95	8.12	0.93
2nd replication Munich sample	20/20	0.80	0.95	6.19	0.92
3rd replication Munich sample	20/20	0.55	0.95	4.17	0.82
4th replication Munich sample	22/20	0.59	0.95	4.03	0.78
Michigan replication sample	17/33	1.0	0.67	9.16	0.98

Again, all Munich samples contain the same healthy elderly control subjects, while differing in the AD subjects. The specificity is thus fixed at the same value for all Munich samples. Sensitivity for all samples was calculated using the optimal threshold (1.05) from the Munich derivation sample to simulate genuine prospective application on a subject-by-subject basis, rather than optimizing the threshold for each sample anew.

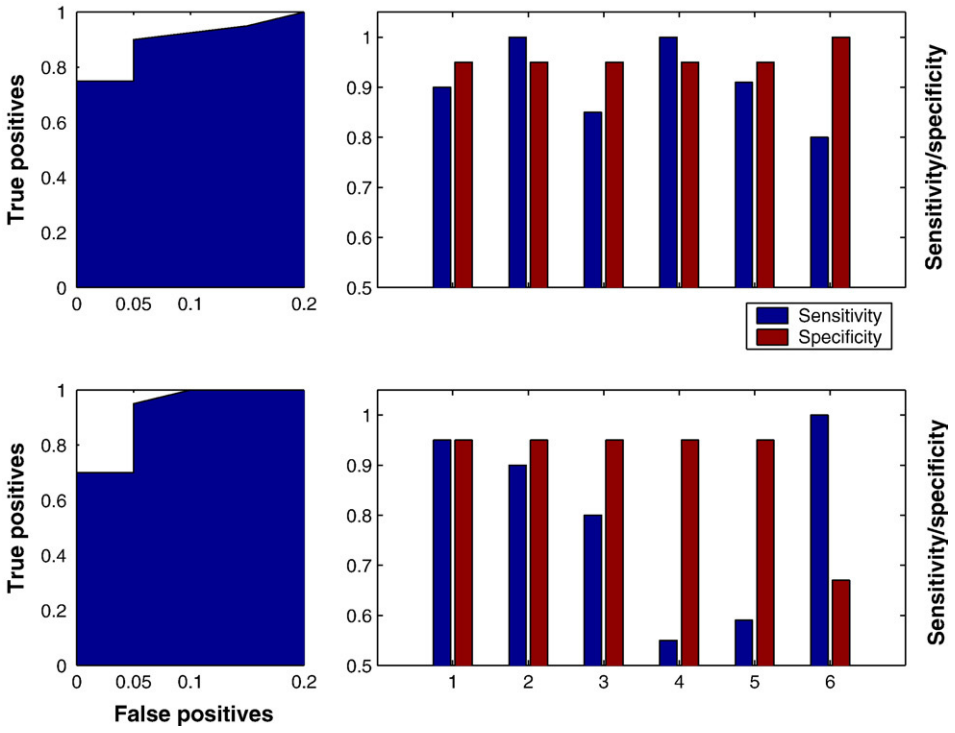

Fig. 1. Illustration of the ROC characteristics of multivariate and univariate group discrimination of AD subjects from normal controls. Left column: ROC-curve for the group discrimination in the derivation sample, using the multivariate (top) and univariate (bottom) biomarker: the ROC-curve looks very good and virtually identical for both markers. Right column: sensitivity/specificity values for the derivation and all replication samples, at the optimum criterion thresholds obtained for the markers in the derivation sample. For the univariate marker (bottom), which was the normalized signal value at location [50 −58 30], every value *below* the threshold 1.05 classified as AD; for the multivariate marker (top), which was the subject expression of the AD-related covariance pattern obtained from the derivation sample, every value *above* −10 was classified as AD. Marked on the *x*-axis are all 6 samples used. '1' = Munich derivation sample, '2' – '5' = Munich replication samples; '6' = Michigan replication sample. One can appreciate visually that the multivariate AD-marker has better average sensitivity than the univariate marker, i.e. generalizes with better robustness from the derivation to the replication samples.

controls therefore fall below the threshold 1.05, resulting in a large fraction (0.33) of false positives. Again, for the multivariate marker this is less of a problem, and despite nominally poorer discriminability as evidenced in *T*-contrast and ROC-area, the sensitivity and specificity at threshold = −10.0 is reasonable at 0.80 and 1.0, respectively.

Supplementary analyses to test the robustness of the multivariate AD-marker in derivation sample

As mentioned before, the multivariate marker is less subject to the corrupting influence of individual brain regions. To illuminate this further, we performed several additional manipulations to test the robustness of the multivariate AD-marker in comparison with the univariate marker.

1. We cut out the parietotemporal area showing the greatest univariate contrast (MNI: [50 −58 30]), by setting all signal values in the $(10\ mm)^3$ cube centered on this coordinate to zero in the derivation sample. We also cut out the $(10\ mm)^3$ neighborhoods of the positive and negative maximum *Z*-values in the AD-related covariance pattern, (MNI: [−28 −50 −32] and [50 −62 26], respectively). (To avoid confusion, we mention here that the location of maximum positive loading in the covariance pattern, MNI: [−28 −50 −32], was not in gray matter tissue and was subsequently not listed in Table 3. For the purpose of supplementary analysis, this made no difference however.) Projecting the covariance pattern into these modified data in the derivation sample, the ROC characteristics of the group discrimination in pattern expression were unchanged. The subject expression values obtained with the modified pattern are very similar to the ones obtained with the original pattern: the average relative difference between both sets of expression values is 3%, the correlation between the two sets is perfect to the fourth decimal place, i.e. discrepancies are less than 0.0001. This demonstrates how multivariate analysis takes into account the interregional correlation structure in the data, and is thus not critically dependent on the inclusion of any particular brain region and can withstand dropping out even the most salient areas in the univariate and multivariate AD-markers.
2. From the Munich-AD subject pool, we chose the 20 least impaired subjects to simulate a prospective application of AD-markers to patients with very early AD. (This sample still shows a significant age difference between AD subjects and controls, which is the reason we did no use it to derive the AD-pattern itself.) The 20 subjects' MMSE statuses were quite high: 18 of the subjects had MMSE = 26, 1 had MMSE = 27 and the remaining 1 had MMSE = 28. Employing our univariate AD-marker, we still find an impressive ROC area of 0.87 for the group discrimination between these high-MMSE individuals and the control subjects from the derivation sample, and a *T*-value of 4.88; however, sensitivity of the AD-diagnosis at the earlier chosen

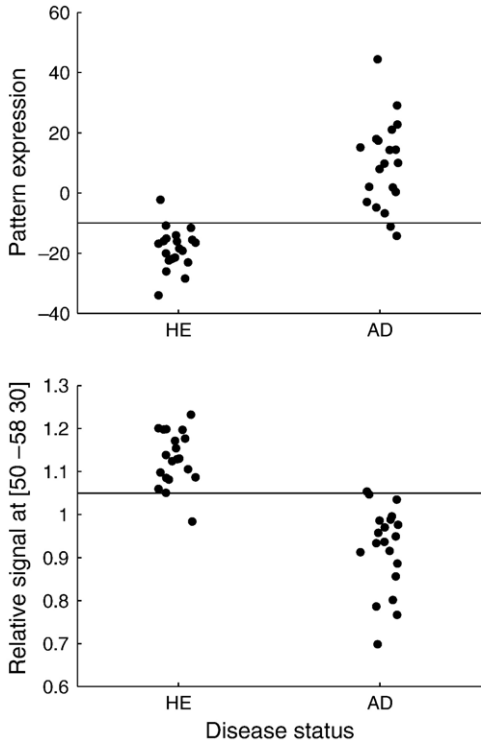

Fig. 2. Demonstration of group discrimination performance between AD patients ('AD') and healthy elderly controls ('HE'), for both multivariate and univariate methods in the Munich derivation sample. Top: multivariate analysis, bottom: univariate analysis, taking normalized signal values at the voxel of maximum *T*-contrast, MNI: [50 −58 30]. The horizontal lines in both plots mark the decision thresholds adopted for the AD diagnosis. The diagnostic algorithm tried to optimize diagnostic sensitivity (=fraction of AD patients correctly classified as AD) while incurring a maximal false-positive rate of 0.05 (=1 in twenty healthy elderly controls). The threshold criterion for AD diagnosis was a pattern expression larger than −10 for the multivariate marker (top) and a relative FDG signal of less than 1.05 at location [50 −58 30] for the univariate marker (bottom). These threshold criteria were applied in the same manner to all replication samples, to simulate a realistic context of clinical application patient-by-patient.

threshold (=1) was markedly decreased at 0.5. Subject expression of our AD-related multivariate pattern, on the other hand, fared much better: the ROC-area was 0.95, and sensitivity at the earlier chosen threshold (=−10) was 0.85.

3. So far, we have only investigated the diagnostic performance of both markers in the replication samples, neglecting the issue of how the *topographic composition* of the markers themselves changes when different derivation samples are used. Both *Z*-maps obtained from our semi-parametric bootstrap procedure as well as the *T*-maps for the univariate comparison are obviously susceptible to sampling error and change when *different* AD subjects are selected for the derivation of the disease marker. We have documented this dependence in Tables 7 and 8. To construct these tables, we substituted the derivation sample of Munich AD subjects with each of the 4 Munich replication samples, and conducted 4 additional analyses to derive both univariate and multivariate markers again. For brevity, we refrained from optimizing decision thresholds or recording ROC-characteristics. The following analysis was restricted to the brain regions that were identified as significant in both *T*- and *Z*-maps. Table 7 shows the global maximum of *T*- and *Z*-maps for all 5 analyses performed.

To provide a more global measure we also recorded how many of the voxels that surpassed the significance threshold in the analysis using the derivation sample also came up as significant in these additional analyses (see Table 8). The results were reported as a percentage. Further, we also reported the number of new voxels that 'popped in' in the additional analysis, as a fraction of all voxels that made the significant threshold in the additional analyses. For an analysis to produce similar results to the derivation sample, the first fraction should thus be close to 100%, while the second fraction should be as small as possible.

Although the results of Tables 6 and 7 are purely descriptive, they convey a sense of how heavily the sample selection influences the final regional composition of the AD-markers. The global maxima shift around and occasional switch laterality, while presumably preserving the gross anatomical structure of their location. However, for replication samples 2 and 4, the global maxima shift dramatically, particular for the univariate *T*-maximum. From Table 7, one can appreciate that the number of super threshold voxels shows similar discordances, affecting the multivariate disease-marker in particular. The major take-home message from these additional derivations is that caution is warranted when localizing the neural correlates of AD to particular brain areas with millimeter precision. Fluctuations due to sampling limitations have a big influence (and might be confounded further by pre-processing choices of spatial normalization and co-registration as well).

However, we like to reiterate that, despite its variability in the regional composition, our multivariate marker still provides superior diagnostic classification beyond the derivation sample. The signal-averaging computation that arrives at the pattern expression of a covariance pattern is more robust than the individual contributing voxel weights, which fluctuate.

Too provide the most accurate univariate and multivariate neural correlates of AD, we formed conjunction masks that identified voxels that were statistically significant in *all* 5 *T*- and *Z*-maps. We then averaged *T*- and *Z*-values at these voxel locations. Reassuringly, the mean *T*- and *Z*- maps have very similar global and local maxima: the area of biggest relative AD-related increase in FDG signal was

Table 7
Derivation of univariate and multivariate biomarkers, using all Munich AD samples in separate analyses

HE–AD comparison	Univariate maximum *T*, MNI coordinates	Multivariate maximum *Z*, MNI coordinates
Derivation sample	8.02, MNI=[50 −58 30]	14.10, MNI=[50 −62 26]
Replication sample 1	10.09, MNI=[54 −62 34]	11.44, MNI=[52 −58 42]
Replication sample 2	9.41, MNI=[10 −42 40]	5.92, MNI=[−46 −74 36]
Replication sample 3	11.70, MNI=[−50 −52 48]	13.43, MNI=[−56 −58 38]
Replication sample 4	6.45, MNI=[2 −30 40]	8.11, MNI=[−46 −46 44]

The left columns show which sample was used. Maximum negative *T* values for univariate comparison (middle column) as well as maximum negative *Z* values for the multivariate comparison (right column), indicating relatively *decreased* blood flow in the AD patients, are shown with their MNI coordinates.

Table 8
Columns 1 and 3 show the fraction of voxels in the derivation sample-derived AD-marker that also were detected in the additional analyses

HE–AD comparison	Fraction of derivation sample voxels that came up again ($T<-3.87$)	Fraction of new additional voxels ($T<-3.87$)	Fraction of derivation sample voxels that came up again ($Z<-3.36$)	Fraction of new additional voxels ($Z<-3.36$)
Derivation sample	100%	N/A	100%	N/A
Replication sample 1	75%	31%	65%	21%
Replication sample 2	36%	11%	23%	19%
Replication sample 3	83%	26%	76%	20%
Replication sample 4	60%	21%	24%	41%

Columns 2 and 4 show the fraction of voxels detected in the additional analyses that were *not* part of the original AD-markers.

found in the right anterior dentate of the cerebellum with only a $(2^2+2^2)^{1/2}=2.8$ mm discrepancy: MNI (*T*-maximum)=[18 −56 −26]; MNI (*Z*-maximum)=[20 −54 −26]. For the areas of relative AD-related decrease, this was similar with absolute maxima in the left inferior parietal lobule (BA 40); MNI (*T*-maximum)=[−50 −54 46]; MNI (*Z*-maximum)=[−50 −58 42]. The second-highest local maxima were found in the posterior cingulate gyrus (BA 31) (MNI(*T*)=[−6 −44 36]; MNI(*Z*)=[−2 −38 40]). For the remaining 3rd region, we have a slight discrepancy with the univariate marker shifting its activation ventrally to toward the parietotemporal junction: MNI(*T*)=[54 −60 36] (angular gyrus, BA 39); MNI(*Z*)=[46 −64 48] (inferior parietal lobule, BA 40). (For all these locations, nearest gray matter locations were reported.)

We can summarize both univariate and multivariate markers to have significant contributions in cerebellar, parietal, temporal and posterior cingulate regions. However, we stress again that for the multivariate marker and its expression's success as a diagnostic marker, the correlative relationships between *all* included voxels are diagnostically important, not just the super threshold ones.

Discussion

We examined the efficacy of multivariate and univariate analytic methods for the diagnosis of early Alzheimer's disease. To provide a more realistic context, we divided the data into a derivation sample and several replication samples. This, however, cannot compensate for any biases in the non-random recruitment that might prevent generalization to the population at large. The derivation sample was used to identify (a) an AD-related covariance pattern with our multivariate analysis, and (b) a brain area in which our univariate analysis shows AD subjects to have the biggest mean deficit in their relative regional cerebral glucose level, compared to the healthy elderly controls. Findings (a) and (b) were then checked prospectively in the derivation sample by in (a) forward-application of the AD-related covariance pattern to the scans of the derivation sample, and in (b) inspection of the relative regional cerebral glucose level at the MNI coordinate of the biggest group-contrast in the derivation sample. For the prospective application, the threshold adopted for the best ROC characteristics in the derivation sample were retained and *not* determined for each replication sample anew. Motivation for this stringent criterion was clinicopractical realism: for a useful neuroimaging marker for AD, application to any new patient with an accurate diagnosis and prediction of discourse is the ultimate goal, *without* any reference to any other patients. This means that not only a suitable marker has to be supplied, providing a dimension along which AD patients differ maximally from healthy elderly controls, but also a threshold criterion for this marker which enables the diagnosis of AD on a subject-by-subject basis.

A survey of the performance of both multivariate and univariate neuroimaging marker showed that they are both quite successful in the sense of yielding impressive ROC-areas, not only in the derivation sample, but in the replication samples as well. The sensitivity of the multivariate marker, however, is noticeably better with a range [0.85,1], whereas the univariate marker occasionally suffers from false negatives with a range [0.55,1].

Limitations of the recorded clinical information in our subject samples prevent us from exploring the differences between our multivariate and univariate AD-markers much further. We can only add some considerations that demonstrate that the multivariate AD-marker might be simultaneously more sensitive as well as more robust than the univariate marker. When recalling that multivariate analysis uses the entire spatial covariance structure of the data, this is not surprising. We performed several supplementary analyses to illustrate this fact: (1) we cut out the parietotemporal area showing the greatest univariate contrast (MNI: [50 −58 30]), by setting all FDG values in the $(10\ \text{mm})^3$ cube centered on this coordinate to zero in the derivation sample; we similarly cut out the most salient areas in the AD-related covariance pattern (MNI: [−28 −50 −32] and [50 −62 26], respectively). When projecting our AD-related covariance pattern into these modified data, the ROC characteristics of the group discrimination in pattern expression were unchanged. (2) From the Munich-AD subject pool, we chose the 20 least impaired subjects to simulate a prospective application of AD-markers to patients with very early AD. Our univariate marker still found an impressive ROC area of 0.87 for the group discrimination between these high-MMSE individuals and the control subjects from the derivation sample, and a *T*-value of 4.88; however, sensitivity of the AD-diagnosis at the earlier chosen threshold (=1.05) was markedly decreased at 0.5. Subject expression of our AD-related multivariate pattern, on the other hand, fared much better: the ROC-area was 0.95, and sensitivity at the earlier chosen threshold (=−10) was 0.85.

These results support the notion that multivariate analysis might be more sensitive than univariate analysis for the early diagnosis of AD early, possibly even before the clinical manifestation of typical symptoms, by utilizing the spatial covariance structure in the data from whole brain recordings.[2] The multivariate techniques do not necessarily rely on underlying "networks" of pathology. If the origins of AD-related glucose level changes are focal, and spread with disease severity, multivariate analysis will probably be more sensitive than univariate analysis as long as atrophy does not cause areas to drop out entirely without any variance contribution, forcing completely new spatial correlative relationships in the data. For the early stage of the disease, when accurate diagnosis is most important, this is unlikely, and disease-related changes in glucose level can be captured along a single dimension (changes in "degree"), before

[2] Normalization to different reference regions, like the Pons, might achieve better univariate group discrimination (observation by N. Foster). We will explore this more thoroughly in a future report.

this dimension breaks down during later disease stages (changes in "kind"). In particular, this implies in the early disease stage that longitudinally increasing disease-severity *within* a subject is interchangeable with increasing disease severity *across* subjects in a cross-sectional sense. The same AD-related covariance pattern should capture these two changes, and should yield an increasing subject expression, regardless of longitudinally or cross-sectionally increasing disease severity. It is conceivable that a sufficient amount of this diagnostic information could be captured by a simplified univariate–multivariate approach of first localizing several key regions in a univariate analysis, with a subsequent linear combination of these areas' signal values to optimize ROC-characteristics for the AD diagnosis. The benefits of such an approach will be investigated more thoroughly in a future report.

We were able to identify disease-related dimension in form of a covariance pattern; however, this does not imply that our AD-marker captures the *only* disease-related dimension: a recent FDG-PET investigation of healthy again, vascular dementia and AD (Kerrouche et al., 2006) pinpointed two dimensions: one that differentiated demented from healthy people in general, and another one that perfectly distinguished all three populations form each other. Further fine-tuning of our marker with additional neuropsychological test–performance data (which itself is multidimensional) to achieve the best possible ROC-characteristics is also possible and represents one of our future research goals.

Our supplementary analysis (3) gave pause for some concern for the goal of precisely mapping the neural correlates of AD-progression. The regional composition of both univariate as well as multivariate disease markers varied substantially as a function of the AD sample selected. Thus trying to glean some insight about the etiology of AD from the brain regions identified might require substantially greater subject numbers than used in our derivations (~20 patients/20 controls). Both *T*- and *Z*-maps were highly variable as a function of the AD sample used in the derivation. This variability, however, did not hamper the diagnostic performance of the multivariate marker beyond the derivation sample, since it is the subject expression of the covariance pattern comprising *all* voxels that is used for the diagnostic classification. As demonstrated in supplementary analysis (1), pattern expression is stable even when the brain regions of largest significance in the *Z*-map are removed. Further, we can surmise that even a substantial amount of noise on the individual voxel weights will leave the pattern expression relatively intact, while changing the number and location of the super threshold voxels in the bootstrap *Z*-map.

Despite this caveat about regional fluctuations for small sample sizes, we can say that both multivariate and univariate disease-markers converged on similar regions when we combined the results of our main derivation sample with all 4 supplementary analyses. Regions that showed a relative increase in FDG were found in the cerebellum, hinting that this structure is relatively preserved in early Alzheimer's disease, while other areas are decreasing. Regions that showed a relative decrease in FDG were found in bilateral parietal and temporal regions (BA 39, 40) and the posterior cingulate gyrus (BA 31).

These regions have been described before and are hardly surprising. For instance, a recent study $H_2^{15}O$–PET study of Mild Cognitive Impairment (Huang et al., 2007) that performed cross-sectional covariance analyses similar to the one in the current article, also identified cerebellum as an areas of relative increase, and bilateral parietal and posterior cingulate areas as relatively decreasing in signal intensity when comparing both severely progressing and stable MCI subpopulations to healthy controls. As in our study, mediotemporal regions were not identified in this cross-sectional comparison, they only contributed in the longitudinal analysis of the progressing MCI population (Huang et al., 2007). The involvement of mediotemporal lobe areas in the metabolic correlates of early Alzheimer's has not been established unanimously: apart from Huang et al., several previous studies also failed to find any supporting evidence (Chetelat et al., 2003a; Ishii et al., 1998; Jagust et al., 2002; Minoshima et al., 1999). [One study (de Leon et al., 2001) found evidence in favor of hippocampal involvement.]

These results could be explained by the fact that neurofibrillary tangles in the ento- and perirhinal cortices cause a lesion of these areas' projections to the posterior cingulate with a subsequent metabolic deficit in the latter, while leaving the metabolism in the cell bodies medial temporal lobe intact. This possibility was verified in an animal model using baboons (Meguro et al., 1999) and could be one candidate mechanism to explain the disjuncture between structure and metabolism observed in early Alzheimer's disease.

In conclusion, we have shown in the current study that FDG–PET imaging combined with multivariate analysis shows some promise as systems-level biomarker for Alzheimer's disease, and might enable early detection of the disease before a clear clinical presentation. It is maybe important to point out what we are *not* claiming:

1. We are not claiming to have found a 'mechanism' biomarker. AD-related FDG–PET patterns are unlikely by themselves to provide more insight about the etiology of Alzheimer's disease. We are making no such claim and recognize fully that we are capturing downstream effects of the disease that are manifesting at the systems' level. This disclaimer, however, casts no aspersions on the efficacy of the covariance pattern as a diagnostic tool.
2. We are not claiming that the specificity of AD-related covariance pattern to Alzheimer's disease has been confirmed yet. In fact, this remains our own research program for the immediate future. The virtue of our approach lies in its reliance on completely standard imaging and analytic technology – without the need for expensive molecular markers or ligands. We showed that AD could be captured with good sensitivity and specificity relative to healthy subjects – taking the clinical judgment, which might be imperfect itself, as the gold standard. We further have to investigate the specificity with respect to other neurodegenerative diseases.

The last point illuminates the wider trade-off that is present for the development of biomarkers in general: one the one end of the continuum we can visualize reductionist strategies that target individual diseases with specially tailored molecular compounds. The development of these compounds might be costly and occasionally fraught with failure, but the successfully ensuing biomarkers are specific to the disease under consideration. Our approach is at the other end of the continuum: we are using widely available imaging and analytic technology, but have to verify our derived systems'-level biomarker is specific to Alzheimer's disease. This specificity is not achieved by a reductionist strategy of honing in on particular molecules, but by the multivariate strategy of simultaneously analyzing the brain-wide correlation structure of blood flow or metabolism, and finding a disease-specific distributed pattern – not only when healthy elderly controls are present, but when data from other neurodegenerative disease are admixed too (Kerrouche et al., 2006).

Following from there, we hope to test whether our covariance pattern can also serve as an outcome biomarker, i.e., can signal successful treatment through a reduction over time in people who respond well to any treatment.

Acknowledgments

We thank Andrew P. Lieberman for reviewing the neuropathology, and David E. Kuhl, Sid Gilman, Gus Buchtel, David Knesper, R. Scott Turner, and Kirk Frey for making images from their research available for this study.

References

Alexander, G.E., Mentis, M.J., Van Horn, J.D., Grady, C.L., Berman, K.F., Furey, M.L., Pietrini, P., Rapoport, S.I., Schapiro, M.B., Moeller, J.R., 1999. Individual differences in PET activation of object perception and attention systems predict face matching accuracy. Neuroreport 10, 1965–1971.

Alexander, G.E., Chen, K., Merkley, T.L., Reiman, E.M., Caselli, R.J., Aschenbrenner, M., Santerre-Lemmon, L., Lewis, D.J., Pietrini, P., Teipel, S.J., Hampel, H., Rapoport, S.I., Moeller, J.R., 2006. Regional network of magnetic resonance imaging gray matter volume in healthy aging. Neuroreport 17, 951–956.

Berg, L., Miller, J.P., Storandt, M., Duchek, J., Morris, J.C., Rubin, E.H., Burke, W.J., Coben, L.A., 1988. Mild senile dementia of the Alzheimer type: 2. Longitudinal assessment. Ann. Neurol. 23, 477–484.

Bozzali, M., Filippi, M., Magnani, G., Cercignani, M., Franceschi, M., Schiatti, E., Castiglioni, S., Mossini, R., Falautano, M., Scotti, G., Comi, G., Falini, A., 2006. The contribution of voxel-based morphometry in staging patients with mild cognitive impairment. Neurology 67, 453–460.

Braak, H., Braak, E., Bohl, J., Bratzke, H., 1998. Evolution of Alzheimer's disease related cortical lesions. J. Neural Transm. 54, 97–106 Suppl.

Brickman, A.M., Habeck, C., Zarahn, E., Flynn, J., Stern, Y., 2007. Structural MRI covariance patterns associated with normal aging and neuropsychological functioning. Neurobiol. Aging 28, 284–295.

Burdette, J.H., Minoshima, S., Vander Borght, T., Tran, D.D., Kuhl, D.E., 1996. Alzheimer disease: improved visual interpretation of PET images by using three-dimensional stereotaxic surface projections. Radiology 198, 837–843.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference. Springer Verlag, New York.

Carbon, M., Trost, M., Ghilardi, M.F., Eidelberg, D., 2004. Abnormal brain networks in primary torsion dystonia. Adv. Neurol. 94, 155–161.

Chetelat, G., Desgranges, B., de la Sayette, V., Viader, F., Berkouk, K., Landeau, B., Lalevee, C., Le Doze, F., Dupuy, B., Hannequin, D., Baron, J.C., Eustache, F., 2003a. Dissociating atrophy and hypometabolism impact on episodic memory in mild cognitive impairment. Brain 126, 1955–1967.

Chetelat, G., Desgranges, B., de la Sayette, V., Viader, F., Eustache, F., Baron, J.C., 2003b. Mild cognitive impairment: can FDG–PET predict who is to rapidly convert to Alzheimer's disease? Neurology 60, 1374–1377.

Chetelat, G., Landeau, B., Eustache, F., Mezenge, F., Viader, F., de la Sayette, V., Desgranges, B., Baron, J.C., 2005. Using voxel-based morphometry to map the structural changes associated with rapid conversion in MCI: a longitudinal MRI study. Neuroimage 27, 934–946.

Davatzikos, C., Fan, Y., Wu, X., Shen, D., Resnick, S.M., 2006. Detection of prodromal Alzheimer's disease via pattern classification of MRI. Neurobiol. Aging.

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/poitron-emission tomography (FDG/PET). Proc. Natl. Acad. Sci. U. S. A. 98, 10966–10971.

Delieu, J., Keady, J., 1996. The biology of Alzheimer's disease: 1. Br. J. Nurs. 5, 162–168.

Devanand, D.P., Habeck, C.G., Tabert, M.H., Scarmeas, N., Pelton, G.H., Moeller, J.R., Mensh, B.D., Tarabula, T., Van Heertum, R.L., Stern, Y., 2006. PET network abnormalities and cognitive decline in patients with mild cognitive impairment. Neuropsychopharmacology 31, 1327–1334.

Drzezga, A., Riemenschneider, M., Strassner, B., Grimmer, T., Peller, M., Knoll, A., Wagenpfeil, S., Minoshima, S., Schwaiger, M., Kurz, A., 2005. Cerebral glucose metabolism in patients with AD and different APOE genotypes. Neurology 64, 102–107.

Efron, B., Tibshirani, R.J., 1994. An Introduction to the Bootstrap. CRC Press, LLC, New York.

Eidelberg, D., Dhawan, V., Moeller, J.R., Sidtis, J.J., Ginos, J.Z., Strother, S.C., Cederbaum, J., Greene, P., Fahn, S., Powers, J.M., et al., 1991. The metabolic landscape of cortico-basal ganglionic degeneration: regional asymmetries studied with positron emission tomography. J. Neurol. Neurosurg. Psychiatry 54, 856–862.

Eidelberg, D., Moeller, J.R., Kazumata, K., Antonini, A., Sterio, D., Dhawan, V., Spetsieris, P., Alterman, R., Kelly, P.J., Dogali, M., Fazzini, E., Beric, A., 1997. Metabolic correlates of pallidal neuronal activity in Parkinson's disease. Brain 120 (Pt 8), 1315–1324.

Engler, H., Forsberg, A., Almkvist, O., Blomquist, G., Larsson, E., Savitcheva, I., Wall, A., Ringheim, A., Langstrom, B., Nordberg, A., 2006. Two-year follow-up of amyloid deposition in patients with Alzheimer's disease. Brain 129, 2856–2866.

Folstein, M.F., Folstein, S.E., McHugh, P.R., 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res. 12, 189–198.

Foster, N.L., Heidebrink, J.L., Clark, C.M., Jagust, W.J., Arnold, S.E., Barbas, N.R., DeCarli, C.S., Turner, R.S., Koeppe, R.A., Higdon, R., Minoshima, S., 2007. FDG–PET improves accuracy in distinguishing frontotemporal dementia and Alzheimer's disease. Brain 130, 2616–2635.

Greicius, M.D., Srivastava, G., Reiss, A.L., Menon, V., 2004. Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. Proc. Natl. Acad. Sci. U. S. A. 101, 4637–4642.

Herholz, K., Salmon, E., Perani, D., Baron, J.C., Holthoff, V., Frolich, L., Schonknecht, P., Ito, K., Mielke, R., Kalbe, E., Zundorf, G., Delbeuck, X., Pelati, O., Anchisi, D., Fazio, F., Kerrouche, N., Desgranges, B., Eustache, F., Beuthien-Baumann, B., Menzel, C., Schroder, J., Kato, T., Arahata, Y., Henze, M., Heiss, W.D., 2002. Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. Neuroimage 17, 302–316.

Higdon, R., Foster, N.L., Koeppe, R.A., DeCarli, C.S., Jagust, W.J., Clark, C.M., Barbas, N.R., Arnold, S.E., Turner, R.S., Heidebrink, J.L., Minoshima, S., 2004. A comparison of classification methods for differentiating fronto-temporal dementia from Alzheimer's disease using FDG–PET imaging. Stat. Med. 23, 315–326.

Huang, C., Eidelberg, D., Habeck, C., Moeller, J., Svensson, L., Tarabula, T., Julin, P., 2007. Imaging markers of mild cognitive impairment: multivariate analysis of CBF SPECT. Neurobiol. Aging 28, 1062–1069.

Hughes, A.J., Ben-Shlomo, Y., Daniel, S.E., Lees, A.J., 1992. What features improve the accuracy of clinical diagnosis in Parkinson's disease: a clinicopathologic study. Neurology 42, 1142–1146.

Hutchinson, M., Nakamura, T., Moeller, J.R., Antonini, A., Belakhlef, A., Dhawan, V., Eidelberg, D., 2000. The metabolic topography of essential blepharospasm: a focal dystonia with general implications. Neurology 55, 673–677.

Ishii, K., Sasaki, M., Yamaji, S., Sakamoto, S., Kitagaki, H., Mori, E., 1998. Paradoxical hippocampus perfusion in mild-to-moderate Alzheimer's disease. J. Nucl. Med. 39, 293–298.

Jagust, W.J., Eberling, J.L., Wu, C.C., Finkbeiner, A., Mungas, D., Valk, P.E., Haan, M.N., 2002. Brain function and cognition in a community sample of elderly Latinos. Neurology 59, 378–383.

Kasama, S., Tachibana, H., Kawabata, K., Yoshikawa, H., 2005. Cerebral blood flow in Parkinson's disease, dementia with Lewy bodies, and Alzheimer's disease according to three-dimensional stereotactic surface projection imaging. Dement. Geriatr. Cogn. Disord. 19, 266–275.

Kerrouche, N., Herholz, K., Mielke, R., Holthoff, V., Baron, J.C., 2006. ^{18}FDG PET in vascular dementia: differentiation from Alzheimer's disease using voxel-based multivariate analysis. J. Cereb. Blood Flow Metab. 26, 1213–1221.

McKeith, I.G., Galasko, D., Kosaka, K., Perry, E.K., Dickson, D.W., Hansen, L.A., Salmon, D.P., Lowe, J., Mirra, S.S., Byrne, E.J., Lennox, G., Quinn, N.P., Edwardson, J.A., Ince, P.G., Bergeron, C., Burns, A., Miller, B.L., Lovestone, S., Collerton, D., Jansen, E.N., Ballard, C., de Vos, R.A., Wilcock, G.K., Jellinger, K.A., Perry, R.H., 1996. Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): report of the consortium on DLB international workshop. Neurology 47, 1113–1124.

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

Meguro, K., Blaizot, X., Kondoh, Y., Le Mestric, C., Baron, J.C., Chavoix, C., 1999. Neocortical and hippocampal glucose hypometabolism following neurotoxic lesions of the entorhinal and perirhinal cortices in the non-human primate as shown by PET. Implications for Alzheimer's disease. Brain 122 (Pt 8), 1519–1531.

Minoshima, S., Cross, D.J., Foster, N.L., Henry, T.R., Kuhl, D.E., 1999. Discordance between traditional pathologic and energy metabolic changes in very early Alzheimer's disease. Pathophysiological implications. Ann. N. Y. Acad. Sci. 893, 350–352.

Moeller, J.R., Strother, S.C., Sidtis, J.J., Rottenberg, D.A., 1987. Scaled subprofile model: a statistical approach to the analysis of functional patterns in positron emission tomographic data. J. Cereb. Blood Flow Metab. 7, 649–658.

Moeller, J.R., Ishikawa, T., Dhawan, V., Spetsieris, P., Mandel, F., Alexander, G.E., Grady, C., Pietrini, P., Eidelberg, D., 1996. The metabolic topography of normal aging. J. Cereb. Blood Flow Metab. 16, 385–398.

Nakamura, T., Ghilardi, M.F., Mentis, M., Dhawan, V., Fukuda, M., Hacking, A., Moeller, J.R., Ghez, C., Eidelberg, D., 2001. Functional networks in motor sequence learning: abnormal topographies in Parkinson's disease. Hum. Brain Mapp. 12, 42–60.

Nakano, S., Asada, T., Yamashita, F., Kitamura, N., Matsuda, H., Hirai, S., Yamada, T., 2006. Relationship between antisocial behavior and regional cerebral blood flow in frontotemporal dementia. Neuroimage 32, 301–306.

Neary, D., Snowden, J.S., Gustafson, L., Passant, U., Stuss, D., Black, S., Freedman, M., Kertesz, A., Robert, P.H., Albert, M., Boone, K., Miller, B.L., Cummings, J., Benson, D.F., 1998. Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. Neurology 51, 1546–1554.

Nichols, L., Pike, V.W., Cai, L., Innis, R.B., 2006. Imaging and in vivo quantitation of beta-amyloid: an exemplary biomarker for Alzheimer's disease? Biol. Psychiatry 59, 940–947.

Pearson, R.C., 1996. Cortical connections and the pathology of Alzheimer's disease. Neurodegeneration 5, 429–434.

Rubin, E.H., Kinscherf, D.A., Morris, J.C., 1993. Psychopathology in younger versus older persons with very mild and mild dementia of the Alzheimer type. Am. J. Psychiatry 150, 639–642.

Scarmeas, N., Habeck, C.G., Zarahn, E., Anderson, K.E., Park, A., Hilton, J., Pelton, G.H., Tabert, M.H., Honig, L.S., Moeller, J.R., Devanand, D.P., Stern, Y., 2004. Covariance PET patterns in early Alzheimer's disease and subjects with cognitive impairment but no dementia: utility in group discrimination and correlations with functional performance. Neuroimage 23, 35–45.

Teipel, S.J., Flatz, W.H., Heinsen, H., Bokde, A.L., Schoenberg, S.O., Stockel, S., Dietrich, O., Reiser, M.F., Moller, H.J., Hampel, H., 2005. Measurement of basal forebrain atrophy in Alzheimer's disease using MRI. Brain 128, 2626–2644.

Thomann, P.A., Wustenberg, T., Pantel, J., Essig, M., Schroder, J., 2006. Structural changes of the corpus callosum in mild cognitive impairment and Alzheimer's disease. Dement. Geriatr. Cogn. Disord. 21, 215–220.

Trollor, J.N., Sachdev, P.S., Haindl, W., Brodaty, H., Wen, W., Walker, B.M., 2005. Regional cerebral blood flow deficits in mild Alzheimer's disease using high resolution single photon emission computerized tomography. Psychiatry Clin. Neurosci. 59, 280–290.

Trost, M., Carbon, M., Edwards, C., Ma, Y., Raymond, D., Mentis, M.J., Moeller, J.R., Bressman, S.B., Eidelberg, D., 2002. Primary dystonia: is abnormal functional brain architecture linked to genotype? Ann. Neurol. 52, 853–856.

Trost, M., Su, S., Su, P., Yen, R.F., Tseng, H.M., Barnes, A., Ma, Y., Eidelberg, D., 2006. Network modulation by the subthalamic nucleus in the treatment of Parkinson's disease. Neuroimage 31, 301–307.

Warkentin, S., Ohlsson, M., Wollmer, P., Edenbrandt, L., Minthon, L., 2004. Regional cerebral blood flow in Alzheimer's disease: classification and analysis of heterogeneity. Dement. Geriatr. Cogn. Disord. 17, 207–214.

Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L.O., Nordberg, A., Backman, L., Albert, M., Almkvist, O., Arai, H., Basun, H., Blennow, K., de Leon, M., DeCarli, C., Erkinjuntti, T., Giacobini, E., Graff, C., Hardy, J., Jack, C., Jorm, A., Ritchie, K., van Duijn, C., Visser, P., Petersen, R.C., 2004. Mild cognitive impairment—beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. J. Intern. Med. 256, 240–246.