

Preserving Individual Differences for Voxel-Based Morphometry (VBM)

Andrew S. Fox, Terrence R. Oakes, Kim Dalton, Brendon Nacewicz, Moo Chung, Andrew L. Alexander, Richard J. Davidson

W.M. Keck Lab for Functional Brain Imaging and Behavior, Waisman Center, University of Wisconsin Madison

Introduction

Neuro-anatomical differences between autism-diagnosed patients and controls have recently been the focus of an increasing amount of research. Recent findings include: increases in total brain volume, decreases in total white matter, increases and decreases in amygdala volume, decreases in hippocampal volume, decreases in thalamus volume, as well as others (Cody et al. 2002, Aylward et al. 1999, Tsatsanis et al. 2003). These findings suggest a large amount of variability within the brain morphometry of autism-diagnosed patients. There is a particular need in voxel-based morphometry (VBM), particularly when dealing with such a high variability of anatomical features, to not dispose of individual differences during coregistration to a particular template.

Current methods used to compute structural differences between groups of subjects using VBM are crucially dependent on the coregistration of the data into a standard space. Each brain must be registered accurately to a single template in order to permit valid comparisons between individuals, yet if the registrations were perfect, all individual differences would (by definition) be removed. Our goal was to adapt the SPM-based method of Good et al. (2001), which coregisters gray- and white-matter probability distributions (referred to as "GWMP"). In our implementation, we used binary brain masks to achieve an accurate coregistration of subject brains into the same sample space while preserving individual differences in the distribution of white and gray matter (referred to as "BRAIN-MASK"). The fundamental idea is to fit individual interior and exterior brain edges to single brain-edge template, thus preserving individual differences in e.g. cortical thickness.

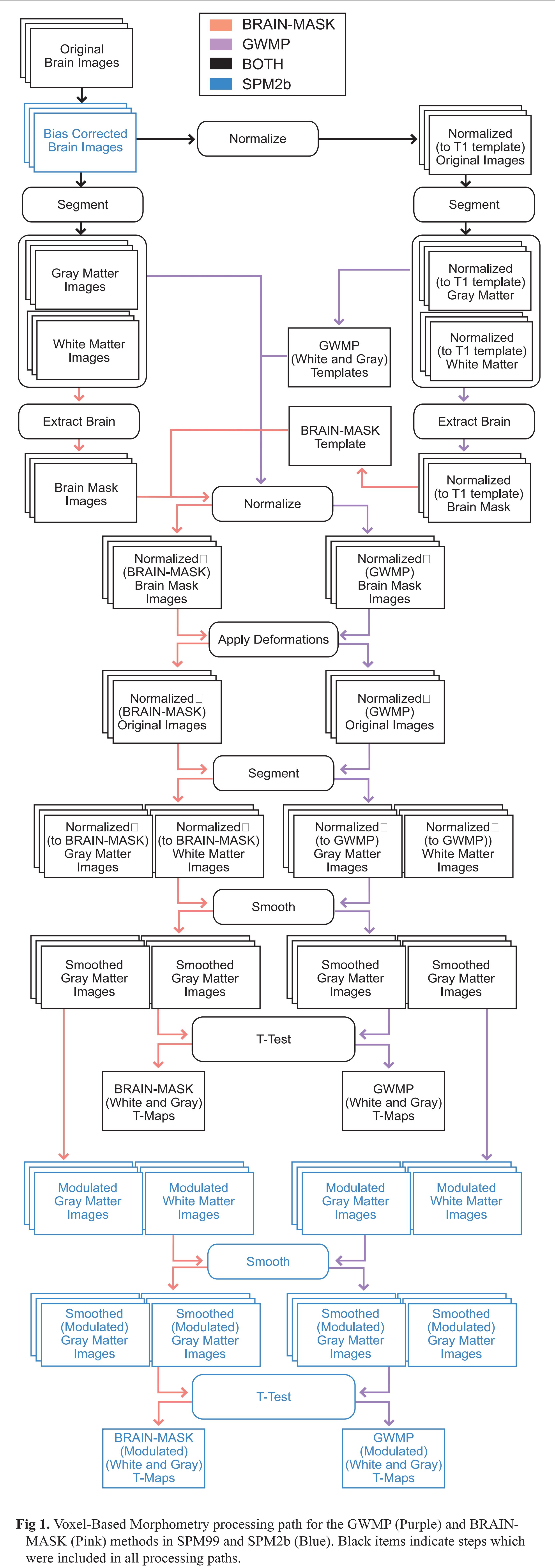


Fig 1. Voxel-Based Morphometry processing path for the GWMP (Purple) and BRAIN-MASK (Pink) methods in SPM99 and SPM2b (Blue). Black items indicate steps which were included in all processing paths.

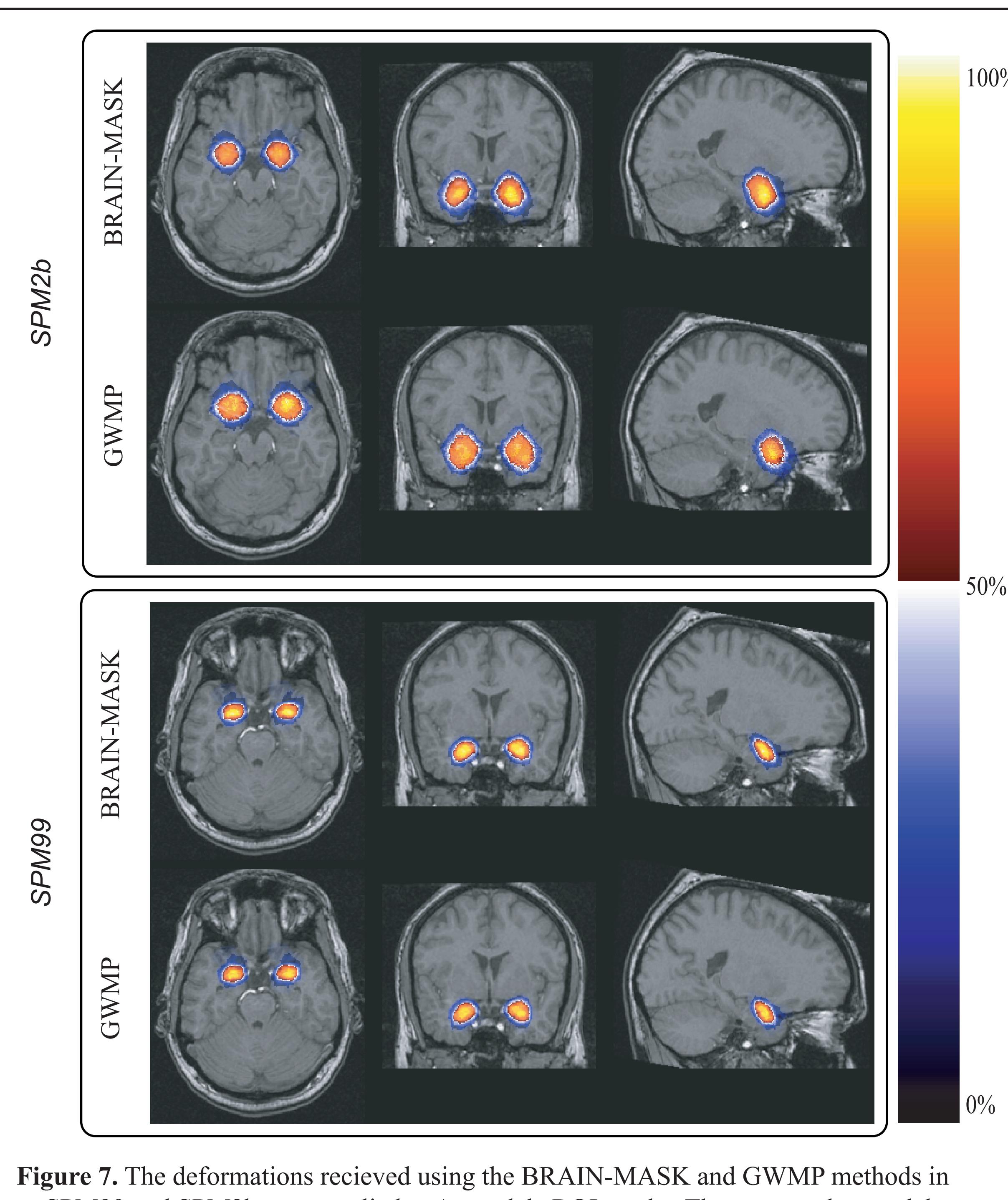


Figure 2. A mask was created, by taking all voxels for each subject that returned a non-zero probability of having brain tissue. These masks were then combined and averaged across subjects to create probability maps of brain tissue after normalization and segmentation according to the GWMP (both gray- and white-matter based normalizations) and BRAIN-MASK methods in both SPM2b and SPM99.

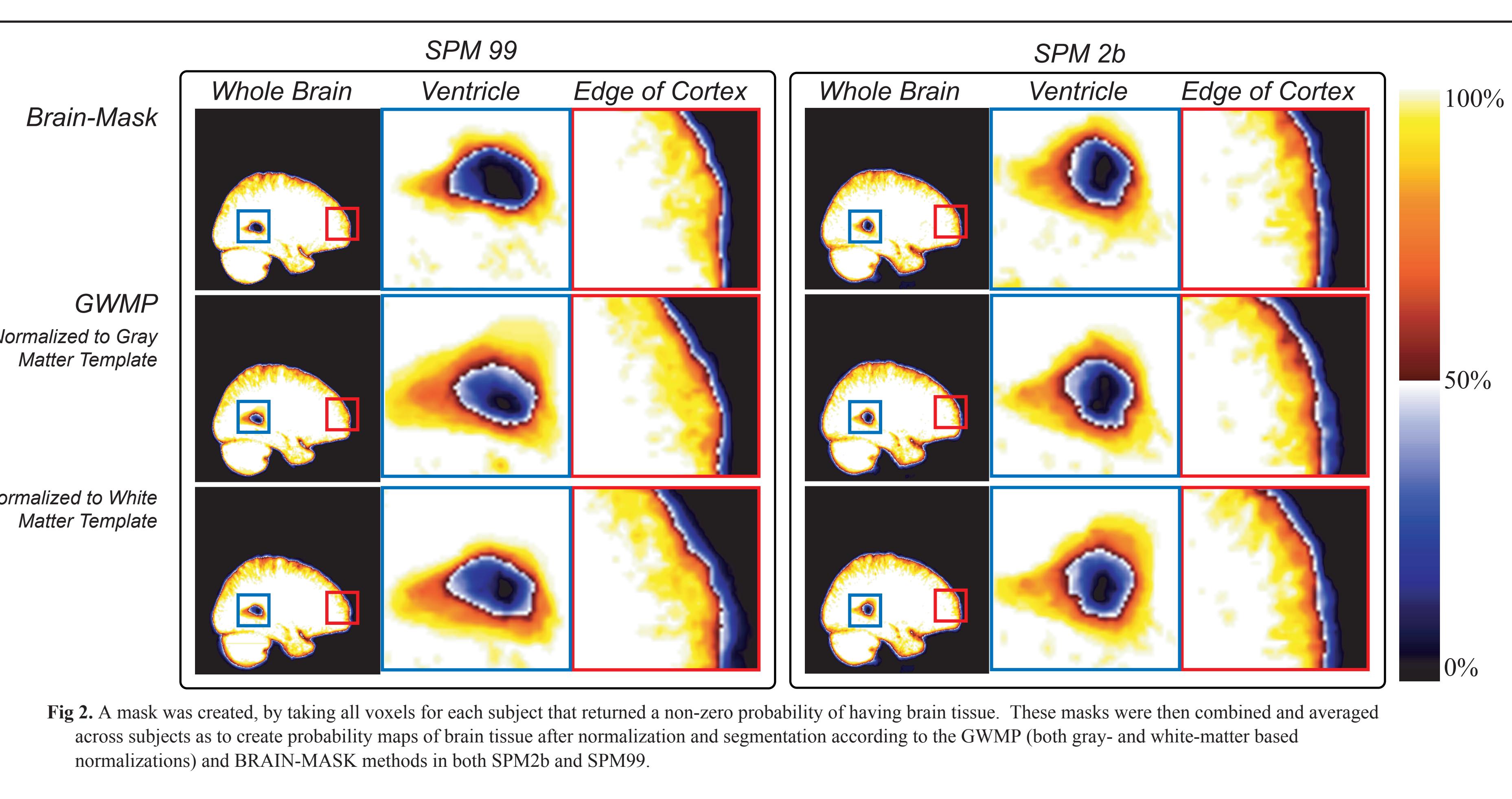


Fig 2. A mask was created, by taking all voxels for each subject that returned a non-zero probability of having brain tissue. These masks were then combined and averaged across subjects to create probability maps of brain tissue after normalization and segmentation according to the GWMP (both gray- and white-matter based normalizations) and BRAIN-MASK methods in both SPM2b and SPM99.

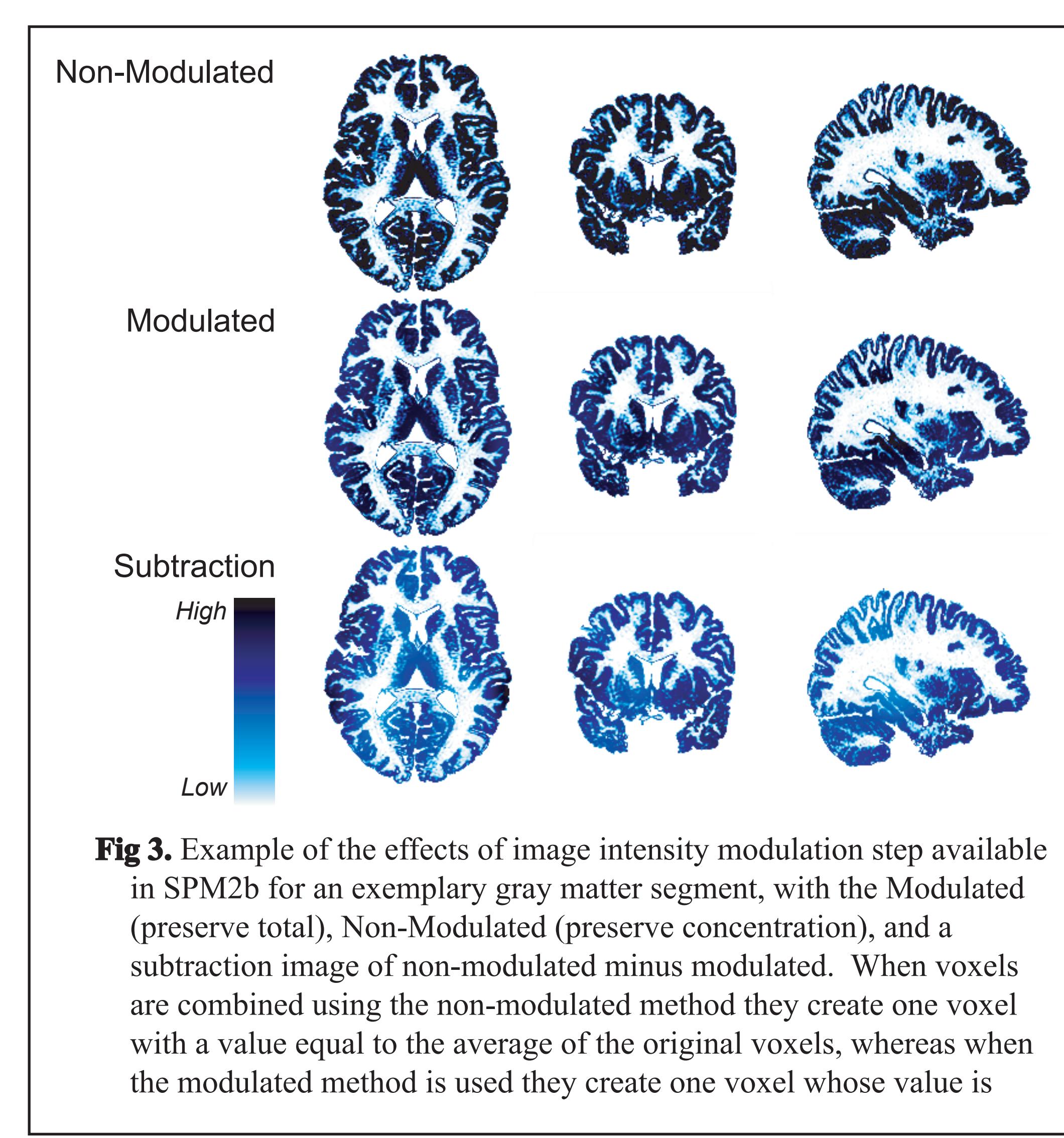


Fig 3. Example of the effects of image intensity modulation step available in SPM2b for an exemplary gray matter segment, with the Modulated (preserve total), Non-Modulated (preserve concentration), and a subtraction image of non-modulated minus modulated. When voxels are combined using the non-modulated method they create one voxel with a value equal to the average of the original voxels, whereas when the modulated method is used they create one voxel whose value is

Methods

For practical reasons, we found the concept of fitting to the brain edges was accurately implemented by fitting a single binary mask of the combined GM/WM distribution to a similar template. In order to thoroughly compare the GWMP and BRAIN-MASK methods we ran all analyses in SPM99 and the new SPM2b. The processing path for GWMP and BRAIN-MASK methods can be seen in the processing path (outlined below) as well as in Figure 1.

The statistical VBM analysis consisted of a t-test comparing autism-diagnosed patients ($n=14$) with control subjects ($n=12$). All autism-diagnosed patients were male, and had previously had a DSM-IV primary diagnosis of autism. Control subjects were all male and approximately age matched (control average age = 17.08, autism-diagnosed=16.06).

In order to assess the preservation of individual differences, an in-depth investigation as to the accuracy of the BRAIN-MASK and GWMP methods was conducted by assessing the effect of the different normalization methods on the amygdala, an area chosen due to its location near to the medial temporal areas reported to normalize particularly well (Salmond et al. 2002). ROI masks were created by manually tracing the left and right amygdala for each subject. The deformations provided by normalization to the study-specific templates (step V below) were applied to these ROIs. The ratio of amygdala volume to whole-brain volume was correlated between the original ROIs and the warped ROIs.

Processing Path

I. Bias Correction

In order to run a comparable analysis to SPM99 in SPM2b, the method used was functionally identical with the addition of the original scans being corrected using a non-linear image intensity normalization (Bias Correction, with 30mm DCT cutoff) before the original segmentation and normalization within SPM2b.

II. Segment

The original T1-MRI scans were first segmented into white, gray and CSF probability maps based on voxel intensities and used a combination of prior probability maps and a modified mixture model cluster analysis technique involving a correction for image intensity nonuniformity associated with MRI images (Ashburner and Friston 2000). Outlying voxels were then deleted from the segments via erosion and conditional dilation.

III. Extract Brain

A binary (0 or 1) mask was created by setting voxels with non-zero probabilities of being gray- or white-matter to one.

IV. Create Templates

a. Normalize to T1
The original T1-MRI scans were normalized to the T1 template in MNI space and the resulting deformations were applied to either gray and white matter probability maps (GWMP), or to a binary brain-mask (BRAIN-MASK). This normalization procedure consists of: (a) determining an optimum 12-parameter affine transformation from a given image to the template; and (b) performing a nonlinear estimation of deformations (Ashburner and Friston 2000).

SPM99: 4x5x4 basis functions

SPM2b: 30mm DCT cutoff

b. Average images

Normalized images were averaged using SPM's mean function to create separate study specific templates for gray matter, white matter and brain-tissue.

c. Smooth Average

The template images were then smoothed using an 8mm FWHM isotropic Gaussian smoothing kernel, in order to match the hard-coded 8mm smoothing applied to the input images in SPM99's spatial normalization algorithms (Ashburner 2002).

V. Normalize to Templates

The original gray matter, white matter and brain-mask images were then normalized to their respective study-specific templates, and the resulting deformations were applied to the original brain scans, which were then segmented and smoothed (12mm FWHM).

SPM99: 8x8x8 basis functions

SPM2b: 30mm DCT cutoff

VI. Statistical analysis on non-modulated data.

VII. Modulate Images (Preserve Total)

Beyond the original SPM99 processing path, the GWMP and BRAIN-MASK methods were also compared using the preserve-total image modulation feature in SPM2b. This function allows for image intensities to be additive for areas that were contracted or expanded during the normalization procedure (see Figure 3.)

VIII. Statistical analysis on modulated data.

Results and Discussion

Within all tested processing paths the BRAIN-MASK and GWMP methods provided different normalizations and different distributions of t-values. The normalizations were compared based on probability maps of brain tissue across subjects (Figure 2), as well as individual brain comparisons (Figure 4). The probability maps suggest a more accurate normalization for SPM99 than SPM2b, though this is most likely due to the parameters used during the normalization procedure in SPM2b (i.e. note the most accurate minimum of 15mm Discrete Cosine Transfer function). The probability maps suggest that within SPM99 the GWMP method less accurately matches the edges of the brain and areas around the ventricles than the brain-mask method, as can be seen by lowered probabilities in those areas.

I. Statistical Results

For the data set used, no processing path found differences above a corrected $p < .05$ significance level. All methods found differences greater than $p < .001$ uncorrected, though in different areas. (CHECK THIS). Though the differences observed (at a uncorrected $p < .001$) could be due to a heterogeneous subject group, it is important to note they could also be explained by random noise created by the normalization procedure using such small sample sizes as can be demonstrated in Figure 4 (SOURCE).

II. SPM Version

The SPM99 and SPM2b processing provided spatially diverse t-maps when comparing the same original scans. In fact, there was more variability associated with the version of SPM than with the different methods (GWMP or BRAIN-MASK). This variability can most likely be associated with the normalization and segmentation changes in SPM2b. Within the normalization procedure SPM2b seems to provide a less consistent match of brain tissue regardless of which method used (see Figure 2). The segmentation feature in SPM2b has been adapted from SPM99 to rely less on prior-probability maps in order to provide more accurate segmentations for abnormal brains. While this change may provide more accurate estimates of the probability of white and gray matter in Autism diagnosed patients, it also provided for high white and gray matter probabilities for some non-brain tissue in four of twelve control brains (see example in Figure 5).

III. Method Used

a. Whole-Brain

Compared to GWMP, BRAIN-MASK normalization consistently provided a more accurate alignment of the edges of the brain and the ventricles as well as the cerebellum using SPM99 (see Probability Maps in Figure 2). These differences failed to replicate in SPM2b, most likely because of a failure to optimize the normalization parameters to achieve the most accurate alignment. Within each version of SPM the methods showed comparable t-map patterns, with differences in t-statistics ranging approximately 1.5 in either direction. The resulting differing t-maps could be attributed either to typical normalization error seen in GWMP or to errors in not fitting individual gyri within BRAIN-MASK, so it is ambiguous which method is more accurate. The image modulation step served to provide t-maps with higher more localized t-values than the non-modulated procedure, though they still provided for the same t-map patterns.

b. Amygdala Volume

The tests on the ratio of amygdala volume showed high correlations between both BRAIN-MASK ($r = .88$) and GWMP ($r = .90$) methods using SPM2. This correlation did not hold however for SPM99 with BRAIN-MASK ($r = .35$) showing lower correlations than GWMP ($r = .65$) in SPM99. This suggests that the normalization may have fit the edges of the brain and ventricles at the expense of areas between. Though this should not have much of an effect on areas such as cortex, it is of particular concern when dealing with the amygdala, which is squeezed tightly between the edges of the brain and the ventricles. These results suggest the BRAIN-MASK method is not optimal in those situations. This may be only of relative interest, as the most prominent differences in amygdala volume and location after normalization occurred between subjects as opposed to between methods. In fact, the normalized amygdala did not have a uniform distribution across subjects, with the amygdala of some subjects failing to overlap at all (see probability map in Figure 7). It is also of note that both BRAIN-MASK and GWMP methods were more accurate in preserving the amygdala volume when dealing with the controls than when dealing with the autism-diagnosed patients, with the lowest correlation for any amygdala of autism diagnosed patients coming from SPM99 (GWMP(left): $r = .70$, (right): $r = .30$; BRAIN-MASK(left): $r = .07$, (right): $r = .24$), and the best coming from SPM2 (GWMP(left): $r = .93$, (right): $r = .53$; BRAIN-MASK(left): $r = .92$, (right): $r = .92$).

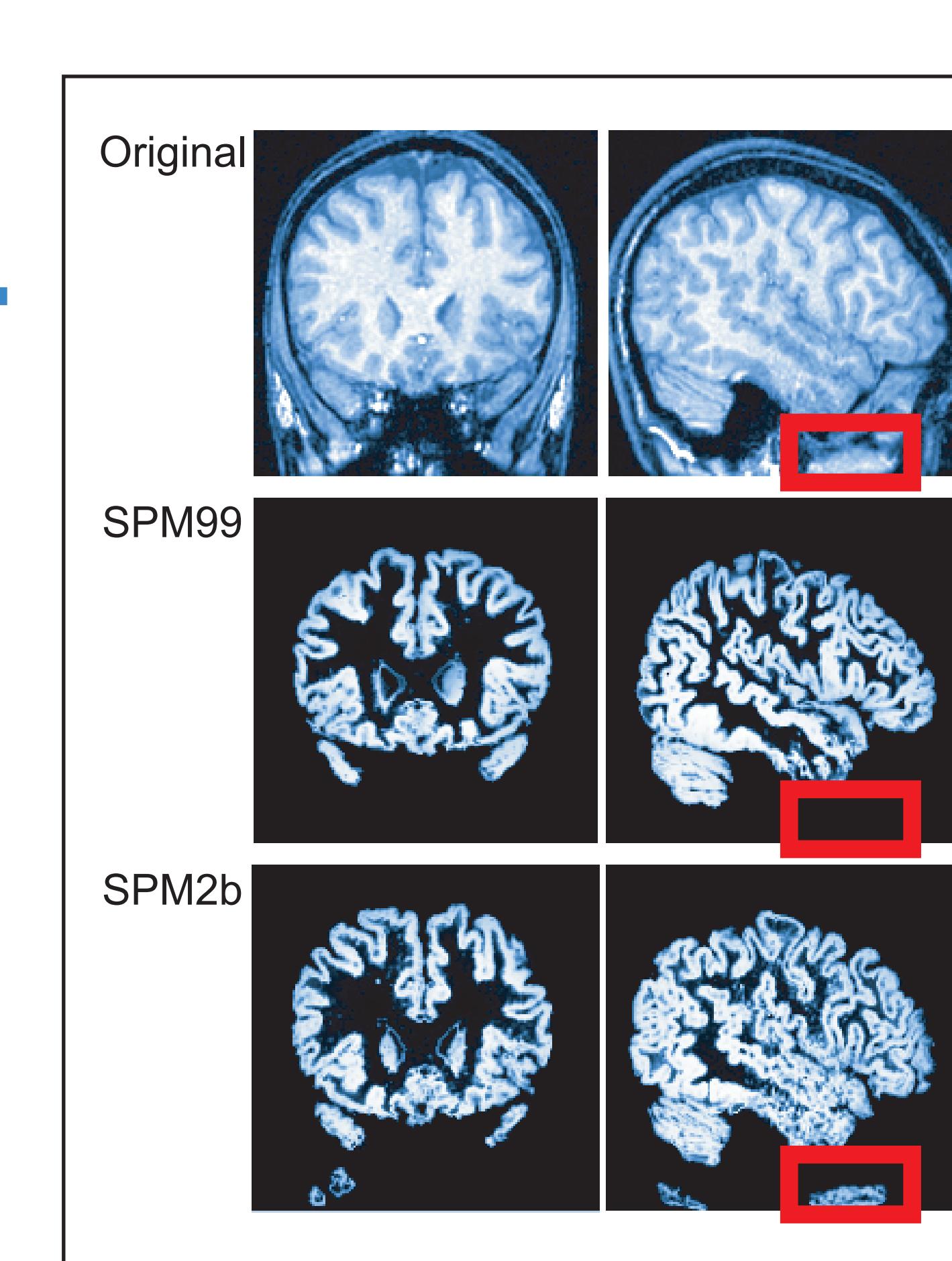


Fig 5. Example of a segmentation error in SPM2b as compared to SPM99. (Based on different Normalizations)

Conclusion

Many of the problems associated with expanding and contracting individual brains, for which BRAIN-MASK was designed to combat, can be theoretically dealt with by modulating segments before statistical analysis. The information provided by the modularized analyses can be useful even when answering questions regarding the concentration of tissue in a particular area, as when compared to the non-modulated analyses it will allow for some assessment of the amount of distortion that occurred during the normalization procedure.

Though the focus of this poster was the normalization procedure, the VBM methodology is equally dependent on the segmentation. Since BRAIN-MASK uses less information to normalize brains into the same sample space it is particularly sensitive to the accuracy of the segmentations and creation of the brain-mask. Researchers need to assure that the segmentations are both accurate and consistent for all subjects to ensure a accurate statistical test.

While additional research is needed to assess the similarities and differences between the two addressed methods, it is unclear that either method provides for an "optimized" VBM. The inconsistent normalizations seen in Figures 2 and 7, and the inaccurate segmentations seen in Figure 5, combined with the statistical experiment noted in Figure 6, suggest inconsistencies in the VBM pre-processing that would lead to a large amount of spatial noise within the segmented images that are being tested, in such a way as to obfuscate the true differences as noted by Bookstein (2001).

When running a VBM experiment it is important for future researchers to carefully inspect both the segmentation and normalization of each image as well as the overlap of both the brain as a whole and structures of interest between subjects, not just looking for errors but also assessing quality.

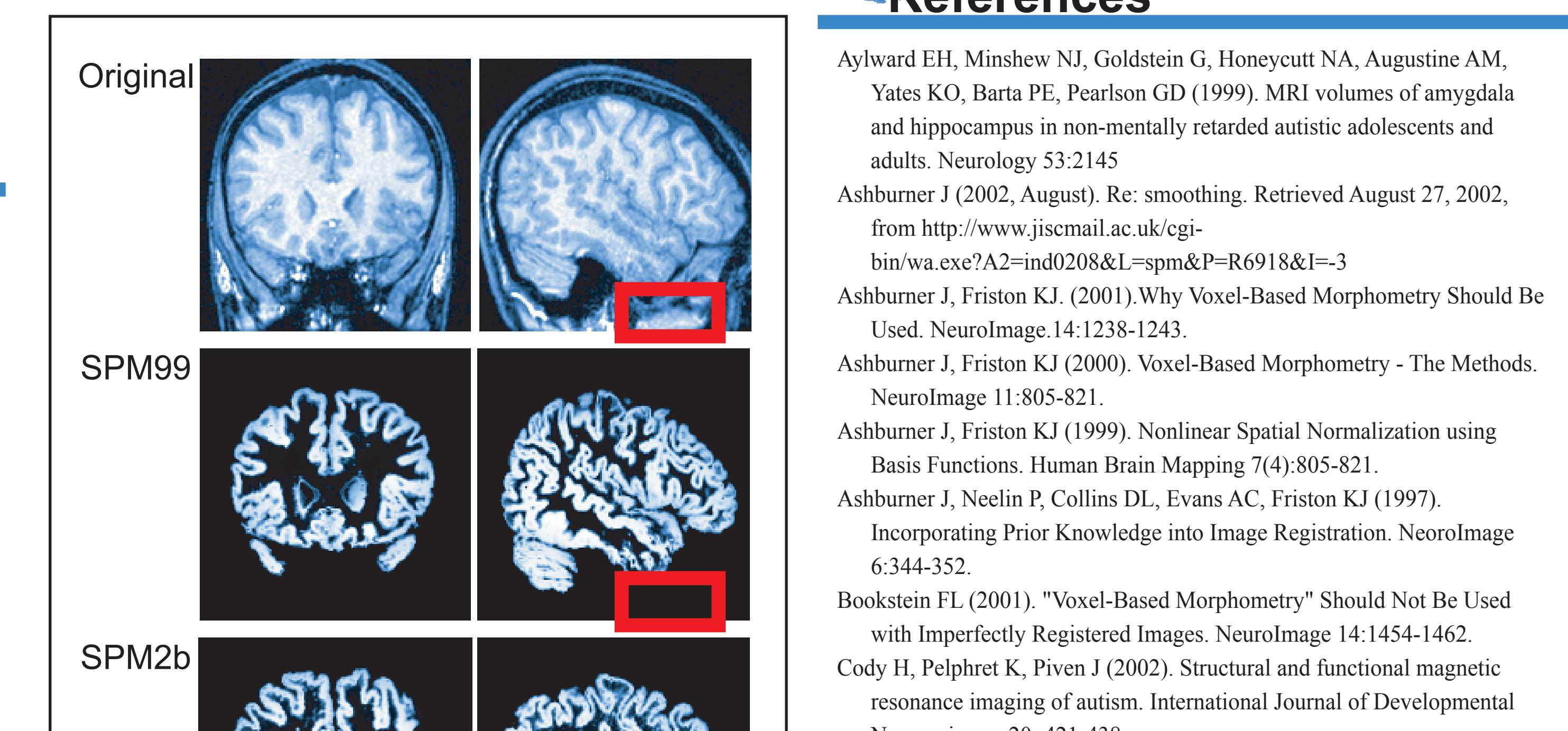


Fig 4. A comparison of gray-matter segments between a template brain and two representative example brains using both GWMP (gray and white) and BRAIN-MASK methods in both SPM99 and SPM2b, where blue signifies tissue only found in the template brain, red signifies tissue only found in the brain being compared to the template, and black signifies overlap.

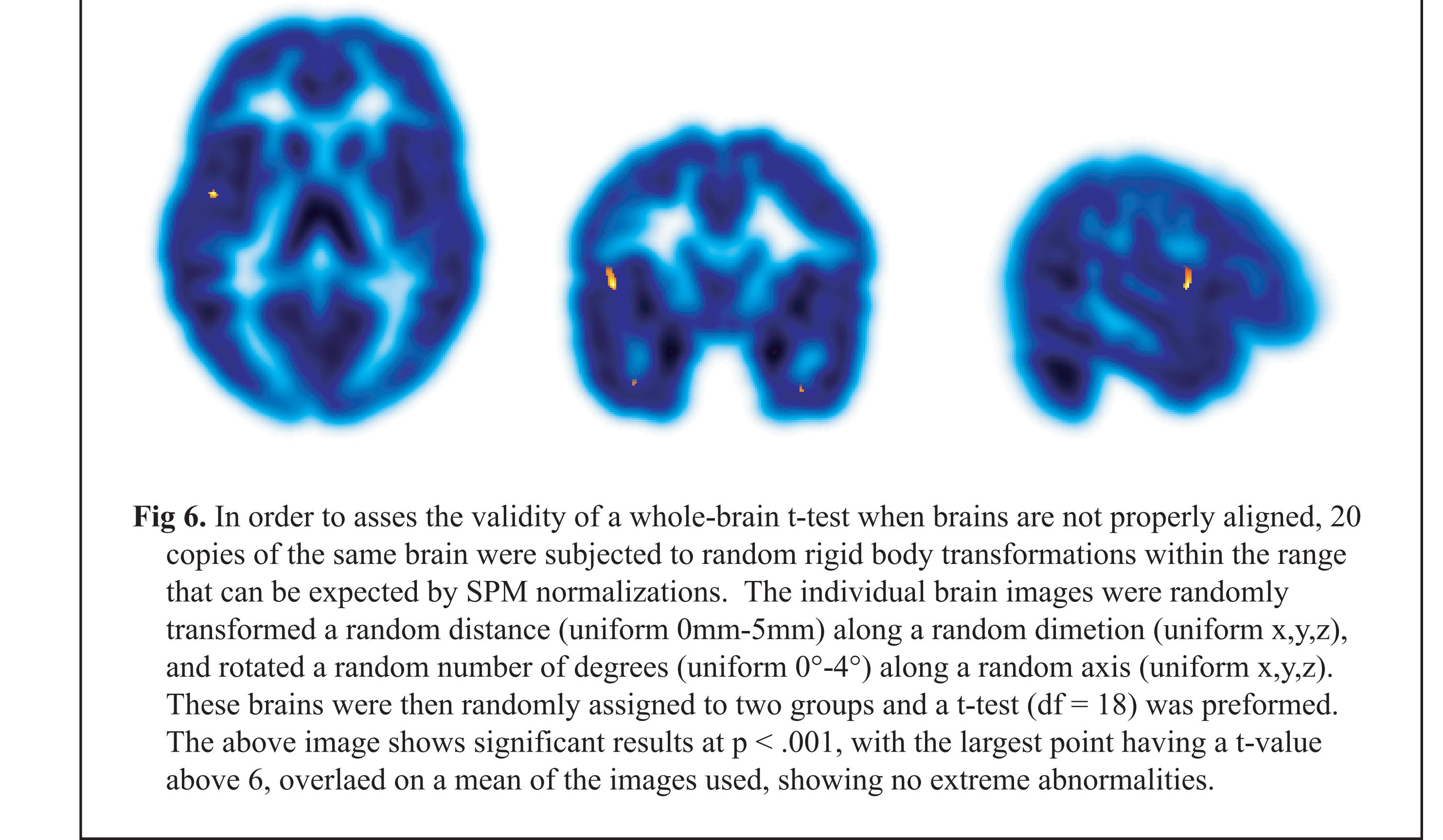


Fig 6. In order to asses the validity of a whole-brain t-test when brains are not properly aligned, 20 copies of the same brain were subjected to random rigid body transformations within the range that can be expected by SPM normalizations. The individual brain images were randomly transformed a random distance (uniform 0mm-5mm) along a random dimension (uniform x,y,z), and rotated a random number of degrees (uniform 0°-4°) along a random axis (uniform x,y,z). These brains were then randomly assigned to two groups and a t-test (df = 18) was performed. The above image shows significant results at $p < .001$, with the largest point having a t-value above 6, overlaid on a mean of the images used, showing no extreme abnormalities.

References

- Aylward EH, Minshew NJ, Goldstein G, Honeycutt NA, Augustine AM, Yates KO, Barr PE, Pearlson GD (1999). MRI volumes of amygdala and hippocampus in non-motorically retarded autistic adolescents and adults. *Neurology* 53:2145.
- Ashburner J (2002, August). Re: smoothing. Retrieved August 27, 2002, from <http://www.jiscmail.ac.uk/cgi-bin/wa.exe?A2=ind2028&L=spm&P=R691&I=-3>
- Ashburner J, Friston KJ (2001). Why Voxel-Based Morphometry Should Be Used. *NeuroImage* 14:1238-1243.
- Ashburner J, Friston KJ (2000). Voxel-Based Morphometry - The Methods. *NeuroImage* 11:805-821.
- Ashburner J, Friston KJ (1999). Nonlinear Spatial Normalization using Basis Functions. *Human Brain Mapping* 7(4):805-821.
- Ashburner J, Neelin P, Collins DL, Evans AC, Friston KJ (1997). Incorporating Prior Knowledge into Image Registration. *NeuroImage* 6:344-352.
- Bookstein FL (2001). "Voxel-Based Morphometry" Should Not Be Used with Imperfectly Registered Images. *NeuroImage* 14:1454-1462.
- Cody H, Pelphet K, Piven J (2002). Structural and functional magnetic resonance imaging of autism. *International Journal of Developmental Neuroscience* 20: 421-438.
- Good CD, Johnstone IS, Ashburner J, Henson RNA, Friston KJ, Frackowiak RSJ (2001). A Voxel-Based Morphometric Study of Aging in 465 Normal Adult Human Brains. *NeuroImage* 14:21-36.
- Salmond CH, Ashburner J, Vargha-Khadem F, Connell A, Gadian DG, Friston KJ (2002). The precision of Anatomical Normalization in the Medial Temporal Lobe using Spatial Basis Functions. *NeuroImage* 17:507-512.
- Tsatsanis KD, Rourke BP, Klin A, Volkmar FR, Cicchetti D, Schultz RT (2003). Reduced Thalamic Volume in High-Functioning Individuals with Autism. *Biological Psychiatry* 53(2):121-129.