Type I Errors in Whole Brain Voxel-Wise Analyses

D. M. Carpenter[1], and C. Y. Tang[1,2]
[1]Radiology, Mount Sinai School of Medicine, New York, New York, United States, [2]Psychiatry, Mount Sinai School of Medicine, New York, NY, United States

Voxel based analyses (VBA), also known as Statistical Parametric Mapping (SPM), are commonly used for exploratory analysis of functional or structural neuroimaging data. The popular technique is relatively easy to use and can survey the entire brain but the method yields inconsistent results (Kanaan 2005). It is difficult to challenge the validity of published works and review submissions because the nature of the immense datasets that underlie VBA results prohibits its presentation in journals. In this abstract we use a simple data set to explore sources of type I errors in areas that often yield positive results and present findings that can serve as a guide for critiquing these SPM presentations.

MATERIALS AND METHODS

Subjects. Two healthy volunteers were recruited and gave informed consent with no history of mental illness. Subjets underwent several imaging sessions over 7 consecutive business days (n=7 for each subject).

Image Acquisition. All imaging sessions were performed on a 3T Allegra MRI scanner (Siemens, Ehrlangen, Germany). The images acquired that are relevant to this study included a turbo spin echo (TSE) T2-weighted Axial (TR = 5380 ms, TE = 99 ms, FOV = 18.3x21 cm, matrix = 512x448, Turbo factor = 11, 28 slices, thickness = 3 mm skip 1 mm); DTI using a pulsed-gradient spin-echo sequence with EPI-acquisition (TR = 4100 ms, TE = 80 ms, FOV = 21 cm, matrix = 128x128, 32 slices, thickness = 3 mm skip 1 mm, b-factor = 1250 s/mm2, 12 gradient directions, 5 averages). Raw DTI data were transferred to an off-line workstation for post-processing. In-house software written in Matlab v2007 (The Mathworks, Natick, MA) was used to compute the anisotropy.

Voxel Based Analysis. Image normalization was performed using SPM2. The non-diffusion weighted image of the DTI was first coregistered to the high resolution anatomical T2 image. Parameters were calculated to normalize the T2 image to the MNI T2 template image that is supplied with SPM2. The normalization parameters were then used to transform the non-diffusion as well as the FA image to the MNI space. The default SPM2 settings were used for both coregistration and normalization. Full width-half maximum Gaussian smoothing kernels of 0mm, 3 mm, 6 mm, 9 mm, and 12 mm were used because smoothing kernel size can vary the results (Jones 2005). Voxel-by-voxel t tests were created to search for significant differences between the subject's FA measures (n = 7 for each subject).

RESULTS

An exploratory VBA t test between the scans of two subjects was performed to investigate how minor misalignment issues can influence the results (Figure 1). The goal of the exploratory analysis is to detect white matter FA differences but a closer visual inspection reveals type I errors. Inspection of areas of high contrast suggested that image misalignments underlie some of the detected differences (see corpus callosum and external capsule of enlarged section of Figure 1). Outlines of the significant clusters were overlaid on the mean of the normalized images from subject A and C (Figure 1 middle and right columns, respectively). The image inspection revealed that the positive findings in these regions were clearly due to image misalignments. This observation remained with a wide range of smoothing kernel sizes. Not all clusters were investigated in this manner but a qualitative inspection is convincing that high contrast regions are susceptible to error using this method.

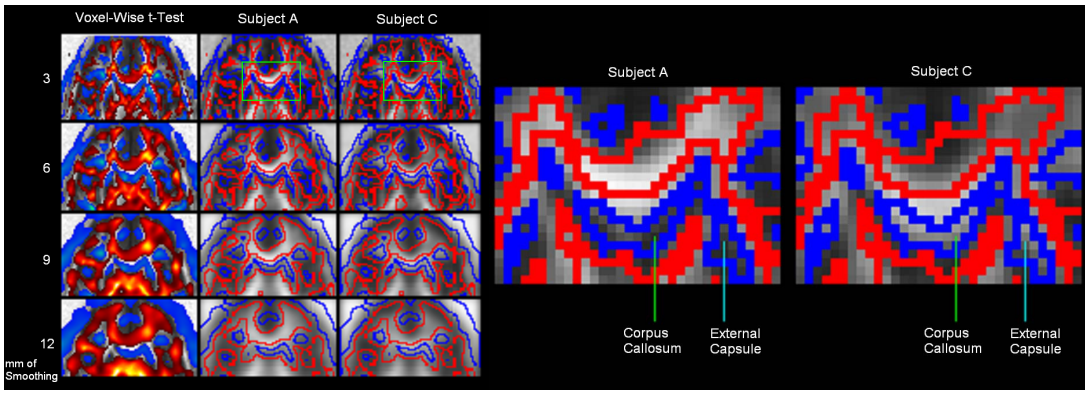

Figure 1. The columns of images (left to right) show t values ($p < 0.01$), and outlines of positive and negative clusters on subject A and subject C. The green boxed areas are enlarged on the right to show misalignments of the corpus callosum and external capsule.

DISCUSSION

In this study we investigated the potential for erroneous conclusions to be drawn from VBA images. The most striking findings were the type I errors in regions that are commonly reported in DTI studies (Kanaan 2005). Several regions were clearly due to image misalignments that resulted in comparisons between white matter and non-white matter regions that, due to tissue differences in FA, yielded positive findings.

This study's findings by no means render all VBA results errors but offer a guide to identify suspicious results. Particular scrutiny should be given to high contrast areas, such as ventricle white matter borders. Special attention should be given to these areas in studies with known anatomical differences between groups, such as ventricular variations in schizophrenia (Keefe 1998). The results presented here should educate authors, readers and reviewers in order to scrutinize findings in the VBA-DTI literature.

References

Kanaan, R.A.A., Kim, J.S., Kaufmann, W.E., Pearlson, G.D., Barker, G.J., McGuire, P.K. Diffusion Tensor Imaging in Schizophrenia. Biol Psychiatry, 58, 921-929

Jones, D. K., Symms, M.R., Cercignani, M., Howard, R.J. (2005). The effect of filter size on VBM analyses of DT-MRI data. NeuroImage, 26, 546-554

Keefe, R. S., and Powchik, P. (1998). Ventricular enlargement in poor-outcome schizophrenia. Biol Psychiatry 43, 783-793.