Comparison of ^{18}F-FDG PET and Optimized Voxel-Based Morphometry for Detection of Alzheimer's Disease: Aging Effect on Diagnostic Performance

Ichiro Matsunari[1], Miharu Samuraki[2], Wei-Ping Chen[1], Daiseke Yanase[2], Nozomi Takeda[1], Kenjiro Ono[2], Mitsuhiro Yoshita[2], Hiroshi Matsuda[3], Masahito Yamada[2], and Seigo Kinuya[4]

[1]The Medical and Pharmacological Research Center Foundation, Hakui-city, Ishikawa, Japan; [2]Department of Neurology and Neurobiology of Aging, Kanazawa University Graduate School of Medical Science, Kanazawa, Japan; [3]Department of Nuclear Medicine, Saitama Medical School Hospital, Saitama, Japan; and [4]Department of Biotracer Medicine, Kanazawa University Graduate School of Medical Science, Kanazawa, Japan

The aim of this study was to compare optimized voxel-based morphometry (VBM) and ^{18}F-FDG PET for discrimination between patients with Alzheimer's disease (AD) and healthy subjects in relation to age. **Methods:** The study population consisted of 2 groups; the first group (27 AD patients and 40 control subjects) was used to determine the locations of significant abnormalities for both PET and VBM using statistical parametric mapping, and the second group (34 AD patients and 50 control subjects) was used to compare the diagnostic performance of PET and VBM. In the second group, a *z*-score map for PET or VBM of each subject was obtained by comparison with the mean and SD of PET or gray-matter MR images of the control subjects. Receiver-operating-characteristic (ROC) curve analysis was then performed to compare the diagnostic performance between PET and VBM. Furthermore, group 2 was divided into the early- and late-onset subgroups, and ROC analysis was performed for each subgroup. **Results:** In the first group, VBM revealed a significant decrease in gray-matter concentration in the hippocampus complex in AD, whereas PET showed a significant reduction in ^{18}F-FDG uptake in the posterior cingulate and parietotemporal lobe. The diagnostic performance of PET (0.988 ± 0.008; mean ± SE), as measured by the area under the ROC curve, was higher than that of VBM with (0.782 ± 0.059) or without (0.832 ± 0.049) modulation. PET yielded a sensitivity, specificity, and overall accuracy of 100%, 92%, and 95%, respectively, whereas for VBM the sensitivity, specificity, and accuracy were 74%, 92%, and 85%. Modulation for the VBM did not improve these values (56%, 94%, and 79%, respectively). When the early- and late-onset subjects were analyzed separately, the superiority of ^{18}F-FDG PET was significant only in the early-onset subgroup. **Conclusion:** The present study indicates that the detection of metabolic alteration by ^{18}F-FDG PET yields a better diagnostic performance for the discrimination between AD patients and healthy control subjects than does the morphologic approach by VBM, particularly in the early-onset subjects.

Key Words: ^{18}F-FDG PET; voxel-based morphometry; Alzheimer's disease

J Nucl Med 2007; 48:1961–1970
DOI: 10.2967/jnumed.107.042820

Because Alzheimer's disease (AD) is the leading form of age-related dementia in many countries, efforts have been made to develop techniques to detect abnormalities associated with the presence and severity of disease (*1*). Of these, ^{18}F-FDG PET combined with automated image analysis—such as statistical parametric mapping (SPM) (*2*) or NEUROSTAT (*3*)—is an established diagnostic imaging technique, applied to distinguish between neurodegenerative disorders according to changes in regional brain glucose use. Using these methods, regional metabolic alteration typically in the precuneus, posterior cingulate gyri, and lateral parietal association cortex is reported to be an early marker of AD (*4–6*).

Apart from the metabolic alterations described, structural MRI studies have demonstrated that atrophy of the medial temporal lobe, including the hippocampal compex, occurs at an early stage of AD (*7–9*). Many such studies have used a manual region-of-interest (ROI) approach to measure gray-matter volumes. Recently, a statistical voxel-based approach using 3-dimensional (3D) T1-weighted MRI, referred as voxel-based morphometry (VBM) (*10*), has emerged as one possible approach for automatic detection of gray-matter volume loss in various neuropsychologic disorders on a voxel-by-voxel basis after anatomic standardization. Recent VBM studies have shown that gray-matter volume loss in the hippocampal complex measured by VBM serves as a reliable

Received Apr. 18, 2007; revision accepted Sep. 10, 2007.
For correspondence or reprints contact: Ichiro Matsunari, MD, PhD, The Medical and Pharmacological Research Center Foundation, Wo32, Inoyama-town, Hakui-city, Ishikawa 925-0613, Japan.
E-mail: matsunari@mprcf.or.jp

marker of AD (*11*) and, thus, has a diagnostic value for detection of AD (*12*). Although VBM, in general, does not allow assessment of structural changes because of some known limitations—such as those encountered in imperfectly registered images (*13*)—it is a popular and practical implementation of morphology that could be compared with PET. To date, however, few data are available that directly compare the diagnostic utility of PET and VBM. According to a recent study by Kawachi et al. (*14*), PET seems to yield a better diagnostic performance for detection of AD than standard VBM does. However, more sophisticated VBM approaches such as optimized VBM (*15*) have been developed for better anatomic standardization and segmentation. Whether these improved VBM techniques would provide diagnostic accuracy comparable with that obtained by PET has, to our knowledge, not yet been investigated.

From a demographic viewpoint, it is well known that the prevalence of AD significantly increases as a function of age (*1*). Although the clinical value of a diagnostic test depends on disease prevalence, age-related changes in comparative diagnostic performance between PET and VBM are unknown.

Thus, the aim of this study was to determine whether the diagnostic performance of optimized VBM for discrimination between AD patients and healthy subjects is comparable with that obtained by ^{18}F-FDG PET in relation to age.

MATERIALS AND METHODS

Study Population

We studied 61 patients with a clinical diagnosis of probable AD (28 men, 33 women; mean age ± SD, 68.5 ± 9.1 y) and 90 age-matched healthy volunteers (44 men, 46 women; mean age, 67.2 ± 8.4 y; not significant [NS] vs. probable AD). Of these, a subset of probable AD patients ($n = 27$) and healthy volunteers ($n = 40$) was randomly selected to determine the location of significant abnormalities for PET or VBM on the basis of voxel-based analysis (group 1). The remaining subjects (34 probable AD patients and 50 healthy volunteers) were used to compare the diagnostic performance of PET and VBM (group 2). This randomization procedure was performed using statistical software (JMP 6.0; SAS Institute Inc.) to keep the AD-to-healthy ratio equivalent between the 2 groups. The clinical characteristics of groups 1 and 2 are summarized in Table 1.

All subjects underwent physical examinations, neuropsychologic examinations, and brain MRI and replied to the questionnaire for clinical information. Cognitive ability was assessed using the Mini-Mental State Examination (MMSE). The brain MRI scans (3D T1-weighted, T2-weighted, MR angiography) were performed on all subjects on the same day as the PET scan, and the 3D T1 images were used for VBM study. The criteria to define "healthy" were no history of neuropsychologic disorders, normal findings on physical and neurologic examinations (MMSE score ≥ 28), and normal findings on brain MRI (no asymptomatic cerebral infarction, hemorrhage, tumor, or brain vessel abnormalities on T1-weighted or T2-weighted MRI or MR angiography). The AD patients were recruited at a memory disorders clinic at Kanazawa University Hospital (Kanazawa, Japan). The diagnosis of probable AD was made by neurologists who were unaware of PET or VBM results according to the criteria of the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (*16*). The mean MMSE score was 22.1 ± 3.5.

The study protocol was approved by the institutional human study committee of The Medical and Pharmacological Research Center Foundation and written informed consent was obtained from all subjects.

PET and MRI

All subjects underwent both ^{18}F-FDG PET and MRI on the same day. The detailed imaging procedure is described elsewhere (*17*). In brief, the subjects fasted for >4 h before the injection of ^{18}F-FDG. Before PET was performed, an intravenous line was established in all subjects with their eyes closed and covered with a mask. Twenty minutes later, 370 MBq of ^{18}F-FDG were injected intravenously and subjects were kept in resting condition (with their eyes closed and covered with a mask) for an additional 40 min; PET was then initiated for 10 min using a full-ring PET scanner (Advance; GE Healthcare). A transmission scan was obtained for 10 min using rotating ^{68}Ge pin sources. Images were reconstructed using an ordered-subset expectation maximization algorithm. The image data matrix was 128 × 128 with a pixel size of 2.11 mm and a slice thickness of 4.25 mm. Both attenuation and scatter corrections were used during image reconstruction.

MRI studies were performed using a 1.5-T system (Signa Horizon; GE Healthcare). A 3D volumetric acquisition of a T1 gradient-echo sequence produced a gapless series of thin transaxial sections using a magnetization-preparation rapid-acquisition gradient-echo sequence (echo time [ms]/repetition time [ms], 2.0/9.2; flip angle, 20°; acquisition matrix, 256 × 192; number

TABLE 1
Clinical Characteristics of Subjects

Characteristic	Group 1		Group 2	
	AD ($n = 27$)	Healthy ($n = 40$)	AD ($n = 34$)	Healthy ($n = 50$)
Sex (M/F)	12/15	20/20	16/18	24/26
Age* (y)	68.6 ± 6.8	66.5 ± 7.6	68.4 ± 10.7	67.7 ± 9.1
Onset of AD* (y)	66.3 ± 6.9	—	65.3 ± 10.9	—
Disease duration* (y)	2.4 ± 1.9	—	3.1 ± 4.5	—
MMSE score*	22.0 ± 3.3	29.2 ± 0.8	22.1 ± 3.8	29.3 ± 0.8

*Data are expressed as mean ± SD.
MMSE = Mini-Mental State Examination.

of slices, 124; pixel size, 0.78 × 1.04; slice thickness, 1.4 mm). T2-weighted MRI and MR angiography were performed according to our routine protocol for screening.

Data Processing and Analysis

Anatomic normalization and statistical processing for PET and MRI were performed using SPM2 (Wellcome Department of Imaging Neuroscience, London, U.K.) in conjunction with MATLAB 7.1 (Mathworks Inc.).

PET

All PET images were spatially normalized in SPM2 to a standard stereotactic space based on the Talairach and Tournoux atlas (*18*), using a 12-parameter linear affine normalization and a further nonlinear iteration algorithm. This was performed using a template for ^{18}F-FDG that had been constructed by averaging ^{18}F-FDG images of group 1 subjects whose deformation was estimated from coregistered MRI. The PET images were then smoothed using a 12-mm full width at half maximum (FWHM) isotropic gaussian kernel.

Theory of Optimized VBM and Modulation

Originally, VBM was developed to identify local differences in the composition of brain tissue (gray or white matter) between 2 or more experimental groups, which is achieved by spatial normalization, segmentation, smoothing, and subsequent statistical analysis (*19*). Spatial normalization fits all structural images into the same stereotactic space. During spatial normalization, SPM requires templates that have contrasts and shapes similar to those of the images to be matched with them. Spatially normalized MR images are then segmented into gray matter, white matter, and cerebrospinal fluid (CSF) using a clustering algorithm. The segmentation procedure involves calculating a Bayesian probability for each voxel to determine whether it belongs to a tissue class on the basis of a priori MRI information; ideally, these tissue probability maps should represent the prior probabilities for the population under investigation. After smoothing, gray-matter intensities should reflect gray-matter concentrations by the partial-volume effect. However, there are cases when the results are influenced by the surrounding tissue, such as the ventricles, as a result of spatial normalization. Optimized VBM has been developed to overcome shortcomings of standard VBM. Figure 1 illustrates the entire preprocessing steps of optimized VBM for gray-matter analysis that was used in the present study. The basic idea of optimized VBM is to improve spatial normalization by the use of gray-matter images and a gray-matter template rather than whole-brain T1 images. This is achieved by segmentation of the original T1 MR images, spatial normalization of segmented images, applying normalization parameters to the original T1 MR images, and segmentation of spatially normalized T1 MR images.

During spatial normalization, brain regions are expanded or contracted. Modulation involves scaling by the amount of expansion or contraction, so that the total amount of gray-matter intensity in the modulated gray matter remains the same as it would be in the original images. Therefore, gray-matter intensity on modulated images should represent tissue volume rather than tissue concentration.

Optimized VBM Procedure

The optimized VBM method with or without modulation was performed for all MRI data as illustrated in Figure 1. The study-specific T1 template and prior gray or white matter and CSF images were generated from group 1 subjects. Each optimally normalized, segmented, and modulated or unmodulated gray-matter image was smoothed with a 12-mm FWHM kernel.

Determination of Regional Metabolic Reduction and Gray-Matter Loss in AD

To identify clusters of significantly reduced ^{18}F-FDG uptake or gray-matter concentration in patients, anatomically normalized PET and gray-matter MRI datasets of the first group (27 AD patients and 40 control subjects) were compared on a voxel-by-voxel basis using the "compare-population 1 scan per subject" routine in SPM2. The "proportional scaling" routine was used to control for individual variation in global mean activity or intensities except for modulated optimized VBM. The results were corrected for multiple comparisons by controlling the "familywise error rate" at $P < 0.05$

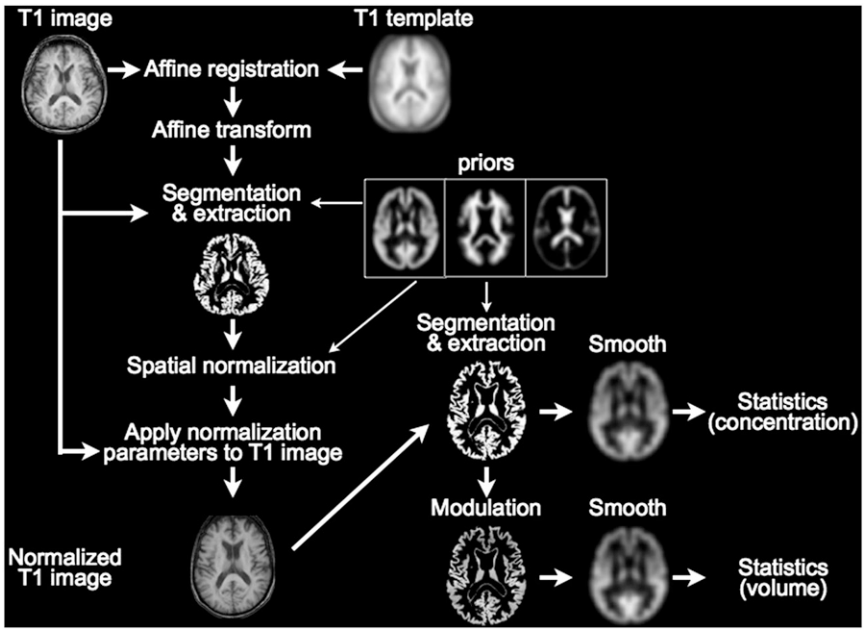

FIGURE 1. Flow chart of optimized VBM. After affine registration and transformation, original T1 images are segmented. Segmented gray-matter images are further spatially normalized using a study-specific gray-matter template (left image of priors). Normalization parameters are then applied to the original T1 images; this is followed by the second segmentation using priors representing tissue probability maps that had been generated from group 1 subjects. Finally, segmented gray-matter images with or without modulation are smoothed for subsequent statistical analysis.

(*20*). The extent threshold was set to 50 voxels. Then, we defined ROI maps for PET and gray-matter MRI obtained from the comparison of the AD patients and healthy control subjects in the first group.

Tomographic *z*-Score Mapping

After preprocessing of ^{18}F-FDG PET and MRI data using SPM2 as described, each ^{18}F-FDG PET or gray-matter image of group 2 subjects (34 AD patients and 50 healthy subjects) was compared with the mean and SD of ^{18}F-FDG PET or gray-matter images of the 40 healthy volunteers of group 1 by voxel-by-voxel *z*-score analysis using a software program developed by Matsuda (*21*); *z* score = ([control mean] − [individual value])/(control SD) as previously reported by Minoshima et al. (*3*) in a PET study. Although this software can implement some SPM functions, we used it only for tomographic *z*-score mapping because optimized VBM was not implemented in the current version. These *z*-score maps were displayed by overlay on tomographic sections. This program registered the SPM[*t*] results for significant decline of ^{18}F-FDG uptake or gray-matter concentrations in patients determined by group comparison in the first group as a specific ROI. Each ^{18}F-FDG PET or gray-matter image of 1 of the 50 healthy subjects was also compared with the averaged ^{18}F-FDG PET or gray-matter image of the remaining 40 healthy volunteers in the same manner as that used for the patients. Using the averaged value of positive *z* scores in the specific ROI in a *z*-score map as the threshold, receiver-operating-characteristic (ROC) analysis was performed for PET and VBM using the DBM-MRMC2.1β program developed at the University of Chicago (*22*) (http://xray.bsd.uchicago.edu/krl). The optimal cutoff was defined as the point to maximize the overall accuracy on the ROC curves.

Assessment of Age of Onset on Diagnostic Performance

The AD patients in group 2 were divided into 2 subgroup of the early- and late-onset AD patients on the basis of the age of onset using a 65-y threshold. The healthy subjects were also assigned into the early- or late-onset subgroup using a 70-y threshold so that the current age of AD patients in each subgroup matched that of healthy subjects. There were 15 AD patients (11 men, 4 women; mean age ± SD, 58.9 ± 7.3 y; mean age of disease onset, 54.7 ± 5.2 y; mean MMSE score, 22.5 ± 4.5) and 23 healthy subjects (13 men, 10 women; mean age, 60.0 ± 7.9 y) included in the early-onset group and 19 AD patients (5 men, 14 women; mean age, 75.9 ± 5.7 y; mean age of disease onset, 73.6 ± 5.5 y; mean MMSE score, 21.7 ± 3.2) and 27 healthy subjects (11 men, 16 women; mean age, 74.3 ± 3.0 y) in the late-onset group. The ROC analysis was performed for each subgroup.

In addition, the diagnostic accuracy of age-specific normal databases and ROIs were also tested. To determine age-specific ROIs, group 1 was divided into the early- and late-onset subgroups in a manner as described for subgrouping group 2. There were 11 AD patients (7 men, 4 women; mean age, 62.1 ± 4.9 y; mean age of disease onset, 59.0 ± 3.9 y; mean MMSE score, 21.2 ± 4.1) and 26 healthy subjects (16 men, 10 women; mean age, 62.0 ± 4.9 y) included in the early-onset group and 16 AD patients (5 men, 11 women; mean age, 73.1 ± 3.3 y; mean age of disease onset, 71.3 ± 2.8 y; mean MMSE score, 22.6 ± 2.6) and 14 healthy subjects (4 men, 10 women; mean age, 75.0 ± 3.0 y) in the late-onset group. ROIs of significant declines in ^{18}F-FDG uptake or gray-matter concentration were determined for each early or late-onset subgroup using SPM2 in a manner as described for the determination of regional metabolic reduction and gray-matter loss in the whole group 1. The results were uncorrected, with $P < 0.001$. Furthermore, healthy subjects in each subgroup served as the age-specific normal database for tomographic *z*-score mapping.

Statistical Analysis

Data are reported as mean ± 1 SD unless specified. Comparisons of 2 mean values were performed using a paired *t* test; a nonpaired *t* test was used where appropriate. When assumptions required for the *t* test were not met, the nonparametric Wilcoxon rank sum test was used. Comparisons of average *z* scores among ^{18}F-FDG PET, VBM, and VBM with modulated VBM were performed using repeated-measures ANOVA, which was followed

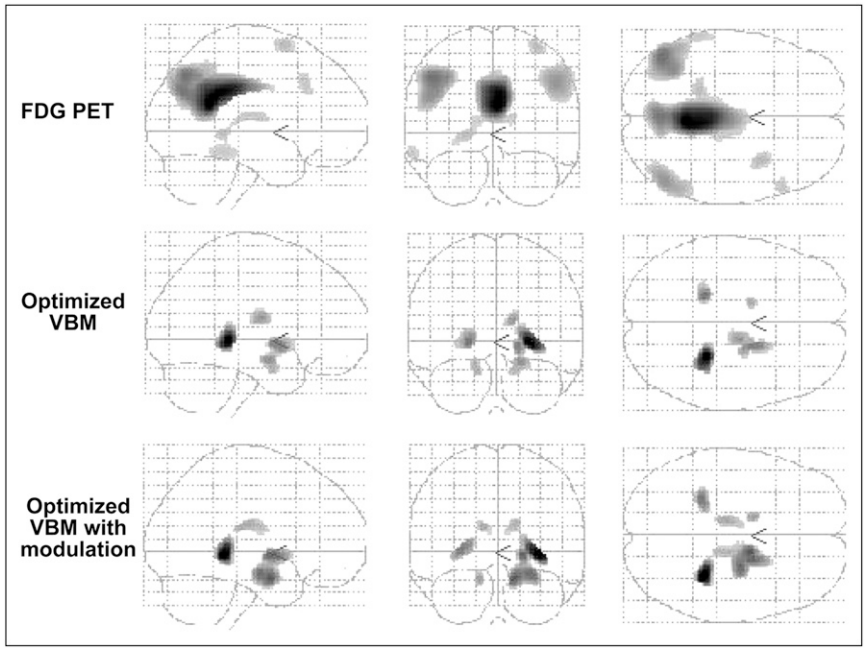

FIGURE 2. Maximum-intensity projections of SPM2 results for significant decline in adjusted ^{18}F-FDG uptake (upper) or gray-matter density by optimized VBM without modulation (middle) and optimized VBM with modulation (lower) in AD patients as compared with age-matched healthy volunteers. Height threshold < 0.05, corrected for multiple comparisons; extent threshold was set to 50 voxels.

by Tukey multiple-comparison tests with correction. The differences in the area under the ROC curves (A_z) were tested by jackknifing and ANOVA in the DBM-MRMC2.1β program. A_z values are expressed as mean ± SE. A P value < 0.05 was considered significant.

RESULTS

Determination of Abnormal Regions by PET and Optimized VBM

The areas of significant declines in ^{18}F-FDG uptake or gray-matter concentration of the first group patients are illustrated in Figure 2 and Table 2. In the PET study, the reductions in ^{18}F-FDG uptake were observed mainly in the posterior cingulate gyrus, the precuneus, and the bilateral angular gyri. In the VBM, reductions in gray-matter concentration were observed mainly in the hippocampal complex irrespective of the use of modulation. These regions with declines in ^{18}F-FDG uptake or gray-matter concentration were delineated as specific ROIs for AD.

Comparison of Diagnostic Accuracy Between ^{18}F-FDG PET and Optimized VBM

Average positive z-score values of ^{18}F-FDG PET and VBM in group 2 are summarized in Table 3. By definition, positive z scores indicate more severe hypometabolism or atrophy. The z-score value of PET or VBM in AD patients was higher than that in healthy subjects. When the z-score values of PET and VBM in AD patients were compared, ^{18}F-FDG PET showed a significantly higher average z-score value than either modulated or unmodulated VBM ($P < 0.001$). The ROC curves for discrimination of AD patients and control subjects were compared between ^{18}F-FDG PET and VBM (Fig. 3). The area under the ROC curve of ^{18}F-FDG PET (0.988 ±

TABLE 3
Average z Scores of ^{18}F-FDG PET and Optimized VBM in Group 2 Subjects

Subject	^{18}F-FDG PET	Optimized VBM	Optimized VBM with modulation	P value*
AD	2.69 ± 1.25†	1.62 ± 0.87	1.25 ± 070	<0.001
Healthy	0.75 ± 0.31	0.74 ± 0.38	0.57 ± 0.37	<0.05
P value‡	<0.0001	<0.0001	<0.0001	

*P value by repeated-measures ANOVA.
†$P < 0.001$ vs. either VBM or VBM with modulation.
‡P value by Wilcoxon rank sum test.

TABLE 2
Location and Peaks of Significant Decrease in Relative ^{18}F-FDG Uptake (Normalized to Global Mean of Whole Brain), Gray-Matter Concentration, and Gray-Matter Volume in AD Patients Compared with Healthy Volunteers

Structure	Coordinate x	y	z	t value
^{18}F-FDG PET				
R/L posterior cingulate gyrus	4	−50	22	12.85
R/L precuneus	0	−70	30	8.81
L angular gyrus	−46	−68	42	8.95
R angular gyrus	50	−68	40	7.57
R precentral gyrus	34	6	68	6.12
R hippocampus	−20	−40	−2	5.84
R thalamus	14	−18	14	5.04
R frontal gyrus	54	22	36	5.57
L temporal gyrus	−62	−38	−16	5.50
Optimized VBM				
R hippocampus (posterior portion)	26	−36	−2	7.45
L hippocampus (posterior portion)	−20	−36	0	6.26
R pallidum	16	0	−2	6.17
R hippocampus (anterior portion)	18	−4	−16	5.94
R thalamus	14	−10	16	5.69
L parahippocampal gyrus	−16	−2	−18	5.66
Optimized VBM with modulation				
R hippocampus	32	−36	−2	6.43
R parahippocampal gyrus	14	−2	−20	5.75
R pallidum	18	0	−2	5.67
R hippocampus	26	−10	−18	5.58
L hippocampus	−28	−40	2	5.51
L parahippocampal gyrus	−14	0	−22	5.18
L thalamus	−12	−20	20	5.10
R caudate	14	−18	22	5.03
R thalamus	14	−28	18	4.94

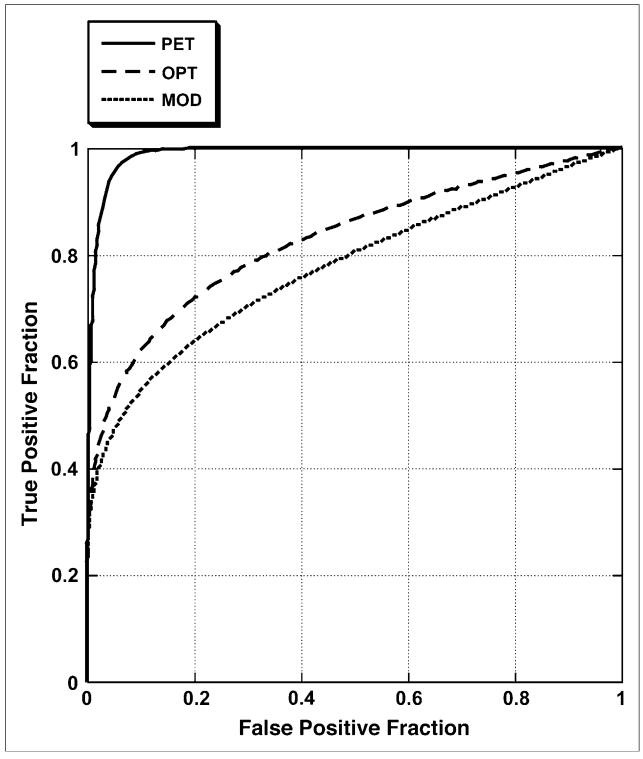

FIGURE 3. ROC curves for discrimination between AD patients and healthy volunteers obtained from PET and VBMs when thresholding at averaged values of positive z scores as measure of hypometabolism or atrophy. Solid line indicates ^{18}F-FDG PET; dashed line indicates optimized VBM without modulation (OPT); fine dot line indicates optimized VBM with modulation (MOD).

0.008 [mean ± SE]) was higher than that of VBM with (0.782 ± 0.059, $P < 0.01$ vs. PET) or without (0.832 ± 0.049, $P < 0.01$ vs. PET) modulation. When the accuracy was maximized, PET yielded a sensitivity, specificity, and accuracy of 100%, 92%, and 95%, respectively, whereas VBM yielded a sensitivity, specificity, and accuracy of 74%, 92%, and 85%. Modulation for VBM did not improve these values (56%, 94%, and 79%, respectively). The optimal cutoff *z*-score value to maximize accuracy was 1.2 for either PET, unmodulated VBM, or modulated VBM.

Effect of Age on Diagnostic Performance of ^{18}F-FDG PET and Optimized VBM

Average positive *z*-score values of ^{18}F-FDG PET and VBM in the early- and late-onset subgroups are summarized in Tables 4 and 5. The *z*-score value of PET or VBM in AD patients was higher than that in healthy subjects in each subgroup. When the *z*-score values of PET or VBM in AD patients were compared between the early- and late-onset subgroups, PET showed a significantly higher average *z*-score value in the early-onset subgroup than that in the late-onset subgroup ($P < 0.001$). In contrast, VBM, with or without modulation, showed a significantly lower average *z*-score value in the early-onset subgroup than that in the late-onset subgroup ($P < 0.001$).

Figure 4 illustrates ROC curves comparing ^{18}F-FDG PET and VBM in the early- and late-onset subgroups. In the early-onset subgroup, the A_z of ^{18}F-FDG PET (0.993 ± 0.008 [mean ± SE]) was significantly larger than that of VBM with (0.782 ± 0.072, $P < 0.05$ vs. PET) or without (0.768 ± 0.124, $P < 0.05$ vs. PET) modulation. In this subgroup, PET yielded a sensitivity, specificity, and overall accuracy of 100%, 96%, and 97%, respectively, whereas VBM yielded a sensitivity, specificity and accuracy of 67%, 91%, and 82%. Modulation for VBM did not improve these values (53%, 87%, and 74%, respectively). The optimal cutoff *z*-score value to maximize accuracy was 1.2 for PET, 1.0 for unmodulated VBM, or 0.7 for modulated VBM. In the late-onset subgroup, however, ^{18}F-FDG PET (0.980 ± 0.015) did not show a higher A_z value than that of VBM with (0.868 ± 0.061) or without (0.884 ± 0.055) modulation (not significant by jackknifing and ANOVA). The sensitivity, specificity, and accuracy in this subgroup were 100%, 89%, and 94% for PET, 84%, 89%, and 87% for VBM, and 79%, 93%, and 87% for VBM with modulation. The optimal cutoff *z*-score value to maximize accuracy was 1.2 for PET, unmodulated VBM, or modulated VBM. Case examples of early- and late-onset AD patients are illustrated in Figure 5.

TABLE 4
Average *z* Scores of ^{18}F-FDG PET and Optimized VBM in Early-Onset Subgroup

Subgroup	^{18}F-FDG PET	Optimized VBM	Optimized VBM with modulation	*P* value*
AD	3.30 ± 1.40†	1.23 ± 0.62	0.87 ± 0.52	<0.001
Healthy	0.69 ± 0.30	0.65 ± 0.40	0.44 ± 0.29	<0.05
P value‡	<0.0001	0.0043	0.0084	

*P value by repeated-measures ANOVA.
†$P < 0.001$ vs. either VBM or VBM with modulation.
‡P value by Wilcoxon rank sum test.

TABLE 5
Average *z* Scores of ^{18}F-FDG PET and Optimized VBM in Late-Onset Subgroup

Subgroup	^{18}F-FDG PET	Optimized VBM	Optimized VBM with modulation	*P* value*
AD	2.20 ± 0.89†	1.93 ± 0.93	1.54 ± 0.69	<0.05
Healthy	0.81 ± 0.31	0.82 ± 0.36	0.67 ± 0.40	NS
P value‡	<0.0001	<0.0001	<0.0001	

*P value by repeated-measures ANOVA.
†$P < 0.05$ vs. VBM with modulation.
‡P value by Wilcoxon rank sum test.

Figure 6 illustrates the areas of age-specific declines in ^{18}F-FDG uptake or gray-matter concentration in the early- and late-onset subgroup. Although these areas were similar to those for the whole group 1 (Fig. 2), ^{18}F-FDG PET showed a broader area of reduced uptake in the early-onset subgroup than that of the late-onset subgroup. Optimized VBM showed more neocortical involvement in the early-onset subgroup, whereas the area of decline was confined to medial temporal lobes in the late-onset subgroup. When age-specific ROIs and normal databases were used for analysis, the A_z of ^{18}F-FDG PET (0.997 ± 0.005) continued to be higher than that of VBM with (0.742 ± 0.104, $P < 0.01$ vs. PET) or without (0.830 ± 0.063, $P = 0.07$ vs. PET) modulation in the early-onset subgroup. The sensitivity, specificity, and accuracy in this subgroup were 100%, 96%, and 97% for PET, 60%, 91%, and 79% for VBM, and 67%, 78%, and 74% for VBM with modulation. In the late-onset subgroup, ^{18}F-FDG PET (0.971 ± 0.021), again, did not show a significantly higher A_z than that of VBM (unmodulated, 0.824 ± 0.078; modulated, 0.912 ± 0.055) (NS). The sensitivity, specificity, and accuracy in this subgroup were 95%, 89%, and 91% for PET, 79%, 81%, and 80% for VBM, and 89%, 89%, and 89% for VBM with modulation.

DISCUSSION

This study directly compares ^{18}F-FDG PET and optimized VBM for the detection of AD in the same individuals. The major finding of this study was that even with the sophisticated technique for MRI, such as optimized VBM, the morphologic approach did not achieve the diagnostic accuracy for the discrimination between AD patients and healthy subjects comparable with that obtained by ^{18}F-FDG PET. This was particularly true in the early-onset subjects.

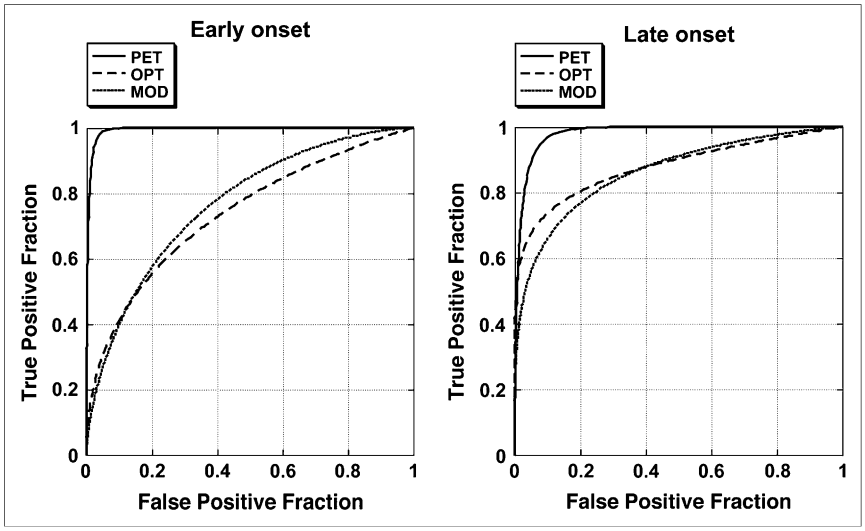

FIGURE 4. ROC curves for discrimination between AD patients and healthy volunteers in younger subgroup (left) and older subgroup (right) obtained from PET and VBM when thresholding at averaged values of positive *z* scores as measure of hypometabolism or atrophy. Solid line indicates ^{18}F-FDG PET; dashed line indicates optimized VBM without modulation (OPT); fine dot line indicates optimized VBM with modulation (MOD).

^{18}F-FDG PET and Optimized VBM in AD Patients

^{18}F-FDG PET is considered to be an established imaging technique to assess metabolic activity in various brain disorders. In particular, its diagnostic efficacy for the detection of AD is well documented in the literature (*1,3,23*). The results of this study confirmed this notion with high diagnostic accuracy, which is in line with these published studies (*1,3,23*). On the other hand, cerebral atrophy in AD patients, as a result of neurofibrillary degeneration and cell loss, is a pathologic feature of AD (*1*). VBM is an automated morphologic image analysis technique that permits the comparison of gray-matter concentration or volume on a voxel-by-voxel basis and has been applied in AD patients. In particular, optimized VBM used in this study minimizes the potential source of errors during spatial normalization by avoiding the influence of nonbrain tissue and, hence, has the effect of reducing the misinterpretation of significant differences relative to standard VBM (*15*). Additionally, modulated VBM enables regional assessment of specific gray-matter volume differences. Thus, the differences in hippocampal areas detected by unmodulated or modulated VBM observed in this study are considered to reflect reduced gray-matter concentration or volume in AD patients. From a clinical perspective, Hirata et al. (*12*) first reported the utility of VBM as a diagnostic tool for detection of AD from healthy subjects with a high diagnostic accuracy of 87.8%, which is essentially similar to the results (85%) of optimized VBM in this study. Furthermore, our data showed that modulation for VBM did not improve diagnostic accuracy, suggesting structural changes in mild AD patients as being more localized rather than global.

To date, few data that relate the use of VBM in comparison with ^{18}F-FDG PET for the detection of AD are available (*24*). A study by Kawachi et al. (*14*) has reported a slightly better diagnostic accuracy for the detection of AD patients as compared with standard VBM. The results of this study indicate that, even with optimized VBM, the morphologic approach may not achieve the overall diagnostic performance as ^{18}F-FDG PET does for discrimination of AD patients from healthy subjects. In particular, the metabolic alteration detected by ^{18}F-FDG PET is a highly sensitive (100%) and specific (92%) diagnostic marker of AD, whereas the reduced gray-matter concentration or volume measured by VBM is a less-sensitive (56%–74%) but, when present, specific (92%–94%) marker of AD. This may be in line with the concept that hypometabolism exceeds structural changes that has been described in early-onset familial AD (*25*).

Effects of Age on Diagnostic Accuracy

Age is known as the strongest risk factor for developing AD (*1*). Because the significance of a diagnostic test depends on disease prevalence, it would be important to assess the aging effect on diagnostic accuracies. In this study, the *z*-score value of PET or VBM in AD patients was higher than that in healthy subjects in both early- and late-onset subgroups. However, the *z* score of PET in AD patients was higher than that of VBM, particularly in the early-onset subgroup, suggesting the higher discrimination power of PET than VBM in this subgroup. Furthermore, PET showed a significantly higher *z*-score value in the early-onset subgroup than that in the late-onset subgroup. This is consistent with the previous finding that early-onset AD has more severe metabolic alteration in temporoparietal areas than does late-onset AD (*26*). In contrast, VBM with or without modulation showed a significantly higher *z*-score value in the late-onset subgroup than that in the early-onset subgroup, which is consistent with an earlier finding showing a greater hippocampal atrophy in late-onset AD (*27*). These observations lead to the hypothesis that the superiority of ^{18}F-FDG PET over VBM would be more significant in the early-onset subjects than in the late-onset subjects. This was confirmed in this study, which showed a higher diagnostic performance of

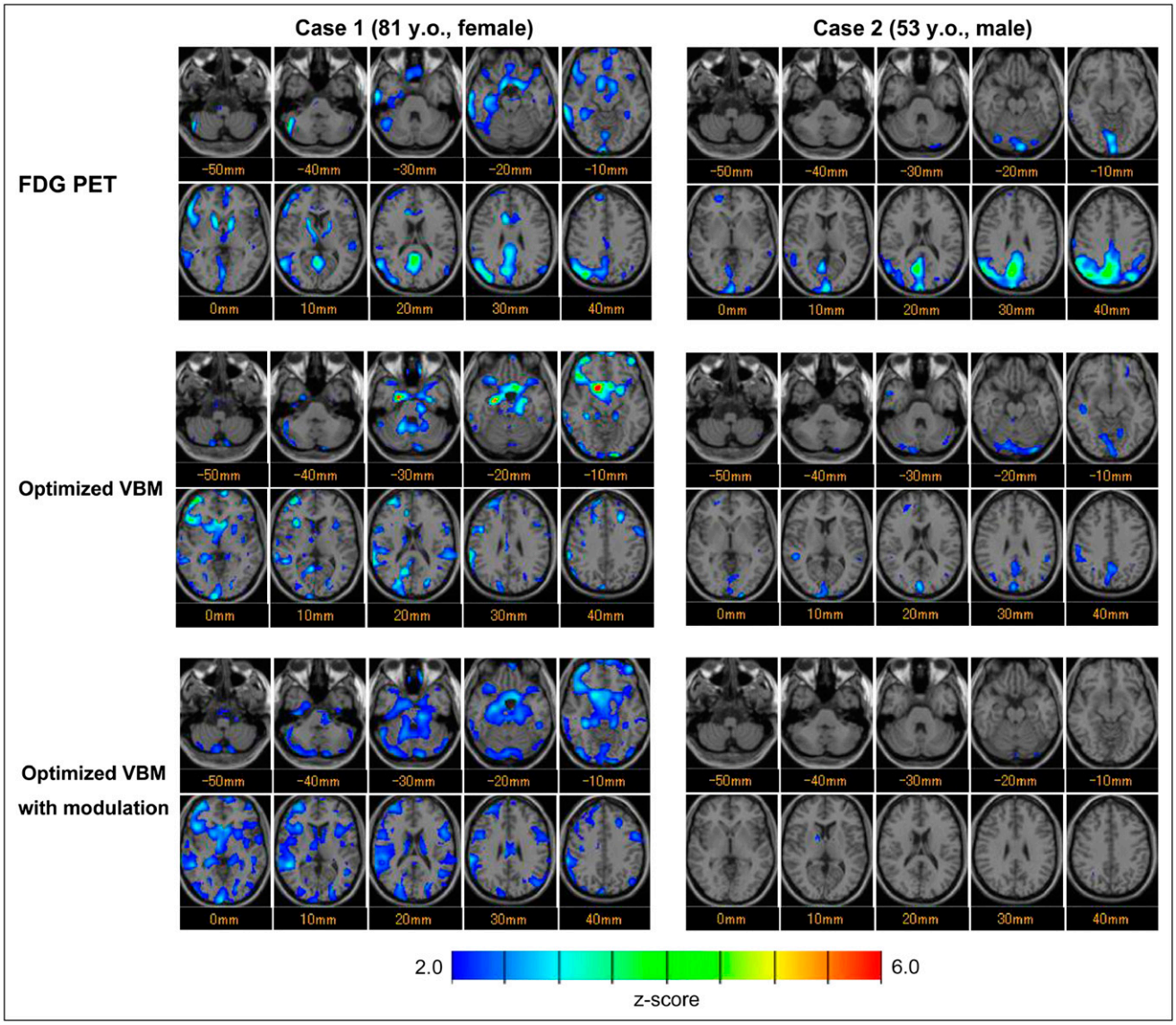

FIGURE 5. Comparison of ^{18}F-FDG PET and VBM *z*-score maps from 81-y-old woman (age of onset, 71 y; MMSE score, 22) and 53-y-old man (age of onset, 50 y; MMSE score, 26) with probable AD. Automated voxel-by-voxel *z*-score analysis was performed by comparison of relative ^{18}F-FDG uptake or gray-matter concentration/volume for patient with mean and SD of ^{18}F-FDG or gray-matter images of healthy volunteers. Color-scaled *z*-score maps ranging from 2.0 to 6.0 were displayed by overlaying on transaxial sections of anatomically standardized MRI of patient. By definition, positive *z* scores indicate more severe hypometabolism in PET or atrophy in VBM. Averaged values of positive *z* scores of the older patient in predefined regions were 2.66 for PET, 3.16 for optimized VBM without modulation, and 3.00 for optimized VBM with modulation, whereas those of the younger patient were 2.85, 0.09, and 0.29, respectively.

^{18}F-FDG PET in the early-onset subgroup than that in the late-onset subgroup. Furthermore, the optimal threshold of the *z* score to maximize the diagnostic accuracy of VBM differed between the early- and late-onset subgroups, indicating that the use of an age-specific reference value may be recommended.

In this study, the early-onset AD patients showed a broader area of reduced ^{18}F-FDG uptake than that of the late-onset patients. This is in line with the literature showing a more prominent metabolic reduction in the early-onset AD (*26*). On the other hand, areas of declines in gray-matter concentration or volume were observed in the medial temporal areas with little neocortical involvement when the whole group 1 was analyzed, which is in line with the literature showing little or no neocortical involvement detected by VBM (*12,24*). However, when the early- and late-onset AD patients were compared, the early-onset AD patients showed more neocortical involvement, whereas the areas of involvement were confined to the medial temporal areas in the late-onset AD patients. This is in line with a recent VBM study showing neocortical involvement, particularly in early-onset AD (*28*).

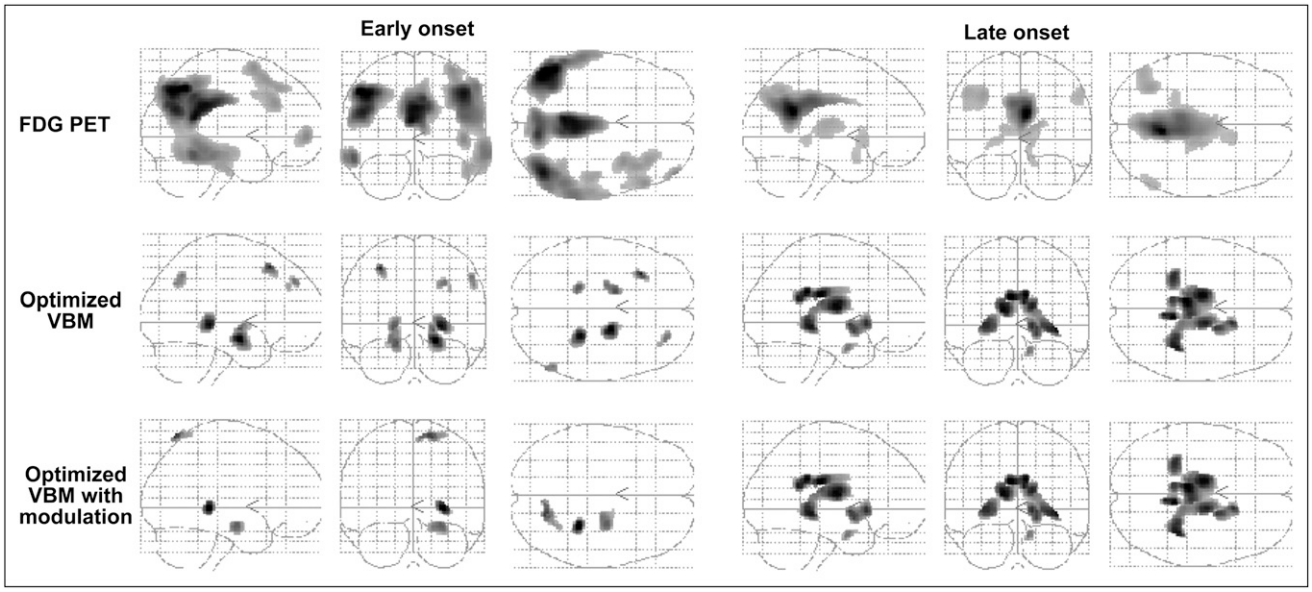

FIGURE 6. Maximum-intensity projections of SPM2 results for significant decline in adjusted ^{18}F-FDG uptake (upper) or gray-matter density by optimized VBM without modulation (middle) and optimized VBM with modulation (lower) in early-onset (left) and late-onset (right) AD patients as compared with age-matched healthy volunteers. Height threshold < 0.001, uncorrected for multiple comparisons; extent threshold was set to 50 voxels.

The use of age-specific ROIs and a normal database for analysis did not alter the results for ^{18}F-FDG PET and VBM. This may not be surprising because the aging effect primarily affects bilateral perisylvian areas for ^{18}F-FDG PET and VBM and not the areas of decline in ^{18}F-FDG uptake or gray-matter volume in AD patients, as demonstrated in our earlier study (*17*).

Limitations

There are limitations of the study to be described. First, VBM is one possible approach and the results of this study do not permit conclusions on ^{18}F-FDG PET versus general morphology. Second, we compared ^{18}F-FDG PET and VBM only to discriminate between AD patients and healthy volunteers. Therefore, our data cannot be applied to other types of diseases, such as frontotemporal dementia. This needs to be addressed in further studies. Third, the diagnosis of probable AD was made on the basis of clinical examinations in this study and, therefore, may differ from that obtained from final pathologic verification—a limitation present in many such studies. However, it has been reported that the diagnostic accuracy can exceed 90% in an academic "memory disorders" clinical setting (*29*), which was the case in this study.

CONCLUSION

The present study indicates that the detection of metabolic alteration by ^{18}F-FDG PET yields a better diagnostic performance for the discrimination between AD patients and healthy control subjects than does the morphologic approach by VBM, particularly in the early-onset subjects.

ACKNOWLEDGMENTS

This study was supported in part by a grant for Development of Advanced Technology for Measurement and Evaluation of Brain Functions, Ishikawa Prefecture Collaboration of Regional Entities for the Advancement of Technological Excellence, from Japan Science and Technology Corporation, Japan, and by a grant for the Knowledge Cluster Initiative (High-Tech Sensing and Knowledge Handling Technology [Brain Technology]) from the Japanese Ministry of Education, Culture, Sports, Science and Technology, Japan. The authors thank the staff in the Department of Neurology of Kanazawa University Hospital for their clinical support.

REFERENCES

1. Petrella JR, Coleman RE, Doraiswamy PM. Neuroimaging and early diagnosis of Alzheimer disease: a look to the future. *Radiology.* 2003;226:315–336.
2. Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp.* 1995;2:189–210.
3. Minoshima S, Frey KA, Koeppe RA, Foster NL, Kuhl DE. A diagnostic approach in Alzheimer's disease using three-dimensional stereotactic surface projections of fluorine-18-FDG PET. *J Nucl Med.* 1995;36:1238–1248.
4. Minoshima S, Giordani B, Berent S, Frey KA, Foster NL, Kuhl DE. Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Ann Neurol.* 1997;42:85–94.
5. Silverman DH, Small GW, Chang CY, et al. Positron emission tomography in evaluation of dementia: regional brain metabolism and long-term outcome. *JAMA.* 2001;286:2120–2127.
6. Herholz K. PET studies in dementia. *Ann Nucl Med.* 2003;17:79–89.
7. De Leon MJ, George AE, Golomb J, et al. Frequency of hippocampal formation atrophy in normal aging and Alzheimer's disease. *Neurobiol Aging.* 1997;18:1–11.
8. Jack CR Jr, Petersen RC, Xu YC, et al. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. *Neurology.* 1999;52:1397–1403.

9. Grundman M, Sencakova D, Jack CR Jr, et al. Brain MRI hippocampal volume and prediction of clinical status in a mild cognitive impairment trial. *J Mol Neurosci.* 2002;19:23–27.
10. Ashburner J, Friston KJ. Voxel-based morphometry: the methods. *Neuroimage.* 2000;11:805–821.
11. Good CD, Scahill RI, Fox NC, et al. Automatic differentiation of anatomical patterns in the human brain: validation with studies of degenerative dementias. *Neuroimage.* 2002;17:29–46.
12. Hirata Y, Matsuda H, Nemoto K, et al. Voxel-based morphometry to discriminate early Alzheimer's disease from controls. *Neurosci Lett.* 2005;382:269–274.
13. Bookstein FL. "Voxel-based morphometry" should not be used with imperfectly registered images. *Neuroimage.* 2001;14:1454–1462.
14. Kawachi T, Ishii K, Sakamoto S, et al. Comparison of the diagnostic performance of FDG-PET and VBM-MRI in very mild Alzheimer's disease. *Eur J Nucl Med Mol Imaging.* 2006;33:803–809.
15. Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, Frackowiak RS. A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage.* 2001;14:21–36.
16. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. 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.* 1984;34:939–944.
17. Yanase D, Matsunari I, Yajima K, et al. Brain FDG PET study of normal aging in Japanese: effect of atrophy correction. *Eur J Nucl Med Mol Imaging.* 2005;32:794–805.
18. Nowinski WL, Fang A, Nguyen BT, et al. Multiple brain atlas database and atlas-based neuroimaging system. *Comput Aided Surg.* 1997;2:42–66.
19. Good CD, Johnsrude I, Ashburner J, Henson RN, Friston KJ, Frackowiak RS. Cerebral asymmetry and the effects of sex and handedness on brain structure: a voxel-based morphometric analysis of 465 normal adult human brains. *Neuroimage.* 2001;14:685–700.
20. Nichols T, Hayasaka S. Controlling the familywise error rate in functional neuroimaging: a comparative review. *Stat Methods Med Res.* 2003;12:419–446.
21. Matsuda H. Role of neuroimaging in Alzheimer's disease, with emphasis on brain perfusion SPECT. *J Nucl Med.* 2007;48:1289–1300.
22. Dorfman DD, Berbaum KS, Metz CE. Receiver operating characteristic rating analysis: generalization to the population of readers and patients with the jackknife method. *Invest Radiol.* 1992;27:723–731.
23. Herholz K, Salmon E, Perani D, et al. Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. *Neuroimage.* 2002;17:302–316.
24. Ishii K, Sasaki H, Kono AK, Miyamoto N, Fukuda T, Mori E. Comparison of gray matter and metabolic reduction in mild Alzheimer's disease using FDG-PET and voxel-based morphometric MR studies. *Eur J Nucl Med Mol Imaging.* 2005;32:959–963.
25. Mosconi L, Sorbi S, de Leon MJ, et al. Hypometabolism exceeds atrophy in presymptomatic early-onset familial Alzheimer's disease. *J Nucl Med.* 2006;47:1778–1786.
26. Sakamoto S, Ishii K, Sasaki M, et al. Differences in cerebral metabolic impairment between early and late onset types of Alzheimer's disease. *J Neurol Sci.* 2002;200:27–32.
27. Frisoni GB, Testa C, Sabattoli F, Beltramello A, Soininen H, Laakso MP. Structural correlates of early and late onset Alzheimer's disease: voxel based morphometric study. *J Neurol Neurosurg Psychiatry.* 2005;76:112–114.
28. Frisoni GB, Pievani M, Testa C, et al. The topography of grey matter involvement in early and late onset Alzheimer's disease. *Brain.* 2007;130:720–730.
29. Rasmusson DX, Brandt J, Steele C, Hedreen JC, Troncoso JC, Folstein MF. Accuracy of clinical diagnosis of Alzheimer disease and clinical features of patients with non-Alzheimer disease neuropathology. *Alzheimer Dis Assoc Disord.* 1996;10:180–188.