◆ Human Brain Mapping 16:228–250(2002) ◆

Comparison of Spatial Normalization Procedures and Their Impact on Functional Maps

Fabrice Crivello,[1]* Thorsten Schormann,[2] Nathalie Tzourio-Mazoyer,[1] Per E. Roland,[5] Karl Zilles,[3,4] and Bernard M. Mazoyer[1,6]

[1]*Groupe d'Imagerie Neurofonctionnelle, UMR 6095, CNRS-CEA LRC36V Université de Caen & Paris 5, Caen, France*
[2]*Institute of Neuroanatomy, Heinrich Heine University, Düsseldorf, Germany*
[3]*Cecil and Oscar Vogt Institute of Brain Research, Heinrich Heine University, Düsseldorf, Germany*
[4]*Institute of Medicine, Research Centre, Juelich, Germany*
[5]*Departement of Neuroscience, Division of Human Brain Research, The Karolinska Institute, Stockholm, Sweden*
[6]*Unité IRM CHU de Caen and Institut Universitaire de France, Caen, France*

Abstract: The alignment accuracy and impact on functional maps of four spatial normalization procedures have been compared using a set of high resolution brain MRIs and functional PET volumes acquired in 20 subjects. Simple affine (AFF), fifth order polynomial warp (WRP), discrete cosine basis functions (SPM), and a movement model based on full multi grid (FMG) approaches were applied on the same dataset for warping individual volumes onto the Human Brain Atlas (HBA) template. Intersubject averaged structural volumes and tissue probability maps were compared across normalization methods and to the standard brain. Thanks to the large number of degrees of freedom of the technique, FMG was found to provide enhanced alignment accuracy as compared to the other three methods, both for the grey and white matter tissues; WRP and SPM exhibited very similar performances whereas AFF had the lowest registration accuracy. SPM, however, was found to perform better than the other methods for the intra-cerebral cerebrospinal fluid (mainly in the ventricular compartments). Limited differences in terms of activation morphology and detection sensitivity were found between low resolution functional maps (FWHM ~10 mm) spatially normalized with the four methods, which overlapped in 42.8% of the total activation volume. These findings suggest that the functional variability is much larger than the anatomical one and that precise alignment of anatomical features has low influence on the resulting intersubject functional maps. When increasing the spatial resolution to approximately 6 mm, however, differences in localization of activated areas appear as a consequence of the different spatial normalization procedure used, restricting the overlap of the normalized activated volumes to only 6.2%. *Hum. Brain Mapping 16: 228–250, 2002.* © 2002 Wiley-Liss, Inc.

Key words: neuroanatomy; spatial normalization; registration; linear and nonlinear warping; probability maps; brain atlas

Contract grant sponsor: European Union Biotechnology; Contract grant number: BT04-CT96-0177.

This work was presented in part at the fifth International Conference on Functional Mapping of the Human Brain, June 1999.

*Correspondence to: Fabrice Crivello, PhD, Groupe d'Imagerie Neurofonctionnelle, UMR 6095, GIP Cyceron, Boulevard Henri Becquerel, BP 5229, 14074 Caen Cedex, France.
E-mail: crivello@cyceron.fr

Received for publication 17 April 2001; accepted 27 March 2002
DOI 10.1002/hbm.10047

INTRODUCTION

Spatial normalization is a necessary and crucial step in the analysis of both structural and functional neuroimaging data because it allows us to bring brain volumes acquired in different individuals in a common neuroanatomical reference space. In this reference space, image statistics may be computed over a sample of subjects, and anatomical labeling performed in a standardized manner across subjects, studies, and laboratories. Given the large variability between individuals in their brain anatomy, both at the microscopic and macroscopic levels, considerable efforts have been devoted over the past years to the development of more and more refined spatial normalization procedures. Several such procedures are now available that perform brain warping with various degrees of accuracy. The original Talairach linear transformation [Fox et al., 1985] translates, reorients, and normalizes the dimensions of a brain volume to fit it in the so-called Talairach box [Talairach and Tournoux, 1988]. This simple linear procedure has been superseded by nonlinear ones that aim at a more sophisticated matching of the structural characteristics of a given brain volume to a target one, using nonlinear deformations based on high-order polynomials [Woods et al., 1997a] or discrete cosine basis functions [Ashburner and Friston, 1999]. These nonlinear procedures, although they may lead to considerable deformations, do not, however, change the topology of the brain volume to which they are applied. More recently, a new class of spatial normalization procedures of a new kind have appeared that aim at realizing a full warp of a brain volume on another one [Christensen et al., 1994; Schormann and Zilles, 1998], or contour/surface-based alignment that cannot fulfill the purpose of a full warp because a lot of grey-value information is not taken into account [Davatzikos et al., 1996; Fischl et al., 1999; Thompson et al., 2000a, 2001]. These are designed to perform high precision cortical matching between different brains. As a consequence, these procedures have the power to modify the topology of the original brain volume.

Considering the conceptual differences between these spatial normalization procedures, the issue of which procedure should be chosen to analyze a neuroimaging dataset is critical, as it was for the European Computerized Human Brain Database (ECHBD) [Fredriksson et al., 2000; Roland and Zilles, 1996], a consortium of European laboratories that joined efforts to gather cytoarchitectonic, MRI, and functional maps to study relationships between microscopically and macroscopically defined functional areas. Although the answer to this question depends on many parameters, including the primary goal of the study, it also relies partly on the relative merits of each procedure. Interestingly, studies comparing normalization procedures are rather scarce in literature [Fischl et al., 1999; Gee et al., 1997; Grachev et al., 1999; Minoshima et al., 1994; Roland et al., 1997; Senda et al., 1998; Sugiura et al., 1999] and have so far dealt with the dispersion metric of selected landmarks, spatial homogeneity of selected anatomical features such as major sulci, or with overlap percentage of restricted volumes of interest. This is probably due to a natural feeling that using a high degree of freedom normalization approach will obviously surpass the alignment accuracy of lowest degree of freedom approaches. If such differences exist, however, it is important to try to quantify them, particularly in the growing field of statistical neuroanatomy. In the present study, conducted as part of the ECHBD project, we performed a cross-comparison of four published normalization procedures, using 3D tissue probability maps derived from high resolution brain MRI datasets acquired in 20 subjects and the Human Brain Atlas as a standard brain [Roland et al., 1994], and inter-subject averaged functional maps derived from the four normalization methods.

MATERIALS AND METHODS

Subjects

High-resolution MRI brain volumes were collected in 20 young healthy volunteers (age: 22.0 $\pm$ 2.4, mean $\pm$ SD, range: [19,28]) as part of the PET brain activation study they participated in. All subjects were free from cerebral abnormality as assessed on their brain T1-weighted MRI images. The protocol was accepted by our local ethic committee and all subjects gave their informed written consent.

MRI data

MRI acquisitions were performed on a 1.5 T Signa Horizon LX scanner (General Electric, BUC, France). High-resolution T1-weighted MRI volumes were acquired using an inversion recovery SPGR-3D sequence (TR = 12.4 msec, TE = 2.2 msec, FA = 10°, TI = 600 msec) sampling the whole subject head in 0.94 $\times$ 0.94 $\times$ 1.5 mm^3 voxels. The image volume was reconstructed as a 256 $\times$ 256 $\times$ 124 matrix.

PET data

Using a standard ^{15}O-labeled water PET activation protocol, nine sequential measurements of the regional cerebral blood flow were obtained from each subject on an ECAT Exact HR + PET camera and reconstructed as a series of 63 contiguous, 2.5-mm-thick cross-sectional images, replicating three times a series of three experimental conditions presented in a randomized order: 1) listening to a factual story (Story); 2) covert generation of the most verbs semantically related to listening nouns (Verb); and 3) a rest condition (Rest) that consisted in lying silently in the dark eyes closed, with no particular instruction except to refrain from moving and avoid a specific mental activity [Papathanassiou et al., 2000]. All conditions were performed in darkness. During the Verb condition, the nouns were delivered through earphones at 0.1 Hz and were of common use, most of them referring to objects and some of them to animals. The subjects were instructed to covertly retrieve as many verbs semantically related to those nouns as possible. Three different noun lists were used for each replication. Only the Verb and Rest conditions were used in this study to compute the Verb *minus* Rest contrast in the functional comparison part of the study.

Standard anatomical format

The standard anatomical format used in this study was the MRI standard brain of the Human Brain Atlas [Roland et al., 1994] and of the European Computerized Human Brain Database [Fredriksson et al., 2000; Roland and Zilles, 1996]. This standard brain is a single young male brain high-resolution T1-weighted MRI volume, having isotropic millimeter resolution, and a $137 \times 181 \times 146$ size. This volume was selected from a sample of 21 male brain MRI, as being the most "representative" in terms of shape, size, and gyral as well as sulcal pattern. This template was scalp edited, oriented and scaled according to the stereotaxic coordinate frame defined by the Talairach atlas [Talairach and Tournoux, 1988].

Intersubject spatial normalization registration

Before any spatial transformation, each subject MRI scan was semi-automatically segmented to remove the extra cranial structures (i.e., scalp, skull, and meninges) using our in-house ATOMIA software [Verard et al., 1997].

The 20 MRI volumes were then spatially normalized using the following four 3D procedures:

1. The classical 12 parameters affine transformation (AFF) as performed by the "alignlinear" routine included in the Automated Image Registration package (AIR 3.0) [Woods et al., 1997b].
2. The fifth order, 168 parameters nonlinear polynomial warp algorithm (WRP), also included in the AIR 3.0 package [Woods et al., 1997b].
3. The "linear & nonlinear" combined procedure implemented in the 1996 version of the Statistical Parametric Mapping software (SPM) [Ashburner and Friston, 1999].
4. A recently proposed multi grid technique based on Navier-Lamé continuum mechanics theory (FMG) [Schormann and Zilles, 1998], first implemented to co-register histological and MRI sections [Schormann et al., 1996].

These four approaches are voxel intensity-based methods, meaning that they are not driven by shape or any model constraint matching, and are all fully automated, thereby ensuring an unbiased comparison of their relative performances. It must be stressed that the four methods were used in their original "default setting," and only options essential for a correct utilization of the algorithms with respect to the characteristics of our MRI datasets were modified. All normalized volumes were resliced using tri-linear interpolation and a $1 \times 1 \times 1$ mm^3 voxels size. Our method's modifiable options are detailed in the four followings paragraphs.

AFF

This procedure was the only linear transformation used in this comparison. The 12 parameters models is a superset of the simplest nine parameters registration method originally defined by Talairach. In addition to the six parameter rigid body rotation and translation, this model allows independent rescaling parameters along three arbitrary chosen coordinate axes of the reference volume. This procedure does not preserve angles and distances but preserves the parallelism of lines after the transformation has been applied. The criterion used for aligning two images is referred to as the ratio image uniformity (RIU) cost function, first introduced by Woods et al. [1992] for intra-subject and intra-modality alignment. The image is divided by the target image on a voxel-by-voxel basis to create a ratio image. The uniformity of this ratio image is then measured by computing the ratio of its standard deviation divided by its mean, thereby providing a normalized cost function value. The minimization of this criterion allows the two images to be registered. The "Initial-sampling" parameter was set to 81, controlling how densely data was sampled during the first iterative

cycle of the algorithm. The "Final-sampling" parameter was set to 1 meaning that the alignment was achieved at the best resolution. Twenty-five iterations were used for each iterative cycle. Final convergence threshold was set to 10^{-5}.

WRP

The initialization step of the complete WRP normalization procedure consists in applying first the AFF method described above so that the linearly aligned image is the starting step of the WRP algorithm. This nonlinear approach proposed by the same authors has been established so as to estimate the model parameters of the spatial transformation in a rapid and reproducible manner. This model can be seen as a natural evolution of the classical linear model, bringing it to higher order polynomial warping models. The fifth order polynomial warp algorithm used in the present study has 168 spatial transformation parameters. The cost function to be minimized by the WRP algorithm is the average across voxels of the squared difference between the resliced and the reference images. An additional parameter is included in the cost function to account for global intensity rescaling of the images relative to one another. Initial-sampling was set to 81 and the final one was set to 9. Fifty iterations were used for each iterative cycle. Final convergence threshold was set to 0.5.

SPM

This approach works by minimizing the average across voxels of the squared difference between the volume to be normalized and a template volume. For the least-squares registration to produce an unbiased estimate of the spatial transformation, the intensity contrast of both images should be similar. The registration simply searches for an optimum solution. The first step of the normalization finds the optimum 12 parameters affine transformation by appropriate weighting of the template voxels. Although this approach does not require scalp editing that discounts the confounding effects of skull and scalp differences, we have used it on the edited reference brain described above, smoothed with an 8 mm FWHM isotropic Gaussian kernel to investigate only the effect of the normalization algorithm (because this procedure begins with smoothing the images to be normalized by the same Gaussian kernel). The affine registration is followed by nonlinear deformations defined by linear combinations of $8 \times 8 \times 8$ 3D discrete cosine transform basis functions [Ashburner and Friston, 1999; Friston et al., 1995a]. This option results in each of the deformation fields being described by 1,536 parameters representing the coefficients of the deformations in three orthogonal directions (8^3 for each one). Sixteen iterations per cycle were used.

FMG

As for the WRP algorithm, this procedure also starts with the AFF linearly aligned image in our study. The nonlinear part appears as follows: the brain image to be normalized is modelled as an elastic medium by applying the principles of Navier-Lamé continuum mechanics theory [Schormann et al., 1998; Schormann and Zilles, 1998]. Then, the theory is extended to a movement model enabling the correction of large spatial differences. In the elastic model, the resulting deformation is only a compromise between the internal strain of the object and external forces. For this reason, only small deformation can be calculated. In this approach a 3D nonlinear deformation field directly results from the solution of a system of coupled partial differential equations describing for each voxel the unique movement (the spatial deformation) of the individual brain onto the reference volume. In this process, the total number of degrees of freedom in one direction (x axis) is equal to the total number of voxels in the volume (for example, normalizing a $137 \times 181 \times 146$ MRI volume generates a deformations field of 10,861,086 parameters). The movement of the source image onto the reference standard brain image is subject to some specific constraints imposed by the continuum theory of movement model. In the present approach, the investigation was focused on minimizing computation time by applying an adapted full multi grid technique (i.e., a multiresolution descending approach) [Schormann et al., 1996]. Information and support for the FMG software are available at thorsten@hirn.uni-dusseldorf.de. Note that in one case, there was a large grey-value difference between both volumes (the template and the volume to normalized), which made the FMG approach fail to converge, so that the total number of normalized MRI volumes was 20 for the AFF, WRP, and SPM procedures and only 19 for the FMG procedure.

Comparison of spatial normalization procedures

To quantify the relative merits of the four spatial normalization procedures previously described, we used the HBA template as a gold standard and computed several images and figures of merits. Note that brains were scalp edited before spatial normalization. This is important because the segmentation performed

in the subsequent analysis was intensity driven, and using scalp edited brains prevented us to assign, for example, meninges to grey matter. Accordingly, only voxels included into the intra-cerebral volume were tissue classified. The comparison of the four procedures was conducted as follows. First, we derived averaged MRI volumes across subjects and visually compared them to the HBA template. Then, we performed a tissue segmentation on both the template and normalized volumes. The degree of spatial overlap between the template and individual MRI volume was quantified for each tissue class (grey matter, white matter, and CSF) and for each procedure. Finally, we evaluated the impact of differences in spatial normalization procedure on functional maps, by applying these procedures on functional maps that were obtained from the same sample of subjects.

Intersubject averaged images

Realigned MRI images were averaged across subjects ($n = 19$ for the FMG procedure and $n = 20$ for the three other ones). For each procedure, the individual spatially normalized MRI volume was first scaled to its maximum value to normalized the intensity of the different MRI volumes. These volumes were then summed across subjects and the resulting voxel values divided by the occurrence of subjects contributing to the average at every location.

Tissue segmentation

The MRI template and individual normalized MRI volumes were tissue classified using the segmentation algorithm included in the SPM99 package. The SPM segmentation of each individual normalized MRI is based on a clustering algorithm identifying voxel intensities of particular tissue types (grey matter, white matters, and CSF), combined with a priori knowledge about the spatial distribution of these clusters in normal subjects [Ashburner and Friston, 1997]. The resulting tissue segmented volumes consist of maps in which a voxel value is the probability for this voxel to belong to the tissue class of interest, ranging from 0 (low probability) to 255 (high probability). Each tissue probability volume was segmented by applying a lower threshold set at 128 (50% probability). Voxels classified as CSF but located outside the outer edge of the cortex were removed using morphological operations (see upper part of Fig. 1). The number of such voxels is highly dependent on the parameters employed during the semi-automated scalp edition, which can vary between subjects, leading to slightly different CSF brain masks even after spatial normalization. In the same vein, voxels inside the ventricles or located in the deepness of sulci that were not classified as CSF, were also assigned to CSF. The lower part of Figure 1 shows that the segmentation procedure was efficient and provided an adequate tissue classification of the HBA template although this segmentation algorithm was originally devoted to segment MRI brain image matched to the MNI template. The proportions of each tissue class of the HBA template are presented in Table I and can be compared to that of the MNI template.

Due to high signal to noise ratio inhomogeneities, when applied to our MRI dataset, this segmentation procedure failed in one case, this was the case for the four normalization used, leading to unreasonable grey and white matter masks. Accordingly, the final analysis steps were applied to the 18 subjects for which both the normalization and segmentation procedures have been successfully completed. To globally quantify the relative merits of spatial normalization procedures at the tissue classification level, proportions of grey matter, white matter, and CSF composing the intracranial brain volume of each subject MRI were computed for each normalization procedure and their differences with corresponding figures for the HBA template averaged across subjects.

Tissue probability maps

To further document this point, we computed two sets of probability images of grey matter, white matter and CSF over the sample of 18 subjects. These probability maps indicate for each procedure and for each tissue class, the proportion of subjects overlapping at each voxel location. The first set of probability maps presents only voxels that were correctly classified by the segmentation procedure, taking the tissue classification performed on the HBA template as a gold standard. This means that for a specific tissue of interest, the boundaries of the "classified" probability maps are the ones of the corresponding HBA tissue template. The second set of probability maps presents the subject overlap corresponding to voxels that were wrongly classified (misclassified) with respect to the HBA tissue template.

The qualitative findings of these two set of maps were quantified using two 3D global criteria, C(i,j,k) and M(i,j,k) referring respectively to classified and misclassified voxels.

For each spatial normalization procedure (indexed by i), subject (indexed by j) and tissue class (indexed by k), we defined $N_{st}(i,j,k)$ as the total number of voxel of subject j belonging to tissue k of the template by

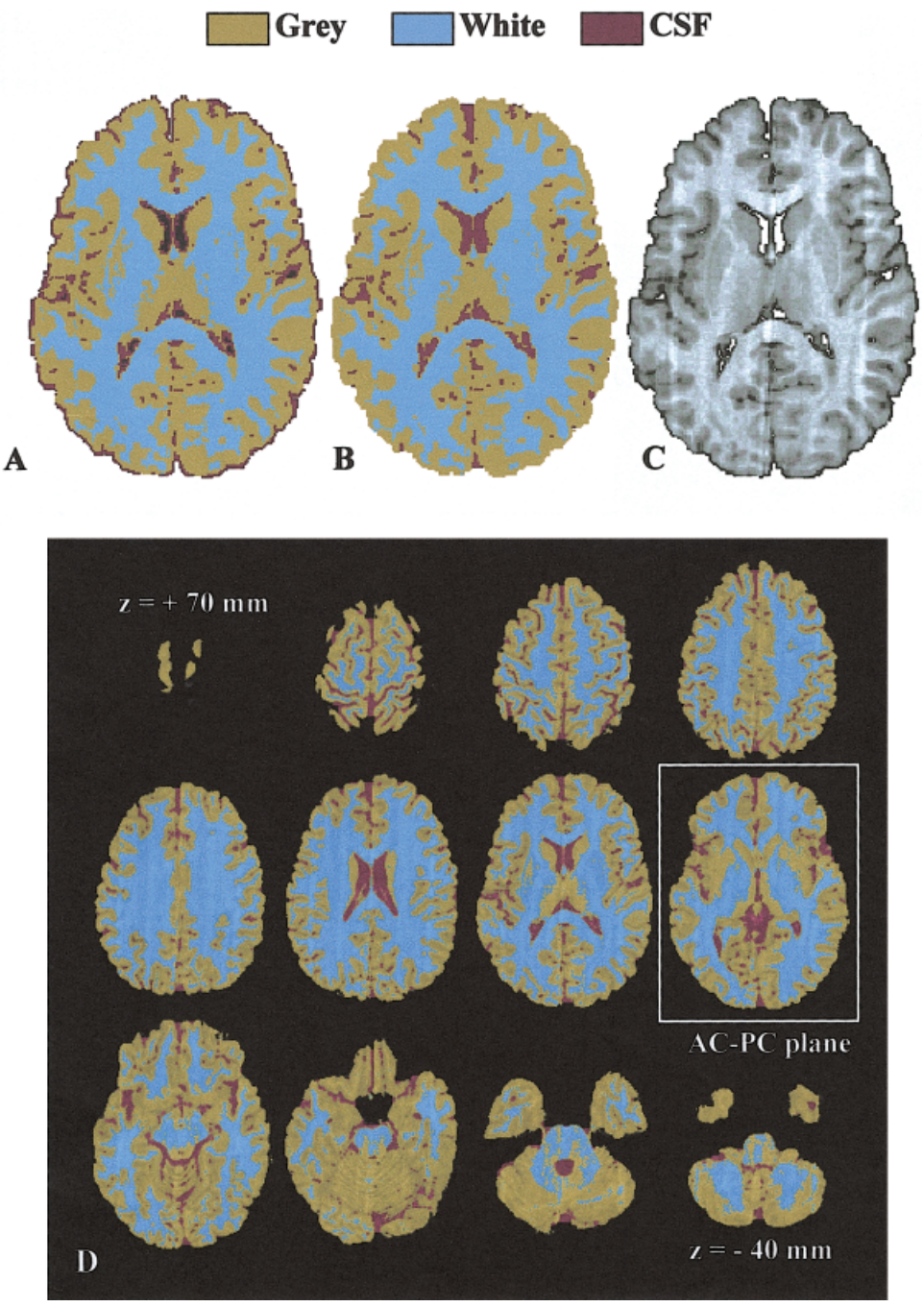

Figure 1.
A: Original axial section of the SPM segmentation applied to the HBA MRI template. **B:** Morphologically processed slice. This process removes the extra cerebral CSF tissue that remains after the semi-automated scalp edition and fills-in the ventricles and deepness of the sulci. **C:** Corresponding MRI slice. **D:** Transverse processed sections of the segmentation of the HBA MRI template from +70 to −40 mm relative to the AC–PC plane.

normalization procedure i, $N_s(i,j,k)$ as the total number of voxel of subject j belonging to tissue k by normalization procedure i and $N_t(k)$ as the total number of voxel of the HBA template for tissue k.

Classified voxels criterion

The percentage of voxels correctly classified by a normalized-segmented MRI volume with respect to the segmented template was computed as:

$$C(i,j,k) = \left(\frac{N_{st}(i,j,k)}{N_t(k)}\right) \times 100,$$

Misclassified voxels criterion

The percentage of voxels misclassified by a normalized-segmented MRI volume with respect to the segmented template was computed as:

TABLE I. Volumes and proportions of grey matter, white matter and CSF of the Human Brain Atlas MRI template as assessed by the Statistical Parametric Mapping segmentation algorithm*

Tissue	Volume (cm^3)	%
Grey matter	861.0 (1016.2)	59.5 (56.8)
White matter	506.7 (560.8)	35.0 (31.3)
CSF	78.6 (212.3)	5.5 (11.9)
Total	1446.3 (1789.3)	100 (100)

* Corresponding figures for the Montreal Neurological Institute MRI brain are given in parentheses.

$$M(i,j,k) = \left(\frac{N_s(i,j,k) - (N_{st}(i,j,k)}{N_s(i,j,k)}\right) \times 100,$$

$$M(i,j,k) = 100 - C(i,j,k)\left(\frac{N_t(k)}{N_s(i,j,k)}\right),$$

It must be noticed that the sum of C(i,j,k) and M(i,j,k) is not equal to 100% because the reference for the C(i,j,k) criterion is the number of voxels in the corresponding class of the template whereas the reference for M(i,j,k) is the number of voxels of the corresponding class of the subject.

Inter-individual criterion variability

The inter-individual variability of these criterion was also investigated. For each procedure, subject, and tissue class, the gain for C(i,j,k) and M(i,j,k) (in %) provided by FMG as compared to the other procedures were assessed by the two following metrics $G_C(i,j,k)$ and $G_M(i,j,k)$ computed as:

$$G_C(i,j,k) = \left(\frac{C(FMG,j,k) - C(i,j,k)}{C(i,j,k)}\right) \times 100,$$

$$G_M(i,j,k) = \left(\frac{M(FMG,j,k) - M(i,j,k)}{M(i,j,k)}\right) \times 100,$$

where C(FMG,j,k) (respectively M(FMG,j,k)) denotes the percentage of correctly classified (respectively misclassified) voxels for the FMG procedure. A positive value of $G_C(i,j,k)$ indicates that for a subject j and a tissue class k, the FMG procedure performs better than the procedure i (AFF or SPM or WRP). In contraxt, a positive value of $G_M(i,j,k)$ indicates that for a subject j and a tissue class k, the FMG procedure performs worse than the procedure i (AFF or SPM or WRP), meaning that the amount of misclassified voxels is higher with FMG.

Overall measures of accuracy based on correct and incorrect classification

To assess the relative classification errors of the four procedures tested, we computed, for each procedure (indexed by i) and subject (indexed by j), the probability of Type I and Type II errors for each tissue class (indexed by k). Suppose an image consist of N voxel classes (here N = 4, because we consider an additional class that contain voxels that were not classified neither in the grey matter, white matter or CSF template class), a confusion matrix A(i,j) of dimension N can be constructed, where each entry $A_{k'k}$ represents the number of class k voxels classified as class k′ by the segmentation algorithm. Two error types (expressed in %) can thus be computed for each tissue class k, which can both be used to described the class-by-class performances of these procedures.

The multi-class Type I error is compute as:

$$MC_I(i,j,k) = \left(\frac{\sum_{k'=1}^{N} Ak'k(i,j) - Akk(i,j)}{\sum_{k'=1}^{N} Ak'k(i,j)}\right) \times 100,$$

where the numerator represents the number of voxels of tissue class k not classified as k and the denominator is the total number of voxels of tissue class k [Zhang et al., 1996]. It must be noted that the Type I error is the dual percentage of the correctly classified voxel criterion defined above ($C(i,j,k) + MC_I(i,j,k) = 100\%$).

The multi-class Type II error is computed as:

$$MC_{II}(i,j,k) = \left(\frac{\sum_{k'=1}^{N} Akk'(i,j) - Akk(i,j)}{\sum_{k'=1}^{N}\sum_{k''=1}^{N} Ak'k''(i,j) - \sum_{k'=1}^{N} Ak'k(i,j)}\right) \times 100,$$

where the numerator represents the number of voxels of other classes called class k. The denominator is the total number of voxels of other classes [Zhang et al.,

1996]. In $MC_I(i,j,k)$ and $MC_{II}(i,j,k)$, each tissue class is weighted equally.

Impact of spatial normalization procedures on functional maps

Finally, to evaluate how spatial normalization affects functional maps, we applied the four procedures on the PET functional data of the 18 subjects for which both the normalization and segmentation procedures have been successfully completed. Using AIR3.0 [Woods et al., 1997a], each individual PET dataset was realigned together for correcting for intra-individual/intra-modality movements, and co-registered to the subject MRI volume for intra-individual intermodality alignment. For each subject, the MRI normalization parameters obtained for each of the four methods were applied to the corresponding PET images to spatially normalize them to the HBA template. Four sets of functional images per subject were thus available. All PET images were resliced using linear interpolation to provide $2 \times 2 \times 2$ mm^3 voxel size volumes. To investigate the cross-effect of smoothness and spatial normalization, we used two versions coming from the original set of functional images. First, the normalized PET images were smoothed with a 3D Gaussian filter of 8 mm full with half maximum (FWHM), leading to a "low resolution" functional dataset. Activation detection was performed on each of the four normalized blood flow difference volumes "Silent verb generation *minus* Rest" using the SPM99 software for the ensuing statistical analysis. Significance level was set at 0.05 corrected for multiple comparisons (degree of freedom = 140). Then, the original normalized PET volumes were re-smoothed with a 4 mm FWHM Gaussian filter to generate a "high resolution" functional dataset. The same statistical analysis was performed for the AFF, SPM, WRP, and FMG functional datasets. Resulting smoothness estimated on the functional maps at the lowest resolution with AFF, SPM, WRP, and FMG through the x, y, and z directions were [8.5 9.4 10.8], [8.6 9.5 10.9], [8.5 9.5 10.7], [8.9 9.5 10.8] mm, respectively. When looking at the highest resolution, these values fell to [5.7 6.4 7.4], [5.8 6.5 7.5], [5.6 6.5 7.3], [6.1 6.5 7.5] mm. Note that in both cases, the different normalization procedures led to very comparable global smoothness values.

RESULTS

The distribution of the global brain volume in the sample of 20 subjects, computed before applying any spatial normalization, was found very homogeneous (1,421 ± 98 cm^3, coefficient of variation = 6.9%) and its average value was not significantly different from that of the Human Brain Atlas template (1,446 cm^3, $P = 0.11$, Student's t-test, degree of freedom = 19), ensuring comparability between the individual MRI datasets and the HBA template.

Intersubject averaged images

Figure 2 shows axial, sagittal, and coronal sections derived from the mean volumes obtained with each procedure. The same sections are also presented for the template volume. The visual inspection of these slices clearly indicates that the three nonlinear approaches dramatically improve the alignment of the normalized volumes as compared to the simplest affine one. Looking at the three nonlinear procedures, the FMG approach improves the delineation of the different sulci and therefore the spatial homogeneity of the observed sample. A better definition of the brain outer edge is observed as well as of internal anatomical landmarks. This improvement is obvious particularly at the level of the superior frontal, precentral and central sulcus on the axial slice, or at the level of the post-central and central sulcus on the sagittal slice. These results demonstrate qualitatively that individual brain anatomy is globally better matched using nonlinear spatial deformation, the highest gain being apparently obtained with the FMG procedure.

Tissue segmentation

Table II presents, the differences (mean and standard deviation) between tissue class proportions obtained for each normalization procedure and the tissue class proportions of the HBA template. Note, first, that all procedures led to overestimation of the size of the grey matter class (in the order of 4%) and to underestimation of both the white matter (−1 to −2.5%) and CSF classes (−2 to −3%). The FMG procedure gave grey matter segmentation figures that were the closest to that of the HBA template. No significant difference for the grey matter tissue volume, however, was found between the four procedures ($P = 0.34$, ANOVA). In contrast, significant differences were found between procedures regarding the white matter and CSF tissue volumes ($P = 0.005$ and $P < 0.0001$, respectively, ANOVA, 17 d.o.f). Post-hoc Student's paired t-test showed that, for the white matter, WRP and FMG were significantly better than SPM and AFF ($P < 0.006$ and $P < 0.04$, respectively), with no differences

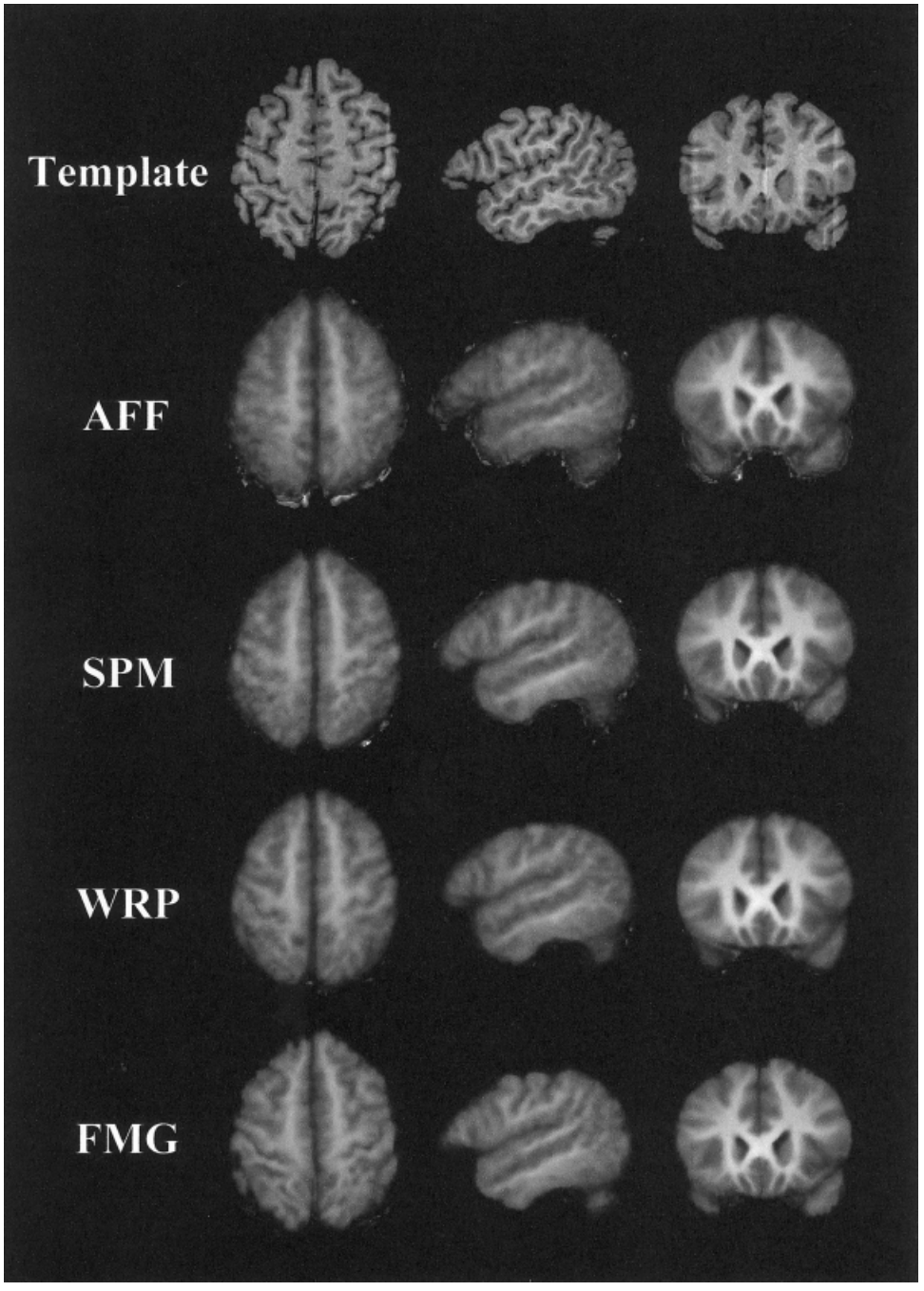

Figure 2.
From left to right: transverse (z = +50 mm), sagittal (x = −44 mm) and coronal (y = +20 mm) sections of the HBA MRI template (top row) and of the inter-subject average volumes computed for the four normalization procedures. Coordinates are relative to the Talairach space. AFF, affine transformation (n = 20); SPM, SPM linear and nonlinear routines of the SPM package (n = 20); WRP, affine and fifth order polynomial normalization of the AIR package (n = 20); FMG, affine and nonlinear fluid transformation (n = 19).

between WRP and FMG (P = 0.15). Note also that the FMG approach again provided the segmentation figures closest to that of the HBA template. No differences were observed between AFF and SPM (P = 0.44). Finally the CSF tissue type exhibited the opposite feature: AFF and SPM were significantly better than WRP and FMG ($P < 0.0001$ for both), with no differences between either AFF and SPM (P = 0.91) or WRP and FMG (P = 0.38).

Tissue probability maps

Figures 3–5 show, for each of the four normalization procedures, probability maps of the grey matter, white matter and CSF classes computed across the subset of 18 subjects. The "Classified voxels" maps give for each voxel and for each procedure, the percentage of subjects for which the tissue class was in accordance with the corresponding tissue

TABLE II. Difference between segmented tissue class proportions computed across the 18 subjects and the corresponding tissue class proportions of the Human Brain Atlas MRI template for the four spatial normalization procedures*

Procedure	Grey matter	White matter	CSF
AFF	4.49 (1.89)	−2.54 (1.81)	−1.96 (0.60)
SPM	4.38 (1.55)	−2.42 (1.48)	−1.96 (0.61)
WRP	4.58 (1.66)	−1.89 (1.46)	−2.69 (0.64)
FMG	3.88 (2.26)	−1.08 (1.95)	−2.79 (0.67)

* Values are mean (SD). AFF, affine transformation; SPM, linear and non-linear routine of the Statistical Parametric Mapping package; WRP, affine and fifth order polynomial normalization of the AIR package; FMG, affine and non-linear fluid transformation.

class of the HBA template. The "misclassified voxels" correspond to voxels found outside the boundaries of the tissue template. Note the better subject overlap obtained with FMG for both the grey and white matter tissues, and the very similar results provided by WRP and SPM. The spatial accuracy visible in the FMG maps indicates that this procedure aligns high-resolution MRI images from different subjects in such a way that internal structures, such as sulci in their deepness, are better aligned to the structural model of the template. For the CSF tissue class, FMG does not appear to be superior to either SPM or WRP, although the extent of overlap in the CSF "misclassified voxels" map seems to be dramatically reduced (see Fig. 5). Nevertheless, Figures 3–5 clearly demonstrate that any of the nonlinear approaches used in this study provides better results than the classical linear AFF procedure.

This is confirmed in a quantitative manner in Table III and Figure 6 for the classified voxels and in Table IV and Figure 7 for the misclassified ones.

Classified voxels

Table III gives, for each tissue class and normalization procedure, the mean and standard deviation of the classified voxels criterion (C(i,j,k), see the Materials and Methods section above for details). The FMG procedure provides the best overlap figures for both grey matter (close to 79%) and white matter (71%). Interestingly, the gain in overlap provided by FMG as compared to the other methods is more pronounced for the white matter class (4–7%) than for the grey matter class (1–5%). Significant differences between procedures were found for all tissue ($P < 0.0001$ for each tissue, ANOVA, 17 d.o.f.). Post-hoc Student's paired t-tests confirm the superiority of the three nonlinear procedures when compared to AFF ($P < 0.0001$ for each procedure) for both grey and white matter. In addition, FMG was found significantly better than WRP for grey matter overlap ($P = 0.025$), and better than both WRP and SPM for white matter overlap ($P = 0.0003$ and $P = 0.0002$, respectively). On the opposite, the best overlap for the CSF class was obtained with SPM ($P < 0.0001$ for all comparisons), the other three methods showing similar performances. Considering the individual variability of this criterion, Figure 6 clearly demonstrates that FMG provided a better alignment of tissue compartment than the three other methods in almost all subjects both for grey and white matter. Note that for these two tissues, SPM and WRP were remarkably similar and that AFF was clearly surpassed by the nonlinear approaches. These findings demonstrate also that the differences pointed in Table III were not driven by one or two subjects but remain valid at the individual level. The lower part of Figure 6 illustrates that SPM performed a better CSF alignment than all other methods for all subjects included in this comparison study.

Misclassified voxels

Table IV and Figure 7 present the average results for the misclassified voxels M(i,j,k) criterion (see the Materials and Methods section). For grey matter, white matter, and CSF, FMG provides the lowest error classification rate, respectively from (0.9–6%), (1–4.9%) and (2–13.3%). As for the classified voxels statistic, significant differences between procedures were observed for all tissues ($P < 0.0001$ for each one, ANOVA, 17 d.o.f.). Post-hoc Student's paired t-tests show the same pattern for the three tissue types. Each nonlinear procedure leads to a lower ratio of misclassified voxels than AFF ($P < 0.0001$ for each procedure). In addition, FMG and WRP were found superior to SPM for the grey matter ($P = 0.0005$ and $P = 0.012$, respectively), for the white matter ($P = 0.020$ and $P = 0.044$, respectively), and for the CSF ($P < 0.0001$ for both). Finally no significant differences were found between FMG and WRP for the grey matter, white matter and CSF ($P = 0.095$, $P = 0.093$, $P = 0.063$, respectively). The investigation of the inter-individual variability of this metric reveals that in almost all subjects, FMG exhibits the lowest rate of classification errors (negative gain for the misclassified), a feature emphasized in the CSF compartment. The individuals results confirm that average results are not driven by

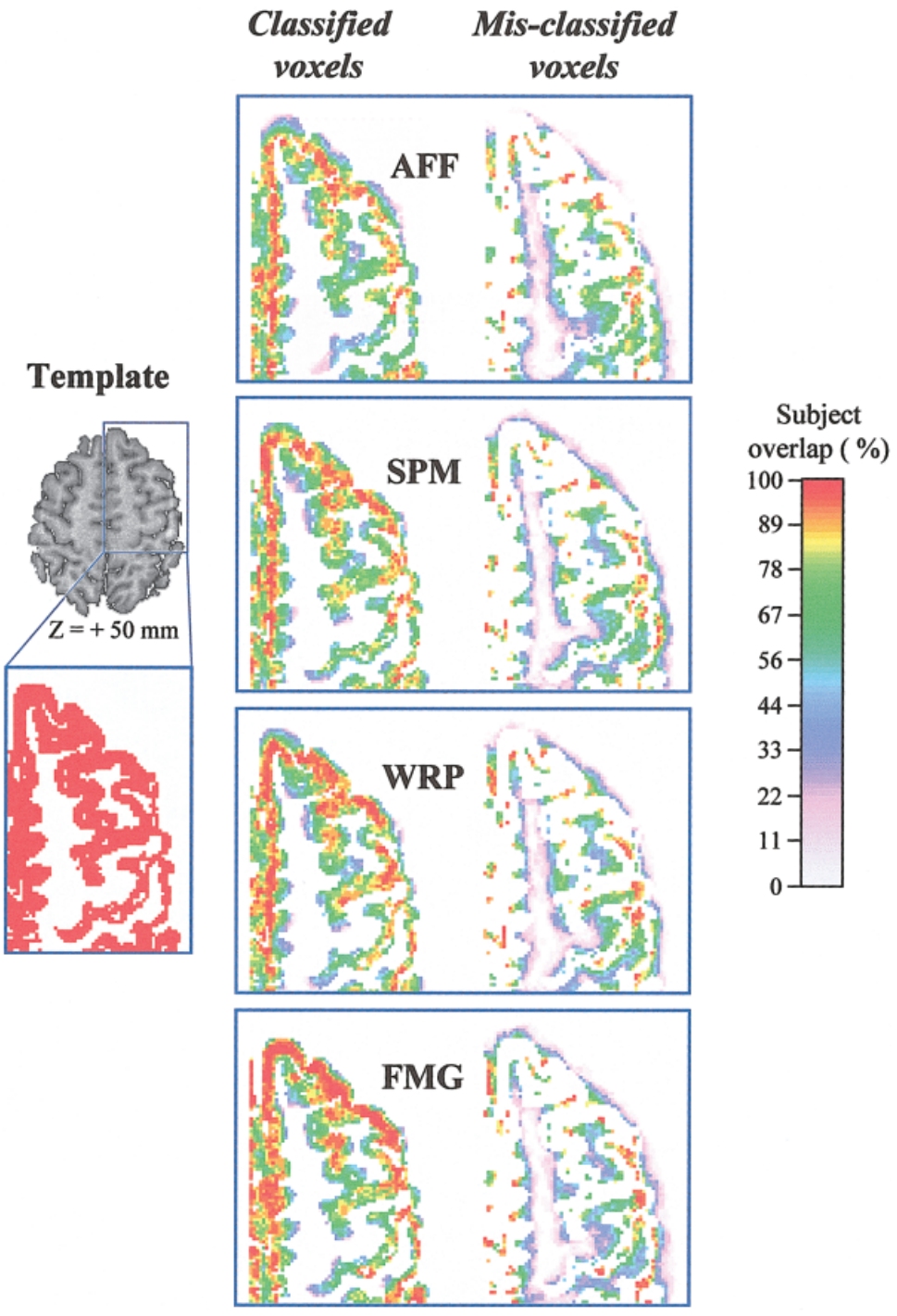

Figure 3.
Left column: Axial section (z = +50 mm) of the HBA MRI template and enlargement of the corresponding segmented grey matter volume focused on the superior frontal gyrus. **Middle column:** Grey matter probability maps (n = 18) computed for the four normalization procedures. These maps are restricted to voxels located inside the corresponding HBA grey matter template and denoted "Classified voxels." **Right column:** Grey matter probability maps obtained for voxels located outside the corresponding HBA grey matter template and denoted "Misclassified voxels." AFF, affine transformation; SPM, SPM linear and nonlinear routines of the SPM package; WRP, affine and fifth order polynomial normalization of the AIR package; FMG, affine and nonlinear fluid transformation.

some outliers but reveal better classification patterns using the FMG normalization procedure.

Overall accuracy based on correct and incorrect classification

Table V gives the results of the Type II errors classification computed for the four normalization procedures. These results shown strong similarities with the misclassified criterion results. The only differences between the M(i,j,k) misclassified criterion and the MC_{II}(i,j,k) Type II error concerns the white matter tissue for which no significant Type II error differences were found between SPM, WRP and FMG. Concerning the CSF tissue class, our re-

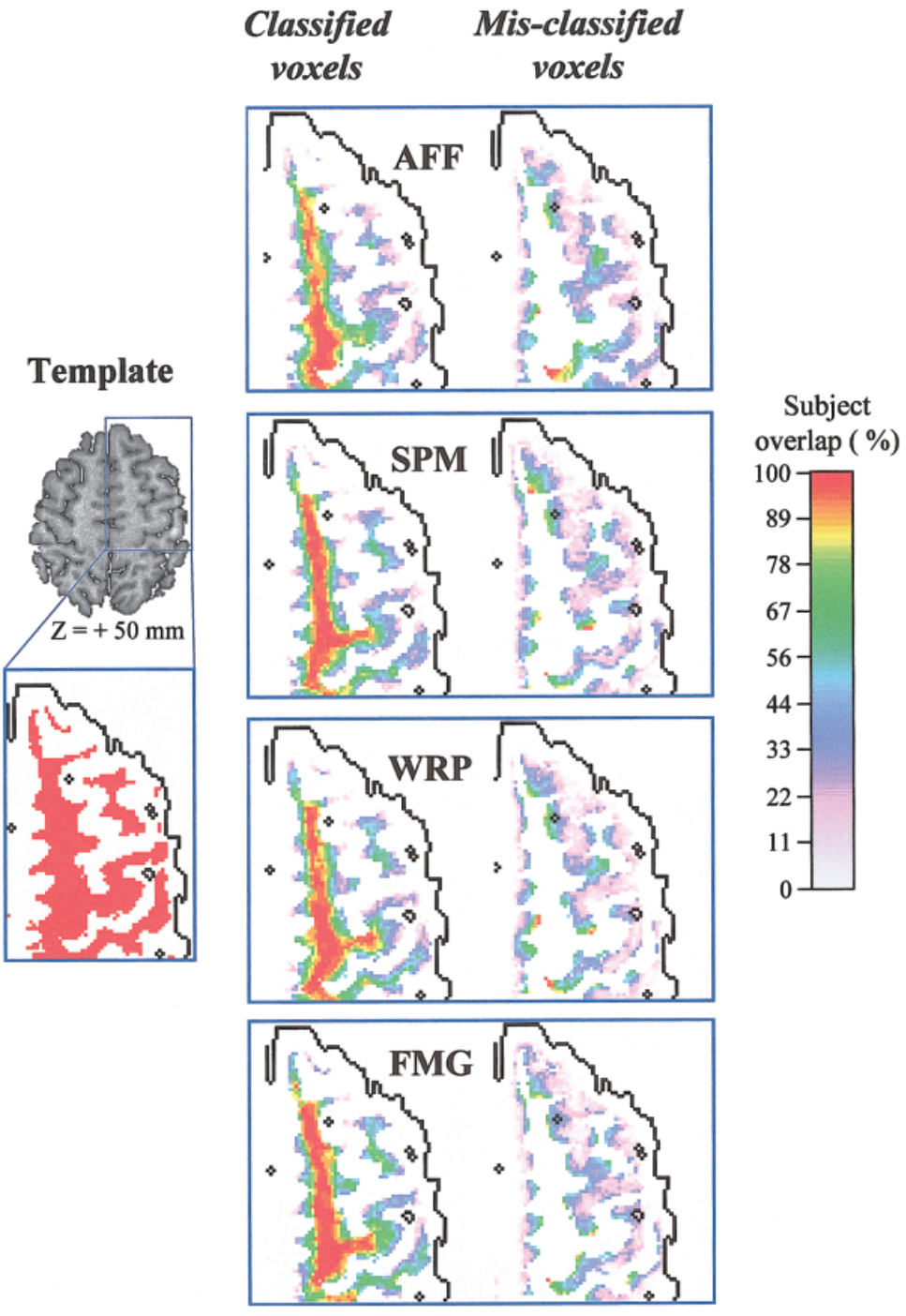

Figure 4.
Left column: Axial sections (z = +50 mm) of the HBA MRI template and of an enlargement of the corresponding segmented white matter volume focused on the superior frontal gyrus. **Middle column:** White matter probability maps (n = 18) computed for the four normalization procedures. These maps are restricted to the voxels located inside the corresponding HBA white matter template and denoted "Classified voxels." **Right column:** White matter probability maps obtained for voxels located outside the corresponding HBA white matter template and denoted "Misclassified voxels." AFF, affine transformation; SPM, SPM linear and nonlinear routine of the SPM package; WRP, affine and fifth order polynomial normalization of the AIR package; FMG, affine and nonlinear fluid transformation.

sults show that the Type II error was significantly higher for SPM than for WRP and FMG ($P < 0001$ for both). This feature was also found for the grey matter tissue class, indicating that sulcal CSF and the surrounding grey cortical ribbon were significantly better aligned with WRP and FMG than with SPM ($P = 0.003$ and $P = 0.0048$, respectively).

Impact of spatial normalization procedures on functional maps

Figures 8 and 9 present the results of the analysis of the functional data at low and high resolution, respectively. The global functional activation volumes obtained with AFF, SPM, WRP, and FMG

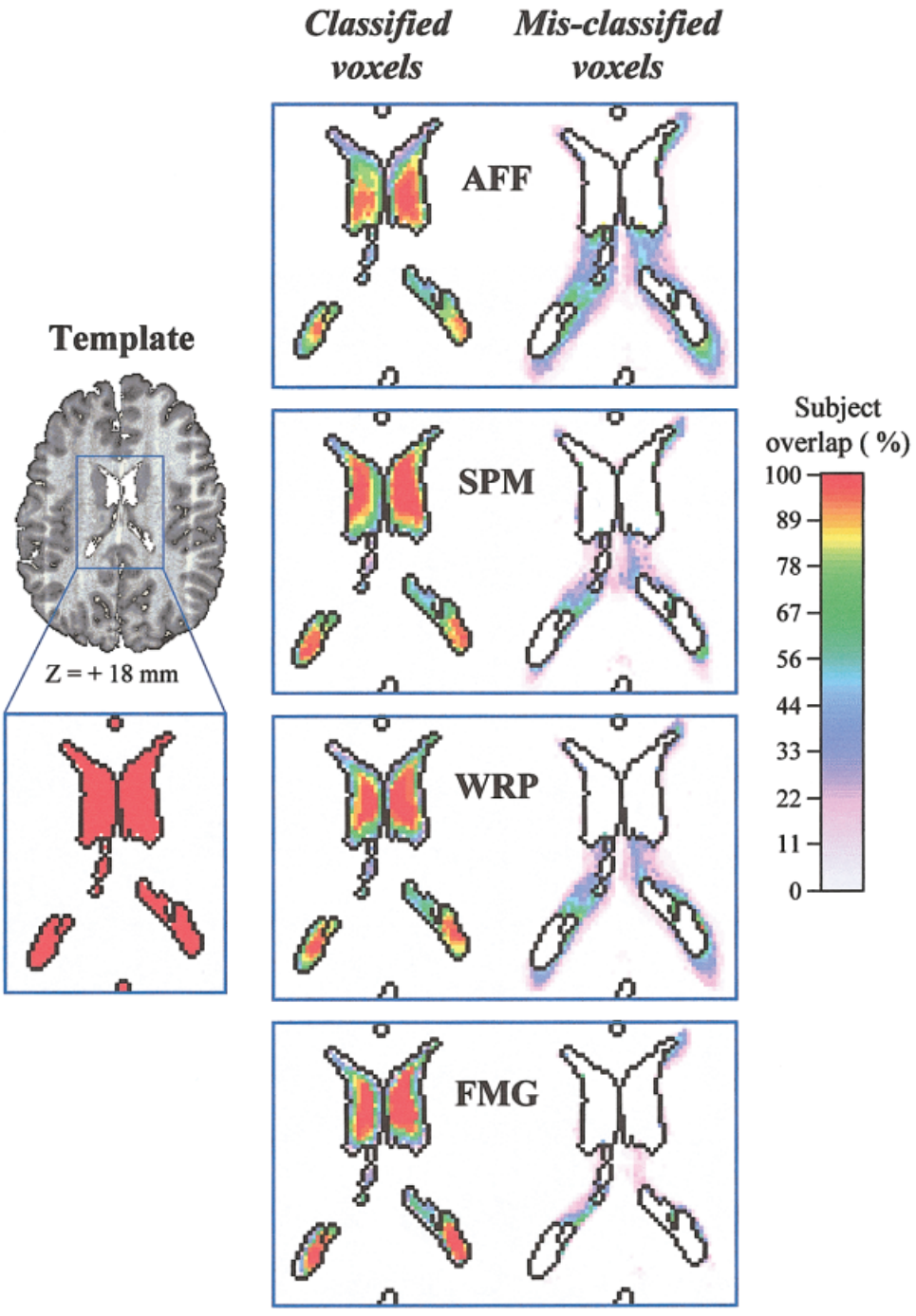

Figure 5.
Left column: Axial sections (z = +18 mm) of the HBA MRI template and of an enlargement of the corresponding segmented CSF volume focused on the ventricles. **Middle column:** CSF probability maps (n = 18) computed for the four normalization procedures. These maps are restricted to the voxels located inside the corresponding HBA CSF template and denoted "Classified voxels." **Right column:** CSF probability maps obtained for voxels located outside the corresponding HBA CSF template and denoted "Misclassified voxels." AFF, affine transformation; SPM, SPM linear and nonlinear routine of the SPM package; WRP, affine and fifth order polynomial normalization of the AIR package; FMG, affine and nonlinear fluid transformation.

normalization of "low resolution" PET blood flow images acquired during the Silent verb generation *minus* Rest contrast were 38.7, 45.3, 42.1, and 42.7 cm^3. Values obtained at "high resolution" were 3.4, 5.3, 3.2, and 4.9 cm^3. A striking feature of these results is that the range of the four activated volumes increases when increasing the resolution (6.6 cm^3 representing 15% of the maximum activated volume obtained at low resolution vs. 1.9 cm^3 representing 45% of the maximum activated volume obtained at high resolution). This indicates that using too small a filter is not suited for an optimal detection of PET activation, even if the criteria that one applied on the smoothing kernel fulfills the assumption of the Gaussian field theory (i.e., the smoothness has to be at least twice the voxel size

TABLE III. Percentage of overlap between segmented tissue classes and the corresponding tissue class of the Human Brain Atlas MRI template, for the four spatial normalization procedures*

Procedure	Grey matter	White matter	CSF
AFF	73.79 (2.12)	64.05 (2.42)	21.61 (3.20)
SPM	77.76 (1.87)	66.90 (2.24)	25.78 (3.98)
WRP	77.36 (2.43)	67.21 (2.16)	22.13 (4.26)
FMG	79.01 (3.99)	71.13 (3.43)	22.65 (4.74)

* Values are mean (SD) averaged across the 18 subjects. AFF, affine transformation; SPM, linear and non-linear routines of the Statistical Parametric Mapping package; WRP, affine and fifth-order polynomial normalization of the AIR package; FMG, affine and non-linear fluid transformation.

[Friston et al., 1995b]). This implies that, independent of the spatial normalization strategy, PET intersubject statistical analysis requires a large enough kernel filter to compensate for the large interindividual functional variability and cannot be performed at very high resolution. At first, examination of the activation volumes at the lowest resolution shows that the three nonlinear methods provided slightly higher activated volumes as compared to the affine procedure, indicating that improved warping of subjects neuroanatomy results in enhanced functional areas volume, despite the fact that the inter-subject averaged functional map resolution (~10 mm) is far more coarse than that of the structural image (~1 mm). The left part of Figure 8 shows that, at low resolution, the four methods gave very similar functional maps. There was no activation found with one method that was not detected with the others. The subtle observed differences concerned the morphology of activation clusters (i.e., outside delineation of the activated areas), and the localization of their local maxima. Table VI shows that the four functional volumes overlap in 42.8% of their union, whereas activated volumes specific to one method represented 26.2% of the same volume. The right part of Figure 8 emphasizes this point and demonstrates that the different functional volumes largely overlap (white areas). These findings indicate that 10 mm spatial resolution functional maps are quite independent of the normalization procedure used, even with a simple linear normalization method such as AFF. Going to higher resolution changed the results quite dramatically. Table VI and Figure 9 highlights this phenomenon, showing that specific activation volumes raised to 62.9% of the total activated volume, the four functional volumes now overlapping in only 6.2% of their union. Actually, the activated volume specific to each procedure (15.1, 20.4, 7.7, and 19.7% for AFF, SPM, WRP, and FMG, respectively) was larger than the four functional volumes overlap (see Table VI, and right side of Fig. 9). This indicates that at high resolution, spatial normalization procedures have a non negligible impact on the localization of activated areas. Actually, the volume of activated areas common to the four methods was preserved mainly for foci located near internal structures (see for example the left insula focus at x = −28 mm).

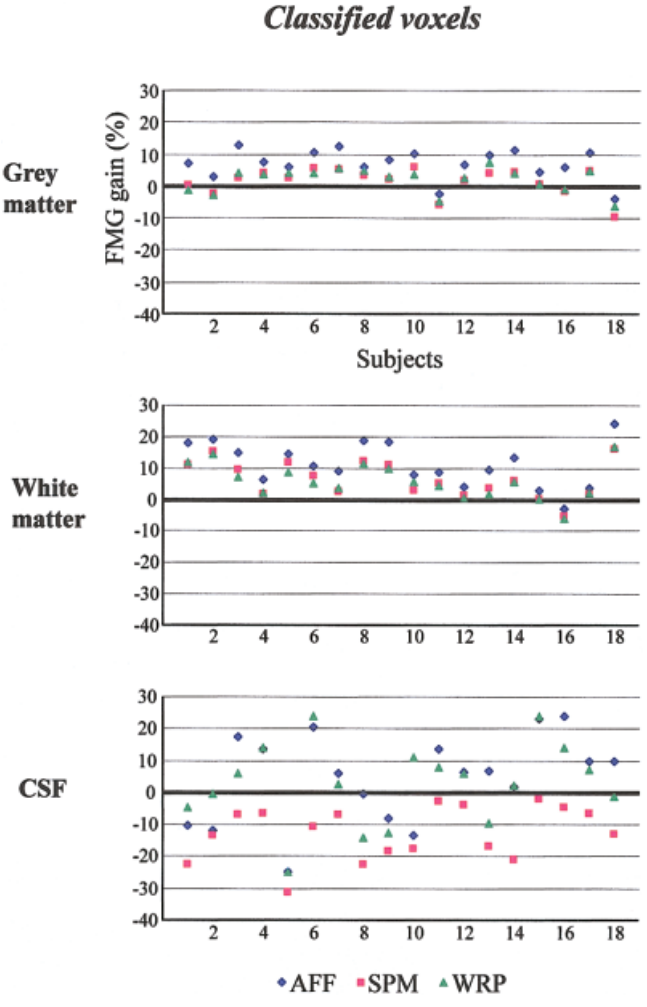

Figure 6.
Gain in tissue class overlap (in %) provided by FMG as compared to AFF, SPM, and WRP, for each subject, the tissue segmented HBA template serving as a reference. Positive values indicate that FMG provides a better overlap of the tissue type. Negative values indicates that FMG provides a worse overlap. AFF, affine transformation; SPM, SPM linear and nonlinear routine of the SPM package; WRP, affine and fifth order polynomial normalization of the AIR package; FMG, affine and nonlinear fluid transformation.

TABLE IV. Percentage of misclassified voxels of a segmented tissue class with respect to the corresponding tissue class of the Human Brain Atlas MRI template, for the four spatial normalization procedures*

Procedure	Grey matter	White matter	CSF
AFF	33.41 (1.75)	32.93 (1.90)	66.98 (3.54)
SPM	29.05 (1.87)	29.60 (1.01)	60.26 (3.26)
WRP	28.26 (0.86)	29.00 (1.26)	55.73 (3.27)
FMG	27.36 (2.48)	28.06 (2.23)	53.72 (4.71)

* Values are mean (SD) averaged across the 18 subjects. AFF: affine transformation; SPM, linear and non-linear routines of the Statistical Parametric Mapping package; WRP, affine and fifth order polynomial normalization of the AIR package; FMG: affine and non-linear fluid transformation.

DISCUSSION

Methodological aspects

Previous comparison studies

We used an original approach for evaluating the spatial normalization performances of four linear and nonlinear inter-subject alignment procedures applied on high-resolution 3D MRI datasets. One of the main goal of our study was to derive a methodology that would be unbiased, from the acquisition of the dataset images to the comparison criterion and its interpretation. This issue remains under debate because at present, standard methods for evaluating the global and regional accuracy of the spatial transformations applied on brain images have not been fully developed. Although many papers have reported important and fundamental methodological improvements in spatial normalization by presenting new core algorithms, their validation has been usually very crude and limited to a comparison with the simplest linear affine transformation [Andersson and Thurfjell, 1997; Ashburner et al., 1998; Ashburner and Friston, 2000; Collins et al., 1994; Davatzikos, 1996; Kochunov et al., 1999; Lancaster et al., 1999; Meyer et al., 1999; Minoshima et al., 1994; Rizzo et al., 1997; Roland et al., 1994; Schormann and Zilles, 1998; Thompson et al., 2000a; Woods et al., 1997b].

Indeed, reports on the comparison between nonlinear warping methods are very scarce, for such studies require tedious processing of large image datasets. Previous authors [Senda et al., 1998; Sugiura et al., 1999], for example, have compared four normalization approaches on the same dataset composed of MRI and PET images: linear scaling [Fox et al., 1985]; SPM 95 [Friston et al., 1995a]; HBA [Roland et al., 1994]; and MICHIGAN [Minoshima et al., 1994]. Using size and contour of the brain, as well as the course of four major sulci, these authors found that SPM and HBA were equivalent from the point of view of anatomical homogeneity, despite the fact that SPM was used to normalize low resolution PET data only. Regarding functional images, these authors concluded to the usefulness of nonlinear procedures, as compared to linear ones, but found that it was better to perform spatial normalization of PET images on a PET template rather than to apply a spatial normalization matrix derived

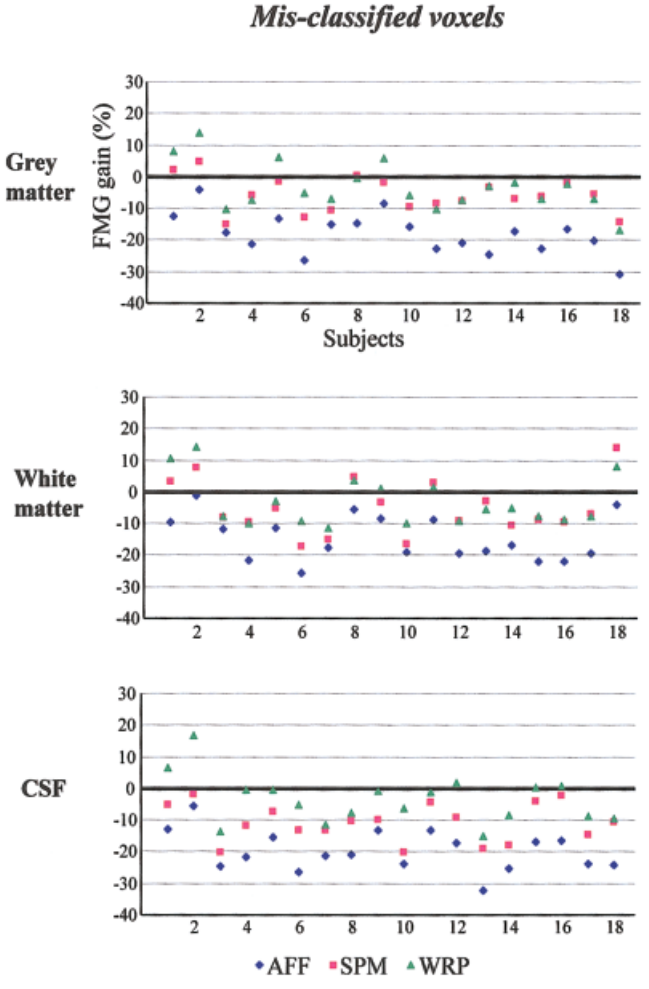

Figure 7.
Gain in misclassification (in %) provided by FMG as compared to AFF, SPM and WRP, for each subject. Negative values indicate that FMG provide a lower percentage of misclassified voxels. Positive values indicates that FMG provides a higher percentage of misclassified voxels. AFF, affine transformation; SPM, SPM linear and nonlinear routine of the SPM package; WRP, affine and fifth order polynomial normalization of the AIR package; FMG, affine and nonlinear fluid transformation.

TABLE V. Multi-class Type II error (%) computed for each class and for the four spatial normalization procedures*

Procedure	Grey matter	White matter	CSF
AFF	11.59 (1.13)	5.14 (0.56)	0.98 (0.18)
SPM	9.97 (1.12)	4.59 (0.34)	0.87 (0.16)
WRP	9.52 (0.65)	4.47 (0.34)	0.63 (0.16)
FMG	9.34 (1.42)	4.54 (0.64)	0.59 (0.17)

*Values are mean (SD), averaged across the 18 subjects. AFF, affine transformation; SPM, linear and non-linear routines of the Statistical Parametric Mapping package; WRP, affine and fifth order polynomial normalization of the AIR package; FMG, affine and non-linear fluid transformation.

from structural images (HBA). These results, however, must be taken cautiously because not all procedures were applied to the two datasets, linear and HBA being applied to MRI data (using an MRI based template) whereas SPM and MICHIGAN were applied to PET data (using a specific PET template). To pool all normalized images in a given modality, the transformation matrix computed in one modality was further used to normalize the other modality image. A second drawback of this study was that not all procedures were automatic, the user interactive version of HBA being used, which biased their comparison. Several studies have indeed compared the performances of manual versus automated warping techniques, reporting a superiority of the latter on the former, except in some cases of brain lesions [Fiez et al., 2000], and a crucial dependence of the manual approach accuracy on the operator skills [Sugiura et al., 1999].

In the present study, we tried to avoid these pitfalls, all normalization procedures being fully automatic and applied to all images acquired in a given modality using the same template. In addition, all steps in the generation of the PET functional statistical maps were automated and rigorously identical except, of course, the MRI spatial normalization: PET–PET intra-subject alignment, PET–MRI intra-subject registration, filtering, and statistical analysis. This ensures that the difference observed in activation foci localization (Figs. 8,9) can truly be ascribed to differences in sulcal patterns of spatially normalized MRI images.

Comparison criterion

Another key point emerging from spatial normalization comparison studies is the comparison criterion used for alignment performance evaluation. To our knowledge, no previous study has used a full 3D criterion for comparing alignment accuracy. Previous comparisons have been based on some measure of the dispersion of anatomical (or functional) landmarks that were placed, either manually or automatically, on both the normalized and template images [Fischl et al., 1999; Grachev et al., 1999; Woods et al., 1997b], thereby restricting the evaluation to an a priori small subset of the 3D brain volumes. In a different vein, a surface based technique has been proposed recently for spatial normalization evaluation [Lohmann et al., 1999], whereby an average distance is computed between the segmented white matter surfaces of each normalized brain and of the template, through random and roughly even sampling of both surfaces. More sophisticated validation criteria have been proposed based on a blurring metric that gives the spatial uncertainty associated with the location of a given anatomical (or functional) area after spatial normalization has been applied [Fischl et al., 1999]. In fact, some studies have used a regional 3D criterion, such as the overlap of volumes of interest [Cardenas et al., 2000; Fiez et al., 2000].

In the present study, 3D probability tissues maps were used to evaluate and compare alignment accuracy across normalizing procedures, allowing both to derive global statistics and to focus on specific tissue or areas of interest. This generic approach has similarities with voxel based methods proposed recently for assessing the difference between samples of structural MRI images, such as "Voxel-based morphometry" (VBM) [Ashburner and Friston, 2000; Woermann et al., 1999], or a method based on the permutation theory [Bullmore et al., 1999]. Both VBM and the permutation method require the image to be smoothed to a certain degree to be theoretically valid (VBM) or computationally tractable (permutation), and consequently cannot be applied to high resolution datasets. In contrast, the high resolution probability tissue maps used in the present study allow one to evaluate the alignment accuracy at the intrinsic resolution of the normalized MR images.

Another point that is worth discussing is how to evaluate the quality of matching an MRI volume to a template when brain regions are missing (i.e., in presence of a lesion or abnormalities such as in Alzheimer disease). It is obvious that the four methods used in our comparison will be inadequate. When brain regions are missing, these automated procedures attempt to reduce image mismatch between the template and the image to be normalized at lesion sites, and this can lead to inappropriate image distortions, especially when nonlinear transformations are used. The higher the degree of freedom of a spatial normalization technique, the worse will be the alignment of a

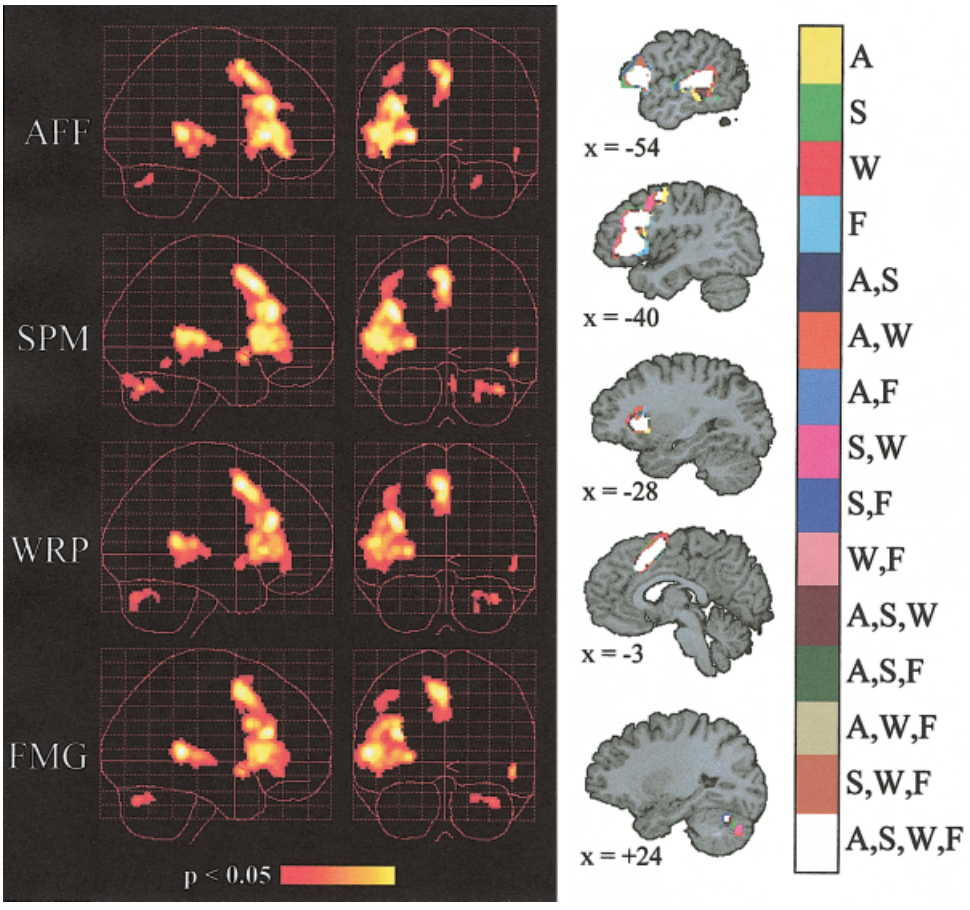

Figure 8.

Impact of spatial normalization procedures on "low resolution" functional activation maps. Stereotactically normalized PET volumes were smoothed with a 3D 8 mm FWHM Gaussian filter. Functional clusters were detected in intersubject averaged (n = 18) PET difference volumes spatially normalized by the four different normalization procedures. **Left side:** Sagittal and coronal orthogonal maximum intensity projection maps of functional volumes thresholded at a 0.05 significance level corrected for multiple comparisons. Note that the contour delineating the brain is derived from the MNI template whereas the statistical analysis was performed on the HBA template. **Right side:** Sagittal sections (x = −54, −40, −28, −3, +24 mm) of the functional volumes superimposed on the corresponding HBA template sections, illustrating activations specific or common to the four functional volumes. Color code indicates which functional volume(s) a given voxel belongs to. A, affine transformation; S, SPM linear and nonlinear routine of the SPM package; W, affine and fifth order polynomial normalization of the AIR package; F, affine and nonlinear fluid transformation.

brain MRI with missing regions. To circumvent this drawback, a method based on the masking of the lesion areas in the calculation of image difference has been developed recently and validated [Brett et al., 2001]. This approach was found superior to the classical approach of this problem. In the case of Alzheimer disease, which is characterized by a diffuse grey matter atrophy, we believe that spatial normalization using a global nonlinear brain match (for example SPM or WRP) are better suited for statistical neuroanatomy studies performed to determine the in vivo cerebral modifications caused by this pathology.

Validation of tissue segmentation

Our data analysis strategy heavily relied on the segmentation scheme, both for each individual normalized MRI volume and for the HBA template, the tissue classification of the latter serving as a gold standard. We have used the segmentation algorithm included in the SPM99 package [Ashburner and Friston, 1997], a well-documented and recently rigorously validated approach [Ashburner and Friston, 2000]. The relatively high resolution of the normalized image (~1 mm) asserts that the partitioned images were not

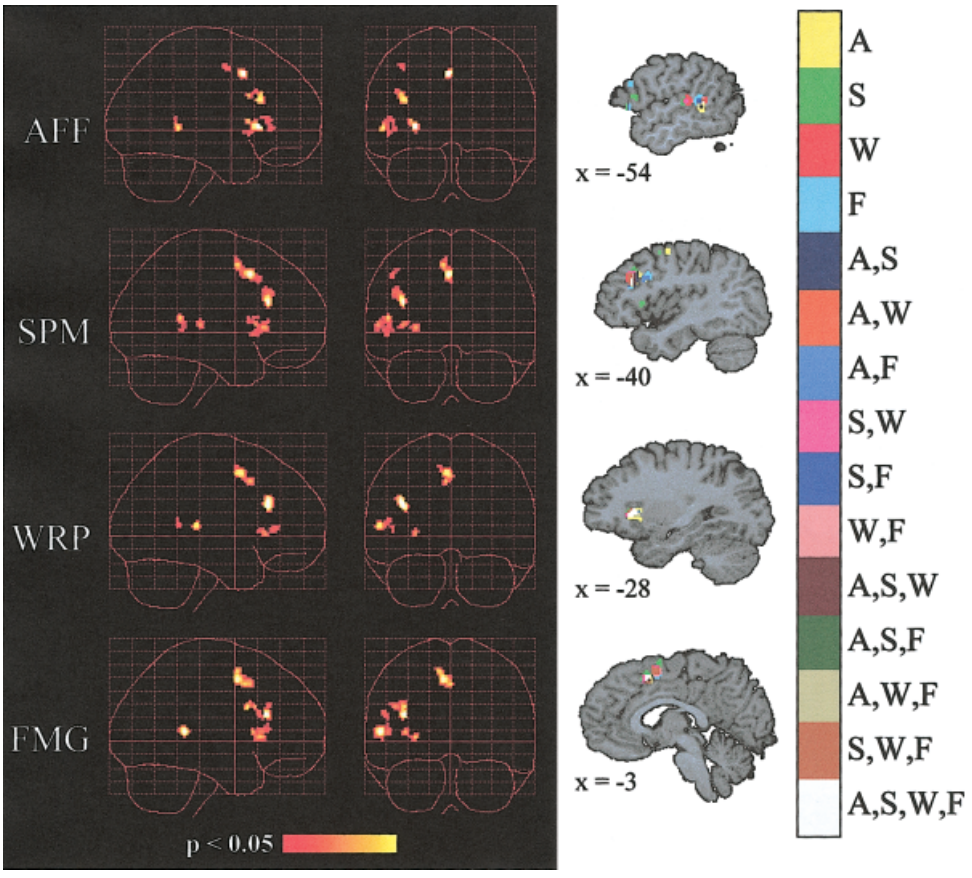

Figure 9.

Impact of spatial normalization procedures on "high resolution" functional activation maps. Stereotactically normalized PET volumes were smoothed with a 3D 4 mm FWHM Gaussian filter. Functional clusters were detected in intersubject averaged (n = 18) PET difference volumes spatially normalized by the four different normalization procedures. Left part of the figure shows sagittal and coronal orthogonal maximum intensity projection maps of functional volumes thresholded at a 0.05 significance level (corrected for multiple comparisons). Note that the contour delineating the brain is derived from the MNI template whereas the statistical analysis was performed on the HBA template. Right part shows sagittal sections (x = −54, −40, −28, −3 mm) of functional volumes superimposed on the corresponding HBA template sections, illustrating activations specific or common to the four functional volumes. Color code indicates which functional volume(s) a given voxel belongs to. A, affine transformation; S, SPM linear and nonlinear routine of the SPM package; W, affine and fifth order polynomial normalization of the AIR package; F, affine and nonlinear fluid transformation.

confounded by partial volume effects in which voxels contained mixture of different tissue types. It must be noticed that the SPM original standard space is the Montreal Neurological Institute (MNI) template [Evans et al., 1994], which has different size and shape compared to the HBA template used in the present study. As shown in Table I and Figure 1, using a priori tissue probability maps designed for the MNI template for segmenting the HBA template was satisfactory, if one excepts the CSF tissue. As a consequence, relatively low overlap was found between CSF probability maps. Our results show that SPM was also superior to the AFF and WRP methods for inter-subject intracerebral CSF alignment. WRP and FMG were both initiated using the "alignlinear" routine of the AIR package that corresponds to the AFF procedure. Meanwhile, SPM contains its own affine transformation that seems to perform better alignment than AFF for internal structures such as the ventricles, and particularly to provide a better starting point for further nonlinear alignment.

FMG vs. other approaches

Unlike the FMG procedure, SPM and WRP are not designed to provide perfect alignment of two different brains. Indeed, because the deformations provided by the discrete cosine basis functions, or polynomial

TABLE VI. Percentage of activated volume specific or common to two or more normalized functional volumes at a 0.05 significance level, corrected for multiple comparisons*

Procedure	Low resolution		High resolution	
A	7.5		15.1	
S	8.8		20.4	
		Σ = 26.2		Σ = 62.9
W	4.9		7.7	
F	5.0		19.7	
A,S	1.0		4.3	
A,W	2.4		2.2	
A,F	1.7		3.0	
S,W	4.0		4.2	
S,F	6.0		6.2	
		Σ = 31		Σ = 30.9
W,F	1.0		1.6	
A,S,W	2.8		1.5	
A,S,F	2.4		3.2	
A,W,F	2.4		0.4	
S,W,F	7.3		4.3	
A,S,W,F	42.8	Σ = 42.8	6.2	Σ = 6.2

* Low (respectively High) resolution was obtained by applying a 8 mm (respectively 4 mm) three-dimensional full with half maximum Gaussian filter to the PET images. A, affine transformation only; S, linear and non-linear routines of the Statistical Parametric Mapping package; W, affine and fifth order polynomial warp; F, affine and non-linear fluid transformation.

warping, are only defined by a few hundred parameters, as opposed to the millions of ones of the FMG procedure, they do not have the potential precision of more sophisticated methods. As a consequence, SPM and WRP offer the advantage of preserving the local relationships among sulci and gyri after the normalization has been applied, and provide a pertinent overall adjustment of global nonlinear size and shape brains. SPM and WRP also have the potential to be superseded by upgraded versions of the same core algorithm (for example using higher order polynomial warps (8, 10, or even 12) or more 3D discrete cosine transform basis functions. Technical and computational limitations, as well as biological constraints, however, restrict their use to global brain matching. In contrast, the fluid mechanical theory implemented in FMG can model any brain onto another. In this method, high-level structural information has to be invoked to guarantee the biological validation of the resulting spatial transformation, which generally implies that the relationship among anatomical landmarks can be lost after the matching step has been performed, depending on the amount of mismatch between the anatomy of the source and the target brains. For example, if the brain atlas provides a peculiar sulcus with two ramifications whereas the brain to be deformed has only one ramification for the same sulcus, the resulting deformed sulcus will finally exhibit two ramifications. This means that the original anatomy is modified in such a way that it is constrained to fit the anatomy of the target one. This specific ability was clearly illustrated in our comparison study, particularly when one looks at the shape of the sulcus after deformation. Furthermore, the FMG algorithm can be adjusted to preserve the local relationship among sulci and gyri as does SPM and WRP, which was done in the present study. By increasing the grid-resolution it is also possible to fully warp one brain to another. Approximate absolute running times on a SUN-SPARC ULTRA 30 were about 15 min for SPM (including the affine transformation), 45 min for AFF+WRP, and 100–120 min for AFF+FMG. Computation time of the actual FMG procedure is thus much larger and can be a drawback when analyzing large datasets.

How far should we remove individual anatomical variability?

Statistical neuroanatomy

Because performing a quasi perfect brain match is by itself a challenge of interest, the question of its necessity in the neuro imagery field has to be discussed.

New computational approaches defined to quantify human neuroanatomical variability require that images to work on be put into a common space. Indeed, group comparison between normal patients and patients, such as schizophrenic or Alzheimer patients, or between young and old subjects, requires the different brains to be similar. But similar to what degree? At present, there are two majors strategies for answering this question. The first one, as mentioned previously, is VBM [Ashburner and Friston, 2000] and involves a voxel-wise comparison of the local concentration of grey or white matter density maps between two or more groups. In this case, one tries to detect differences in the regional concentration of brain tissues at a local scale, and if the spatial normalization is perfect, then all the segmented tissue images would appear identical and no differences would be detected. A global nonlinear brain match (SPM, WRP) is almost satisfactory, therefore, and there is no need for a more precise normalization of the brain images. VBM, using SPM as the gold standard spatial normalization procedure, has been used already to demonstrate local

structural brain differences among various populations [Abell et al., 1999; Good et al., 2001; May et al., 1999; Paus et al., 2000; Sowell et al., 1999; Woermann et al., 1999; Wright et al., 1995]. To our knowledge, however, there has been no comparison of the relative merits of SPM and WRP to determine the differences observed in classical VBM statistical neuroanatomy. The second class of methods aims at identifying macroscopic anatomical shape and size difference among the brains of different population. Here, statistics are not carried out directly on the normalized volumes, but are applied to the parameters describing the ensuing estimated nonlinear deformations. These techniques, called deformation-based and tensor-based morphometry, compare the transformation fields corresponding to different populations. The more precise is the mapping between homologous features of a individual brains to a target one, the more relevant is the displacement field characterizing the difference in the shape and the size of these specific features [Ashburner et al., 1998; Davatzikos, 1996; Gaser et al., 1999]. For this purpose, sophisticated methods are of interest and are highly recommended to reduce significantly the anatomical residual variability between subjects after spatial normalization. The use of these high-dimensional elastic normalization approaches have permitted us to investigate, for example, longitudinal studies of the corpus callosum growth patterns in children [Thompson et al., 2000b] or to map precisely morphological differences of the same structure in schizophrenic patients [Narr et al., 2000]. The results obtained from our comparison study have shown that the FMG algorithm provide a slightly better alignment of anatomical features between subjects. These results, however, indicate that SPM, WRP, or FMG procedures are not accurate enough, both in terms of warping the individual to the template, to be used in deformation- or tensor-based morphometry analysis. The use of one of these procedures combined with displacement measures can blur the identification of subtle and localized effects of brain morphology between populations, due to the important residual variability remaining after spatial normalization (Tables IV,V). In a recent work, Good et al. [2001] used the smooth dimensional deformation field obtained from the SPM99 normalization procedure to investigate the effect of age on the volume reduction (or growth) of anatomical structures in a sample of 465 subjects. This was done using a "modulation step" preserving the volume of a particular tissue within a voxel by multiplying the voxels values in the segmented images by the Jacobean determinants derived from the spatial normalization step. As the authors said, "…while this current approach cannot provide exact matches between small cortical or deep grey matter structures, it provides additional information in a practical way for large subject cohorts." The anatomical residual variability was compensated in this case by the very large number of subjects of the study, acting as a supplementary image smoothing process. This improvement has also been implemented by Davatzikos et al. [2001]. These authors use a very high dimensional elastic transformation to preserve the volumetric measurements of each tissue, both locally and globally. This study demonstrated that the combination of high dimensional elastic transformation and VBM can be used to accurately quantify very localized atrophy and volumetric changes.

Regarding the normalization procedures compared in our study, we confirm that smooth normalization procedure such as SPM, WRP, or FMG provide an amount of residual variability that restricts their use to traditional VBM analysis of tissues class densities. VBM analysis does not measure any higher order shape characteristics and are less sensitive to spatial normalization errors in comparison with methods based on displacement fields.

Functional neuroanatomy

Another field of investigation where the accuracy of spatial normalization is of relevance is the investigation of the neural basis of higher cognitive functions. Indeed, a functional parcellation of the brain consists of segregating and precisely localizing cerebral functional fields. A basic hypothesis underlying cognitive neuroanatomy studies is that functional fields in different individuals are located at similar anatomical locations [Watson et al., 1993]. As a consequence, the use of warping algorithms that perform the best registration of different brains should result in more accurate alignment of functional areas and therefore enhance the intrinsic detection sensitivity of these functional fields. This characteristic has been found in previous studies that have related the effect of spatial normalization on functional datasets and have concluded that the normalization procedure with higher localization accuracy of structural features provide the most accurate and sensitive functional results [Andersson and Thurfjell, 1997; Downs et al., 1999; Fischl et al., 1999; Gee et al., 1997; Senda et al., 1998].

Our findings moderate these assertions. We did not investigate the differences in terms of sensitivity between functional maps obtained with different normalization procedure (activated volumes, extent, and overlap in a common space). Using conservative sig-

nificance level and traditional smoothing kernel for classical PET inter-subject analysis (~10 mm final 3D spatial autocorrelation), our results show that there were only subtle differences between inter-subject maps, either linearly or nonlinearly stereotactically normalized. This suggests that the blurring introduced by the inter-individual functional variability dominates the anatomical variability that can be reduced by precise warping strategies. Although anatomical features such as sulci, were aligned with millimeter resolution, the underlying functional fields identified by inter-subject PET activations (mainly the extent of these functional sites) were spread around the anatomical center of interest. This, of course, does not imply that the activation centers after spatial normalization are located exactly at the same coordinates, but it shows that conclusions derived from the analysis of such PET experiments are independent of the normalization scheme used to move the functional volumes in the standard space. It should be noted that the Gaussian kernel size we used was about 4 mm lower than most filter sizes used in the functional neuroimaging literature. One can easily hypothesize that using a more traditional larger kernel size will emphasize the homologies between functional maps produced with different normalization strategies and enhance the overlap of activated volume across methods. Tuning the significance threshold to ultra high conservative thresholds should probably leave the maximum intensity activation peak, allowing the comparison of their relative localization in the stereotaxic space. This was achieved in our study by improving the spatial resolution of the functional study to ~6 mm, which corresponds to the resolution of many fMRI maps. No differences were observed in terms of the number of activated areas but non-negligible variations appeared in terms of their localizations, the activation volume overlap between methods being restricted to a few percents of the total activated volume. These variations can be of importance when researchers have to distinguish precise activation areas located in the same anatomical features, but the problem of selecting the method that gives the best "theoretical" results still remains.

Relationships between microscopic and macroscopic areas

A last topic of interest, in which a complete warp is of crucial importance, is the analysis of the precise relationships between macroscopic functional and cytoarchitectonic associative areas involved in higher cognitive function such as language or mental imagery. It has been demonstrated that the macroscopic anatomical variability is highly related to the microstructural one in primary cortices [Geyer et al., 1996], but this tight relationship seems to disappear for associative areas [Amunts et al., 1999]. The large variability of the architectural organization of the cerebral cortex that exists at every scale requires us to approach this issue through probabilistic schemes and atlases rather than through single individual observations (as was the case for the pioneer, Brodmann's map). Such studies are based on the hypothesis of sulcal anatomical standardization and therefore require all datasets to be accurately normalized in a probabilistic space at a macroscopic level. The most accurate brain is likely to provide the most relevant conclusions regarding the relative positions of functional landmarks and cytoarchitectural areas. This innovative approach imposes the use of quasi perfect warp between brains and will be a prerequisite for the analysis of very large neuroimaging databases such as the European Computerized Human Brain Database (ECHBD) [Roland and Zilles, 1996] or the International Consortium for Brain Mapping [Mazziotta et al., 2000]. The main goal of the ECHBD, now followed by the NEUROGENERATOR project [Roland et al., 2001] was to investigate the spatial congruence of probabilistic map of cytoarchitectonic fields and functional PET and fMRI areas. These cytoarchitectonic maps are useful and can offer precise stereotaxic information of interindividual variability of microscopic areas only if the surrounding macroscopic structural variability of the post-mortem subjects is completely reduced using a quasi perfect alignment procedure. This is what was done in the ECHBD using a very high dimensional elastic version of the FMG procedure. The results observed in our study concerning the impact of normalization procedures on high resolution functional maps (Table VI, Fig. 9) suggest that using this high dimensional procedure, the observed differences on the location and extent of functional activations will be more striking than the ones observed with the actual version of the FMG procedure. To build an inter-subject average functional map with the lowest between subjects residual variability, one has to use the must accurate normalization procedure to align the different subject brains. Our results lead us to conclude that new database projects integrating high resolution functional datasets with micro- or macroscopic structural informations should be based on normalization procedure providing the highest degree of accuracy, such as FMG, to investigate close relationships between structure and function.

ACKNOWLEDGMENTS

This study was supported in part by European Union Biotechnology grant BT04-CT96-0177, "European Computerized Human Brain Database." The authors would like to thank D. Papathanassiou for his help in acquiring the anatomical and functional datasets used in this study. The authors would also like to thank the anonymous reviewers who helped us to improve the original manuscript.

REFERENCES

Abell F, Krams M, Ashburner J, Passingham RE, Friston KJ, Frackowiak RSJ, Happé F, Frith CD, Frith U (1999): The neuroanatomy of autism: a voxel-based whole brain analysis of structural scans. Neuroreport 10:1647–1651.

Amunts K, Schleicher A, Bürgel U, Mohlberg H, Uylings H, Zilles K (1999): Broca region revisited: cytoarchitecture and intersubject variability. J Comp Neurol 412:319–341.

Andersson JLR, Thurfjell L (1997): Implementation and validation of a fully automatic system for intra- and interindividual registration of PET brain scans. J Cereb Blood Flow Metab 21:136—144.

Ashburner J, Friston KJ (1997): Multimodal image coregistration and partitioning: a unified framework. Neuroimage 6:209–217.

Ashburner J, Friston KJ (1999): Nonlinear spatial normalization using basis functions. Hum Brain Mapp 7:254–266.

Ashburner J, Friston KJ (2000): Voxel-based morphometry: the methods. Neuroimage 11:805–821.

Ashburner J, Hutton C, Frackowiak RSJ, Johnsrude I, Price CJ, Friston KJ (1998): Identifying global anatomical differences: deformation-based morphometry. Hum Brain Mapp 6:348–357.

Brett M, Leff AP, Rorden C, Ashburner J (2001): Spatial normalization of brain images with focal lesions using cost function masking. Neuroimage 14:486–500.

Bullmore ET, Suckling J, Overmeyer S, Rabe-Hesketh S, Taylor E, Brammer MJ (1999): Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brains. IEEE Trans Med Imaging 18:32–42.

Cardenas VA, Truran D, Weiner MW (2000): MRI warping for automatic structure identification. Neuroimage 11:S499.

Christensen GE, Rabbitt RD, Miller MI (1994): 3D brain mapping using a deformable neuroanatomy. Phys Med Biol 39:609–618.

Collins DL, Neelin P, Peters TM, Evans AC (1994): Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. J Comput Assist Tomogr 18:192–205.

Davatzikos C (1996): Spatial normalization of 3D brain images using deformable models. J Comput Assist Tomogr 20:656–665.

Davatzikos C, Vaillant M, Resnick SM, Prince JL, Letovsky S, Bryan RN (1996): A computerized approach for morphological analysis of the corpus callosum. J Comput Assist Tomogr 20:88–97.

Davatzikos C (1998): Mapping image data to stereotaxic spaces: applications to brain mapping. Hum Brain Mapp 6:334–338.

Davatzikos C, Genc A, Xu D, Resnick SM (2001): Voxel-based morphometry using the RAVENS maps: methods and validation using simulated longitudinal atrophy. Neuroimage 14:1361–1369.

Downs JH, Lancaster JL, Fox PT (1999): Surface-based spatial normalization using convex hulls. In: Toga AW, editor. Brain warping. San Diego: Academic Press. p 263–282.

Evans AC, Kamber M, Collins DL, MacDonald D (1994): An MRI-based probabilistic atlas of neuroanatomy. In: Shorvon S, editor. Magnetic resonance scanning and epilepsy. New York: Plenum. p 263–274.

Fiez JA, Damasio H, Grabowski TJ (2000): Lesion segmentation and manual warping to a reference brain: intra- and interobserver reliability. Hum Brain Mapp 9:192–211.

Fischl B, Sereno MI, Tootell RBH, Dale AM (1999): High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum Brain Mapp 8:272–284.

Fox PT, Perlmutter S, Raichle ME (1985): A stereotactic method of anatomical localization for positron emission tomography. J Comput Assist Tomogr 9:141–153.

Fredriksson J, Roland PE, Svensson P, Amunts K, Cavada C, Hari R, Cowey A, Crivello F, Geyer S, Kostopoulos G, Mazoyer BM, Popplewell D, Schleicher A, Schormann T, Seppa M, Uylings H, de Vos K, Zilles K (2000): The European Computerized Human Brain Database. Neuroimage 11:S906.

Friston KJ, Ashburner J, Frith CD, Poline JB, Heather JD, Frackowiak RSJ (1995a): Spatial registration and normalization of images. Hum Brain Mapp 2:165–189.

Friston KJ, Holmes AP, Worsley KJ, Poline JB, Frith CD, Frackowiak RSJ (1995b): Statistical parametric maps in functional imaging: a general approach. Hum Brain Mapp 2:189–210.

Gaser C, Volz P, Kiebel SJ, Riehemann S, Sauer H (1999): Detecting structural changes in whole brain based on nonlinear deformations: application to schizophrenia research. Neuroimage 10: 107–113.

Gee JC, Alsop DC, Aguirre GK (1997): Effect of spatial normalization on analysis of functional data. In: Hanson KM, editor. Medical imaging 1997. Bellingham: SPIE. p 550–560.

Geyer S, Ledberg A, Schleicher A, Kinomura S, Schormann T, Bürgel U, Klingberg T, Larsson J, Zilles K, Roland PE (1996): Two different areas within the primary motor cortex of man. Nature 382:805–807.

Good C, Johnsrude I, Ashburner J, Friston KJ, Frackowiak RSJ (2001): A Voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage 14:21–36.

Grachev ID, Berdichevsky D, Rauch SL, Heckers S, Kennedy DN, Caviness VS, Alpert NM (1999): A method for assessing the accuracy of intersubject registration of the human brain using anatomic landmarks. Neuroimage 9:250–268.

Kochunov PV, Lancaster JL, Fox PT (1999): Accurate high-speed spatial normalization using an octree method. Neuroimage 10: 724–737.

Lancaster JL, Fox PT, Downs H, Nickerson DS, Hander TA, El Mallah M, Kochunov PV, Zamarripa F (1999): Global spatial normalization of human brain using convex hulls. J Nucl Med 40:942–955.

Lohmann G, von Cramon DY, Steinmetz H (1999): Sulcal variability of twins. Cereb Cortex 9:754–763.

May A, Ashburner J, Büchel C, McGonigle DJ, Friston KJ, Frackowiak RSJ, Goadsby PJ (1999): Correlation between structural and functional changes in brain in an idiopathic headache syndrome. Nat Med 5:836–838.

Mazziotta JC, Toga AW, Evans AC, Fox PT, Lancaster JL, Woods RP (2000): A probabilistic approach for mapping the human brain. In: Toga AW, Mazziotta JC, editors. Brain mapping: the systems. San Diego: Academic Press. p 141–156.

Meyer J, Gunn RN, Myers R, Grasby PM (1999): Assessment of spatial normalization of PET ligand images using ligand-specific templates. Neuroimage 9:545–553.

Minoshima S, Koeppe RA, Frey KA, Kuhl DE (1994): Anatomic standardization: linear scaling and nonlinear warping of functional brain images. J Nucl Med 35:1528–1537.

Narr KL, Thompson PM, Sharma T, Moussai J, Cannestra AF, Toga AW (2000): Mapping morphology of the corpus callosum in schizophrenia. Cereb Cortex 10:40–49.

Papathanassiou D, Etard O, Mellet E, Zago L, Mazoyer BM, Tzourio-Mazoyer N (2000): A common language network for comprehension and production: a contribution to the definition of language epicenters with PET. Neuroimage 11:347–357.

Paus T, Zijdenbos A, Worsley KJ, Collins DL, Blumenthal J, Giedd JN, Rapoport JL, Evans AC (2000): Structural maturation of neural pathways in children and adolescents: in vivo study. Science 283:1908–1911.

Rizzo G, Scifo P, Gilardi M-C, Bettinardi V, Grassi F, Cerutti S, Fazio F (1997): Matching a computerized brain atlas to multimodal medical images. Neuroimage 6:59–69.

Roland PE, Graufelds CJ, Wahlin J, Ingelman L, Andersson M, Ledberg A, Pedersen J, Akerman S, Dabringhaus A, Zilles K (1994): Human Brain Atlas: for high resolution functional and anatomical mapping. Hum Brain Mapp 1:173–184.

Roland PE, Zilles K (1996): The developing European computerized human brain database for all imaging modalities. Neuroimage 4:S39–S47.

Roland PE, Geyer S, Amunts K, Schormann T, Schleicher A, Malikovik A, Zilles K (1997): Cytoarchitectural maps of the human brain in standard anatomical space. Hum Brain Mapp 5:222–227.

Roland PE, Svensson G, Lindeberg T, Risch T, Baumann P, Dehmel A, Frederiksson J, Halldorson H, Forsberg L, Zilles K (2001): A database generator for human brain imaging. TINS 24:562–564.

Schormann T, Henn S, Zilles K (1996): A new approach to fast elastic alignment with applications to human brains. Lecture Notes Comput Sci 1131:337–342.

Schormann T, Henn S, Zilles K (1998): Method for computing and displaying 2D- and 3D-spatial differences of structures. International Patent [PCT/EP99/04442].

Schormann T, Zilles K (1998): Three-dimensional linear and nonlinear transformation: an integration of light microscopical and MRI data. Hum Brain Mapp 6:339–347.

Senda M, Ishii K, Oda K, Sadato N, Kawashima R, Sugiura M, Kanno I, Ardekani BA, Minoshima S, Tatsumi I (1998): Influence of ANOVA design and anatomical standardization on statistical mapping for PET activation. Neuroimage 8:283–301.

Sowell ER, Thompson PM, Holmes CJ, Batth R, Jernigan TL, Toga AW (1999): Localizing age-related changes in brain structure between childhood and adolescence using parametric mapping. Neuroimage 9:587–597.

Sugiura M, Kawashima R, Sadato N, Senda M, Kanno I, Oda K, Sato K, Yonekura Y, Fukuda H (1999): Anatomic validation of spatial normalization methods for PET. J Nucl Med 40:317–322.

Talairach J, Tournoux P (1988): Co-planar stereotaxic atlas of the human brain. Three-dimensional proportional system: an approach to cerebral imaging. Stuttgart: Georg Thieme Verlag.

Thompson PM, Woods RP, Mega MS, Toga AW (2000a): Mathematical/computational challenges in creating deformable and probabilistic atlases of the human brain. Hum Brain Mapp 9:81–92.

Thompson PM, Giedd JN, Woods RP, MacDonald D, Evans AC, Toga AW (2000b): Growth patterns in the developing brain detected by using continuum mechanical tensor map. Nature 404:190–193.

Thompson PM, Mega MS, Woods RP, Zoumalan CI, Lindshield CJ, Blanton RE, Moussai J, Holmes CJ, Cummings JL, Toga AW (2001): Cortical change in Alzheimer disease detected with a disease-specific population-based brain atlas. Cereb Cortex 11: 1–16.

Toga AW (1999): Brain warping. San Diego: Academic Press.

Verard L, Allain P, Travère JM, Baron JC, Bloyet D (1997): Fully automatic identification of AC and PC landmarks on brain MRI using scene analysis. IEEE Trans Med Imaging 16:610–616.

Watson JDG, Myers R, Frackowiak RS, Hajnal JV, Woods RP, Mazziotta JC, Shipp S, Zeki S (1993): Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. Cerebral Cortex 3:79–94.

Woermann FG, Free SL, Koepp M, Ashburner J, Duncan JS (1999): Voxel-by-voxel comparison of automatically segmented cerebral grey matter: a rater-independent comparison of structural MRI in patients with epilepsy. Neuroimage 10:373–384.

Woods RP, Cherry SR, Mazziotta JC (1992): Rapid automated algorithm for aligning and reslicing PET images. J Comput Assist Tomogr 16:620–633.

Woods RP, Grafton ST, Holmes CJ, Cherry SR, Mazziotta JC (1997a): Automated Image Registration: I. General methods and intrasubject, intramodality validation. J Comput Assist Tomogr 22:139–152.

Woods RP, Grafton ST, Watson JDG, Sicotte NL, Mazziotta JC (1997b): Automated Image Registration: II. Intersubject validation of linear and nonlinear models. J Comput Assist Tomogr 22:153–165.

Wright IC, McGuire PK, Poline JB, Travère JM, Murray RM, Frith CD, Frackowiak RSJ, Friston KJ (1995): A voxel-based method for statistical analysis of grey and white matter density applied to schizophrenia. Neuroimage 2:244–252.

Zhang (1996): Pattern recognition 29:1335–1346.