Published in final edited form as:

Neuroimage. 2005 June ; 26(2): 600–608. doi:10.1016/j.neuroimage.2005.02.005.

Comparison of Different Methodological Implementations of Voxel-Based Morphometry in Neurodegenerative Disease

Matthew L. Senjem, B.S., Jeffrey L. Gunter, Ph.D., Maria M. Shiung, B.S., Ronald C. Petersen, M.D., Ph.D., and Clifford R. Jack Jr., M.D.

Mayo Clinic and Foundation Department of Radiology 200 First Street SW Rochester, Minnesota 55905

Abstract

Voxel-based morphometry (VBM) is a popular method for probing inter-group differences in brain morphology. Variation in the detailed implementation of the algorithm, however, will affect the apparent results of VBM analyses and in turn the inferences drawn about the anatomic expression of specific disease states. We qualitatively assessed group comparisons of 43 normal elderly control subjects and 51 patients with probable Alzheimer's disease, using five different VBM variations. Based on the known pathologic expression of the disease, we evaluated the biological plausibility of each. The use of a custom template and custom tissue class prior probability images (priors) produced inter-group comparison maps with greater biological plausibility than the use of the Montreal Neurological Institute (MNI) template and priors. We present a method for initializing the normalization to a custom template, and conclude that, when incorporated into the VBM processing chain, it yields the most biologically plausible inter-group differences of the five methods presented.

INTRODUCTION

Voxel-based morphometry (VBM) (Ashburner and Friston, 2000) is increasingly used to investigate differences in brain morphology among various types of patient groups and groups of control subjects. The method provides an estimate of inter-group differences in grey matter (GM) density and/or volume on a voxel-wise basis in a standardized space. VBM is particularly popular in evaluating the pathology of various neurodegenerative disorders, under the assumption that a close relationship exists between specific anatomic patterns of cerebral atrophy and the resulting cognitive disturbances. Previous VBM studies investigated, for example, Alzheimer's Disease (AD) (Baron et. al., 2001; Rombouts et. al., 2000; Frisoni et. al., 2002; Keller et. al., 2004; Testa et. al., 2004), Semantic Dementia (SD) (Rosen et. al., 2002; Mummery et. al., 2000), Frontal Temporal Lobe Dementia (FTLD) (Rosen et. al., 2002), and temporal lobe epilepsy (Keller et. al., 2004). Many of these studies, while focusing on similar disease states, have varied significantly with respect to the particular sequence of image processing steps employed. Thus, comparing the results of one study to another can be problematic. This is particularly true if the results are inconsistent.

A variety of options exist in the implementation of VBM. It is not clear which set of image processing options produce more plausible or more accurate results. Indeed, it is even difficult

Corresponding Author: Clifford R. Jack, Jr., M.D. Mayo Clinic Department of Radiology 200 First Street SW Rochester, MN 55905
Phone: 507-284-8548 FAX: 507-284-2405 jack.clifford@mayo.edu.

²This step involves iteratively updating a nonlinear transformation that best maps the segmented GM image to the MNI GM prior, then applying the transformation to the image. As a preparatory step to Method 5, we recorded the final iteration of the nonlinear transformation parameters, before they were applied to the image. These parameters were then used to initialize the normalization of the subject's image to the custom template in Method 5.

to define accuracy, as there is rarely a gold standard for comparison. It has been argued that imperfect spatial normalization may affect the validity of VBM results (Bookstein, 2001; Ashburner and Friston, 2001), underscoring the need for a comparison of different VBM methods. Bookstein pointed out that imperfect registration may lead one to misinterpret group differences as phenomena characteristic of a specific disease state, when in fact they could be due to alignment and subsequent comparison of dissimilar structures. Particularly in regions with large gradients, he argues, inferences made from group comparisons can be problematic.

The purpose of this project was to systematically evaluate the effects of varying certain elements of the VBM image processing chain in the context of differentiating cognitively normal elderly subjects from patients with AD. This was achieved by qualitative comparisons of the final output of five different VBM methods, implemented using the SPM2 software package (<http://www.fil.ion.ucl.ac.uk/spm/>). In addition to evaluating the standard (Ashburner and Friston, 2000) and optimized (Good et. al., 2001) VBM methods, we evaluated the effects of incorporating custom template and custom tissue class prior probability images (priors), hand edited brain masks to aid in segmentation, and an in-house scheme for initializing normalization to a custom template. This initialization technique partially addresses, but does not eliminate, concerns about registration raised by Bookstein (Bookstein, 2001).

Correlation of various VBM implementations with the known anatomic distribution of AD pathology

As will be seen, the various VBM processing methods produce different results when a group of normal elderly control subjects and AD patients are compared. In the absence of a precise quantitative method for comparing VBM implementations, a useful approach is to qualitatively compare results of each method against well-known concepts about the anatomic distribution of AD pathology (Hyman, 1997; Braak and Braak, 1991). The two cardinal pathologic hallmarks of AD are senile plaques (SP) and neurofibrillary tangles (NFT). Neurofibrillary pathology begins in the medial temporal lobe limbic cortex and spreads outward in a stereotypical fashion, to adjacent temporal lobe neocortical areas and then to frontal temporal and parietal neocortical association cortex. Senile plaques have a less stereotypical anatomic distribution, and do not display the orderly anatomic progression seen with NFT. Most patients with clear cut clinical AD, however, will demonstrate extensive SP and NFT deposition both in the limbic cortex as well as neocortical association areas at pathology.

Clinical symptoms in AD are attributed to loss of neurons, synapses, and dendrites. This pathology is observed as progressive cerebral atrophy. Therefore, when contrasting normal elderly control subjects versus AD subjects with VBM, one would expect a corresponding reduction in GM density, and/or volume. The most profound group differences are expected in the limbic cortex and neocortical association areas, with sparing of the sensory motor areas, cerebellum, and primary visual cortex. The ventricular system, of course, enlarges ex vacuo in patients with cerebral atrophy.

We evaluated the various VBM processing methods by focusing on the plausibility of results with respect to the known anatomic distribution of atrophy in AD. For example, if Method A produced more widespread areas of significant difference between patients and controls in appropriate limbic and neocortical association areas, we considered the results of Method A to be more plausible than those of Method B. Conversely, if Method B produced nonsensical inter-group differences, for example in the cerebellum or in primary sensory-motor areas, and Method C did not; we considered the results of Method C to be more plausible than those of Method B.

MATERIALS AND METHODS

Human Subjects and MRI Acquisition

Fifty-one probable AD and 43 cognitively normal elderly control subjects were selected from the Mayo Alzheimer's Disease Research Center (ADRC) and Alzheimer's Disease Patient Registry (ADPR) (Table).

These are IRB approved prospective longitudinal studies of aging and dementia, which incorporate MRI in the protocol. MR studies were performed at 1.5T with a standardized imaging protocol that included a coronal T1-weighted 3-dimensional volumetric sequence with 124 contiguous partitions, 1.6 mm slice thickness, 22×16.5 cm FOV, 192 views, and 25 degree flip angle.

MR Image Processing

The 51 AD and 43 control MR images were processed using five different variants of the VBM image processing algorithm. Our general approach in moving from Method 1 through Method 5 was to add an enhancement at each step which we thought might improve results. In the process, we deviated progressively more from standard VBM. Each of the five VBM methods included various combinations of normalization, segmentation, and smoothing, followed by statistical comparisons. All of these steps were implemented using the SPM2 software package (<http://www.fil.ion.ucl.ac.uk/spm/>). These algorithms are described extensively elsewhere, but for completeness, we provide short descriptions of the segmentation and normalization algorithms.

Segmentation algorithm

Each of the five VBM methods investigated in this study involved segmentation of MR images into white matter (WM), grey matter (GM), and cerebrospinal fluid (CSF). The segmentation routine incorporated into the SPM2 package utilizes a modified Gaussian Mixture Model (Ashburner and Friston, 2003). The algorithm uses tissue class probability images, which are stereotactically registered to template space, to improve the accuracy of the classification of each voxel into one of the three tissue types. A step for correcting intensity non-uniformity is also included, making the algorithm more robust to images with smooth intensity variations.

Normalization algorithm

The normalization algorithm of SPM2 begins by performing a 12 DOF affine registration of an individual patient's scan to a template, which seeks to minimize the least-squares distance of the two images. A number (16 in our experiments) of nonlinear iterations are then performed, using the discrete cosine transform to remove global nonlinear differences between the images (Ashburner and Friston, 1999). Also, a volume-preserving "modulation" step is employed, wherein each voxel is multiplied by the local volumetric expansion or contraction factor derived from the transformation¹.

Method 1

Standard VBM with the Montreal Neurological Institute (MNI) template: Method 1 is the "standard" VBM method (Ashburner and Friston, 2000), depicted schematically in Figure 1. Method 1 proceeds as follows: (i) perform a 12 degrees of freedom (DOF) affine registration between each subject's MRI image and the MNI template, (ii) perform a nonlinear normalization of the affine-registered subject image to the MNI template, (iii) segment the

¹The local expansion or contraction factor is defined as the determinant of the Jacobian of the deformation field at the specified voxel location.

normalized image using the MNI priors, (iv) apply Jacobian modulation (Ashburner and Friston, 2000) to the segmented GM image, (v) apply a 12 mm full width at half maximum (FWHM) spatial smoothing to the modulated GM image. The final step is statistical comparison, wherein the smoothed, modulated GM images are compared between patient and control populations using a two-sided T-test.

Method 2

Optimized VBM with the MNI template and priors: In optimized VBM, spatial normalization of each individual volume is based on matching the initial GM segmentation with the GM prior (Good et. al., 2001), whereas in standard VBM, spatial normalization is performed on the volume proper, using the MNI whole head template as a target for normalization (Ashburner and Friston, 2000). Optimized VBM, as originally described in (Good et. al., 2001) includes the use of a customized template; however, we have loosened the term to include all VBM methods which base the non-linear warping on matching GM to a GM prior, regardless of which template is actually used. This is in an effort to describe what was done without having to invent new terms for the various combinations of template and algorithm used.

Method 2 is depicted schematically in Figure 2, with the template option set to MNI, and the optional brain mask *not* included. The method proceeds as follows: (i) perform a 12 DOF affine registration between each subject's MRI image and the MNI template, (ii) segment the affine-registered image using the MNI priors, (iii) perform a nonlinear normalization of the segmented GM image to the MNI GM prior, applying the parameters obtained to the original whole head image, (iv) segment the normalized whole head image using the MNI priors, (v) apply Jacobian modulation to the segmented GM image, (vi) apply 12 mm FWHM spatial smoothing to the modulated GM image, and proceed to the statistics stage.

Method 3

Optimized VBM using custom template and priors: Method 3 is depicted schematically in Figure 2, with the template option set to custom, and the hand-edited brain mask (dotted box) *not* included. Method 3 follows the same algorithm as Method 2, using the custom template and priors instead of the MNI template and priors. Before describing Method 3 in detail, we must describe the creation of our customized template and priors.

Creation of Custom Template and Priors: The algorithm used for construction of our custom template and priors is depicted schematically in Figure 3, and proceeds as follows: (i) perform a 12 DOF affine registration between each subject's MRI image and the MNI template, (ii) segment the affine registered image using the MNI priors, (iii) perform a nonlinear normalization of the segmented GM image to the MNI GM prior³, (iv) segment the normalized subject image using the MNI priors, (v) average together all the normalized whole head, GM, WM, and CSF images to obtain the custom template and GM, WM, and CSF priors, respectively; (vi) apply an 8 mm FWHM Gaussian smoothing kernel to the custom template and priors. For comparison purposes, Figure 4 displays sagittal views of the MNI GM prior adjacent to our customized version.

Method 3 proceeds as follows: (i) perform a 12 DOF affine registration between each subject's MRI image and the custom template, (ii) segment the affine-registered image using the custom priors, (iii) perform a nonlinear normalization of the segmented GM image to the custom GM prior, applying the parameters obtained to the original whole head image, (iv) segment the normalized whole head image using the custom priors, (v) apply Jacobian modulation to the segmented GM image, (vi) apply 12 mm FWHM spatial smoothing to the modulated GM image, and proceed to the statistics stage.

Method 4

Optimized VBM using a custom template, plus the use of manually edited brain masks for each patient: In order to assess the potential benefit of utilizing expert manual editing to “cleanup” errors in automated segmentation, brain masks for each subject were created using the brain extraction, hole filling, and tracing tools in the Analyze software package (Robb, 2001). Method 4 is depicted schematically in Figure 2, with the template option set to custom, and the optional brain mask (depicted in a dotted box) *included*. Method 4 was identical to Method 3, except that the hand-edited brain mask was used as a weighting image during both segmentation steps.

Method 5

Optimized VBM using a custom template plus a re-initialization routine: Method 5 is depicted schematically in Figure 5. This method follows the same algorithm as Method 3, except for the re-initialization routines (explained below).

In creating the custom template and priors, the original images are normalized to the MNI template, and then segmented using the MNI priors. If a given image has been normalized to the MNI template during the process of creating a custom template, the MNI-normalized image should be closer to the custom template than the original. Therefore, the normalization routine may benefit by using the MNI-normalized image as a starting point from which to normalize to the custom template. Using the normalized image as a starting point, however, may introduce additional interpolation error. Instead of applying the MNI template normalization to each image, we use the parameters to initialize the custom template normalization. We accomplished this by slightly modifying the code in a few of the SPM2 routines associated with normalization. The modified version of the normalization routine allows one to use previous warping parameter estimates as initial estimates for subsequent normalizations.

Similarly, the initial segmentation using the custom priors can be initialized with the affine component of the MNI-normalization, which is also computed during creation of the custom template and priors. Except for these two initialization steps, Method 5 was the same as Method 3, and required no extra computation time beyond the time required by Method 3.

Statistics: In all five of the VBM methods, the segmented GM images were adjusted to represent actual volume, using the Jacobian modulation method of SPM2 (Ashburner and Friston, 2000). The modulated GM images were smoothed with a 12 mm FWHM Gaussian kernel, and then compared using a two sample T-test within the general linear model framework of SPM2.

The T statistic maps produced from the above methods were displayed both uncorrected ($p < 0.001$), and corrected ($p < 0.05$) using the family wise error (FWE) correction for multiple comparisons (Worsley et. al., 1996).

RESULTS

Results of group effect (AD vs. control) voxel-wise two-sided T-test analyses of GM volume are illustrated in the tri-planar glass brain displays of Figures 6–10. For each figure, the uncorrected ($p < 0.001$) results are displayed on the left side, with the FWE corrected ($p < 0.05$) results displayed on the right. Figures 6–10 display results obtained from Methods 1–5, respectively.

Comparison of Methods 1 and 2 illustrates that optimized VBM clearly identifies inter-group differences over a greater anatomic extent than standard VBM. Using the criteria of plausibility, results were mixed. Reduction in GM volume in AD patients compared to controls is present

over a noticeably greater extent of the lateral temporal neocortex, particularly on the left. This would indicate better plausibility for Method 2 vs. Method 1; however, other implausible differences were detected with Method 2. Inter-group differences in the cerebellum and the ventricles appear in Method 2, which were not present in Method 1. These findings do not match the known distribution of AD pathology. Though the lateral ventricles on average are certainly larger in AD patients than in controls, there is obviously no GM inside the ventricular system. The inter-group differences identified within the ventricular system by Method 2 are most likely due to misinterpreting differences in the position of the lateral ventricular wall as differences in GM volume. Karas et. al. reported similar findings, which they attributed to imperfect registration of voxels within the ventricles (Karas et. al., 2003). An inter-group difference in cerebellar GM volume in Method 2 is implausible, as the cerebellum is unaffected in general by AD pathology. These results underscore the need for a better normalization algorithm.

Comparison of Methods 2 and 3 illustrates that, by using custom template and prior images, inter-group differences in the extent of GM loss in patients compared to controls in appropriate neocortical association areas are detected over a greater anatomic extent than when using the MNI template and prior images. Both methods identify significant inter-group differences in GM volume within the lateral ventricular system, and do so to approximately the same extent. As discussed above, the findings in the ventricular system are most likely an effect of imperfect normalization. Both methods also identify inter-group differences in cerebellar GM, which we interpreted as an error. Overall, however, the results of Method 3 illustrate a more plausible match with the expected distribution of GM loss in AD patients relative to control subjects.

Comparison of Methods 3 and 4 reveals little difference between the two, implying the additional information provided by hand-edited brain masks is negligible.

Comparison of Method 5 with Methods 3 and 4 reveals several interesting differences. First, although Method 5 does identify a difference in GM volume within the ventricular system, this artifact is not present after correction for multiple comparisons (Worsley et. al., 1996), unlike either Methods 3 or 4. This suggests that the initialization routine employed in Method 5 resulted in a better alignment of the images, especially near strong image gradients. Also, no inter-group cerebellar differences are identified with Method 5, another feature in favor of greater plausibility for Method 5 versus Method 3 or 4. The anatomic extent of significant inter-group differences is approximately the same with Method 5 vs. Methods 3 or 4, in the absence of correction for multiple comparisons. When corrected for multiple comparisons, however, a greater anatomic extent of inter-group differences in appropriate neocortical association areas is seen with Method 5. Overall, therefore, when the criteria of anatomic plausibility is used, a reasonable qualitative ranking of the results would be Method 5 > Method 3 = Method 4 > Method 2 > Method 1.

DISCUSSION

We drew four main conclusions from the results of this study.

(1) Changes in the image processing chain of VBM noticeably influence the results of inter-group morphometric comparisons. Inferring fundamental principles of biology from apparent inter-group differences in VBM analyses should be approached judiciously, i.e. one must be careful not to mistake differences due to implementation details for biological differences. Other authors (Davatzikos, 2004; Bookstein, 2001) have made similar points. Bookstein pointed out that imperfect alignment of homologous structures, particularly in areas with large image gradients, could lead one to make incorrect inferences about anatomical variation between groups of patients and controls. Davatzikos (Davatzikos, 2004) suggested that a better

approach to characterizing group differences would employ a multivariate framework, rather than the mass-univariate framework inherent in SPM analyses. A multivariate approach would be more robust to detecting changes in cortical networks, he argues, and would therefore be advantageous in analyzing neurodegenerative disease progression.

(2) Optimized VBM produces different results than those obtained with standard VBM. Furthermore, the use of custom template and prior images improves the plausibility of inter-group comparisons, presumably due to improved segmentation and spatial normalization. Although our study does not directly measure the effects varying the template and priors separately, most investigators believe that the use of study-specific templates and priors produces a superior result (Good et. al., 2001).

When constructing a study-specific template using MNI space as a first approximation, there may be a residual bias toward MNI space carried over to the custom template, and subsequently, to the custom normalized scans. Other groups have developed unbiased iterative procedures for determining optimal templates (e.g. Davis et. al., 2004), and cortical unfolding algorithms for basing normalizations on sulcal and gyral pattern matching (e.g. Thompson and Toga, 1996; Van Essen and Drury, 1997). A comparison of these normalization methods to the methods used in this paper (esp. Method 5) would be useful in investigating the issue of template bias. For example, Salmond et. al. examined the effect of normalization on anatomical localization of specific structures in the brain (Salmond et. al., 2002). Such an analysis, however, is outside the scope of this paper.

Admittedly, the qualitative visual comparison of VBM results to the known anatomic distribution of pathology in AD is not the ideal method for comparing various VBM image processing algorithms. If a particular method, however, does not produce qualitatively reasonable results, it stands little chance of producing quantitatively reasonable results.

(3) The use of manually edited brain masks in the segmentation and normalization procedures does not significantly affect group comparison results. Based on the results of Method 4, the additional time and effort required for this step does not appear warranted.

(4) The additional step (in Method 5) of using previous estimates to initialize the normalization and segmentation routines gave more plausible results, was relatively easy to implement, and required no extra computation time beyond the time required by Method 3. Based on these results, our current VBM approach is to use method 5.

One might expect Methods 3 and 5 to produce the same results, since they are identical except for the initializations of the segmentation and normalization steps. This, in fact, is not the case. The initialization step of Method 5 results in an improvement, because the normalization essentially addresses the following question: after removing the variance of each scan from MNI space, what linear combination of nonlinear basis functions best removes the variance from the population mean?

Critiques of VBM have questioned the validity of the mass-univariate approach (Davatzikos, 2004), and the potential problems associated with imperfect registration of anatomic structures near large image gradients (Bookstein, 2001). Visual inspection of results from Methods 1–4 reveals significant group wise differences within the ventricles. This is a direct effect of the imperfect nature of the normalization and segmentation algorithms used. The results from Method 5 reveal less significant group differences inside the ventricles. Furthermore, after correction for multiple comparisons (Worsley et. al., 1996) is applied, the significance of these voxels is lowered below our threshold ($p < 0.05$). This suggests that the initialization steps in Method 5 resulted in better normalization and segmentation near strong image gradients, and

the ensuing results are indeed attributable to real morphometric differences, or at least more so than the results of the other methods.

Acknowledgments

Grant Support:

- The National Institute on Aging - AG16574, AG06786, AG11378.
- Robert H. and Clarice Smith and Abigail Van Buren Alzheimer's Disease Research Program.

REFERENCES

1. Ashburner J, Friston KJ. Nonlinear spatial normalization using basis functions. *Human Brain Mapping* 1999;7(4):254–266. [PubMed: 10408769]
2. Ashburner J, Friston KJ. Voxel-based morphometry - the methods. *Neuroimage* 2000;11:805–821. [PubMed: 10860804]
3. Ashburner J, Friston KJ. Why voxel-based morphometry should be used. *Neuroimage* 2001;14(6):1238–1243. [PubMed: 11707080]
4. Ashburner, J.; Friston, KJ. Image segmentation. In: Frackowiak, RSJ.; Friston, KJ.; Frith, C.; Dolan, R.; Friston, KJ.; Price, CJ.; Zeki, S.; Ashburner, J.; Penny, WD., editors. *Human Brain Function*. Vol. 2nd edition. Academic Press; 2003.
5. Baron JC, Chetelat G, Desgranges B, et al. In Vivo mapping of grey matter loss with voxel-based morphometry in mild Alzheimer's disease. *Neuroimage* 2001;14:298–309. [PubMed: 11467904]
6. Bookstein FL. “Voxel-based morphometry” should not be used with imperfectly registered images. *Neuroimage* 2001;14:1454–1462. [PubMed: 11707101]
7. Braak H, Braak E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathology* 1991;82:239–259.
8. Davatzikos C. Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. *Neuroimage* 2004;23:17–20. [PubMed: 15325347]
9. Davis B, Lorenzen P, Joshi S. Large deformation minimum mean squared error template estimation for computational anatomy. *Proceedings, International Symposium on Biomedical Imaging (ISBI)* 2004:173–176.
10. Frisoni GB, Testa C, Zorzan A, et al. Detection of gray matter loss in mild Alzheimer's disease with voxel-based morphometry. *J. Neurol. Neurosurg. Psychiatry* 2002;73:657–664. [PubMed: 12438466]
11. Friston KJ, Ashburner J. Generative and recognition models for neuroanatomy. *Neuroimage* 2004;23:21–24. [PubMed: 15325348]
12. 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. [PubMed: 11525331]
13. Hyman BT. The neuropathological diagnosis of Alzheimer's disease: Clinical-pathological studies. *Neurobiology of Aging* 1997;18:S27–S32. [PubMed: 9330982]
14. Karas GB, Burton EJ, Rombouts SARB, van Schijndel RA, O'Brien JT, Scheltens P. h. McKeith IG, Williams D, Ballard C, Barkhof F. A comprehensive study of gray matter loss in patients with Alzheimer's disease using optimized voxel-based morphometry. *Neuroimage* 2003;18:895–907. [PubMed: 12725765]
15. Keller SS, Wilke M, Wieshmann UC, Sluming VA, Roberts N. Comparison of standard and optimized voxel-based morphometry for analysis of brain changes associated with temporal lobe epilepsy. *Neuroimage* 2004;23:860–868. [PubMed: 15528086]
16. Mummery CJ, Patterson K, Price CJ, Ashburner J, Frackowiak RSJ, Hodges JR. A voxel-based morphometry study of semantic dementia: relationship between temporal lobe atrophy and semantic memory. *Ann. Neurology* 2000;47:36–45.

17. Robb RA. The biomedical imaging resource at Mayo Clinic. *IEEE Trans Med Imaging* 2001;20(9): 854–867. [PubMed: 11585203]
18. Rombouts SA, Barkhof F, Witter MP, Scheltens P. Unbiased whole-brain analysis of grey matter loss in Alzheimer's disease. *Neuroscience Letters* 2000;285:231–233. [PubMed: 10806328]
19. Rosen HJ, Gorno-Tempini ML, Goldman WP, Perry RJ, Schuff N, Weiner M, Feiwell R, Kramer JH, Miller BL. Patterns of brain atrophy in frontotemporal dementia and semantic dementia. *Neurology* 2002;58:198–208. [PubMed: 11805245]
20. Salmond CH, Ashburner J, Vargha-Khadem F, Connelly A, Gadian DG, Friston KJ. The precision of anatomical normalization in the medial temporal lobe using spatial basis functions. *Neuroimage* 2002;17:507–12. [PubMed: 12482103]
21. Testa C, Laakso MP, Sabattoli F, Rossi R, Beltramello A, Soininen H, Frisoni G. A comparison between the accuracy of voxel-based morphometry and hippocampal volumetry in Alzheimer's disease. *J. Mag. Res. Imaging* 2004;19:274–282.
22. Thompson P, Toga AW. A Surface-based technique for warping 3-dimensional images of the brain. *IEEE Trans. Med. Imaging* 1996;15(4):1–16.
23. Thompson PM, Hayashi KM, de Zubicaray G, Janke AL, Rose SE, Semple J, Hong MS, Herman D, Gravano D, Dittmer S, Doddrell DM, Toga AW. Dynamics of gray matter loss in Alzheimer's disease. *Journal of Neuroscience* 2003;23(3):994–1005. [PubMed: 12574429]
24. Van Essen DC, Drury HA. Structural and functional analyses of human cerebral cortex using a surface-based atlas. *Neuroscience* 1997;17(18):7079–7102. [PubMed: 9278543]
25. Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC. A unified statistical approach for determining significant voxels in images of cerebral activation. *Human Brain Mapping* 1996;4:58–73.

Fig. 1. shows a schematic flow diagram of Method 1.

Fig. 2. shows schematic flow diagrams of Methods 2–4. Method 2 used the MNI template and priors. Methods 3–4 used the custom template and priors. The brain mask was used only for Method 4.

Fig. 3. shows a schematic flow diagram for creation of the custom template and priors. Note the registration and normalization parameters are saved for future use in Method 5.

Fig. 4. shows sagittal views of the MNI GM prior (left), and our customized GM prior (right). Note the morphological differences, especially the difference in ventricle size.

Fig. 5. shows a schematic of Method 5.

Fig 6.
(Method 1) Group effect analysis with standard VBM before (left), and after (right) family-wise error correction.

Fig 7.
(Method 2) Group effect analysis with optimized VBM using MNI template and priors before (left), and after (right) FWE correction.

Fig 8.
(Method 3) Group effect analysis with optimized VBM using custom template and priors before (left), and after (right) FWE correction.

Fig 9.
(Method 4) Group effect analysis with optimized VBM using custom template and priors, plus hand-edited brain masks before (left), and after (right) FWE correction.

Fig 10.

(Method 5) Group effect analysis with optimized VBM using custom template and priors, plus the re-initialization scheme described in the text, before (left), and after (right) FWE correction.

Table

Demographic Information

	Normals	AD
Number of Subjects	43	51
Sex (F/M)	25/18	30/21
Age (mean, range)	80.1, 57–100	75.3, 49–92
Education (mean # years)	13.6	13
MMSE (mean)	28.7	21.8