AJNR

Optimization of DARTEL Settings for the Detection of Alzheimer Disease

J. Komatsu, I. Matsunari, M. Samuraki, K. Shima, M. Noguchi-Shinohara, K. Sakai, T. Hamaguchi, K. Ono, H. Matsuda and M. Yamada

Optimization of DARTEL Settings for the Detection of Alzheimer Disease

J. Komatsu, I. Matsunari, M. Samuraki, K. Shima, M. Noguchi-Shinohara, K. Sakai, T. Hamaguchi, K. Ono, H. Matsuda, and M. Yamada

ABSTRACT

BACKGROUND AND PURPOSE: Although Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL) has been introduced as an alternative to conventional voxel-based morphometry, there are scant data available regarding the optimal image-processing settings. The aim of this study was to optimize image-processing and ROI settings for the diagnosis of Alzheimer disease using DARTEL.

MATERIALS AND METHODS: Between May 2002 and August 2014, we selected 158 patients with Alzheimer disease and 198 age-matched healthy subjects; 158 healthy subjects served as the control group against the patients with Alzheimer disease, and the remaining 40 served as the healthy data base. Structural MR images were obtained in all the participants and were processed using DARTEL-based voxel-based morphometry with a variety of settings. These included modulated or nonmodulated, nonsmoothed or smoothed settings with a 4-, 8-, 12-, 16-, or 20-mm kernel size. A z score was calculated for each ROI, and univariate and multivariate logistic regression analyses were performed to determine the optimal ROI settings for each dataset. The optimal settings were defined as those demonstrating the highest χ^2 test statistics in the multivariate logistic regression analyses. Finally, using the optimal settings, we obtained receiver operating characteristic curves. The models were verified using 10-fold cross-validation.

RESULTS: The optimal settings were obtained using the hippocampus and precuneus as ROIs without modulation and smoothing. The average area under the curve was 0.845 (95% confidence interval, 0.788−0.902).

CONCLUSIONS: We recommend using the precuneus and hippocampus as ROIs without modulation and smoothing for DARTEL-based voxel-based morphometry as a tool for diagnosing Alzheimer disease.

ABBREVIATIONS: AD = Alzheimer disease; AUC = area under the curve; CDR = Clinical Dementia Rating; DARTEL = Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra; HS = healthy subjects; MMSE = Mini-Mental State Examination; ROC = receiver operating characteristic; VBM = voxel-based morphometry

Statistical neuroimaging analysis techniques such as voxel-based morphometry (VBM) have been widely used to evaluate structural MR imaging data, but in recent years, such methods have been suggested as a diagnostic aid for the early detection of Alzheimer disease (AD).[1,2] However, conventional VBM has often been criticized due to its imperfect registration of individual images of the standard brain.[3] Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL) has been introduced as an alternative method to conventional VBM due to novel abilities allowing precise segmentation and normalization of images.[4] Some studies have reported that DARTEL-based VBM provides a greater diagnostic accuracy for the detection of AD than conventional VBM methods.[5] Modulation, which involves the scaling of images depending on the extent of expansion or contraction, is considered an important processing step for

DARTEL-based VBM. The smoothing before statistical image analysis is also an important factor that can affect VBM results. However, there are scant data available regarding the optimal image-processing settings for DARTEL-based VBM.

Shima et al[6] previously demonstrated that a subset of patients with AD demonstrate atrophy in neocortical areas such as the posterior cingulate gyri and parietal lobe rather than in traditional areas such as the hippocampus, which is particularly true in early-onset AD. Therefore, it would be conceivable to hypothesize that optimal ROI settings for the detection of AD might vary according to age.

The purpose of this study was to optimize image-processing and ROI settings for the discrimination of patients with AD from age-matched healthy subjects using DARTEL-based VBM.

MATERIALS AND METHODS

Subjects

This prospective study was performed as a part of the Ishikawa Brain Imaging Study, which included any research to seek and develop imaging biomarkers for early and objective assessment of AD and other forms of neurodegenerative diseases using PET and MR imaging.[7,8] The study protocol was approved by the Medical and Pharmacological Research Center Foundation ethics committee, and written informed consent was obtained from all subjects before participation in the study. Of the 594 consecutive patients who were examined by neurologists and who underwent MR imaging (3D-T1-weighted, T2-weighted, MR angiography) at our memory clinic between May 2002 and August 2014, we selected 240 patients with a clinical diagnosis of probable AD at an early stage. Of the MR imaging scans, 3D-T1-weighted scans were used for both screening and analysis, whereas T2-weighted MR imaging and MR angiography were used for screening. Diagnosis of AD was based on the criteria of the National Institute of Neurologic and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association.[9] Forty-three patients were excluded from the study on the following grounds: 1) evidence of moderate-to-severe cognitive disturbance: grade 2 or more on the Clinical Dementia Rating (CDR),[10] with evidence of severe language, attentional, or behavioral disturbances that might complicate neurologic assessment; 2) uncontrolled major systemic disease or other neurologic disorders; and 3) evidence of focal brain lesions determined by MR imaging. We also excluded 39 patients with AD older than 80 years of age because it was impossible to find suitable age-matched healthy controls. Finally, 158 patients with AD were enrolled for analysis.

Regarding the generation of a control group for the construction of a healthy data base and for comparison against patients with AD, healthy subjects (HS) were recruited in response to advertisements. The following criteria were used to define healthy: 1) no history of brain trauma, psychiatric or neurologic disorders, or uncontrolled major systemic diseases and no current use of centrally acting drugs; 2) no abnormalities following general and neurologic examination; 3) a Mini-Mental State Examination (MMSE)[11] score of ≥28 and no clinical evidence of dementia; and 4) no evidence of asymptomatic cerebral infarction or brain vessel abnormalities on MR imaging. From the criteria, 704 subjects were determined to be HS among the 1369 recruited volunteers. From the 704 subjects, 198 age-matched subjects were selected; of these, 158 served as the control group against the patients with AD and the remaining 40 served as the healthy database.

Imaging Procedure

Both patients with AD and HS underwent structural MR imaging analysis. The structural MR imaging studies were performed using a 1.5T system (Signa Horizon; GE Healthcare, Milwaukee, Wisconsin). A 3D volumetric acquisition of a T1-weighted gradient-echo sequence produced a gapless series of thin transaxial sections using a magnetization-prepared rapid acquisition of gradient-echo sequence (TE/TR, 2.0/9.2 ms; flip angle, 20°; acquisition matrix, 256 × 192; number of slices, 124; pixel size, 0.78 × 1.04; slice thickness, 1.4 mm).

Image Processing

MR imaging data were analyzed with DARTEL-based spatial normalization with SPM8 software (http://www.fil.ion.ucl.ac.uk/spm/software/spm12).[4] MR images from 158 patients with AD and 198 HS were used to create templates for the DARTEL-based normalization technique. 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. Thus, gray matter intensity on modulated images should represent tissue volume rather than tissue concentration on nonmodulated images. During processing, both modulated and nonmodulated gray matter images were obtained for DARTEL-based VBM analysis. Modulated and nonmodulated gray matter images were nonsmoothed or smoothed, which is an image blurring using a function with a 4-, 8-, 12-, 16-, or 20-mm full width at half maximum Gaussian kernel, respectively, to investigate the effect of smoothing kernel size on DARTEL-based VBM. Twelve image datasets for each subject were generated.

Tomographic Z Score Mapping

After the preprocessing of MR imaging data, gray matter MR images were compared with the mean and SD of normal database gray matter images using a voxel-by-voxel z score analysis with a software program developed by Matsuda et al[12]: Z Score = ([Control Mean] − [Individual Value]) / (Control SD). This comparison was performed so that a positive z score value would represent reduced gray matter concentration or volume. These z score maps were displayed by overlay on tomographic sections. In each z score map, WFU Pick atlas-based ROIs (Department of Radiology of Wake Forest University School of Medicine, Winston-Salem, North Carolina; fmri.wfubmc.edu)[13] were drawn on the amygdala, hippocampus, parahippocampus, posterior cingulate gyrus, precuneus, frontal lobe, occipital lobe, parietal lobe, and temporal lobe.

Analysis

For each dataset, we investigated the diagnostic ability of each technique to discriminate patients with AD from HS. We per-
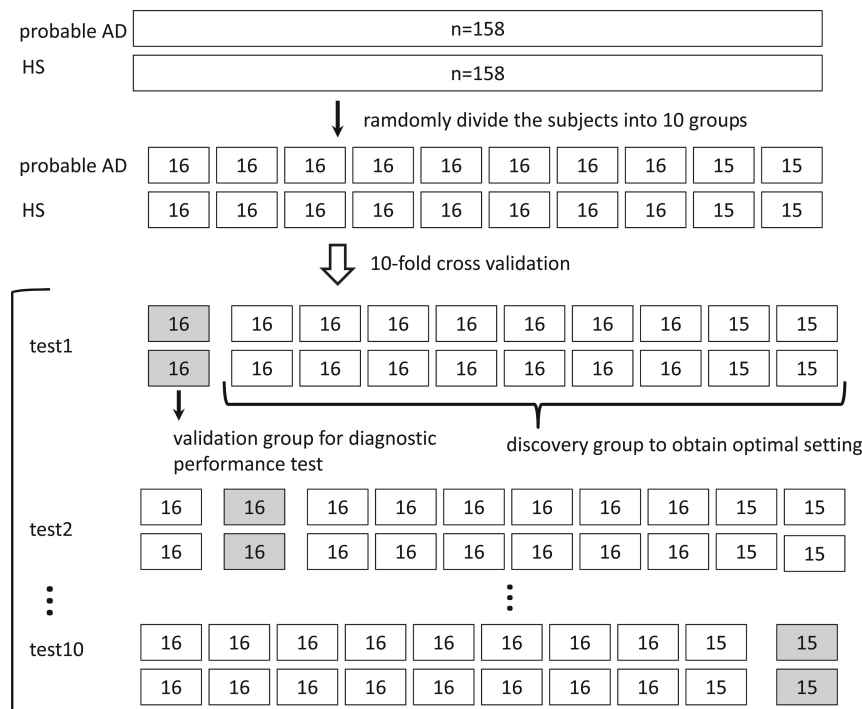
FIG 1. The scheme of 10-fold cross-validation. One hundred fifty-eight subjects with AD and 158 HS were randomly divided into 10 folds, with the same number of subjects with AD and HS in each fold. In each iteration, 9 of the folds were used for discovery (optimal setting determination) and 1 fold was used for validation (diagnostic performance test).

Table 1: Clinical characteristics of AD group and HS group

	AD (n = 158)	HS (n = 158)	NDB (n = 40)	P Value
Age (yr)	69 ± 8	68 ± 8	69 ± 6	.881[a]
Onset age (yr)	67 ± 8	–		
Sex (M/F)	80:78	87:71	20:20	.692[b]
MMSE	22.7 ± 3.6	29.0 ± 0.7	29.6 ± 0.6	<.001[a]
CDR (0.5/1)	20/138	–		
Term of education (yr)	11.7 ± 2.6	12.2 ± 2.9	12.1 ± 2.3	.196[c]

Note:—NDB indicates normal data base.
[a] ANOVA.
[b] χ^2 test.
[c] Kruskal-Wallis test.

formed 10-fold cross-validation to optimize image-processing and ROI settings and test the diagnostic performance. The subjects were randomly divided into 10 folds, with the same number of patients with AD and HS in each fold. In each iteration, 9 of the folds were used for discovery (optimal setting determination) and the remaining one was used for validation (diagnostic performance test). The scheme of 10-fold cross-validation is illustrated in Fig 1.

On the basis of the z score, the optimal ROIs for the discrimination of AD under each processing condition were determined using univariate and multivariate logistic regression analyses referring likelihood ratio χ^2 test statistic. Although our patients with AD and healthy subjects were matched for age and number as a whole, they were not individually matched. Therefore, we used regular logistic regression rather than the conditional logistic regression to analyze our datasets. The optimal settings were defined as the settings demonstrating the highest χ^2 test statistics in the multivariate logistic regression analyses. Multivariate logistic regression analyses

were performed using a stepwise backward elimination procedure. To assess the aging effect, age was entered as a variable for both univariate and multivariate analysis.

The validation group was used to estimate the diagnostic performance of the optimal setting with receiver operating characteristic (ROC) analysis. Diagnostic accuracy was assessed by the area under the curve (AUC). Data were expressed as mean ± SD. Comparisons of mean values were performed with ANOVA. When assumptions required for the ANOVA were not met, the nonparametric 2-sided Kruskal-Wallis test was used. The proportional difference among the groups was assessed using a χ^2 test. Statistical significance was defined as $P < .05$ (2-sided). All the statistical analyses were performed using a statistical software package (JMP 10; SAS Institute. Cary, North Carolina).

RESULTS
Clinical characteristics of patients with AD, HS, and the healthy database are summarized in Table 1. Clinical characteristics of the discovery and validation groups in the tests 1–10 are shown in On-line Tables 1A–10A. The results of univariate and multivariate analyses for the discrimination of patients with AD from HS in the tests 1–10 are shown in On-line Tables 1B–1G, 2B–2G, 3B–3G, 4B–4G, 5B–5G, 6B–6G, 7B–7G, 8B–8G, 9B–9G, and 10B–10G. The summaries of the optimal ROIs and diagnostic performance expressed as AUCs for each image-processing setting are shown in On-line Tables 1H–10H. Finally, the summary of 10-fold cross-validation is shown in Table 2.

Determination of Optimal Image-Processing and ROI Settings
When the smoothing kernel size was set to 8–20 mm, the amygdala and parietal lobe ROIs mainly contributed to the discrimination of AD regardless of the use of modulation (On-line Tables). When images were modulated and the smoothing kernel size was set to 0–4 mm, the amygdala and posterior cingulate gyrus ROIs were the main contributor. When images were nonmodulated and the smoothing kernel size was set to 0–4 mm, the hippocampus and precuneus ROIs were the predominant contributor (On-line Tables).

The results of multivariate analysis are summarized in Table 2. In 8 of 10 tests, the highest χ^2 statistic was obtained when images were nonmodulated and nonsmoothed. Furthermore,

Table 2: Summary of 10-fold cross-validation[a]

| | Optimized Image-Processing Condition | | | Likelihood Ratio | | |
	Modulation	Smoothing Kernel Size (mm)	Optimal ROI	χ^2 Test Statistic	P Value	AUC
Test 1	−	Nonsmoothed	Hippocampus	111.92	<.001	0.830
Test 2	−	Nonsmoothed	Hippocampus Precuneus	113.55	<.001	0.945
Test 3	−	Nonsmoothed	Hippocampus Precuneus	120.76	<.001	0.869
Test 4	−	Nonsmoothed	Hippocampus Precuneus	126.04	<.001	0.756
Test 5	−	4	Hippocampus Precuneus Parahippocampus	123.10	<.001	0.871
Test 6	+	8	Amygdala Frontal lobe Parietal lobe	118.83	<.001	0.738
Test 7	−	Nonsmoothed	Hippocampus Precuneus	124.73	<.001	0.836
Test 8	−	Nonsmoothed	Hippocampus Precuneus	119.94	<.001	0.871
Test 9	−	Nonsmoothed	Hippocampus Precuneus	126.26	<.001	0.851
Test 10	−	Nonsmoothed	Hippocampus Precuneus	127.45	<.001	0.787

[a] Likelihood ratio χ^2 test statistic and P values were determined by multivariate logistic regression analyses. Optimal image-processing condition and ROI and diagnostic performance are expressed as area under the curve for each validation group.

the optimal ROIs for the above settings mostly included the hippocampus and precuneus. Age did not survive as a variate to discriminate patients with AD from HS in any test. Thus, the optimal settings for the discrimination of patients with AD from HS were obtained when ROIs were set to the hippocampus and precuneus without modulation and smoothing in 7 of the 10 tests.

Most interesting, following modulation, the z score of the amygdala was increased, and conversely, that of the hippocampus was reduced. Examples are presented in Fig 2.

Diagnostic Performance of DARTEL-Based VBM Using the Optimal Image-Processing Settings and ROIs

Using the optimal image-processing and ROI settings determined by multivariate analyses of the training groups, we performed ROC analyses in each test group to assess the diagnostic ability as shown in Table 2. The AUC ranged from 0.738 to 0.945. When only the results of the ROI settings of the hippocampus and precuneus without modulation and smoothing (tests 2, 3, 4, 7, 8, 9, and 10) were summarized, the average of the AUC was 0.845 (95% confidence interval, 0.788–0.902). Additionally, there was a general trend for the ROC results to get worse along with smaller χ^2 statistics as the smooth function kernel size increased, particularly in nonmodulation settings (On-line Tables).

DISCUSSION

The major finding of this study was that the diagnostic ability of DARTEL-based VBM was highest when MR images were non-modulated and nonsmoothed with ROIs set to the hippocampus and precuneus.
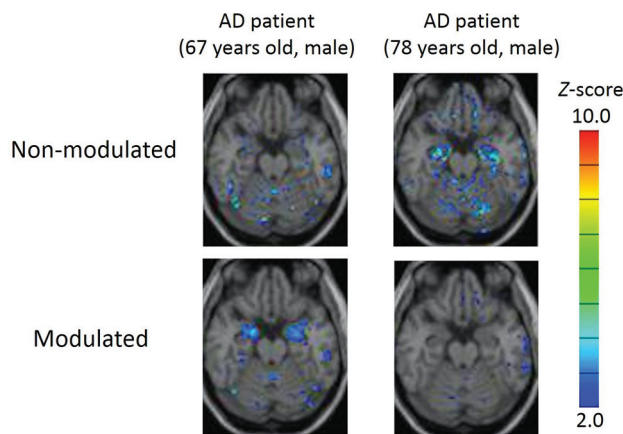

AD patient (67 years old, male) **AD patient (78 years old, male)**

Non-modulated

Modulated

Z-score 10.0 ... 2.0

FIG 2. Two sample cases of z score maps with and without modulation from patients with AD. Automated voxel-by-voxel z score analysis was performed by comparison of the relative gray matter concentration of patients with the mean and SD of HS. Color-scaled z score maps were displayed by overlaying them on the spatially normalized transaxial MR imaging. Smoothing was not applied. The modulated MR image of the patient with AD (67-year-old man; MMSE score, 19; CDR, 1) showed a high z score in the amygdala. The finding became obscure when the MR image was nonmodulated. Meanwhile, the nonmodulated MR image of another patient with AD (78-year-old man; MMSE score, 21; CDR, 1) showed a high z score in the hippocampus. The finding disappeared when the MR image was modulated.

Impact of Modulation on the Discrimination of AD in DARTEL-Based VBM

In the present study, modulation significantly influenced the diagnostic performance of DARTEL-based VBM for AD. Following modulation, volumetric differences in the amygdala became increasingly visible, and conversely, those of the hippocampus became more obscure. In theory, perfect spatial normalization

would result in no detectable differences between the individual gray matter images unless modulated.

The negative effect of modulation on the hippocampus might suggest that the detected differences between patients with AD and healthy subjects are likely to reflect imperfect registration between images rather than true volume differences. By contrast, the positive effect of modulation on the amygdala might suggest that spatial normalization was successful. These findings indicate that the effects of modulation influence each brain region differently, likely dependent on the structural complexity of each area.

Optimal Smoothing Kernel Size for the Discrimination of AD by DARTEL-Based VBM

In statistical image analysis, smoothing is routinely applied to reduce noise, normalize the distribution, and compensate for imperfect image registration.[14] A previous modulated DARTEL-based VBM study reported that optimal kernel size varied according to the group size.[15] The present study indicated that optimal kernel size varied depending on the use of modulation and variation in the ROI. If we took into account the above settings, the optimal setting was without smoothing the images. This result could be due to an increased ability to detect localized abnormalities by not smoothing. However, this anatomic difference between patients with AD and HS would be lost with smoothing, resulting in a smaller χ^2 and AUC, particularly in nonmodulation settings. Like the modulation effects as aforementioned, the effects of smoothing could also be region-dependent. Furthermore, our observation might be related to the specific programs used in this study during data processing. Therefore, one should interpret our results with caution, taking these circumstances into account.

Optimal ROI for the Discrimination of VAD by DARTEL-Based VBM

To determine a suitable ROI for the diagnosis of AD, previous VBM studies analyzed corresponding areas of gray matter volume for patients with AD and HS using group comparison analyses.[1,2,4,5,16,17] In the present study, we extracted areas of significant correlation for the discrimination of AD using multivariate logistic regression analysis and determined such areas as ROIs. As a result, in addition to medial temporal structures, the precuneus was designated as an optimal ROI for the detection of AD. This is in line with a previous study by Shima et al[6] demonstrating that some patients with AD show atrophy in neocortical areas such as the posterior cingulate gyrus and precuneus rather than in the medial temporal structures, particularly in young patients. Thus, the diagnostic performance of VBM for the discrimination of AD could be improved by combining the neocortical areas with medial temporal structures as ROIs.

Effects of Age on the Diagnostic Accuracy of AD with DARTEL-Based VBM

We did not find aging effects on the choice of VBM parameters when age was included as a predictor to discriminate AD from HS in our regression model. This was perhaps because most of our patients had late-onset AD as reflected by the mean onset age (67 ± 8 years) and the age range not being wide enough to show significant effects. This needs to be addressed in further studies focusing on young patients with AD.

Limitations

There are limitations to the study. First, the diagnosis of probable AD was made on the basis of clinical examinations and therefore may differ from that obtained with final pathologic verification, a limitation present in many such studies. However, it has been reported that diagnostic accuracy can exceed 90% in an academic memory disorders clinic setting.[18] Second, we investigated the optimal settings of DARTEL-based VBM to only discriminate between patients with AD and healthy subjects. Therefore, our data cannot be applied to other types of dementia. Finally, the data would be more robust if we could use our optimized model parameters applied on outside AD datasets such as the Alzheimer's Disease Neuroimaging Initiative data (http://adni.loni.usc.edu/). However, this proposal is beyond the scope of the current study and should be addressed in future.

CONCLUSIONS

For the discrimination of VAD from HS using DARTEL-based VBM, we recommend using the precuneus and hippocampus as ROIs without modulation and smoothing. The use of optimized ROIs and image-processing settings can provide a high level of diagnostic accuracy in the discrimination of AD.

REFERENCES

1. Hirata Y, Matsuda H, Nemoto K, et al. **Voxel-based morphometry to discriminate early Alzheimer's disease from controls.** *Neurosci Lett* 2005;382:269–74 CrossRef Medline
2. 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;33:801–09 CrossRef Medline
3. Bookstein FL. **"Voxel-based morphometry" should not be used with imperfectly registered images.** *Neuroimage* 2001;14:1454–62 CrossRef Medline
4. Ashburner J. **A fast diffeomorphic image registration algorithm.** *Neuroimage* 2007;38:95–113 CrossRef Medline
5. Matsuda H, Mizumura S, Nemoto K, et al. **Automatic voxel-based morphometry of structural MRI by SPM8 plus diffeomorphic anatomic registration through exponentiated lie algebra improves the diagnosis of probable Alzheimer disease.** *AJNR Am J Neuroradiol* 2012;33:1109–14 CrossRef Medline
6. Shima K, Matsunari I, Samuraki M, et al. **Posterior cingulate atrophy and metabolic decline in early stage Alzheimer's disease.** *Neurobiol Aging* 2012;33:2006–17 CrossRef Medline
7. Chen WP, Samuraki M, Shima K, et al. **Effect of an age-mismatched and sex-mismatched normal database on the diagnostic performance of ^{18}F-FDG PET for Alzheimer's disease: the Ishikawa Brain**

Imaging Study. *Nucl Med Commun* 2011;32:1128–33 CrossRef Medline

8. Samuraki M, Matsunari I, Chen WP, et al. **Glucose metabolism and gray-matter concentration in apolipoprotein E ε4 positive normal subjects.** *Neurobiol Aging* 2012;33:2321–23 CrossRef Medline

9. McKhann G, Drachman D, Folstein M, et al. **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;34:939–44 CrossRef Medline

10. Hughes CP, Berg L, Danziger WL, et al. **A new clinical scale for the staging of dementia.** *Br J Psychiatry* 1982;140:566–72 CrossRef Medline

11. Folstein MF, Folstein SE, McHugh PR. **"Mini-mental state": a practical method for grading the cognitive state of patients for the clinician.** *J Psychiatr Res* 1975;12:189–98 CrossRef Medline

12. Matsuda H, Mizumura S, Nagao T, et al. **Automated discrimination between very early Alzheimer disease and controls using an easy Z-score imaging system for multicenter brain perfusion single-photon emission tomography.** *AJNR Am J Neuroradiol* 2007;28:731–36 Medline

13. Maldjian JA, Laurienti PJ, Kraft RA, et al. **An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets.** *Neuroimage* 2003;19:1233–39 CrossRef Medline

14. Jones DK, Symms MR, Cercignani M, et al. **The effect of filter size on VBM analyses of DT-MRI data.** *Neuroimage* 2005;26:546–54 CrossRef Medline

15. Shen S, Sterr A. **Is DARTEL-based voxel-based morphometry affected by width of smoothing kernel and group size? A study using simulated atrophy.** *J Magn Reson Imaging* 2013;37:1468–75 CrossRef Medline

16. Frisoni GB, Testa C, Zorzan A, et al. **Detection of grey matter loss in mild Alzheimer's disease with voxel based morphometry.** *J Neurol Neurosurg Psychiatry* 2002;73:657–64 CrossRef Medline

17. Matsunari I, Samuraki M, Chen WP, et al. **Comparison of 18F-FDG PET and optimized voxel-based morphometry for detection of Alzheimer's disease: aging effect on diagnostic performance.** *J Nucl Med* 2007;48:1961–70 CrossRef Medline

18. Rasmusson DX, Brandt J, Steele C, et al. **Accuracy of clinical diagnosis of Alzheimer disease and clinical features of patients with non-Alzheimer disease neuropathology.** *Alzheimer Dis Assoc Disord* 1996;10:180–88 CrossRef Medline