Neuroradiology (2010) 52:203–213
DOI 10.1007/s00234-009-0600-1

DIAGNOSTIC NEURORADIOLOGY

An enhanced voxel-based morphometry method to investigate structural changes: application to Alzheimer's disease

Xingfeng Li · Arnaud Messé · Guillaume Marrelec · Mélanie Pélégrini-Issac · Habib Benali

Received: 14 May 2009 / Accepted: 14 September 2009 / Published online: 6 October 2009

Abstract
Introduction When characterizing regional cerebral gray matter differences in structural magnetic resonance images (sMRI) by voxel-based morphometry (VBM), one faces a known drawback of VBM, namely that histogram unequalization in the intensity images introduces false-positive results.
Methods To overcome this limitation, we propose to improve VBM by a new approach (called eVBM for enhanced VBM) that takes the histogram distribution of the sMRI into account by adding a histogram equalization step within the VBM procedure. Combining this technique with two most widely used VBM software packages (FSL and SPM), we studied GM variability in a group of 62 patients with Alzheimer's disease compared to 73 age-matched elderly controls.
Results The results show that eVBM can reduce the number of false-positive differences in gray matter concentration.
Conclusion Because it takes advantage of the properties of VBM while improving sMRI histogram distribution at the same time, the proposed method is a powerful approach for analyzing gray matter differences in sMRI and may be of value in the investigation of sMRI gray and white matter abnormalities in a variety of brain diseases.

Keywords Voxel-based morphometry · Structural magnetic resonance image · Alzheimer's disease · Gray matter · Histogram matching

Electronic supplementary material The online version of this article (doi:10.1007/s00234-009-0600-1) contains supplementary material, which is available to authorized users.

X. Li · A. Messé · G. Marrelec · M. Pélégrini-Issac · H. Benali
Laboratoire d'Imagerie Fonctionnelle, and LINeM,
GHU Pitié-Salpêtrière, Inserm, UPMC Univ Paris 06,
UMR_S 678, 91 bd de l'Hôpital,
75634 Paris, Cedex 13, France

X. Li · A. Messé · G. Marrelec · M. Pélégrini-Issac · H. Benali
LINeM, International Laboratory of Neuroimaging and Modeling,
IUGM, Université de Montréal,
4565 chemin Queen-Mary,
Montréal H3W 1W5, Canada

X. Li · A. Messé · G. Marrelec · M. Pélégrini-Issac · H. Benali
CRM, Université de Montréal,
CP 6128, Succursale Centre-ville,
Montréal H3C 3J7, Canada

H. Benali (✉)
UMR_S 678 Inserm/UPMC, Laboratoire d'Imagerie
Fonctionnelle, Faculté de Médecine, CHU Pitié-Salpétrière,
91 Bd de l'Hôpital,
75634 Paris Cedex 13, France
e-mail: Habib.Benali@imed.jussieu.fr

Introduction

Voxel-based morphometry (VBM) is a method that characterizes brain differences in vivo using structural magnetic resonance images (sMRI). It has been successful in identifying structure differences since it was established in 1995 [1–4]. The method has been applied in a wide variety of studies, including schizophrenia, developmental and congenital disorders, temporal lobe epilepsy, aging, and Alzheimer's disease (AD) [3, 5–7]. For example, VBM finds differences in gray matter (GM) concentration between age-matched elderly controls and AD subjects [8, 9].

Although conventional VBM method is a powerful method to study sMRI, it may overestimate some structure differences or detect false-positive regions. The reason for this is that the information of the sMRI may be distorted in the processing steps of the conventional method (brain extraction, non-uniformity correction, segmentation, registration, smoothing and modulation, and statistical multivar-

iate comparison). This information distortion can be due to false positives in the statistical comparison in the last step [10]; but it can also be due to unequalization of the image histograms of different subjects in the beginning of the preprocessing steps. For example, due to biological variability in morphology across subjects [11], it is generally observed that the GM histogram distribution of sMRI is different across subjects. Furthermore, the GM histogram distribution is often mixed with that of white matter (WM) even within the same subject, and this could be one of the reasons for the difficulties encountered when segmenting WM from GM in the same sMRI.

To overcome the limitation of image histogram unequalization between subjects, we suggest an enhanced VBM (eVBM) method, which is based on the VBM method. The basic idea of the approach is to enhance the image histogram in conventional VBM before groupwise comparison. All image histograms are adjusted according to the histogram distribution of a template sMRI so that all histograms are matched with each other before further analysis. In this way, the method reduces the drawback of having big histogram differences in the same group data.

In the following, the value of eVBM compared to the conventional VBM method is assessed by studying GM variability in a group of control subjects compared to a group of patients with AD. Extensive research has been focusing on the application of sMRI to study how the disease begins and progresses [9, 12–15]. VBM is a method of choice to evaluate the cortical atrophy of the GM in AD. Atrophy in AD has been demonstrated by various VBM studies, but the regional distributions of atrophy found in these studies are not consistent with one another. For example, some studies [9, 13, 16] found occipital and cerebellum differences, but others [17, 18] did not. This inconsistency may result from the fact that VBM overestimates GM concentration difference in some regions and underestimates it in others due to the between-subjects variability in histogram distributions. Therefore, eVBM could be a valuable method to investigate the inconsistent results found in the literature of AD studies. AD is an application of choice to compare eVBM with conventional VBM.

The paper is organized as follows. Firstly, we describe the histogram matching method used in the study. Then, we concentrate on the application of this method for comparing AD patients and age-matched controls and we compare the results of conventional VBM with the eVBM method using both FSL-VBM (for details see FSL-VBM v1.1: http://www.fmrib.ox.ac.uk/fsl/ [19]) and statistical parametric mapping (SPM5; http://www.fil.ion.ucl.ac.uk/spm/) with VBM5.1 (http://dbm.neuro.uni-jena.de/vbm/vbm5-for-spm5/) software packages (SPM5-VBM5.1). We conclude the study by discussing the advantages and the limitations of the eVBM method.

Materials and methods

eVBM and histogram matching algorithm

The basic idea of eVBM is to reduce the unequalization of sMRI histograms in the conventional VBM method. This is obtained by transforming each sMRI with a histogram matching algorithm, which is a generalization of histogram equalization. Conventional VBM usually successively involves brain extraction, non-uniformity correction, segmentation, registration, smoothing and modulation, and statistical inference for groupwise comparison. The proposed eVBM approach then consists of introducing the histogram matching step just after brain extraction.

The goal of the histogram matching algorithm is to adjust the histogram distribution of a sMRI to a template histogram distribution. For example, the source images in Fig. 1 a, b are transformed to a template image histogram distribution (such as that shown in Fig. 1 c) with a desired brightness distribution over the whole image gray scale, leading to Fig. 1 d, e, respectively.

More formally, let $\{x\}$ be a discrete grayscale image to match to the template histogram distribution, and n_i be the number of occurrences of gray level i. The probability of an occurrence of a voxel of level i in the image $\{x\}$ is:

$$p_x(i) = p(x = i) = \frac{n_i}{n}, 0 \leq i < L,$$

where L and n are the total number of gray levels in the image and the total number of voxels in the image, respectively. p_x is the histogram of the image, normalized to [0,1]. The cumulative distribution function (CDF) cdf_x corresponding to p_x is defined as:

$$cdf_x(i) = \sum_{j=0}^{i} p_x(j),$$

which is also the image's accumulated normalized histogram. Let us design a transformation of the form $y=T(x)$ to produce a new image $\{y\}$, such that its CDF cdf_y is linearized across the whole range of values, i.e.

$$cdf_y(i) = iK$$

for some constant K. The properties of the CDF allow to perform the following transformation:

$$y = T(x) = cdf_x.$$

Matching the histogram of $\{x\}$ to a given histogram template $\{z\}$ then consists of transforming $\{x\}$ to $\{\widetilde{x}\}$ according to [20]:

$$\widetilde{x}(i) = cdf_z^{-1}\{cdf_x(i)\},$$

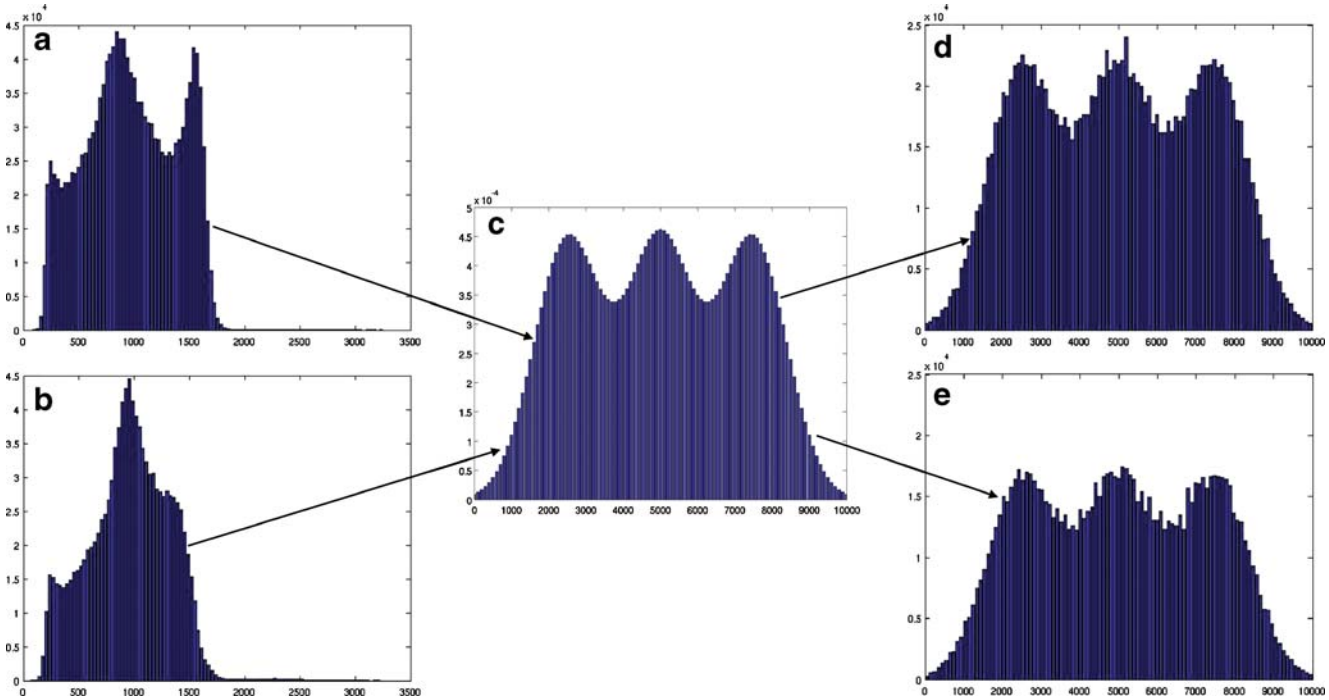

Fig. 1 *a* and *b* Histograms of two typical sMRIs from control subjects after removing the background and the brain skull of the images. The histogram in *a* shows two peaks, while that in *b* shows one peak. *c* Template histogram. *d* and *e* The matched histogram of *a* and *b*, respectively

where cdf_z represents the CDF of the template histogram $\{z\}$, and cdf_z^{-1} denotes the inverse transformation of cdf_z.

A Gaussian mixture template histogram distribution (Fig. 1 c) was adapted as a reference distribution for histogram matching, with three identical Gaussian laws (from 0 to 10,000 gray levels, of mean 2,500, 5,000, and 7,500, respectively; the standard deviation was 900) regularly spaced to enhance contrast. The MATLAB® (The MathWorks, Inc.) computer code histo.m was used to realize the histogram matching algorithm and is freely available from: http://optica.csic.es/projects/tools/steer/1.matlabPyrTools/index.html.

In the following sections, we detail how this method was compared to conventional VBM using both FSL-VBM and SPM5-VBM5.1 software packages to study GM concentration differences between age-matched controls and AD patients.

Subjects

All the data were obtained from the Open Access Structural Imaging Series (OASIS, http://www.oasis-brains.org/) database, generously contributed by Dr. Randy Buckner [21]. The study was approved by a regional ethics committee. One hundred and sixty-nine right-handed subjects over the age of 60 were included at the beginning of the study. Only 135 subjects (62AD patients) were eventually used in the analysis for both FSL-VBM and SPM5-VBM. The age of AD patients (25 males) was 76.63±7.27, and the age of control subjects (19 males) was 75.71±8.97. The age of AD patients was not significantly different from that of controls (t=0.6558, non-significant at p<0.05, two-tailed t test). The mini mental state examination (MMSE) score was 24.66±3.88 for AD patients and 29.08±1.10 for control subjects. MMSE of controls was significantly different from that of AD patients (t=8.6810, p<0.05). The AD subjects were clinically diagnosed with very mild to moderate AD, and AD patients were divided into two groups: clinical dementia rating (CDR)=0.5, very mild dementia (45 subjects); CDR=1, mild dementia (17 subjects) [22].

Structural MRI acquisition

All sMRI were collected with a 1.5 T scanner (Vision, Siemens, Erlangen, Germany). Structural images were acquired with a transmit-receive circularly polarized head coil, and a T_1-weighted magnetization prepared rapid gradient-echo sequence (TR [recovery time]=9.7 ms; TE [echo time]=4 ms; flip angle=10°), giving 128 (gap 1.25 mm) sagittal slices of 256×256 image voxels with a voxel size of 1×1×1.25 mm. For each subject, three to four individual T_1-weighted MRI scans were obtained in single scan sessions, and individual sMRI was averaged to increase signal-to-noise ratio. No neuroimaging evidence of focal lesions such as brain tumor was found, and neither cortical nor subcortical vascular lesions were visible on the structural images.

Structural MRI processing

Data processing followed the steps described in Fig. 2, including (for eVBM) or not (for conventional VBM) the histogram matching step. To better localize, interpret, and compare regional differences in the results, we chose the 116-labeled regions of the AAL template [23] to label regions in the resulting statistical maps.

SPM5-VBM5.1 and SPM5-eVBM5.1 data processing chains

The VBM analysis (Fig. 2a) with SPM was conducted as follows. The averaged structural images of each of the initial 169 subjects were first registered to the Talairach space by the *mritotal* function provided by MINC tools (http://www.bic.mni.mcgill.ca/software/) to improve the registration in SPM5 (Fig. 2a). Twenty-one subjects were mis-registered by the MINC tools and were therefore excluded from the study. The averaged structural images were resampled to $1\times1\times1$ mm by using the *mincresample* function in MINC tools. Then, brain images were normalized to the Montréal Neurological Institute (MNI) space using the tool provided by SPM5, with default parameters. This tool includes bias correction (correction for the intensity inhomogeneities in images), unified segmentation [24], and registration to the MNI space (Fig. 2a). Thirteen subjects (seven controls and six AD patients) were further discarded due to the mis-segmentation of GM, even when trying different segmentation parameters. For better later comparison, only the remaining 135 subjects (62 AD patients (25 males) and 73 age-matched healthy elderly controls (19 males) aged 60 to 96) subjects were used in all experiments.

Finally, a Gaussian kernel (full width at half maximum (FWHM)=7.05 mm) was used for smoothing the remaining normalized and modulated GM images. A two-sample *t* test with assumption of equal variance was selected to address whether GM concentration was different between two groups of subjects. The significant difference level with the family-wise error (FWE) corrected threshold was set to be $p<0.05$ for all our experiments.

The SPM5-eVBM5.1 approach differed from the conventional SPM5-VBM5.1 pipeline in that after registration to the Talairach space, the BET method of FSL [25] was employed to extract the brain from the average structural image of each subject (Fig. 2a dashed gray block). BET is used only for eVBM and not for the conventional VBM, and a histogram matching step was added before unified segmentation (gray block in Fig. 2a).

Besides, since the GM of 13 out of 148 subjects could not be segmented correctly no matter what parameters we used in SPM5-VBM5.1, we also investigated the segmentation procedure further for these subjects in a separate analysis and assessed whether the histogram matching procedure improved segmentation results (see the "Discussion" section).

FSL-VBM and FSL-eVBM data processing chains

So that FSL can be used with default settings, the averaged structural images were resampled to $2\times2\times2$ mm by using the mincresample function in MINC tools. The FSL-VBM analysis (Fig. 2b) was then conducted as follows. First, the BET method [25] was employed to extract the brain from the averaged structural image for each of the 135 subjects that were also processed with SPM5-VBM5.1. Next, non-uniformity correction was carried out, and FAST4 [26] was used to segment tissues according to their type. The segmented GM partial volume images were then aligned to the MNI standard space (MNI152) by applying the affine registration tool FLIRT [27, 28] and nonlinear registration FNIRT [29, 30] methods, which use a B-spline representation of the registration warp field [31]. The registered images (before smoothing) were averaged to create a study specific template, and the native GM images were then nonlinearly re-registered to the template image. Visual check was performed to control the quality of brain image extraction, segmentation, and registration for each averaged structural image. Mis-extracted, mis-segmented, and mis-registered images were then processed again by using different parameter values until the results looked visually satisfactory. The registered GM partial volume images were then modulated (to correct for local expansion or contraction) by dividing them by the Jacobian of the warp field. The segmented and modulated images were then smoothed with an isotropic Gaussian kernel with a standard deviation of 3 mm (FWHM=7.05 mm).

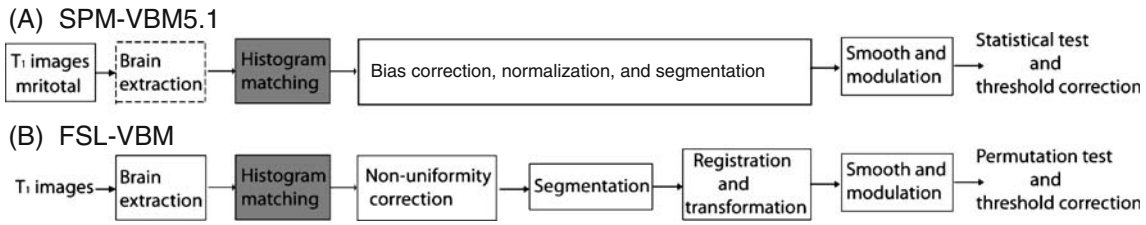

Fig. 2 eVBM data analysis protocol proposed in this study. The *gray block* shows the additional histogram matching step incorporated into the conventional VBM method. The *block diagram* in **a** shows SPM5-VBM5.1 steps, while that in **b** shows FSL-VBM steps

Finally, permutation-based non-parametric testing (1,000 permutations) was used in a voxel-wise general linear model [32] for comparison of different groups. FWE [33] was adopted to correct the threshold for multiple comparisons across space. Threshold-free cluster enhancement was employed to assess cluster significance [10]. The significant difference level with the FWE corrected threshold was set to be $p<0.05$ for all our experiments. Since one difference between FSL-VBM and SPM5-VBM5.1 is that FSL-VBM uses a permutation test for the statistical comparison in the last step of VBM analysis (comparing Fig. 2b with a) while SPM5-VBM5.1 uses a two-sample t test, for further comparison of the results obtained by the two methods, the modulated and smoothed GM images obtained from FSL were also analyzed with a two-sample t test (i.e., the same method as in SPM5-VBM5.1) thresholded at $p<0.05$ (FWE correction for multiple comparisons).

The FSL-eVBM approach differed from the conventional FSL-VBM pipeline in that a histogram matching step was added before non-uniformity correction (gray block in Fig. 2b).

To further compare FSL results with SPM results, the final statistical maps were resampled to $1\times1\times1$ mm by using the mincresample function in MINC tools.

Evaluation of the eVBM method

In order to assess the value of histogram matching, we conducted five experiments to compare eVBM with conventional VBM.

1. The first experiment consisted of comparing control subjects with AD patients using FSL-VBM and FSL-eVBM (processing steps are shown in Fig. 2b).
2. The second experiment consisted of comparing control subjects with control subjects by using FSL-VBM and FSL-eVBM. This control versus control experiment was designed to address the issue that the histogram distribution unequalization in sMRI leads to false-positive results. Since no ground truth is available, the idea of this experiment is based on the assumption that when comparing two groups of control subjects, one should find no significant difference in GM concentration or, put it differently, the differences one would find are likely to be false positives due to histogram variability in the sMRI. For the "control versus control" experiments, the control data were split into two groups (37 subjects for one group and 36 subjects for the other group; the mean age of the subjects in the two groups was 78.15±9.28 and 73.97±8.32, non-significant difference, $t=1.80$). The criterion for assigning the subjects to one or the other group was that the histograms of the sMRI in the first group (seven males) were similar to the histogram shown in Fig. 1 b (one mode only, i.e., gray matter), while the histograms of the sMRI in the second group (13 males) were similar to that shown in Fig. 1 a (two modes for GM and WM, respectively). In this way, using conventional VBM, we expect to find differences only due to histogram variability (hence, false positives in the sense that they do not reflect an actual change in GM concentration or an atrophy phenomenon), while eVBM is expected to reduce these differences (therefore, reduce the number of false positives) since histogram matching reduces histogram variability.
3. The third experiment also aimed at comparing the control subjects with AD patients, but the difference was that histogram matching was performed *after* non-uniformity correction (and not before, see Fig. 2b) in the FSL-VBM processing pipeline. MRI image intensity non-uniformity is an artifact usually caused by radio-frequency field inhomogeneity, eddy currents, and so on. The purpose of non-uniformity correction step in VBM is to remove the bias due to intensity non-uniformity. Histogram matching is also a bias correction method that reduces histogram variability in sMRI. We therefore wished to study whether the non-uniformity correction in the pre-processing steps affected histogram matching in eVBM or not.
4. The fourth experiment consisted of comparing control subjects with AD patients by using SPM5-VBM5.1 (Fig. 2a).
5. The fifth experiment consisted of comparing control subjects with control subjects (with the same groups as in the second experiment, as described above) by using SPM5-VBM5.1 (Fig. 2a).

Results

Comparing AD patients with control subjects by using FSL-VBM and FSL-eVBM (first experiment)

The results obtained by using VBM and eVBM for detecting significant differences between controls and AD patients are given in Fig. 3a, b, respectively. Colored areas represent regions in which GM probability was significantly higher in controls than in patients ($p<0.05$, corrected), superimposed on a structural image in the MNI space. The color bar denotes the magnitude of the difference in GM concentration as measured by the T values. Figure 3a shows the results of the conventional FSL-VBM method. Regions of the hippocampus and parahippocampus gyrus, amygdala region, fusiform gyrus, occipital cortex, frontal and temporal lobes, middle cingulum cortex, and caudate

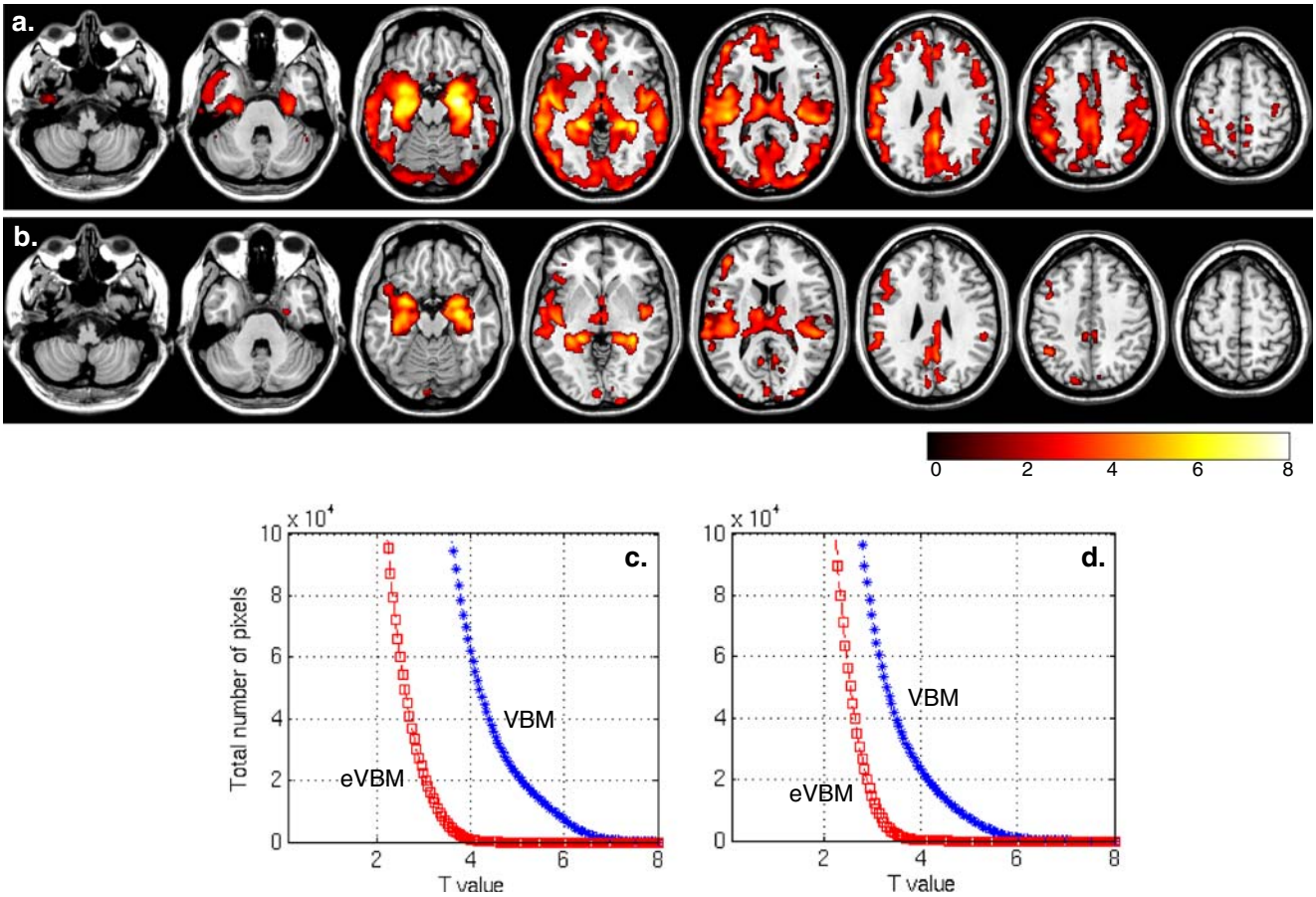

Fig. 3 Comparison of FSL-VBM and FSL-eVBM methods. **a**: *T*-value maps obtained using the conventional FSL-VBM method. **b** *T*-value maps obtained using the FSL-eVBM method. *Colored regions* show where the GM probability was significantly higher in controls than in patients ($p<0.05$, FWE corrected threshold). **c** and **d** Total number of voxels above the threshold plotted against the *T* value using the two methods. **c** The results of the comparison between controls and AD patients; **d** the results of the comparison between the first group of controls and the second group of controls. The *star-dashed curve* denotes the conventional VBM results, and the *square-dashed curve* represents the results of the eVBM method

nucleus were found to be significantly different in AD patients compared with controls. Figure 3b shows the results obtained with the FSL-eVBM method. The most significant GM differences between patients and controls were found in the hippocampus and parahippocampus gyrus and amygdala regions bilaterally, while only parts of the insula and precuneus regions showed significant differences. Other regions in which a significant difference was found included the caudate nucleus, the temporal cortex, and the lingual cortex. These results exhibit conservative structure differences and lower *T* values than those shown in Fig. 3a. Although both FSL-VBM and FSL-eVBM methods detected a significant difference in the temporal cortex, the size of this region was different, suggesting that FSL-eVBM detected less differences in the results. We also compared the controls with AD patients using the two methods as shown in Fig. 3c, which plots the total number of voxels above the threshold against the *T* value for both methods. At the same *T*-value level, FSL-VBM found more voxels above that level than the FSL-eVBM method.

Figure 4a (respectively Fig. 4b) shows the results of the comparison between AD patients and controls using FSL-VBM (respectively FSL-eVBM) when a two-sample *t* test was used for statistical inference instead of the permutation test. Compared with Fig. 3, less significant differences in GM concentration were found for VBM (Fig. 4a compared with Fig. 3a) as well as for eVBM (Fig. 4b compared with Fig. 3b).

Comparing control subjects with control subjects by using FSL-VBM and FSL-eVBM (second experiment)

To better illustrate the behavior of the FSL-eVBM method compared with the conventional VBM, we compared the first group of control subjects (one peak in the histogram as shown in Fig. 1 b) with the second group of controls (two peaks in the histogram, Fig. 1 a) using FSL-VBM and FSL-

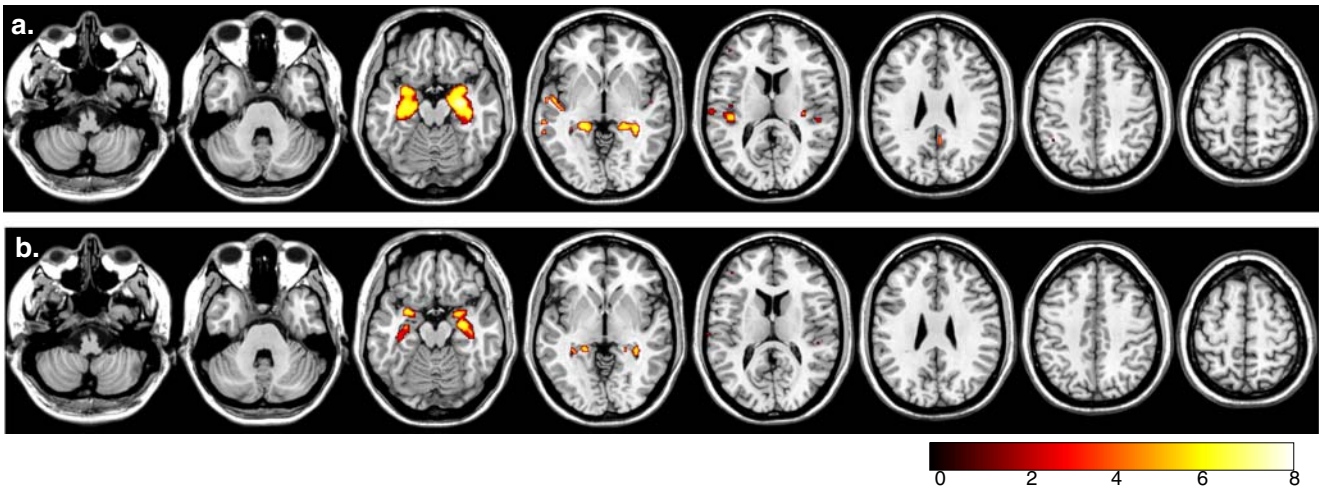

Fig. 4 FSL-VBM and FSL-eVBM with a two-sample *t* test for statistical inference instead of a permutation test. **a** *T*-value maps for the conventional VBM method. **b** *T*-value maps for the eVBM method. *Colored regions* show where the GM probability was significantly higher in controls than in patients ($p<0.05$, FWE corrected threshold)

eVBM. Figure 3d plots the total number of voxels above the threshold against the *T* value for both methods. Conventional VBM yielded more supra-threshold voxels than eVBM, suggesting that the conventional method did introduce false-positive results in the analysis.

FSL-eVBM with histogram matching performed after non-uniformity correction (third experiment)

When histogram matching was performed after non-uniformity correction, the results obtained when comparing control subjects with AD patients were very similar to those obtained when histogram matching was performed before non-uniformity correction (results not shown here), suggesting that the histogram matching algorithm is insensitive to non-uniformity correction.

Comparing AD patients with control subjects by using SPM5-VBM5.1 and SPM5-eVBM5.1 (fourth experiment)

Figure 5a shows the results obtained with SPM5-VBM5.1 when comparing control subjects with AD patients, while Fig. 5b shows the corresponding results obtained with SPM5-eVBM5.1. Figure 5a shows that the most significant GM differences between patients and controls were found in the hippocampus and parahippocampus gyrus and part of the temporal cortex, while SPM5-eVBM5.1 found fewer significant regions between the two groups, as shown in Fig. 5b.

Figure 5c plots the total number of voxels above the threshold against the *T* value for both methods. The results in Fig. 5c show that more voxels were above the same *T*-value threshold using conventional VBM than using eVBM method.

Comparing control subjects with control subjects by using SPM5-VBM5.1 and SPM5-eVBM5.1 (fifth experiment)

We also compared the first group of control subjects (one peak in the histogram as shown in Fig. 1 b) with the second group of controls (two peaks in the histogram, Fig. 1 a) using both methods. As observed with the FSL processing pipeline, conventional VBM yielded more supra-threshold voxels than eVBM, suggesting that the conventional method did introduce false-positive results in the analysis.

Discussion

In this study, we have developed an eVBM method to analyze sMRI data. The eVBM method is based on the histogram adjustment of the structural images. The idea is based on the fact that there is considerable histogram distribution variability in sMRI (Fig. 1 a, b for example). This histogram unequalization is thus likely to lead to false-positive results in the VBM analysis (Figs. 3d and 5d). We investigated the consequences of histogram unequalization and adapted a histogram matching algorithm to adjust the histogram distribution. This adjustment leads to a better (in terms of Gaussian distributions) histogram distribution of the image intensities (Fig. 1 c, d). And therefore, it allows for areas of lower local contrast to achieve a high contrast without affecting the global contrast. Histogram equalization accomplishes this by effectively spreading out the most frequent intensity values. As a result, this is expected to avoid false-positive significant differences when comparing two groups. We investigated this issue in VBM by using the FSL-VBM and SPM5-VBM5.1 software packages.

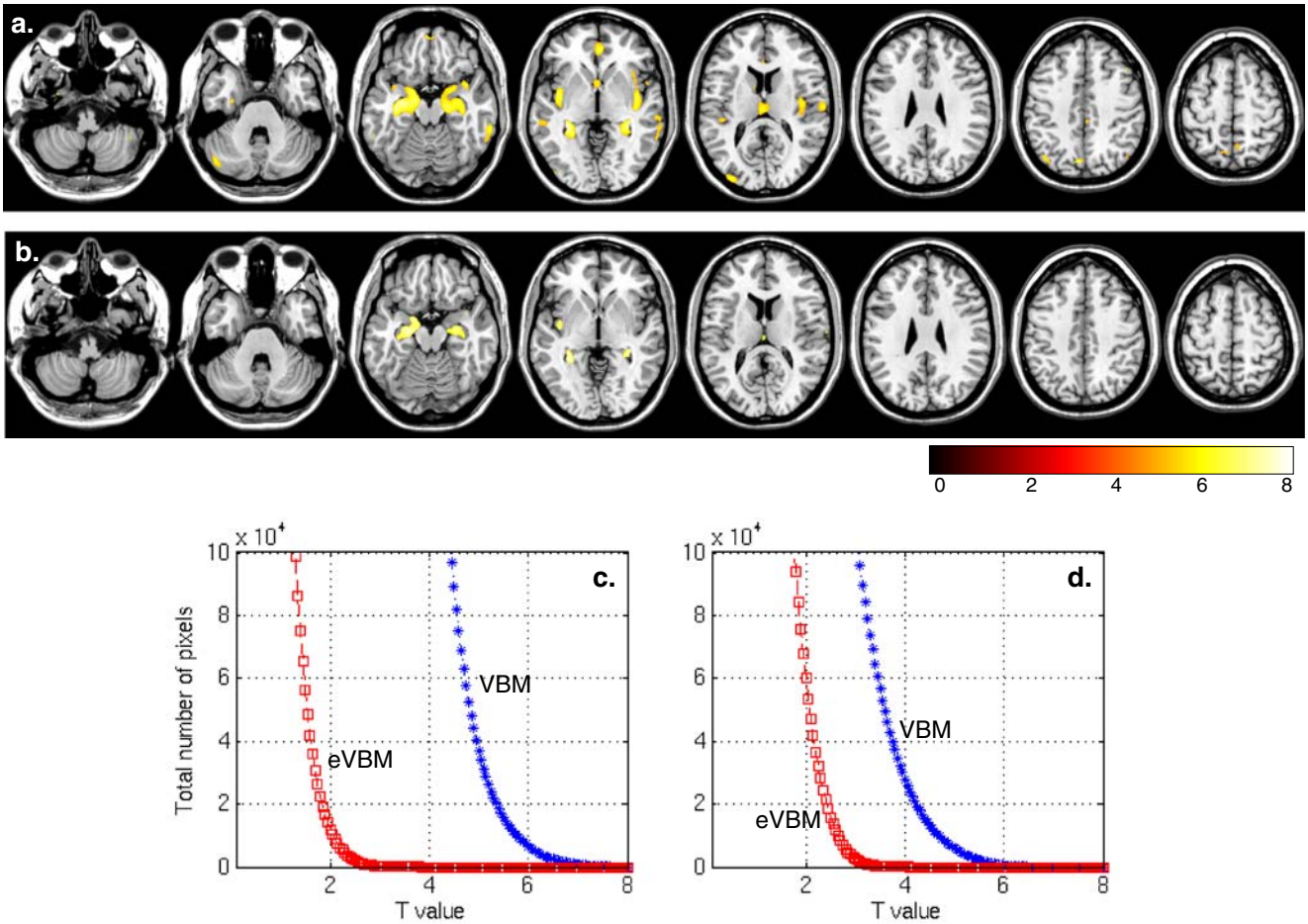

Fig. 5 Comparison of SPM5-VBM5.1 and SPM5-eVBM5.1 methods. **a** *T*-value maps obtained using the conventional SPM5-VBM5.1 method. **b** *T*-value maps obtained using the SPM5-eVBM5.1 method. *Colored regions* show where the GM probability was significantly higher in controls than in patients ($p<0.05$, FWE corrected threshold). **c** and **d** Total number of voxels above the threshold plotted against the *T* value. The *star-dashed curve* denotes the conventional VBM results, and the *square-dashed curve* represents the results of the eVBM method. **c** The results of the comparison between controls and AD patients; **d** the results of the comparison between the first group of controls and the second group of controls

FSL-VBM and SPM5-VBM5.1 results

Gray matter loss in AD was studied extensively [9, 13, 15–18, 34–38]. Generally, these findings involved GM loss in entorhinal cortices (EC), frontal lobe, precuneus, insula, caudate nucleus, and temporal lobe.

We conducted a FSL-VBM analysis to compare AD patients with age-matched controls, and the results were in agreement with previous studies in general [11, 18, 34, 35]. Using FSL-eVBM, we detected regions that were significantly different between controls and AD patients, these regions including EC and temporal lobes. Parts of these regions were found by previous studies [15, 16, 38]; these regions include cingulum, caudate nucleus, and thalamus [12, 17]. Our eVBM results are more in agreement with studies [18, 34, 35]. Unlike a previous study [15], we did not find any difference in a large part of the uncus cortex. While we found differences in caudate nucleus and thalamus as previous studies [15, 16], we also found differences in occipital cortex of AD patients with the eVBM method, as in studies [9, 13, 17]. The eVBM approach also proved to be able to detect regions in parahippocampus and hippocampus cortices by SPM5-VBM5.1.

In addition, Figs. 3d and 5d show that there were more voxels above a given significant threshold for FSL-VBM than for eVBM. This trend became more obvious (Figs. 3c and 5c) when comparing controls with AD patients. The differences in GM detected by using eVBM method had lower *T* values compared with conventional VBM, suggesting that this method reduces false-positive voxels in the results. This could be due to the gray matter differences between controls and AD patients; this could also be caused by the considerable variability in the native image histograms in controls and AD patients. Enhanced VBM is a method to overcome the limitation of histogram variability in the pre-processing step of conventional

VBM. The results show that the image histogram matching method should be used for the VBM analysis to obtain conservative results because it can reduce histogram variability effects that may enlarge the difference when comparing two or more groups.

The two methods gave the same conclusion, namely that eVBM yields less false-positive results, as shown in Figs. 3d and 5d. Furthermore, both FSL and SPM5 detected that the GM concentration in the hippocampus, the parahippocampus, and the temporal cortices in AD patients was significantly different from that in controls.

It should be noted that FSL-VBM and SPM5-VBM5.1 are different from the very beginning (see the block diagrams in Fig. 2). For instance the algorithms for segmentation and registration differ: FSL uses a Markov random field [26] method to perform the segmentation, while SPM5 uses a unified segmentation [24] algorithm that involves Gaussian mixtures. The final statistical procedures are also different: permutation tests [32] in FSL-VBM vs. a two-sample *t* test in SPM5-VBM5.1. Nevertheless, the significant different regions were similar if both methods employed a two-sample *t* test for comparing controls with AD patients (compare Fig. 3a with Fig. 4a and Fig. 3b with Fig. 4b, respectively). This suggests that the differences in pre-processing between the two conventional VBM methods, or between the two eVBM methods, have less influence on the results than the differences in statistical procedure for group comparison.

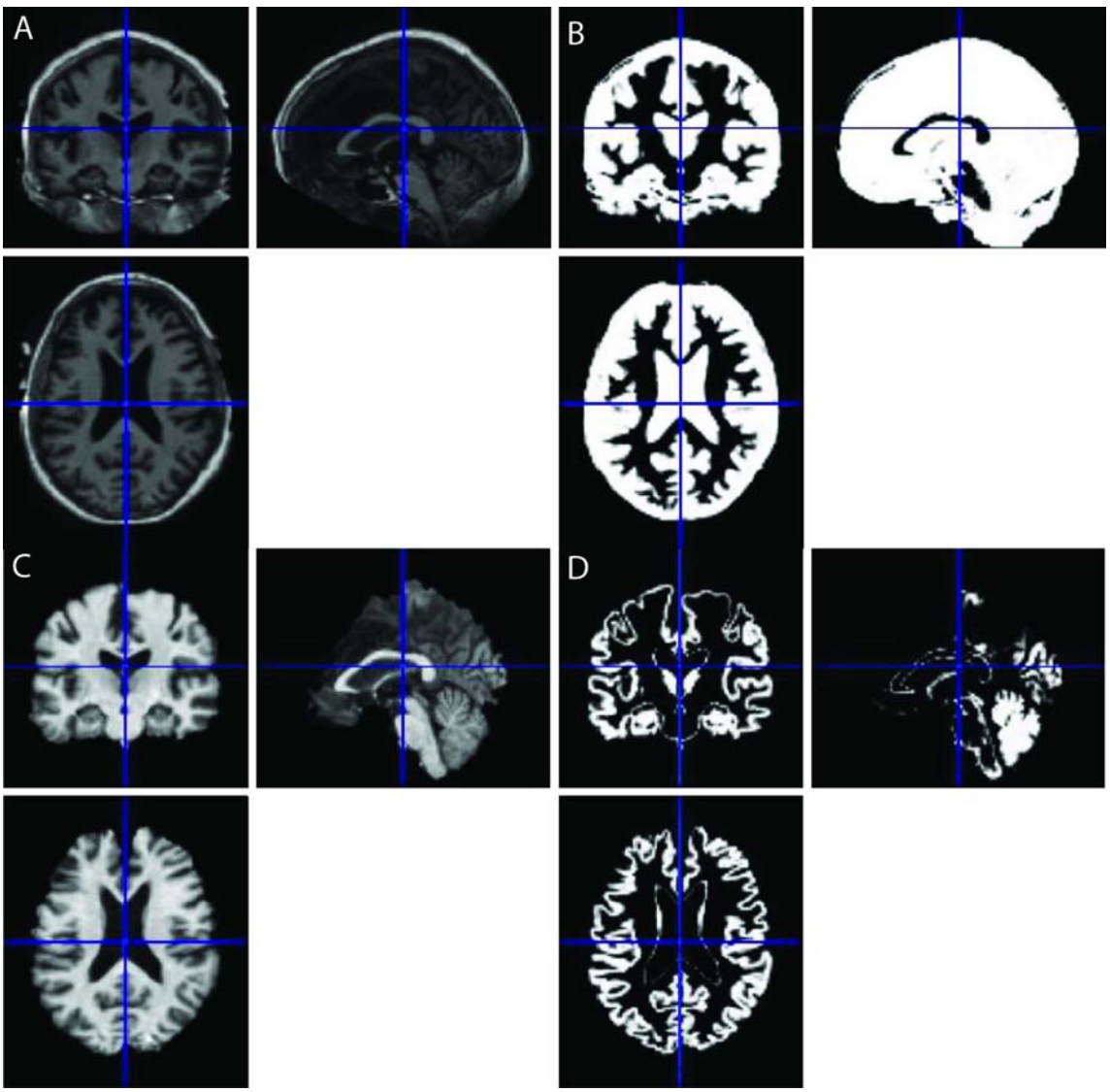

Fig. 6 SPM5-VBM5.1 GM segmentation results for one AD patient. **a** The original image; **b** the corresponding GM segmented image, which shows that the segmentation procedure failed since GM and CSF are mixed in a single image. **c** and **d** Segmentation results after histogram matching for the same subject (using the same parameters in SPM5 as in **a** and **b**). **c** The brain extracted and histogram matched image; **d** the corresponding GM segmented image, which is now correct

Histogram matching improves tissue segmentation

In addition to reduce the unequalization effect, the histogram matching operation enhances image contrast and consequently improves the performance of the subsequent processing steps such as tissue segmentation. We found that the suggested histogram matching method can also improve gray matter segmentation at least with SPM5. Without histogram matching, 13 (seven controls and six AD patients) out of the 148 subjects that were correctly registered to the Talairach space could not be segmented correctly by the default Gaussian kernel [2 2 2 4]. We then set the Gaussian kernel to [1 1 1 4], [3 3 3 4], [4 4 4 4], and [6 6 6 4], and the results were still not satisfactory. However, when histogram matching was performed before segmentation, all these subjects could be segmented with the default Gaussian kernel. This could be due to the fact that SPM5 uses a unified segmentation algorithm with a Gaussian mixture model to segment the GM, WM, and corticospinal fluid (CSF) [24]. After histogram matching, the distributions of GM, WM, and CSF were more similar to a Gaussian mixture, thus making the segmentation easier. For example, Fig. 6 illustrates the segmentation results with or without histogram matching on one subject. The results show that histogram matching can improve the segmentation in SPM5-VBM5.1 software package.

Advantages and limitations of the proposed method

The suggested eVBM approach has at least two advantages. Firstly, this method is easy to implement and apply to sMRI data. In addition, eVBM is not sensitive to the image histogram difference, this difference including both within- and between-group differences. It overcomes the within-group difference, which could cause false-positive results. Histogram matching can be regarded as one method of pre-processing in VBM analysis. In that respect our results show that pre-processing can have a major impact on the VBM results, in agreement with a recent VBM study [39].

One limitation is that this method needs a histogram distribution template for the matching although it is easy to obtain a mixture Gaussian distribution as shown in Fig. 1 c.

On the other hand, while our results show an increase in specificity, i.e., a decrease in false-positives voxels, they give no information about a possible decrease in sensitivity, i.e., about a possible increase in false-negative voxels. However, there is no direct way to address the issue of sensitivity since no ground truth is available. Some information can be obtained if the same subjects are scanned several times so that different datasets can be compared, as suggested in [11]. The subjects from the OASIS database were scanned three to four times each, but we chose to use their average anatomical image to maximize signal-to-noise ratio in the context of this study. We plan in the near future to conduct similar experiments using the initial structural images of the subjects, which would allow us to better study sensitivity aspects and further assess the robustness of the eVBM results.

Conclusion

We proposed an eVBM method to analyze structural MR images. We applied this method to study GM loss in AD patients compared with age-matched healthy elderly controls, by using conventional VBM using two different softwares widely used in the community. We found that eVBM had the advantage of reducing the number of false-positive results, suggesting that this method may be a better approach than conventional VBM to analyze sMRI.

Acknowledgments X. Li is funded by a "poste vert" from Inserm. This work was supported by the International Laboratory on Neuroimaging and Modeling (Inserm—UPMC Univ Paris 06—Université de Montréal). The authors thank Pr. S. Lehéricy and Dr. V. Perlbarg for their helpful comments.

The authors thank Dr. Randy Buckner and his colleagues for making their OASIS data available to us.

Conflict of interest statement We declare that we have no conflict of interest.

References

1. Wright IC, McGuire PK, Poline JB, Travere JM, Murray RM, Frith CD, Frackowiak RS, Friston KJ (1995) A voxel-based method for the statistical analysis of gray and white matter density applied to schizophrenia. Neuroimage 2:244–252
2. Ashburner J, Friston KJ (2000) Voxel-based morphometry—the methods. Neuroimage 11:805–821
3. Ashburner J, Friston KJ (2001) Why voxel-based morphometry should be used. Neuroimage 14:1238–1243
4. Bookstein FL (2001) "Voxel-based morphometry" should not be used with imperfectly registered images. Neuroimage 14:1454–1462
5. Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, Frackowiak RS (2001) A voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage 14:21–36
6. Mechelli A, Price CJ, Friston KJ, Ashburner J (2005) Voxel-based morphometry of the human brain: methods and applications. Current Medical Imag Reviews 1:1–9
7. Celone K, Calhoun V, Dickerson B, Atri A, Chua EF, Miller SL, DePeau K, Rentz DM, Selkoe DJ, Blacker D, Albert MS, Sperling RA (2006) Alterations in memory networks in mild cognitive impairment and Alzheimer's disease: an independent component analysis. J Neurosci 26:10222–10231
8. Baron JC, Chetelat G, Desgranges B, Perchey G, Landeau B, de la Sayette V, Eustache F (2001) In vivo mapping of gray matter loss with voxel-based morphometry in mild Alzheimer's disease. Neuroimage 14:298–309
9. Karas GB, Burton EJ, Rombouts SARB, Schijndel RAV, O'Brien JT, Scheltens PH, McKeith IG, Williams D, Ballard C, Barkhof F (2003) A comprehensive study of grey matter loss in patients with

Alzheimer's disease using optimized voxel-based morphometry. Neuroimage 18:895–907
10. Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage 44(1):83–98
11. Tardif CL, Collins DL, Pike GB (2009) Sensitivity of voxel-based morphometry analysis to choice of imaging protocol at 3 T. Neuroimage 44:827–838
12. Scahill R, Schott J, Stevens J, Rossor M, Fox N (2002) Mapping the evolution of regional atrophy in Alzheimer's disease: unbiased analysis of fluid-registered serial MRI. Proc Natl Acad Sci USA 99:4703–4709
13. Ramani A, Jensen J, Helpern J (2006) Quantitative MR imaging in Alzheimer disease. Radiology 241:26–44
14. Chetelat G, Degranges B, Sayette VDL, Viader F, Eustache BJC (2002) Mapping grey matter loss with voxel-based morphometry in mild cognitive impairment. NeuroReport 13:1939–1943
15. Frisoni GB, Testa C, Zorzan A, Sabattoli F, Beltramello A, Soininen H, Laakso MP (2002) Detection of grey matter loss in mild Alzheimer's disease with voxel based morphometry. J Neurol Neurosurg Psychiatry 73:657–664
16. Thompson P, Hayashi K, Zubicaray GD, Janke AL, Rose SE, Semple J, Herman D, Hong MS, Dittmer SS, Doddrell DM, Toga AW (2003) Dynamics of grey matter loss in Alzheimer's disease. J Neurosci 23:994–1005
17. Hirata Y, Matsuda H, Nemoto K, Ohnishi T, Hirao K, Yamashita F, Asada T, Iwabuchi S, Samejima H (2005) Voxel-based morphometry to discriminate early Alzheimer's disease from controls. Neurosci Lett 382:269–274
18. Paola MD, Macaluso E, Carlesimo GA, Tomaiuolo F, Worsley KJ, Fadda L, Caltagirone C (2007) Episodic memory impairment in patients with Alzheimer's disease is correlated with entorhinal cortex atrophy: a voxel based morphometry study. J Neurol 254:774–781
19. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy R, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23(S1):208–219
20. Gonzales RC, Woods RE (1993) Digital image processing. Addison-Wesley, USA, pp 173–182
21. Marcus DS, Wang TH, Parker J, Csernansky JG, Morris JC, Buckner RL (2007) Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. J Cogn Neurosci 19:1498–1507
22. Morris JC (1993) The clinical dementia rating (CDR): current version and scoring rules. Neurology 43:2412b–2414b
23. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labelling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15:273–289
24. Ashburner J, Friston KJ (2005) Unified segmentation. Neuroimage 26:839–851
25. Smith SM (2002) Fast robust automated brain extraction. Hum Brain Mapp 17:143–155
26. Zhang Y, Brady M, Smith SM (2001) Segmentation of brain MR images through a hidden Markov random field model and the expectation maximization algorithm. IEEE Trans Med Imag 20:45–57
27. Jenkinson M, Smith SM (2001) A global optimisation method for robust affine registration of brain images. Med Image Anal 5:143–156
28. Jenkinson M, Bannister PR, Brady JM, Smith SM (2002) Improved optimisation for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17:825–841
29. Andersson JLR, Jenkinson M, Smith SM (2007) Non-linear optimisation. FMRIB technical report TR07JA1 from www.fmrib.ox.ac.uk/analysis/techrep.
30. Andersson JLR, Jenkinson M, Smith SM (2007) Non-linear registration, aka spatial normalisation. FMRIB technical report TR07JA2 from www.fmrib.ox.ac.uk/analysis/techrep.
31. Rueckert D, Sonoda LI, Hayes C, Hill DLG, Leach MO, Hawkes DJ (1999) Non-rigid registration using free-form deformations: application to breast MR images. IEEE Trans Med Imag 18:712–721
32. Nichols TE, Holmes AP (2002) Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum Brain Mapp 15:1–25
33. Nichols TE, Hayasak S (2003) Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat Methods Med Res 12:419–446
34. Dickerson BC, Goncharova I, Sullivan MP, Forchetti C, Wilson RS, Bennett DA, Beckettl A, DetoleDO ML (2001) MRI-derived entorhinal and hippocampus atrophy in incipient and every mild Alzheimer's disease. Neurobiol Aging 22:747–754
35. Cardenas VA, Du AT, Hardin D, Ezekiel F, Weber P, Jagust WJ, Chui HC, Schuff N, Weiner MW (2003) Comparison of methods for measuring longitudinal brain change in cognitive impairment and dementia. Neurobiol Aging 24:537–544
36. Delbeuck X, Linden VD, Collette F (2003) Alzheimer's disease as a disconnection syndrome? Neuropsychol Rev 13:79–92
37. Rémya F, Mirrasheda F, Campbellb B, Richter W (2005) Verbal episodic memory impairment in Alzheimer's disease: a combined structural and functional MRI study. Neuroimage 25:256–266
38. Hämäläinen A, Tervo S, Grau-Olivares M, Niskanen E, Pennanen C, Huuskonen J, Kivipelto M, Hànninen T, Tapiola M, Vanhanen M, Hallikainen M, Helkala EL, Nissinen A, Vanninen R, Soininen H (2007) Voxel-based morphometry to detect brain atrophy in progressive mild cognitive impairment. Neuroimage 37:1122–1131
39. Acosta-Cabronero J, Williams GB, Pereira JMS, Pengas G, Nestor PJ (2008) The impact of skull-stripping and radio-frequency bias correction on grey-matter segmentation for voxel-based morphometry. Neuroimage 39:1654–1665