Comparison of Methods for Automated Lesion Identification in Stroke Patients

Anil Ramlackhansingh

MSc. in Clinical Neurology

July 2007

Supervisor: Prof. Cathy Price

Word count (excludes Reference List and Appendices): 9,804

MSc Clinical Neurology
2006/07

UMI Number: U593848

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI U593848
Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Acknowledgements

I would like to thank to Prof. Cathy Price and Dr. Mohamed Seghier for their excellent guidance and support during this entire project. I would like to especially thank Dr. Seghier for his help though out.

Special thanks also to the Education Unit and all my colleagues at Queen Square. My time at the Institute of Neurology has been of immense benefit and will stay with me for the rest of my life.

Special thanks also to my wife, family and friends who have given me the strength and reason to carry on studying.

Table of Contents

1 Abstract

1.1 Introduction

Lesion – Symptom mapping forms the foundation to our understanding of the function of different parts of the brain. As the science of neuro-imaging has developed, the detail and quality of images produced on scanning patients have improved. The result is that a single scan can produce a wealth of data. The manual analysis and interpretation of this data have thus become a time consuming and laborious affair. Manual analysis also has the drawback of being highly subjective and user dependent. This has resulted in the development of automated methods for the interpretation of such data. Before data could be analysed, it must be pre-processed into a format that can be optimally used by these methods. This step involves the normalisation, segmentation and spatial smoothing of patient scans. Scans must be pre-processed before automated analysis can be done. The use of automated methods is supposed to make data interpretation quick, reliable and reproducible.

1.2 Aim

In this study the smoothing aspect of pre-processing will first be looked at. Here different smoothing levels will be used to try and decide on a value for optimal smoothing. Once this is decided analyses using three automated lesion identification methods will be done. The methods include Voxel Based Morphometry (VBM), Posterior Probability Mapping (PPM) and Fuzzy Clustering with fixed Prototypes (FCP). Parameters for the optimal functioning

of these methods will be determined and suggestions will also be made as to which method may be the best and under what circumstances.

1.3 Results

The optimal Full Width at Half Maximum (FWHM) value for smoothing was found to be 8mm. When the methods were looked at the following could be said:

VBM:

- Analyses produced results in all 5 patients and
- Results only produced one level of information.

PPM:

- The optimal probability threshold was determined to be 0.5
- Analysis provides two levels of information.
- Lesions were identified with sharper borders and
- Analyses required the longest computer processing time.

FCP:

- An α value of 0.5 provides the best results.
- Lesions identified with less sharp (i.e. fuzzy) borders – method possibly more sensitive and
- Analyses required the least computer processing time.

1.4 Conclusion

This study provides further evidence that smoothing must be carried out on all scans to enable accurate and reliable lesion identification. An optimal FWHM of 8mm was determined.

The study also determines an optimal value of 0.5 for the probability threshold in PPM analysis. An optimal α value of 0.5 was also found for FCP analysis. VBM proved to be the easiest method to use while PPM estimated the confidence to declare a tissue as abnormal and FCP was the most sensitive to lesion presence. Further work, however, should be done to further investigate the probability threshold for PPM and α value for FCP.

2 Introduction

2.1 Background

Lesion – Symptom mapping studies have been integral to the development of insight into the functioning of the human brain (Bates et al. 448-50). With the advent of computer based imaging techniques (Computerised Tomography and Magnetic Resonance Imaging) the ability to investigate patients in vivo was realised. This was again forwarded with the development of functional MRI. MRI allowed non-invasive imaging at a high spatial resolution with lesions and other types of impaired tissue shown as atypical changes in signal density.

The localization and sizing of lesions was first achieved by manually tracing the volume of interest (VOI) by a human reporter (Mehta et al. 1438-54). This manual approach has many drawbacks. It is a labour intensive and time consuming task. It is highly subjective and user dependent. It also requires a detailed knowledge of neuroanatomy.

In the 1990's computational programs were developed that could identify lesions automatically. At the Wellcome Trust Centre for Neuroimaging (http://www.fil.ion.ucl.ac.uk/spm/), Statistical Parametric Mapping (SPM) was developed to enable just that. This program was written using the **mat**rix **lab**oratory (MATLAB) programming platform. Currently the most up-to-date version of SPM is version 5.

Using SPM, different automated lesion identification methods have been developed. In order for these methods to function, images must first be prepared or pre-processed into a format that can be analysed by the program (SPM in this case) (Ashburner and Friston 1238-43). It should be noted however that pre-processing is not perfect. It can result in the introduction of error into an image and a compromise must be met to limit this effect. In order to limit error, smoothing has also been introduced into the pre-processing process.

2.2 Automated methods

Automated methods were developed in an attempt to make analysis easier, standard and more reliable. The methods rely on the fact that a lesion can be thought of as essentially abnormal tissue. This abnormal tissue composition means that its magnetic properties differ from that of normal brain tissue. Thus when the brain is imaged with MRI, the voxels corresponding to the lesion differ to some extent from voxels corresponding to normal brain tissue. It is this difference that automated methods detect and display. Distinguishing between abnormal and normal voxels can however prove difficult. This can result in potentially normal voxels being classed as abnormal and visa versa. This is most evident around the borders of a lesion where the difference between normal and abnormal can become blurred.

Automated methods can be classed in two categories according to the nature of MRI images used as input:

1. multi-channel or multi-spectral methods that necessitate several weighted MRI images, including T1, T2, PD, FLAIR images with or without contrast agents (Wu et al. 1205-15).
2. mono-channel or mono-spectral methods that deal with lesion identification on only one type of image (e.g. if only T1 weighted MRI images are available) (Stamatakis and Tyler 167-77).

In this study, three different automated methods of lesion identification in a mono-spectral mode are compared. Thus only one image type is used for analyses with one T1 weighted MRI image per patient.

The three methods to be compared include:

- Voxel Based Morphometry (VBM),
- Posterior Probability Mapping (PPM) and
- Fuzzy Clustering with fixed Prototypes (FCP).

They each quantify lesions in a different manner. The aim of this study was to compare these different methods in an attempt to decide which method is most reliable and / or gives the most information in a given situation. Factors that affect the sensitivity and specificity of each method were also varied in an attempt to find the best possible conditions for the functioning of the different methods. A secondary aim of the study was to determine what level of

smoothing would be most useful for lesion identification. Smoothing is part of data pre-processing and will be discussed further below.

Further explanation of the science of lesion identification and the pre-processing of data will now be considered.

2.3 Data pre-processing

Before images are analysed attempts are made to make them as standard as possible. This step involves the pre-possessing of images. This process limits the influence of inter-individual anatomical variability and corrects for sampling error created during image acquisition (Mechelli et al. 1-9;Bates et al. 448-50). It will be discussed further now.

2.3.1 Spatial Normalisation

Everyone's brain is different. There are small differences in anatomy and volumes clearly vary with head and body size. However gross anatomy generally remains the same. When using automated methods, the patient's brain is compared to an age-matched bank of controls. Thus it is important that only similar parts of the brain are compared (Ashburner and Friston 805-21). Thus, for example, the patient's right cerebellar hemisphere is compared to the controls' right cerebellar hemispheres and not the pons. It can thus be seen that if part of a patient's brain not properly aligned, artefactal differences will arise. Conversely if the brains of control subjects are not aligned they will increase inter-subject variability. This will result in false positives (Mehta et al. 1438-54;Crinion et al.).

During spatial normalisation, the scan is registered on to a standard template aligning the scan and making comparison possible. This template is obtained from the average of several scans. The presence of lesions however makes normalisation more difficult (Crinion et al.). For example, the presence of a stroke prevents subject scans from lining up well with the template. Templates are thus chosen that match up with the subject scan as much as possible (Davatzikos et al. 1361-69). Several templates may also be used in the normalisation of one scan. The registration algorithm (which is responsible for normalisation) selects the best possible templates to use for each study. Thus different templates may be used in the normalisation of different studies (e.g. adult, child, Alzheimer's disease templates are all available). Once a template is chosen however, it must be held constant within a study.

In this study the unified segmentation algorithm for normalisation was used. This method has been shown to be more effective than previous normalisation approaches (Crinion et al.). At the end of normalisation the subject's scan is scaled and warped. To limit this effect, it is thus important that the best possible method and templates are used for normalisation. Normalisation is not meant to be perfect but rather minimises the anatomical differences between subjects. It is also true that without normalisation these automated methods would not function.

2.3.2 Segmentation

Segmentation aids normalisation by further increasing the accuracy and sensitivity of the automated methods (Mehta et al. 1438-54;Ashburner and

Friston 805-21). During segmentation the patient's scan is broken up into its constituent parts. These include (Davatzikos et al. 1361-69;Ashburner and Friston 805-21):

1. Grey Matter,
2. White Matter and
3. Cerebrospinal Fluid.

The voxel signal intensity of each of these constituent parts varies on MRI. The separation of the scan into these constituents is based on this variation. Modern segmentation techniques also correct for signal intensity variation that arise because of the non-uniformity of conditions during MR image acquisition. Over time, several segmentation algorithms have been developed. Most utilise a Gaussian mixture model based on global and / or local image intensity for tissue classification (Mehta et al. 1438-54;Ashburner and Friston 805-21). This study used the unified segmentation algorithm which iteratively combines the segmentation and normalisation process.

The same procedure is necessary for both patients and controls. Thus it ensures that, for example, the grey matter of the patient's right cerebral hemisphere is compared to the grey matter of the controls' cerebral hemisphere. In this manner, unified segmentation goes one step beyond normalisation, with comparisons limited to voxels that should have similar signal intensity. It also allows the likelihood of abnormal tissue to be calculated.

Thus segmentation generally results in a single scan being split into three. Different methods / algorithms have developed for segmentation (Bezdek, Hall, and Clarke 1033-48;Clark et al. 730-42).This study focused on the Grey and White matter constituents produced by the unified segmentation algorithm. Before comparing these images, the last step of pre-processing requires spatial smoothing which will now be discussed in more detail.

2.3.3 Smoothing

Smoothing is normally the last pre-processing step. After smoothing important patterns in the data set are captured while noise is minimised. It serves to minimise the error introduced during normalisation and segmentation. It reduces noise while retaining as much detail as possible in the image. It functions to:

1. Render the data so that it becomes more normally distributed. This helps to ensure that the assumptions underlying the theory of Gaussian random field are met thereby increasing the validity of the parametric statistical tests.
2. Help compensate for the inexact nature of spatial normalisation.
3. Ensure that each voxel in the image is an average of the grey or white matter from around the voxel in the unsmoothed image (Mechelli et al. 1-9).
4. Improve the signal to noise ratio.
5. 'Cushions' against the imperfections in spatial normalisation introduced within the data set (Jones et al. 546-54).

Smoothing is typically achieved with a Gaussian random filter in SPM5. The Gaussian filter is a 3D convolution operator. It acts as a mean filter using a Gaussian hump as a kernel (Fisher et al.). An example of such a hump is shown in figure 1. The calculation of this Gaussian distribution is actually performed using the voxels of the scan. The next step in smoothing is to decide on the size of the filter to be used when smoothing with the Gaussian kernel. The size of the filter is expressed in terms of the Full Width at Half Maximum (FWHM) of the kernel. FWHM is an expression of the extent of any given function. Figure 2 illustrates what is meant by FWHM. This approach has been used for some time in engineering and biological applications.

Data are first processed with a single low pass filter (using the Gaussian kernel). The FWHM of this filter can be varied. Previously, studies using diffusion tensor MRI images have determined the affect that varying filter size has on image analysis (Foong et al. 333-36). Filter sizes as large as 14 mm and 16mm have been used (Wu et al. 1205-15;Jones et al. 546-54). For MRI scans however a general 'rule of thumb' exists. It implies that the smoothing kernel should be at least 2-3 times the voxel dimension. In this study, unsmoothed images and three different filter sizes were compared. The selected filter sizes were 4mm, 8mm and 12mm. An example of the result of smoothing is shown in figure 3. These images demonstrate the increasing fuzziness that occurs with smoothing.

2.3.4 Modulation

Modulation attempts to correct for the volume changes that occur during normalisation (Mehta et al. 1438-54). Unmodulated images represent local

grey matter densities while modulated images represent the absolute amount of grey pattern present. Modulation, however, is an optional step in pre-processing. Modulation may cause difficulties with further analysis and can be associated with poorer lesion detection in stroke patients (Mehta et al. 1438-54). In this study, unmodulated images where thus used.

With pre-processing completed, analysis of image localisation can proceed. It is appropriate now to consider how these different methods work.

2.4 Automated Image Analysis Methods

2.4.1 Voxel Based Morphometry (VBM)

In VBM a voxel-wise statistical analysis is performed (Ashburner and Friston 805-21;Ashburner and Friston 1238-43). A general linear model is employed for this. This is a flexible framework allowing different statistical tests to be performed (Mechelli et al. 1-9). The t test and F test are commonly used in statistical routines of SPM5. The VBM technique using SPM software was the first to be developed and is thus referred to as classical. The result of the analysis is a statistical parametric map (SPM) of t or F statistics. It shows where in the brain the null hypothesis can be rejected at a given confidence.

Graphical representation of the SPM uses coloured voxels on an MRI brain 'slice.' Different colours or gradients of colours can be used to identify areas with different statistical significance.

2.4.2 Posterior Probability Mapping (PPM)

PPM relies on Bayesian inferences when deciding on the probability that a given voxel is abnormal. In Bayesian inference, two levels of calculations are performed. The evidence or observation of the first calculation is used to update the subsequent calculation. It is the results of this second calculation that tells us the probability that a given hypothesis is true. Thus the results of the first calculation act as the priors in the second calculation (Friston et al. 465-83;Friston and Penny 1240-49).

PPM analysis thus takes part in two stages. In the first stage VBM analysis is carried out on all voxels. This initial analysis identifies areas where potentially abnormal voxels may be found. This is followed by the second analysis. During this analysis a probability threshold is set. It is this probability threshold that distinguishes PPM analysis from VBM analysis. This probability threshold represents the minimum probability that a voxel is abnormal. It can be varied between analyses. The second analysis is again a VBM analysis with the results of the first analysis acting as priors. Thus while in classical VBM analysis the threshold is fixed, in PPM, the threshold is varied according to the variability identified in the first analysis. For example, in areas of high prior variability, the classical thresholds are relaxed avoiding false negatives. The result of both stages of calculation is the generation of a probability distribution map (Friston et al. 465-83). This map displays the abnormal voxels for a given probability threshold.

Analysis can then be repeated using another probability threshold. The results can again be displayed on MRI brain 'slices' with different colours representing different probability thresholds. Results can thus be easily interpreted. This technique also provides two levels of information. One; it locates statistically abnormal voxels and two; it can calculate the probability that these voxels are actually abnormal.

2.4.3 Fuzzy Clustering with fixed Prototypes (FCP)

FCP functions differently from VBM and PPM because the analysis does not rely directly on statistical test. Several concepts need to be explained before FCP is fully described. First the idea behind clustering is considered. In clustering, similar data elements are grouped together into clusters or classes. Clusters are thus as dissimilar as possible. In fuzzy clustering, data elements can belong to more than one group. Each data element must have a membership level associated with it. The membership level defines the degree of membership that data element has to the group.

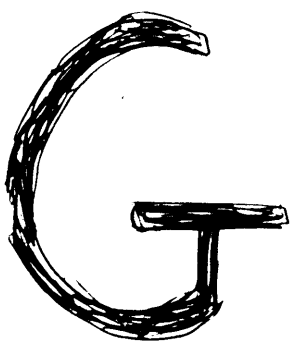
Secondly we have to consider the idea of being an outlier. A group of data elements will have a mean. If one of the data elements varies significantly from the others, the effect is to move the mean towards that data element. This data element can be thought of as an 'outlier.' Outliers have the effect of inflating the variance of a data group and moving the mean towards the outlier. Outliers can thus skew the true group effect and various techniques have been employed in the past to remove outliers. FCP however relies on the identification of these outliers (Seghier, Friston, and Price). For this study

outlier voxels were considered as belonging to the stroke lesion of patient scans. Outlier voxels will thus show less signal intensity when compared to the corresponding voxels in normal subjects. In calculations related to this study the amount (or distance) that the outlier lies away for the mean effect of all voxels is represented by D. U represents the membership value of a voxel to the cluster. The entire process is outlined in figure 4.

At this stage the concept of α is introduced. Alpha acts as a 'tuning' parameter. It is one of the parameters involved in the calculation of D. Thus varying α will have a direct effect on the value of D. Increasing α results in the "smoothing" of values of D. Alpha can thus be thought of as a "smoothing" parameter varying the range of values for D. Increasing α increases the smoothness of the distribution of D while decreasing α has the opposite effect. By increasing the smoothness of distribution of D the distance from the mean effect of the voxels is decreased (see graph on figure 4). As recalled, in FCP, the detection of outliers relies on the value of D being very different from the mean effect of all voxels. With α increasing, the ability to pick up such large variations decreases. Thus the sensitivity of the method decreases. However the specificity of the method will increase as we can be certain that the outliers identified when α smoothes D do in fact lie far from the mean. Varying α thus provides a convenient way for varying the sensitivity or specificity of FCP. This can possibly prove to be one of the strengths of this method. For example when analysing a patient with a small stroke, it may be useful to use a value of α that increases sensitivity. However patients with large lesions can be investigated using larger values of α making the analysis more specific.

2.5 Summary of Methods used

These three methods represent a range of techniques that are currently used for lesion identification. There is no as yet published data comparing all three methods. Thus the aim is to make some suggestions as to which method may be the best. It may however become apparent that the different methods may be most suitable under different conditions but this too needs to be determined.



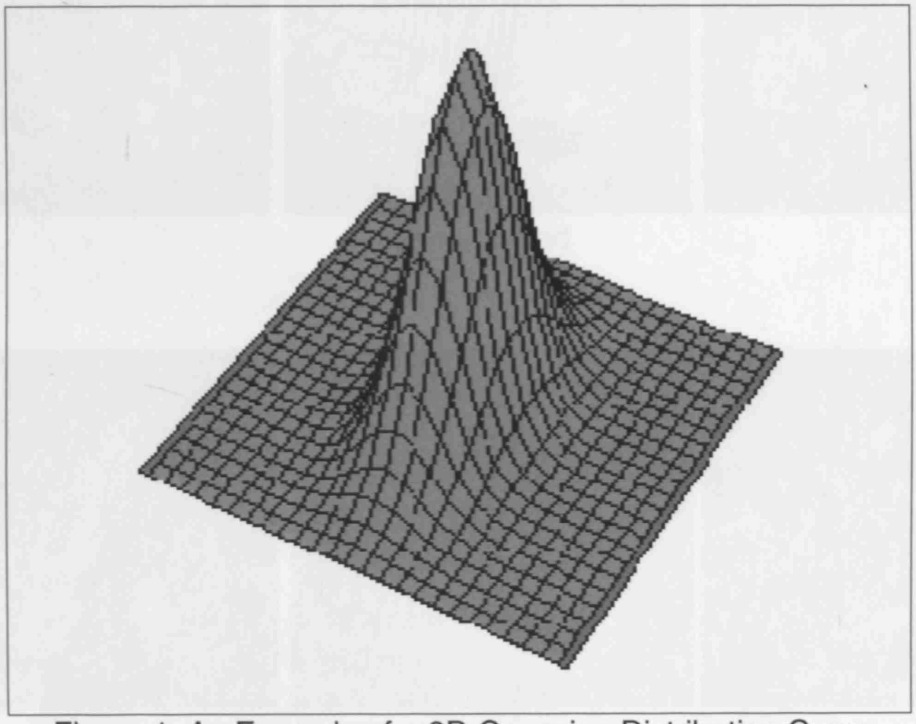


Figure 1: An Example of a 3D Gaussian Distribution Curve

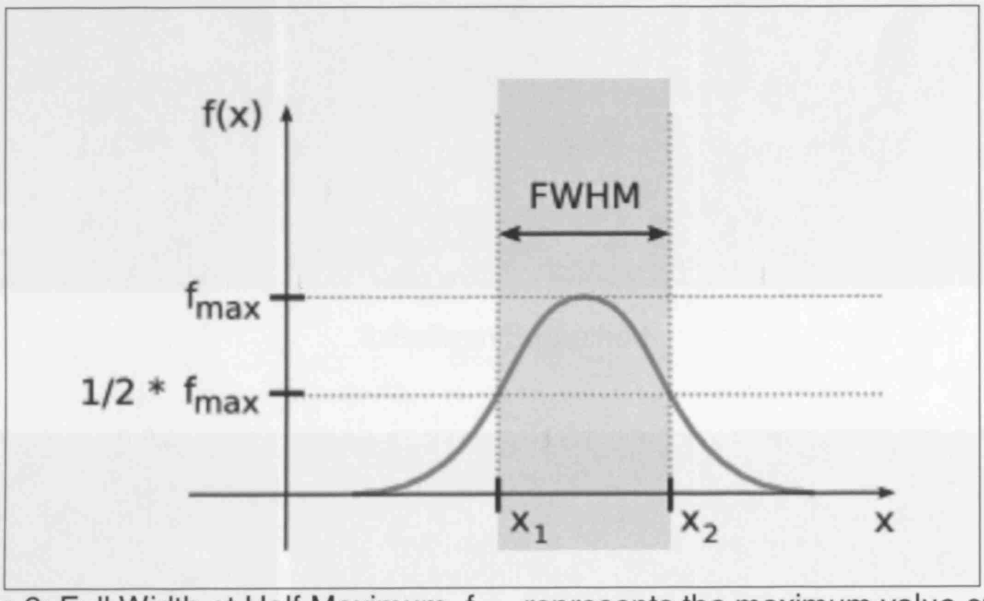
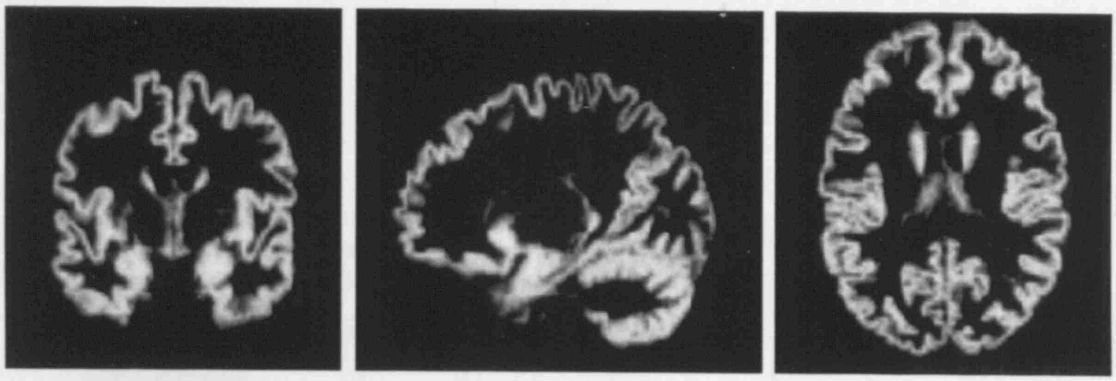
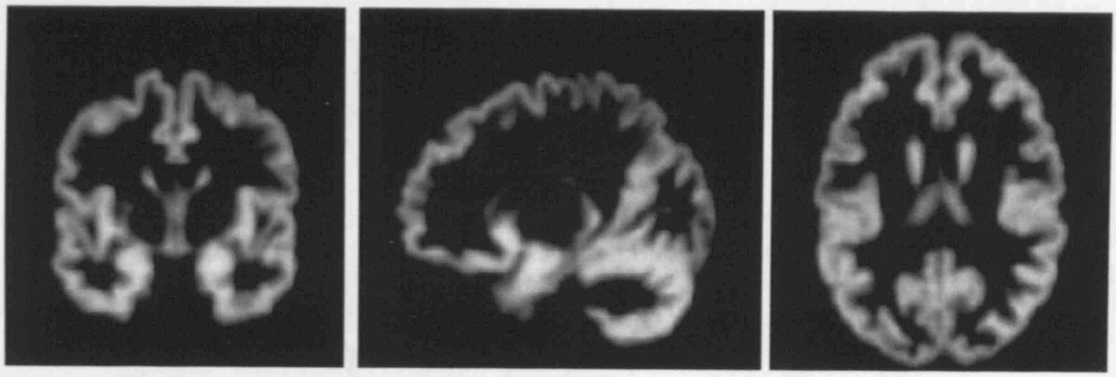


Figure 2: Full Width at Half Maximum, f_{max} represents the maximum value of the distribution curve, x_1 and x_2 are the limits of the FWHM filter.

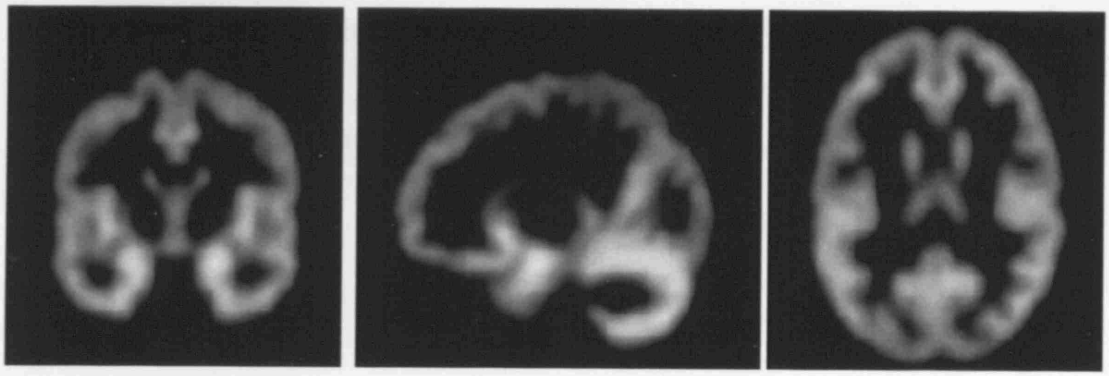
Figure 3: Effect of Smoothing with a Gaussian random filter on T1 segmented Grey Matter



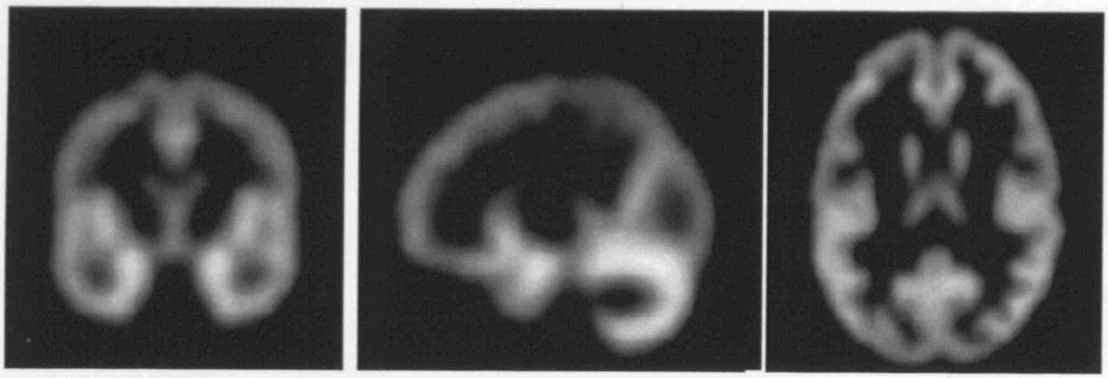
Unsmoothed



4x4x4mm Smoothed



8x8x8mm Smoothed



12x12x12mm Smoothed

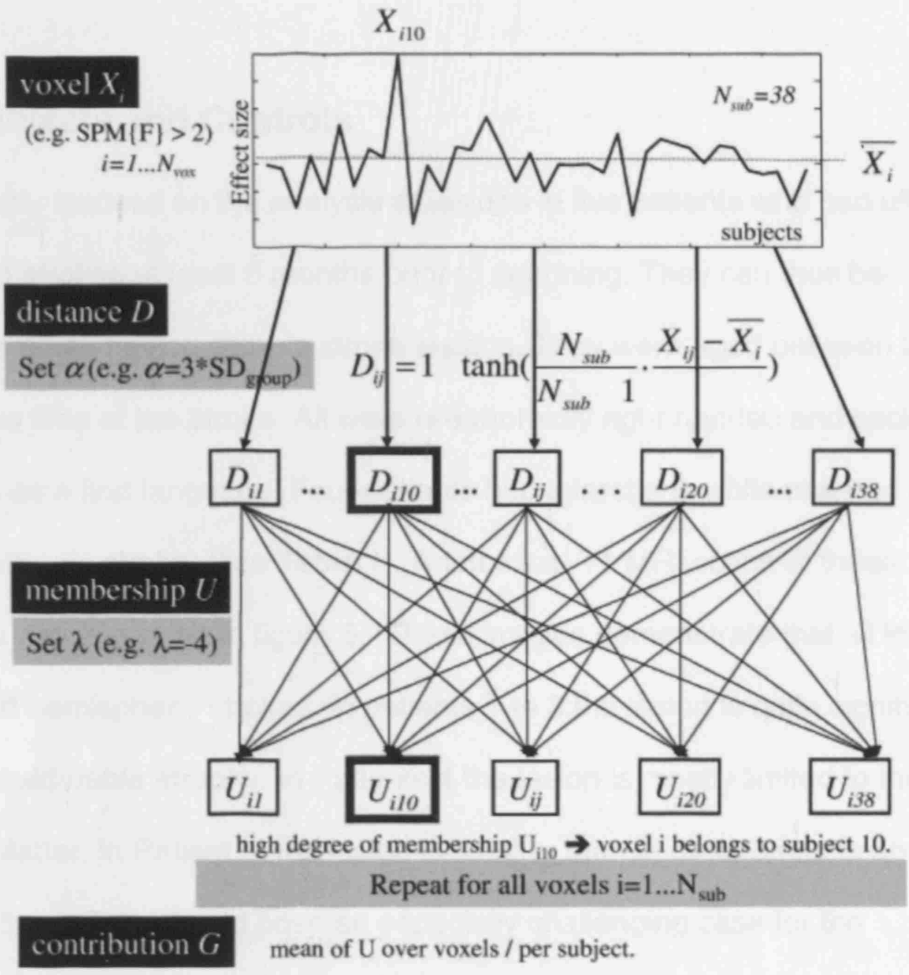


Figure 4: An overview of analysis used in FC
P to identify abnormal voxels (Seghier, Friston, and Price)

3 Methods

3.1 Subjects and Controls

This study focused on the analysis of lesions in five patients who had all suffered strokes at least 6 months prior to scanning. They can thus be considered as having chronic stroke lesions. They were aged between 23 and 68 at the time of the stroke. All were pre-morbidly right handed and spoke English as a first language. Four patients had infarctions while one had a haemorrhagic stroke, (see Table 1). Anatomical T1 MRI scans of these patients are illustrated in figure 5. These images demonstrate that all lesions were left hemispheric strokes. In patients 1 to 3 the lesion is quite significant with considerable atrophy. In Patients 4 the lesion is mostly limited to the White Matter. In Patient 5, the lesion is smaller than in other patients and less well defined. This should pose an especially challenging case for the automated methods. It should be noted however that posterior circulation strokes were not represented in this group.

Sixty-four controls were obtained from a bank of MRI scans from the Wellcome Trust Centre for Neuroimaging. These controls were age-matched, normal healthy right-handed individuals with no neurological disease. Individuals were aged 21 – 75 years and all spoke English as their first language.

3.2 MRI acquisition and pre-processing

Image acquisition was from a 1.5 Tesla Sonata Siemens MRI scanner using a 3D GRE T1-weighted sequence (176 axial slices, 1mm thick with no gap, 1x1mm in-plane resolution).

Image pre-processing was conducted using SPM version 5. Normalisation and segmentation were based on the unified segmentation algorithm. The resulting Grey Matter and White Matter were used for analysis.

Before investigation of the automated methods, the effect of smoothing FWHM on VBM analysis was assessed. Smoothing was thus carried out on the Grey Matter images of all patients at three FWHM values:

1. 4mm x 4mm x 4mm,
2. 8mm x 8mm x 8mm and
3. 12mm x 12mm x 12mm.

The resultant images were used in this first stage of analysis.

At the end of this stage a consensus would be met on an FWHM that is optimal for smoothing. Using this value, the White Matter images were also smoothed. The result of this stage of pre-processing was thus Grey and White Matter images smoothed to an optimal FWHM. These images were then used in the investigation of the automated lesion identification methods.

3.3 Data analysis

All analysis was performed using the SPM version 5 software package. Analysis, as noted above, involved two stages.

First analyses used the Grey Matter images produced from smoothing at different FWHM. VBM analysis was performed on these Grey Matter images at each FWHM of smoothing. During this analysis voxels of the Grey Matter images of each patient are compared to anatomically comparable voxels of all 64 controls. Comparison takes the form of a two-sample t-test. For the t-test a threshold of $p<0.001$ (uncorrected) was applied. The result was the generation of t maps by assessing the contrast "patient < controls." The results were displayed onto 'glass brains' and on the standard T1 image of SPM5. The result was the generation of the images of lesions at different levels of smoothing for each patient.

These results were then reviewed to determine the optimal FWHM of smoothing necessary to give satisfactory results. Using this value, further analyses were conducted using VBM on White Matter. Grey and White Matter analyses were then carried out using PPM and FCP.

The VBM analysis of White Matter was completed in a similar fashion to Grey Matter analysis but using only the optimal smoothing.

PPM analysis followed the VBM analysis. As noted in the introduction, PPM analyses can use different thresholds or levels of probability. Using the

optimal smoothing, the images were analysed at three thresholds – high (0.9 or 90% probability), middle (0.5 or 50% probability) and low (0.3 or 30% probability).

FCP was then performed. As noted in the introduction, α is a tuning parameter. It is integral to the calculation of D which in turn is needed in the identification of outlier voxels. Increasing α deceases the sensitivity (while increasing the specificity) of the FCP method. Thus, using the optimal smoothing, analysis was repeated at three levels of α. Two of the values of α were calculated according to the standard deviation (σ) of the whole data (voxel) set. The following values were thus used:

1. $\alpha = 2.5\ \sigma$,
2. $\alpha = 3.0\ \sigma$ and
3. $\alpha = 0.5$.

The result of these analyses was a bank of Statistical Parametric Maps for both White and Grey Matter. This bank consisted of:

1. Different FWMH of smoothing for VBM Grey Matter
2. Using optimal smoothing, VBM White Matter, PPM and FCP analyses
3. Three different probably thresholds for PPM analysis, and
4. Three different values of α for FCP analysis.

3.4 Results comparison

The next step was the comparison of the images obtained from the analyses.

Firstly the issue of smoothing was resolved by reviewing images at different FWHM. As indicated these images were created using VBM analysis and Grey Matter images. The result of this first stage was a decision on an optimal smoothing value.

Secondly the results of the different methods were compared.

For VBM analysis, Grey Matter maps were displayed on the standard T1 MRI brain image of SPM5 at different levels of smoothing. White Matter maps were displayed on T1 MRI brain slices.

For PPM, Grey Matter images were again displayed on the standard T1 MRI brain image of SPM5 and White Matter displayed on T1 MRI brain slices. The true strength of the PPM analysis however lies in the ability to vary probability. For each patient the images produced at different thresholds of probability were thus displayed onto a single MRI brain (for Grey Matter) or MRI slices (for White Matter). This enabled us to visualise the effect that varying probability had on results.

If there was difficulty with analysis at a certain probability threshold, this probability was excluded. The images created thus consisted of only two levels of probability.

FCP images could not be displayed on the SPM T1 MRI 'glass brain.' MRI slices were thus used. The images were created using the optimal smoothing value and the three different values of α. In order to decide on an optimal value of α, results from each patient were displayed onto a single MRI brain slice in a similar fashion as in the PPM analysis. It was thus possible to visualise the effect that varying α had on the analysis.

The resultant images were then visually compared. As noted earlier VBM results helped decide on an appropriate FWHM of smoothing. PPM probability images were used to determine the best possible threshold that could be used for this method. FCP images were used to determine the best value for α.

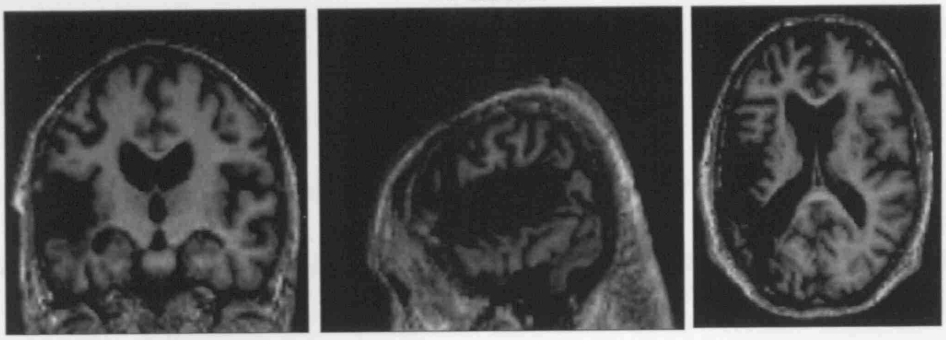
Finally the results of the different methods were compared. Their ability to locate lesions with as little noise as possible was determined visually. The best possible method for analysis could then be suggested. It should be noted however that different methods might hold different areas of strength. This too could be suggested by how the methods coped with different lesions.

Table 1: Characteristics of subjects scanned

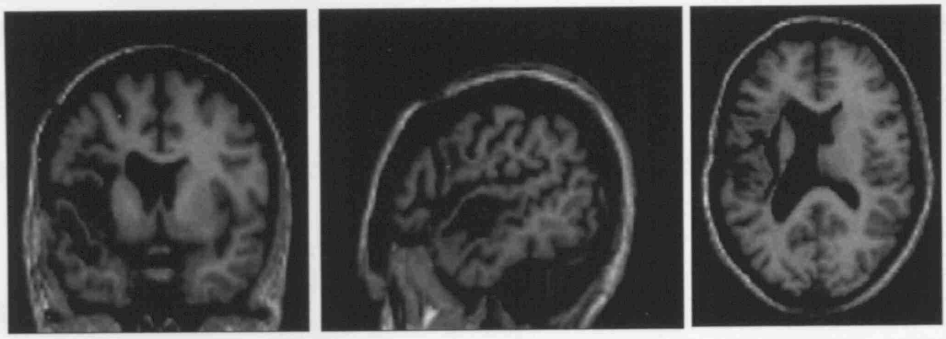
Sex	Stroke type	Age at stroke (years)	Timing scanning after stroke (years)
M	Infarction	68	8
F	Infarction	34	7
M	Infarction	47	0.5
F	Infarction	23	10
M	Haemorrhagic	60	4

Figure 5: Anatomical images of Patients 1 to 5, T1 MRI scans

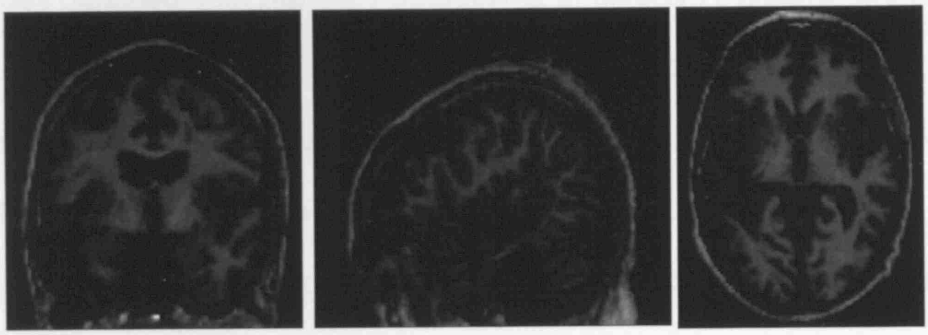
Patient 1



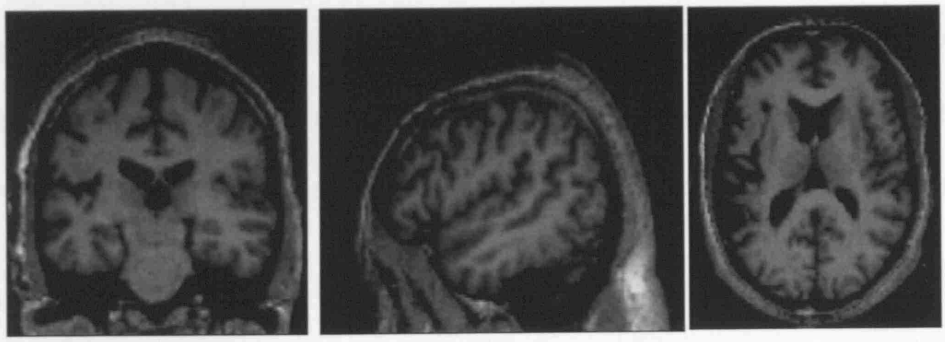
Patient 2



Patient 3



Patient 4



Patient 5

4 Results

The results will be presented under two headings:

1. Smoothing and
2. Findings of lesion identification methods.

4.1 Smoothing

The aim of the first part of this study was to determine the optimal FWHM for smoothing. This initial phase used Grey Matter segmented images and VBM.

The results of patient 1 will be presented in detail followed by a brief mention of the results from other patients. After the description of these images a consensus will be made on an appropriate FWHM for smoothing.

4.1.1 Patient 1

Images obtained from Patient 1 smoothing, at different FWHM's, are shown in figure 6. Within the range of FWHM values assessed (i.e. 0 to 12mm), as smoothing increased the area identified as abnormal also increased. As FWHM increased, images also become 'cleaned up.' Small areas of abnormality were removed as larger areas increased in size. The central area of these larger lesions also became more deeply coloured. This can be expected as smoothing increases the fuzziness of an image. Thus smaller areas of abnormality are merged into the background image. Larger areas however spread into the surrounding area and the borders of the lesion

became increasingly blurred. The effect was (in this case) to increase the size of the lesion.

This effect can be confirmed by review of the differences in t value between Patient 1 and controls. These results are shown on Table 2. This table illustrates the effect that increasing smoothing has on two different voxles and on the maximum t score of all voxels. Voxel 1 (co-ordinates -10 14 -4) is near the ventricles and not involved in the lesion. Voxel 2 on the other hand (co-ordinates -36 0 -2) is involved in the lesion. The maximum t score, T, gives us an idea of what is happening throughout the lesion. The t score of Voxel 1 steadily decreased as the FWHM of smoothing increased while the trend was an increase in the t score of Voxel 2 as the FWHM of smoothing increased. This would indicate that Voxel 1 is becoming less and less significant as the FWHM of smoothing increased. The trend for Voxel 2 however, was to become more and more significant as the FWHM increased. It should be noted however that between 8mm and 12mm FWHM of smoothing, the t score of Voxel 2 actually dropped. This suggests that significance was actually lost between 8mm and 12mm of smoothing. The maximum t score (T) is seen to be decreasing as the FWHM increases. T can be thought of as representing the maximum difference between Patient 1 and controls. A larger the value of T thus signifies a greater difference between the analysis for Patient 1 and controls. As expected this maximum difference decreases as smoothing increases and images become fuzzier.

Overlap images were also created in an attempt to better demonstrate how smoothing affected lesion size. These images (figure 7) demonstrate that while there is a substantial increase in lesion size from unsmoothed to 4mm and from 4mm to 8mm the increase in size from 8mm to 12mm is not so dramatic.

The images from this analysis should also be compared to the anatomical images provided in figure 5. The lesions identified at an FWHM of 8mm, closely resemble the lesions as seen on the anatomical scans.

4.1.2 Other Patients

It must be clear that these results were not limited to Patient 1. Figure 8 illustrates overlay images at 8mm and 12mm smoothing for Patient 2 though to 5.

These images all support the findings of Patient 1. Smoothing removed smaller areas of abnormality as images became fuzzier. However significant areas steadily increased in size. Again the magnitude of this increase drops off after 8mm making 12mm analysis similar to 8mm. Again images can be compared to the anatomical images of the patients in figure 5. Here too, 8mm smoothing produced lesions of similar size as the lesions seen on these scans.

The findings between patients are also interesting. Analysis generally identified a large area which can be thought of as the area of interest. For

Patients 4 and 5 however, several small areas of abnormality were still detected at 8mm smoothing.

Considering these initial results a conclusion was reached that 8x8x8mm FWHM smoothing would provide a good compromise between high lesion detectability, low artefact contribution and accurate lesion delineation. These initial results also gave us an impression of what to expect for the results of further analysis.

Table 2: Effect of the FWHM of smoothing on the t scores of 2 voxels and the maximum t score of all voxels (T). Voxel 1 is not involved in the lesion while voxel 2 is involved.

FWHM	Voxel 1 (co-ordinates -10 14 -4) not involved in the lesion	Voxel 2 (co-ordinates -36 0 -2) involved in the lesion	Maximum t score (T)
0mm (unsmoothed)	181.4	8.1	181.4
4x4x4mm	35.00	18.0	35.00
8x8x8mm	12.6	19.8	19.80
12x12x12mm	8.1	15.8	17.00

4.2 Findings of lesion identification methods

The second part of this study presents the findings of each lesion identification method. VBM will first be presented followed by PPM and FCP.

4.2.1 VBM

This classical analysis is based on the use of t test with a $p < 0.001$ (corrected for multiple comparisons). Analysis produced results in all five patients in the study.

Figure 9 (A – Grey Matter and B – White Matter) illustrates the results of these analyses. These images can be compared with the overlap 8mm and 12mm smoothed VBM images shown in figure 7 and 8.

It should be noted here that these images only provide us one level of information. Voxels are identified as abnormal only if they differ significantly from control voxels. As stated previously this analysis uses the t test with $p < 0.001$. Thus we know nothing more than the highlighted voxels differ significantly from control voxels.

4.2.2 PPM

In PPM analysis we posed two questions:

1. What threshold or probability would be best in the analysis of scans and
2. How does PPM compare to the two other analysis methods?

With respect to the first issue, 8mm smoothed images were first analysed using probability thresholds of 0.3, then 0.5 and then 0.9. As noted in the introduction, however, it is important to remember that the exact Probability Threshold used in calculation varies according to the priors used in the

Bayesian calculation. Therefore it was not possible to define objectively the PPM's threshold (i.e. the expected posterior probability) for use in calculation. The results of the analysis are shown on figure 10 as overlap images of probability thresholds 0.3, 0.5 and (for all except patient 5) 0.9.

All scans, except Patient 5, produced results at thresholds 0.3, 0.5 and 0.9. As would be expected the volume of abnormal voxels steadily decreased as probability increased. So much so that in the case of Patient 5, a probability threshold of 0.9 did not pick up any abnormal voxels. As noted above, VBM analysis showed the lesion to be small and limited mostly to the Grey Matter. This can be confirmed by cross-referencing the anatomical images in figure 5. Thus, as this scan contained a small lesion it is possible that (at 8mm of smoothing) it was difficult to define using such a high probability. Comparison with the anatomical images in figure 5 also showed that the 'true' lesion matched the lesion detected by PPM analysis most closely when a probability threshold of 0.5 was used.

It was thus decided that an empirical threshold level 0.5 would be acceptable. Finally, it should be noted that PPM analysis took more computing time than the other methods. Generally the analysis of a single Grey Matter / White Matter T1 weighted MRI scan took 10 minutes. It must be remembered however that the PPM analysis provides two levels of information. Firstly, it identifies statistically significant voxels (as compared to control voxels) and secondly, a probability threshold that the voxel is actually abnormal is built into the analysis.

4.2.3 FCP

Next the data were analysed using the FCP method. Again we posed two questions:

1. What is the optimal value for the 'tuning' parameter α and
2. How dose FCP compare to the other analysis methods?

With respect to α, three different values were chosen for analysis, $\alpha = 0.5$, $\alpha = 2.5\sigma$ and $\alpha = 3.0\sigma$. Images produced during FCP analysis unfortunately could not be displayed onto the 3D image of SPM5. Sections are thus illustrated in figures 11 and 12 with abnormal Grey Matter voxels in red and abnormal White Matter voxels in yellow. Figure 11 illustrates the results of analysis for Patient 1 using different values of α.

They illustrate that the volume of voxels detected as abnormal steadily increased moving from $\alpha = 0.5$ to $\alpha = 2.5\sigma$ to $\alpha = 3.0\sigma$. These results were compared visually to the anatomical images in figure 5. This shows that the anatomic lesion matches the lesion detected by FCP analysis best when $\alpha = 0.5$. This finding was similar for all 5 patients analysed. By this rational, it was thus decided that $\alpha = 0.5$ would in fact be the optimal value of α for analysis. Figure 12 illustrates the analysis of the remaining patients with $\alpha = 0.5$ and 8mm smoothing.

It should be noted that lesions identified by FCP analysis were less well defined (i.e. had less sharp boundaries). FCP analysis also detected a greater

number of small areas of abnormality as compared to VBM and PPM analysis. This would suggest that FCP is more sensitive but less specific than VBM or PPM. Finally is should be pointed out that FCP proved to be the fastest of the three methods generally completing analysis within 1 minute. This may have implications on computing time.

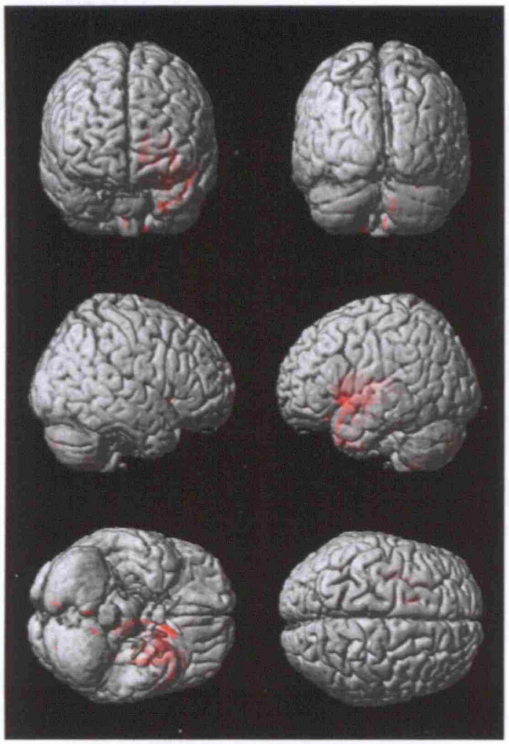
4.3 Summary

A summary is the results are shown below in Table 3.

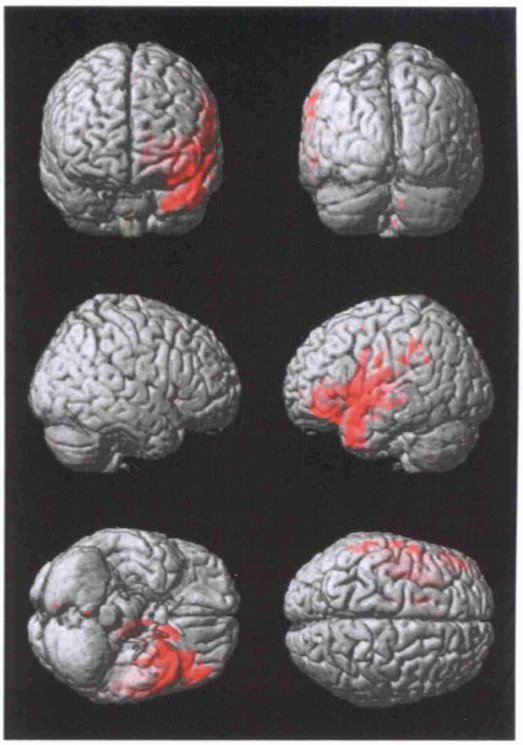
Table 3: Main results of analyses.

Effect / Method assessed	Conclusions
Smoothing	Images must be smoothed prior to analysis. An FWHM of 8x8x8mm was found to be the optimal level of smoothing.
VBM	Reliable providing results for all patients. Proves only one level of information.
PPM	A probability threshold of 0.5 gives the best results. Provides two levels of information. Lesions identified with sharper borders. Requires the longest computer processing time.
FCP	An α value of 0.5 provides the best results. Lesions identified with less sharp borders – method possibly more sensitive Requires the least computing time.

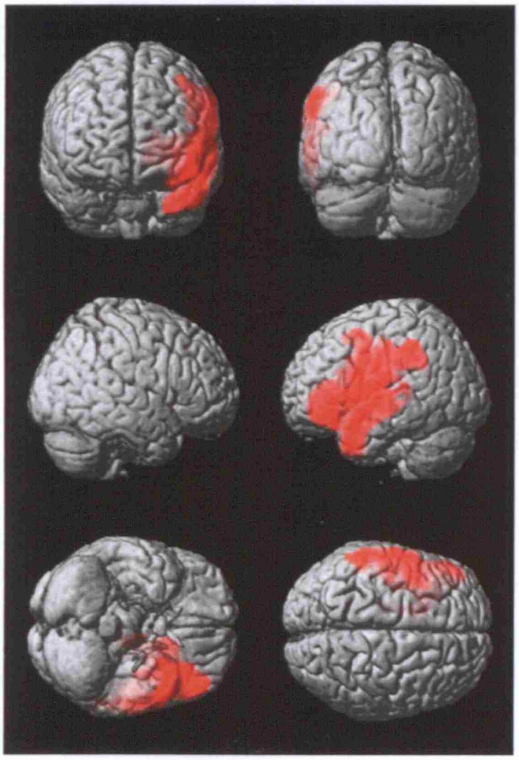
Figure 6: Patient 1, Smoothed and unsmoothed images, Grey Matter



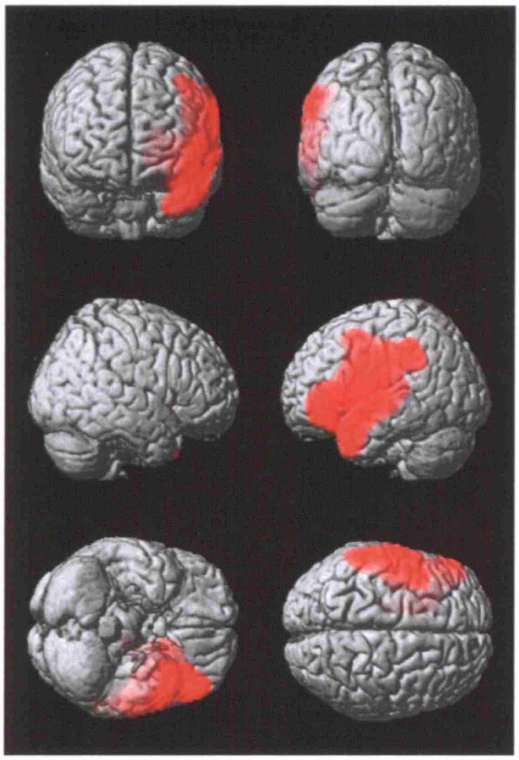
Unsmoothed



4x4x4mm smoothed

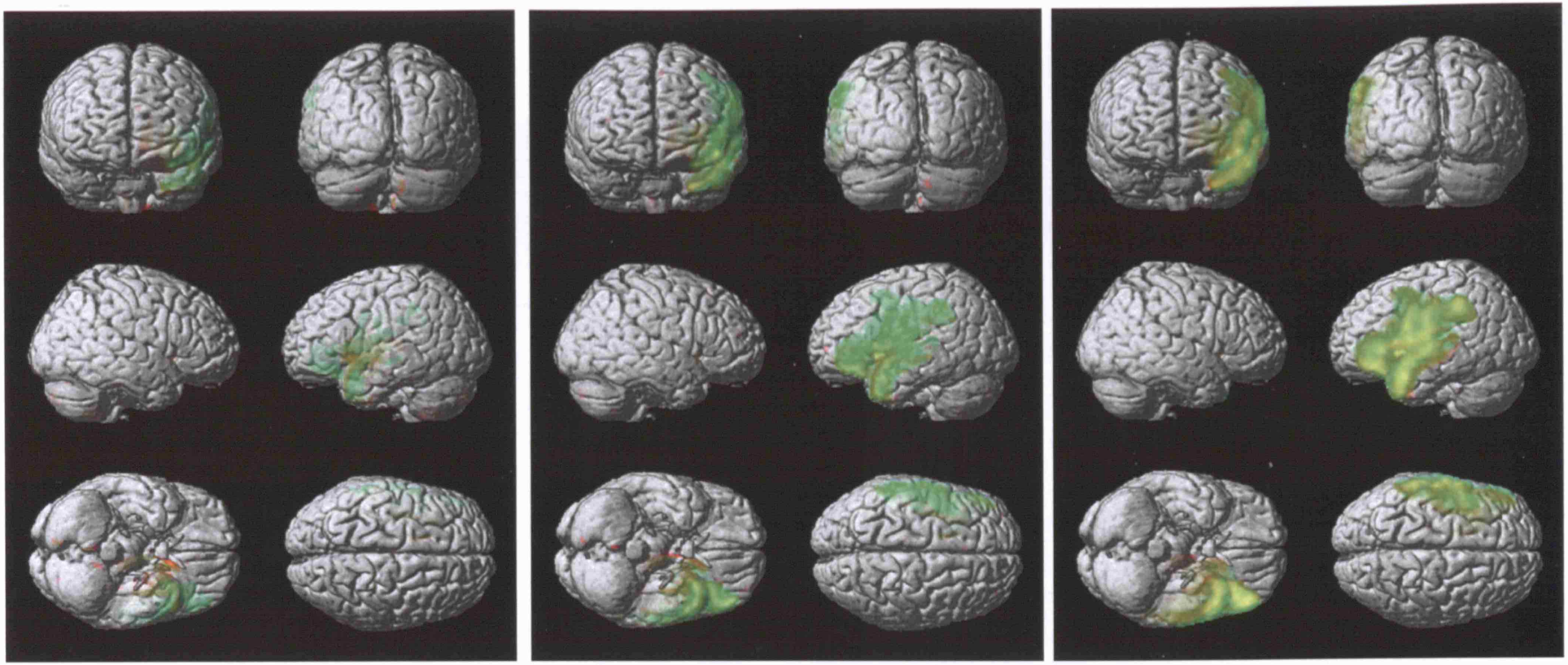


8x8x8mm smoothed



12x12x12mm smoothed

Figure 7: Overlap images for Patient 1, unsmoothed, 4mm, 8mm and 12mm smoothing



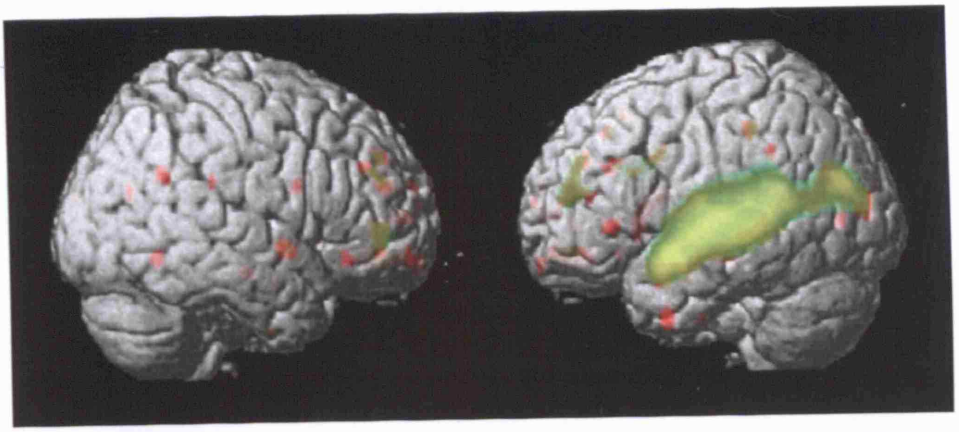
Overlap of unsmoothed and 4x4x4mm smoothed images

Overlap of 4x4x4 smoothed and 8x8x8mm smoothed images

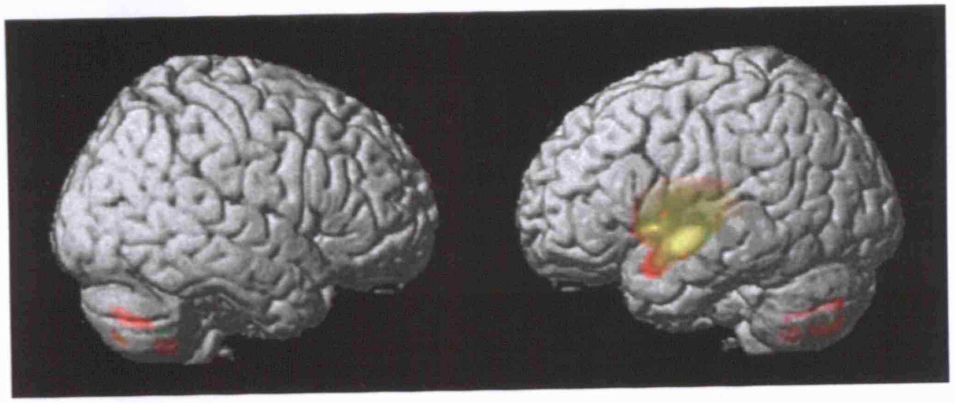
Overlap of 8x8x8mm smoothed and 12x12x12mm smoothed images

Figure 8: Overlay images of Patient 2 to 5 at 8mm and 12mm smoothing Grey Matter only.

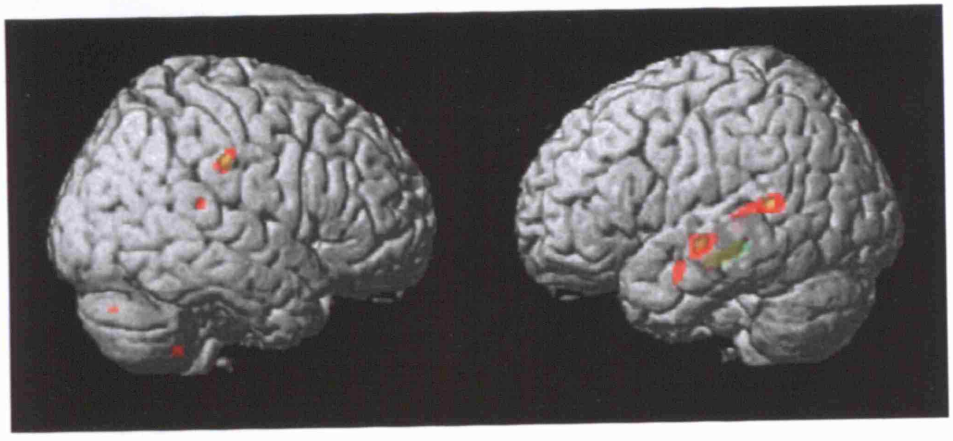
Patient 2



Patient 3



Patient 4



Patient 5

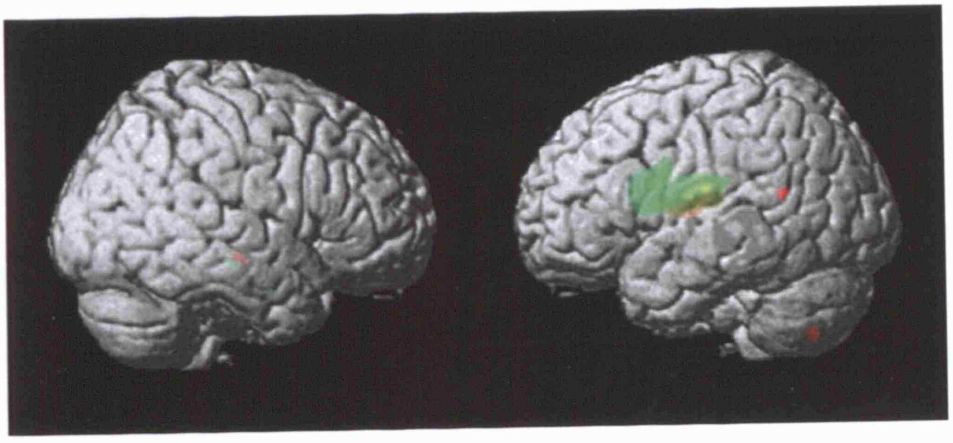


Figure 9: Summary for all 5 patients for VBM analysis, 8mm smoothing. A) Grey Matter

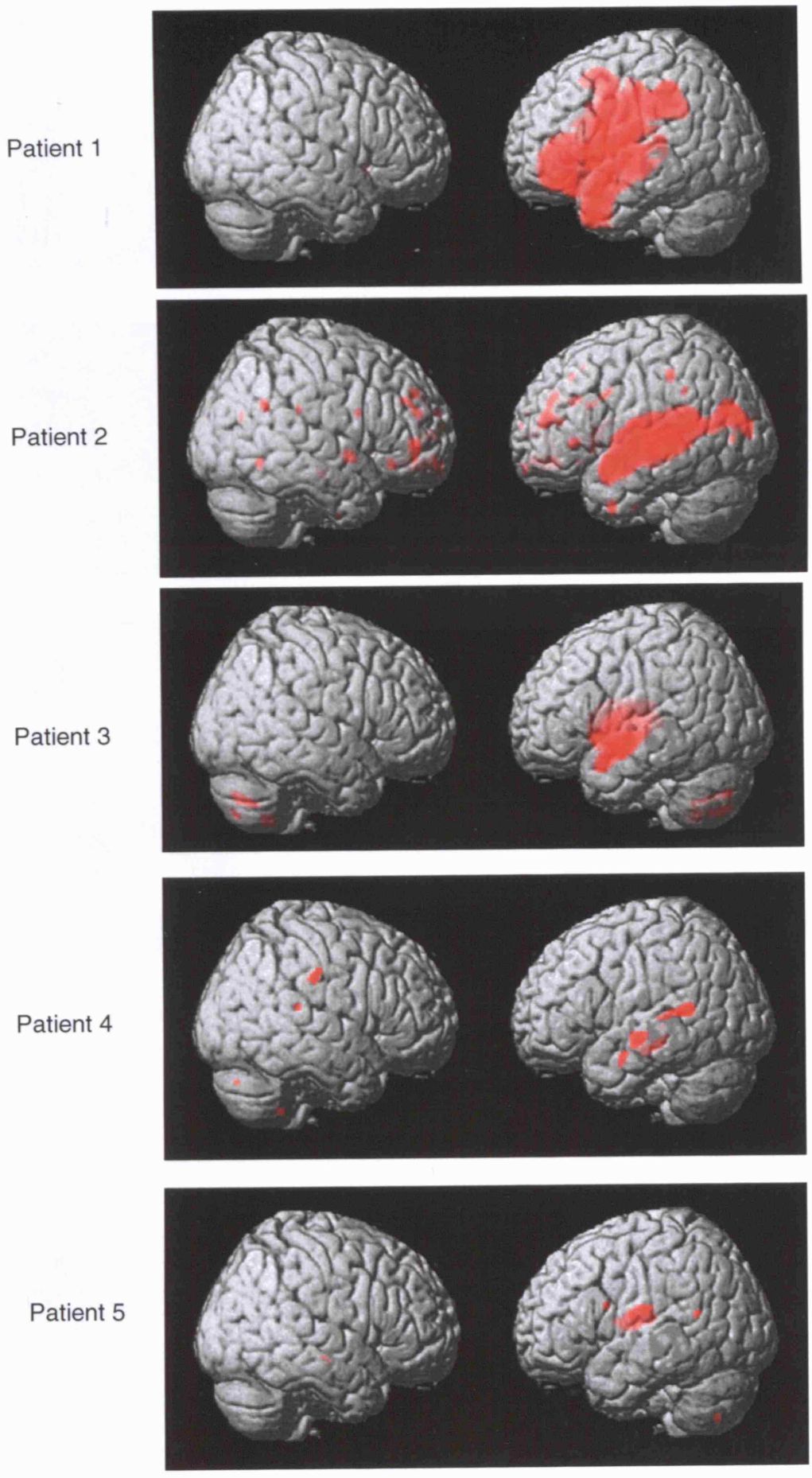
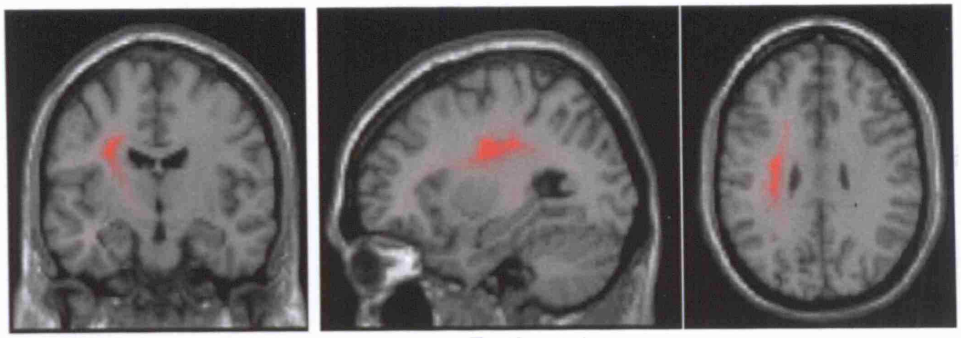
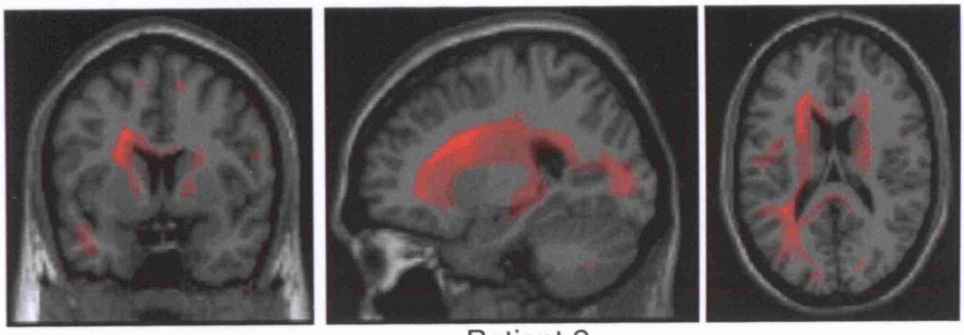


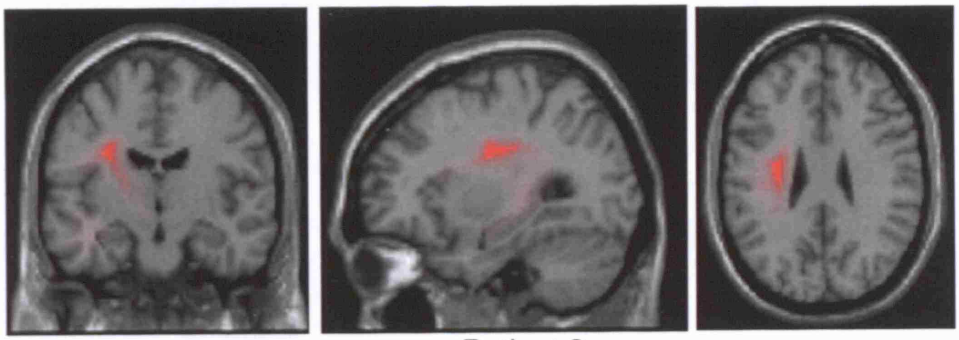
Figure 9: Summary for all 5 patients for VBM analysis.
8mm smoothing. B) White Matter



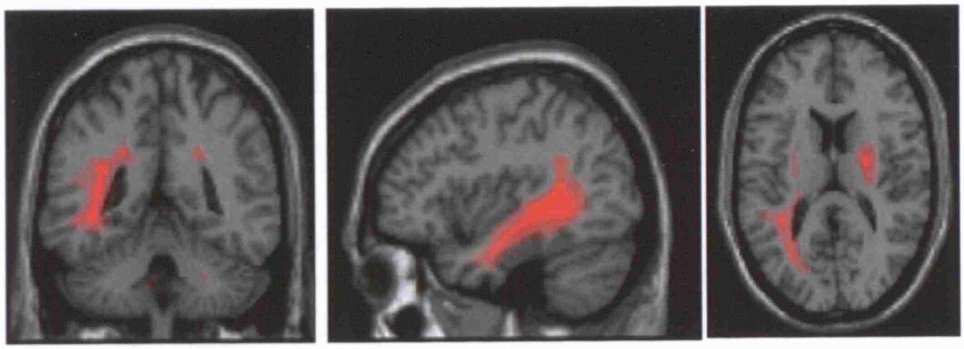
Patient 1



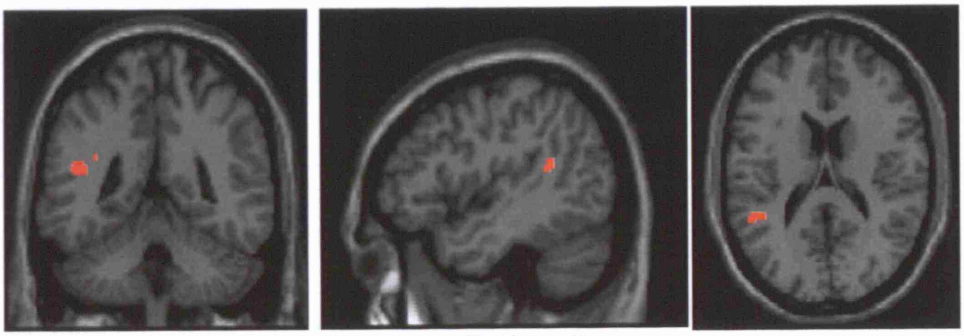
Patient 2



Patient 3



Patient 4



Patient 5

Figure 10: The effect of thresholding on PPM analysis, A) Grey Matter

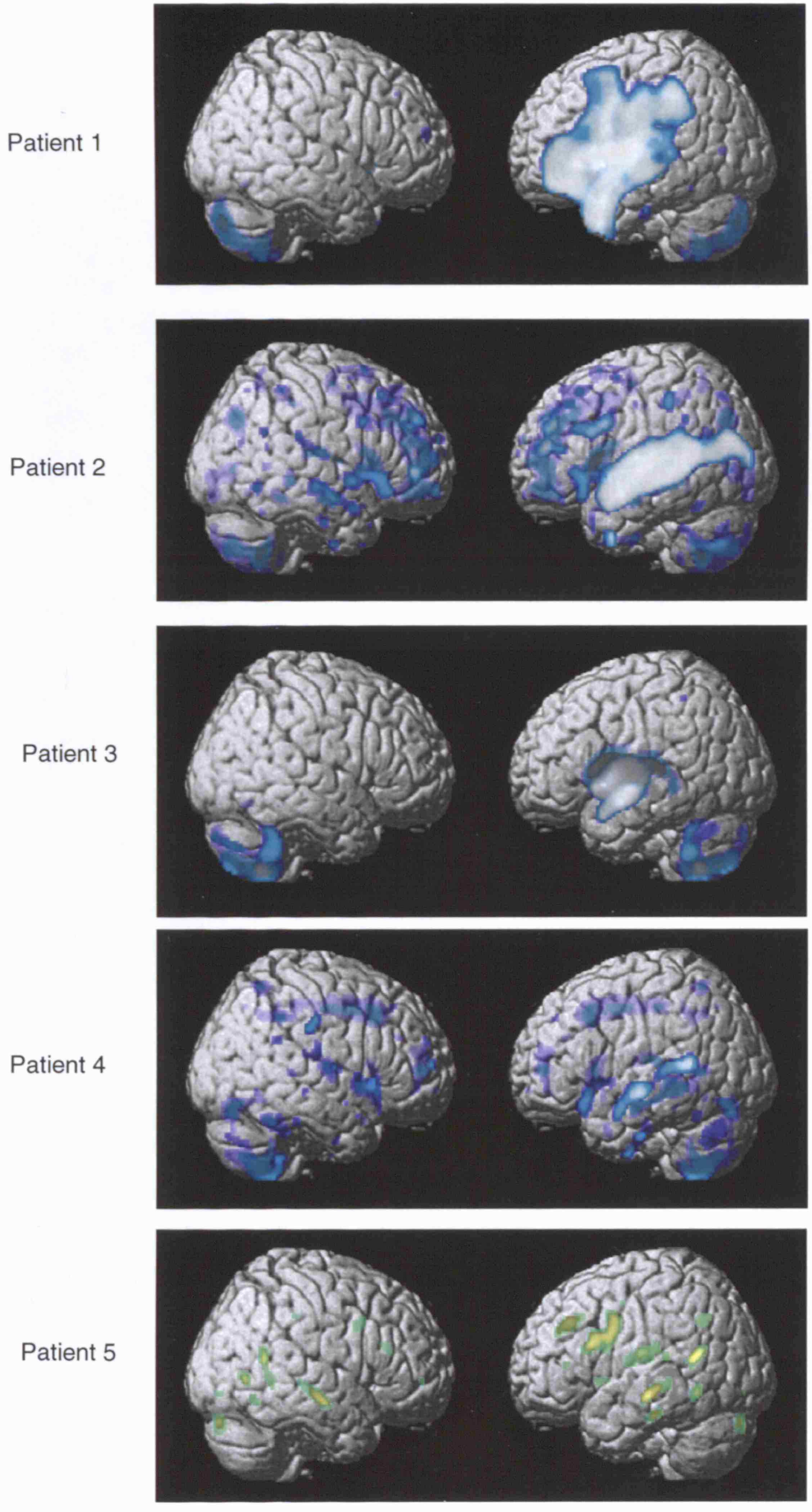


Figure 10: The effect of thresholding on PPM analysis, B) White Matter

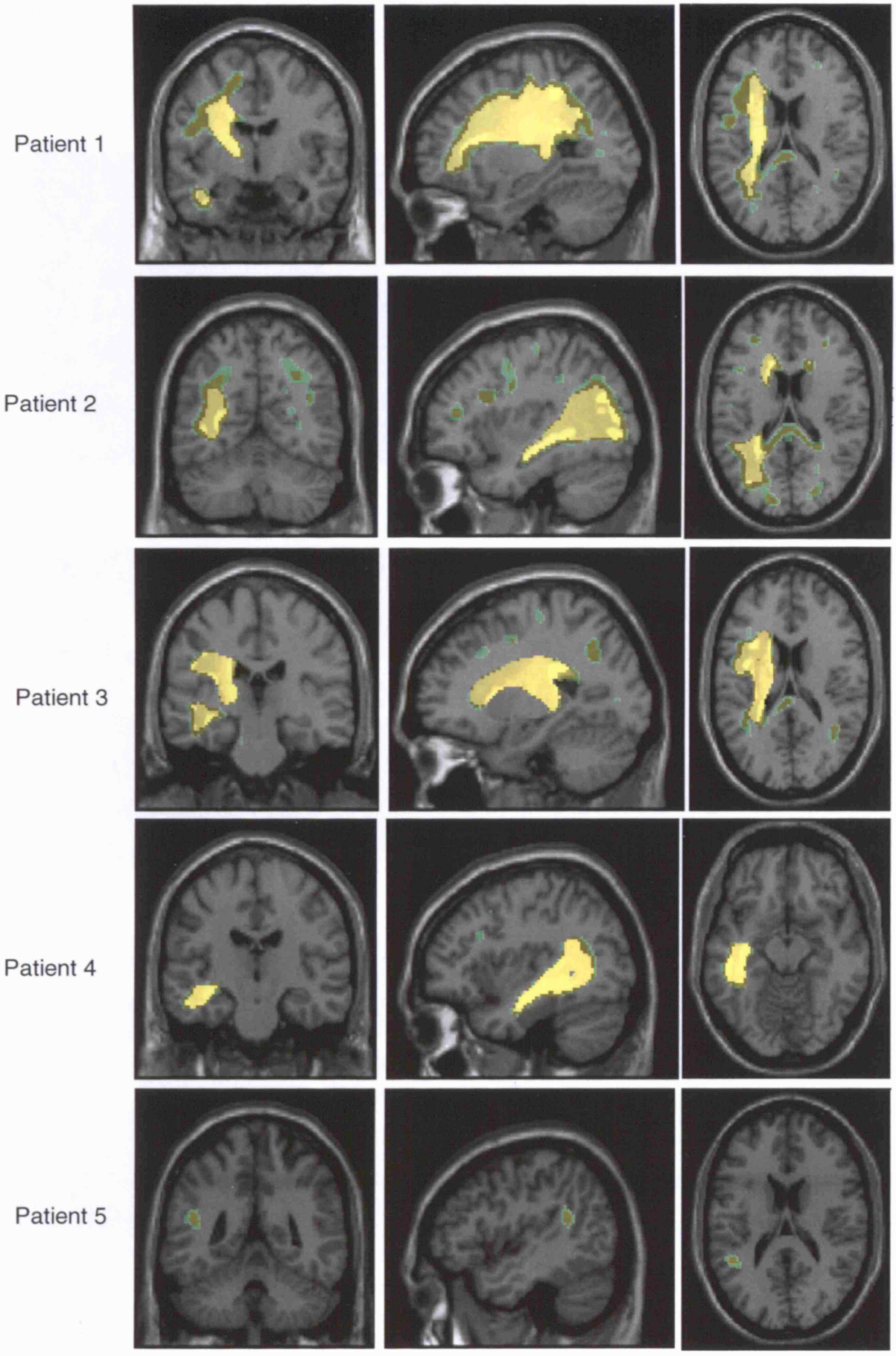
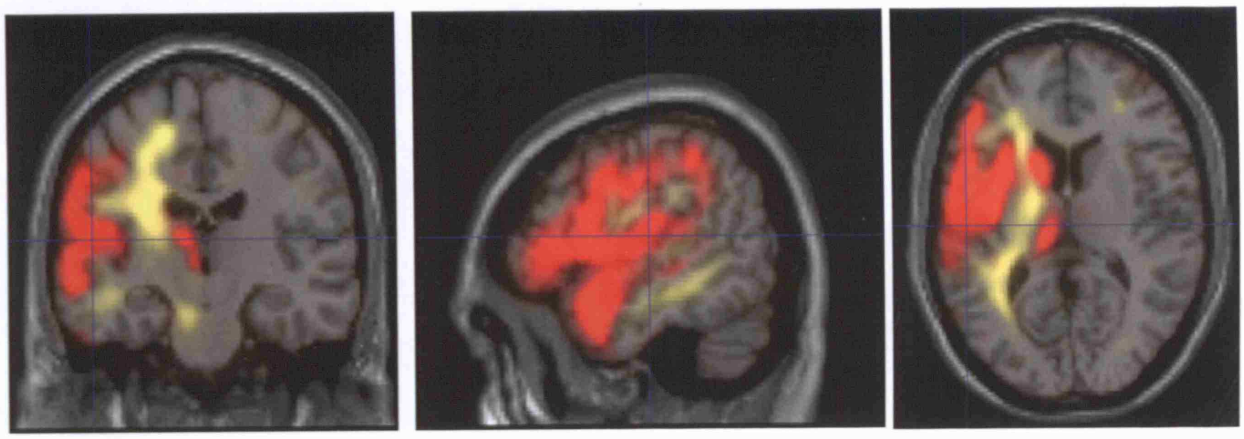
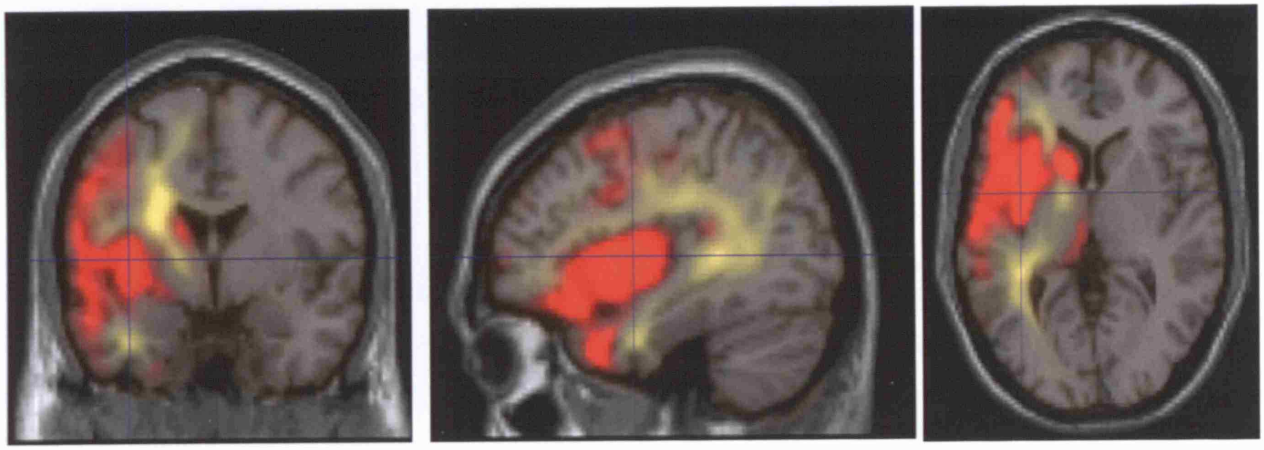


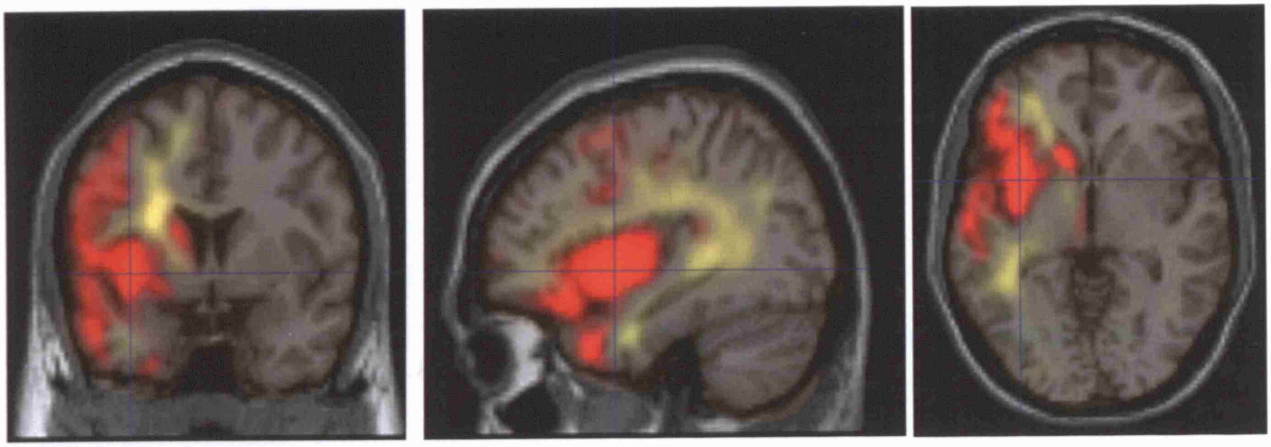
Figure 11: The effect of varying α on FCP analysis. All images for Patient 1, 8 mm smooth
A) Grey Matter shown in Red



$\alpha = 0.5$

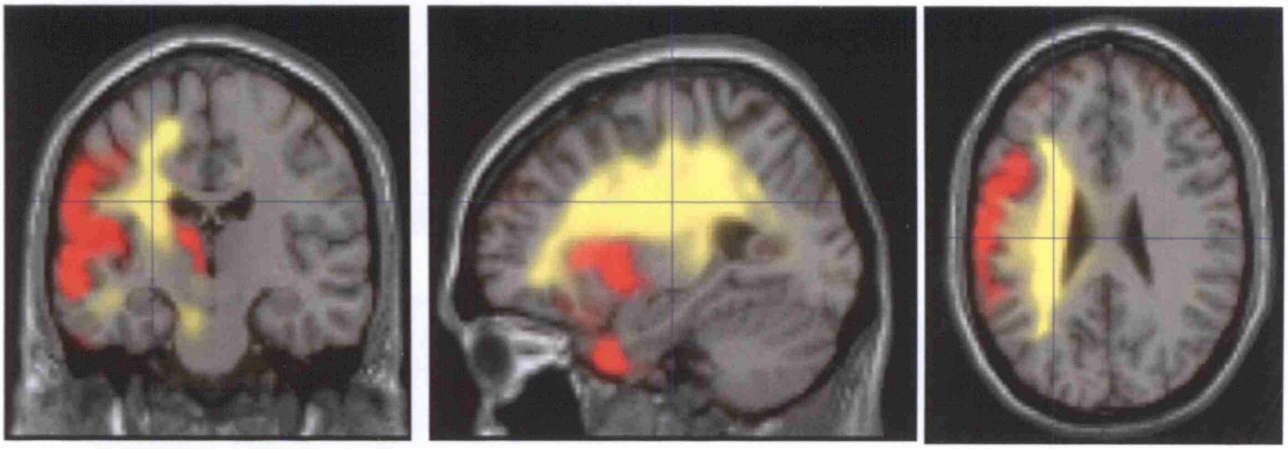


$\alpha = 2.5\sigma$

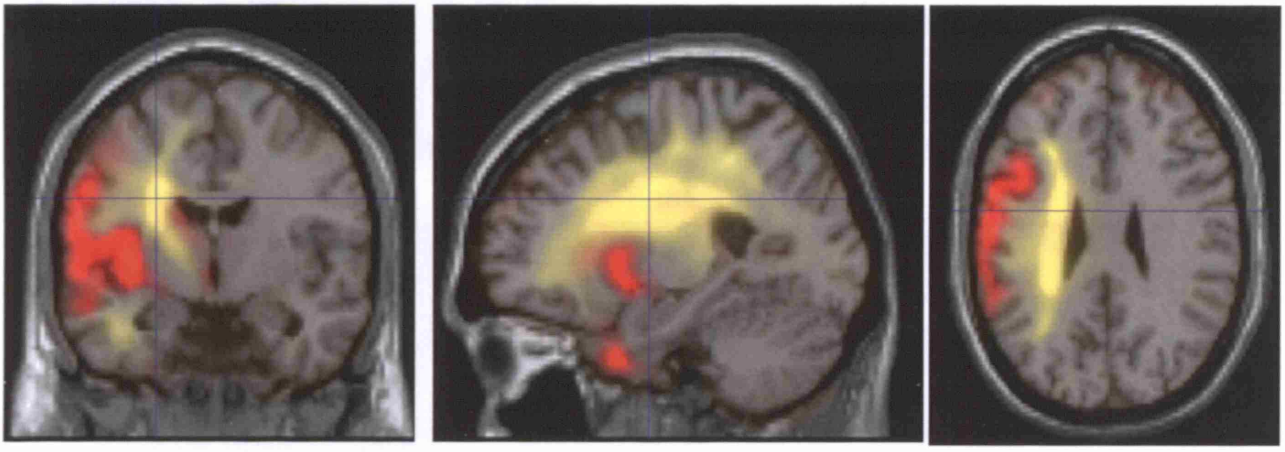


$\alpha = 3.0\sigma$

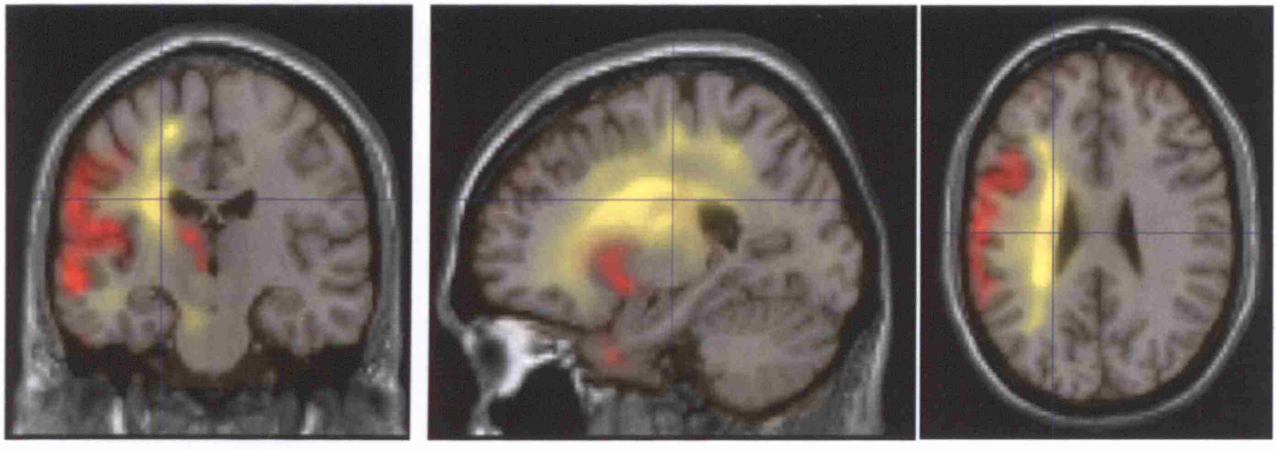
Figure 11: The effect of varying α on FCP analysis. All images for Patient 1, 8 mm smoothed
B) White Matter shown in Yellow



$\alpha = 0.5$

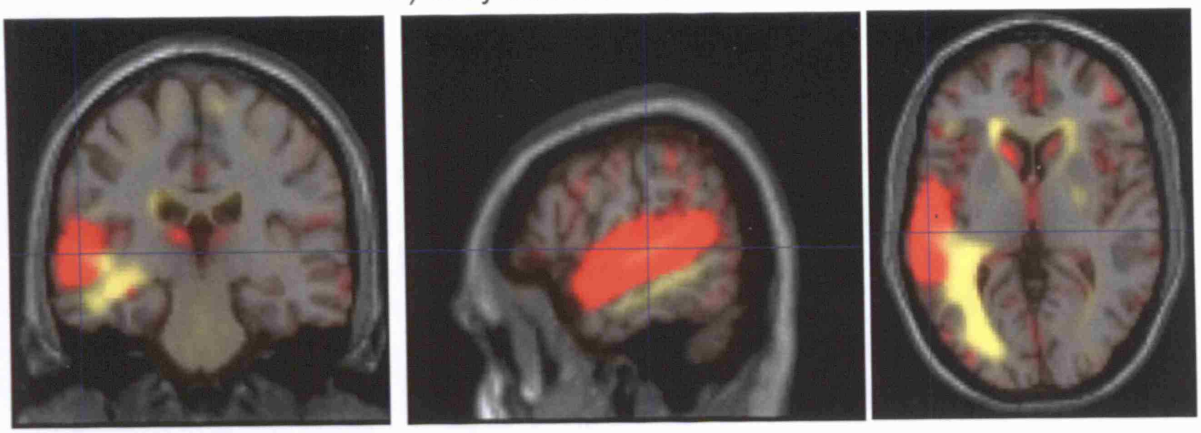


$\alpha = 2.5\sigma$

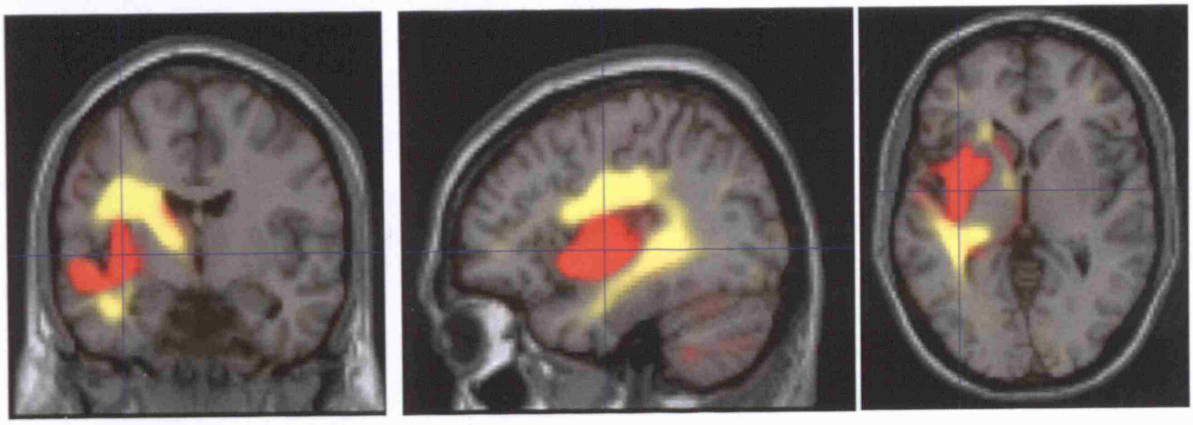


$\alpha = 3.0\sigma$

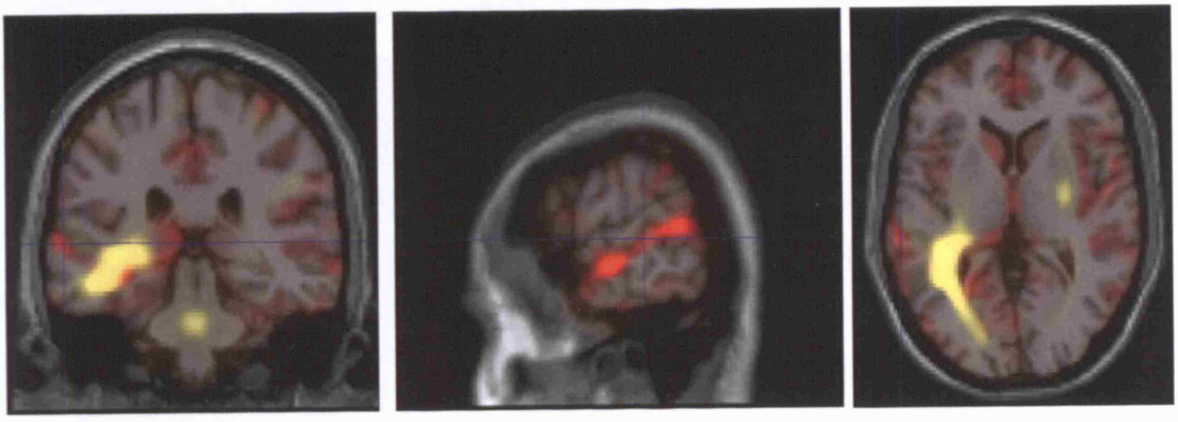
Figure 12: FCP analysis, $\alpha = 0.5$, Patients 2 to 5, 8mm smooth
A) Grey Matter in Red



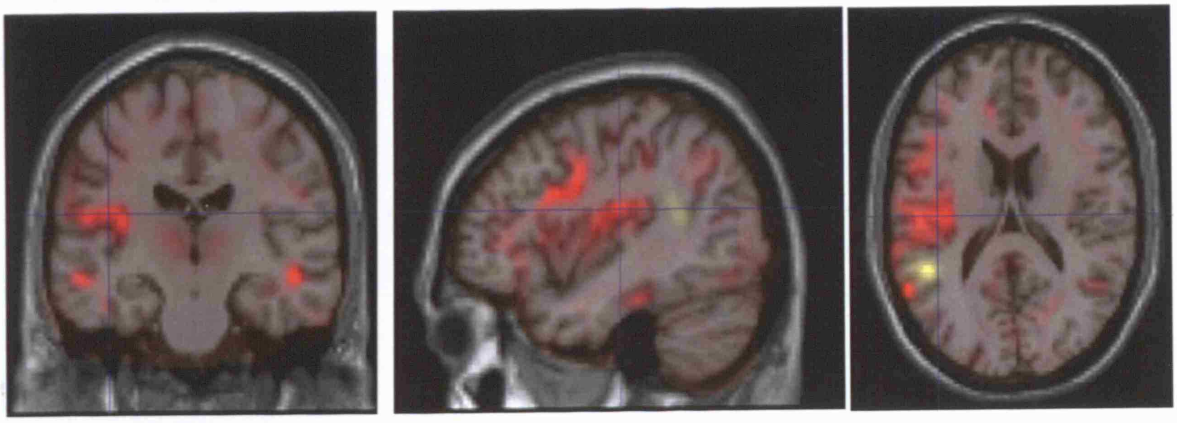
Patient 2



Patient 3

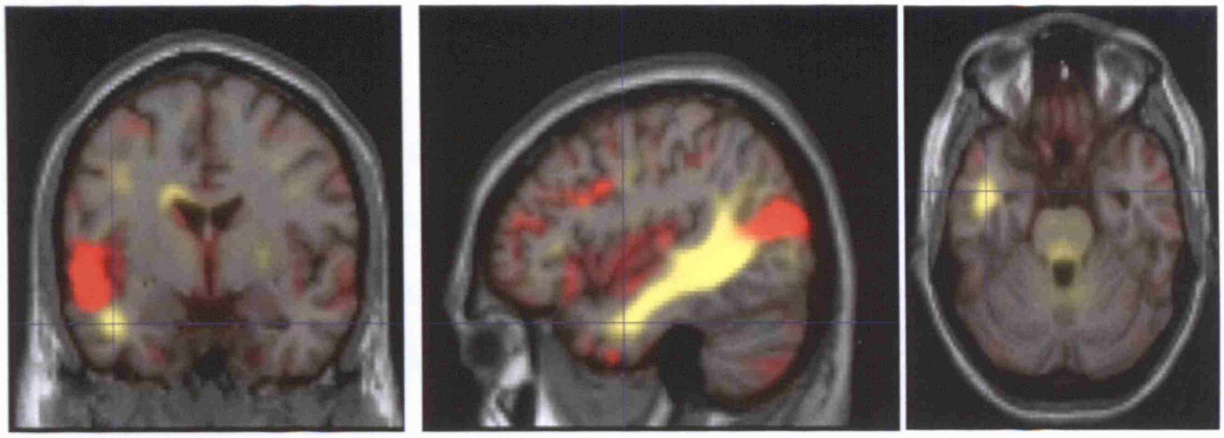


Patient 4

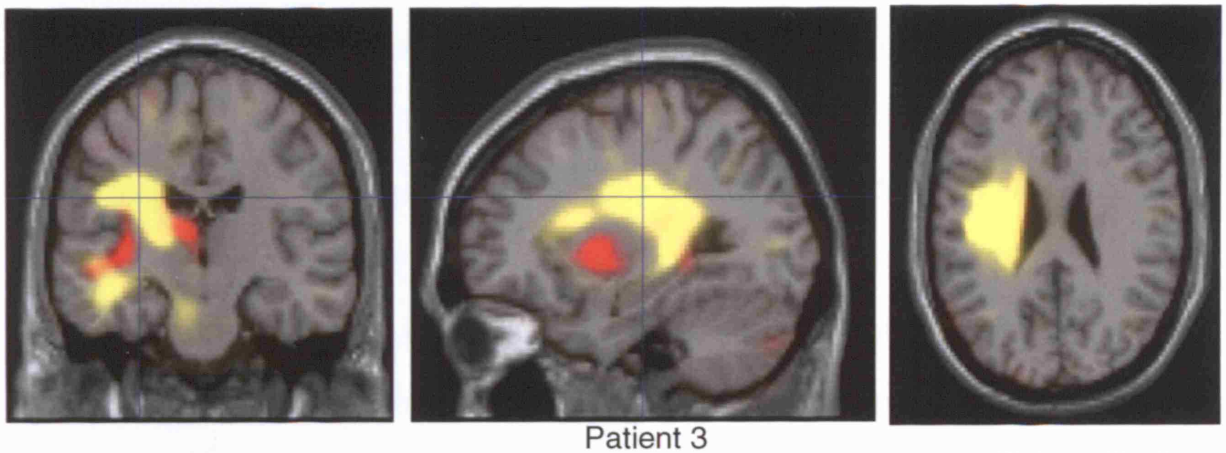


Patient 5

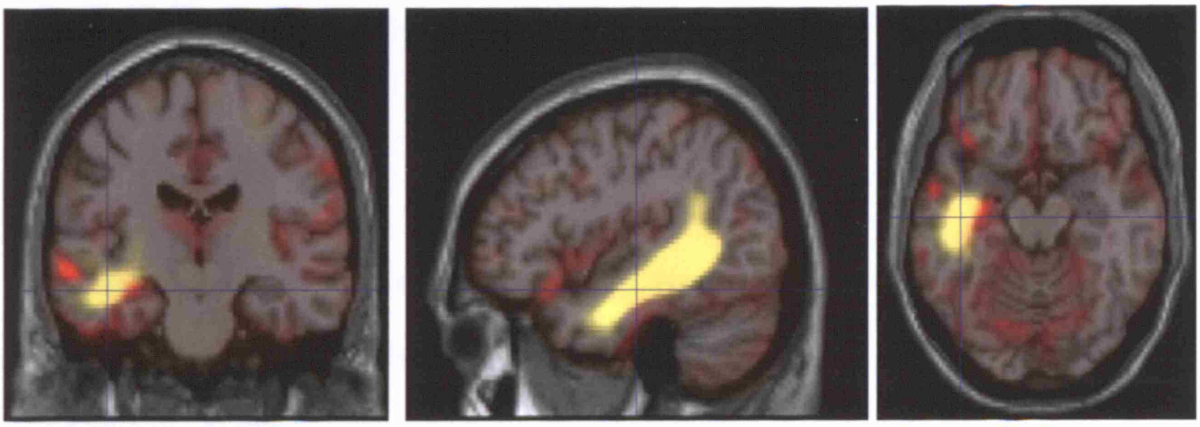
Figure 12: FCP analysis, $\alpha = 0.5$, Patients 2 to 5, 8mm smooth
B) White Matter in Yellow



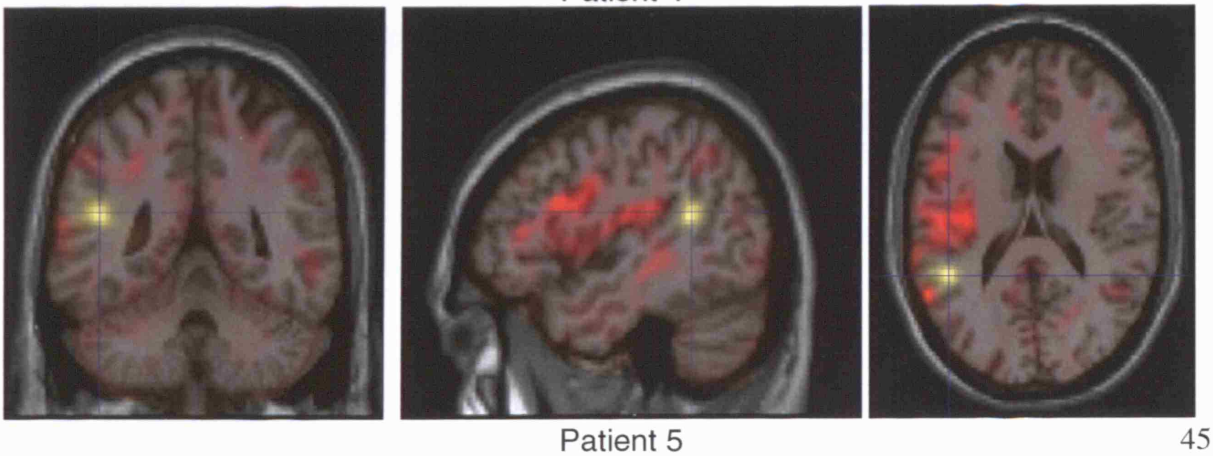
Patient 2



Patient 3



Patient 4



Patient 5

5 Discussion

As the science of Neuro-imaging has developed, more and more data have been produced as spatial resolution increases with thinner and thinner anatomical slices. Automated sequences have thus become important for the reliable, quick and robust analysis of lesions.

Images however had to be moulded and shaped into usable data sets that could be analysed by programs. Thus the science of image pre-processing developed hand in hand with these automated methods. The development of these technologies has been rapid with many different methods being tried and tested. In this study we have looked at the smoothing aspect of pre-processing and three different types of lesion identification methods. The results of the effect of smoothing will be discussed initially followed by the lesion identification methods.

5.1 Smoothing

Smoothing is the last stage of pre-processing. Normalisation and segmentation are both imperfect and smoothing helps to even out this imperfection. In this study we analysed data that was unsmoothed and smoothed using increasing FWHM. It is interesting that all data sets produced results.

However three effects were noticed as the FWHM of smoothing increased:

1. The larger lesions of interest became larger in size.
2. The smaller lesions decreased in size.
3. As the FWHM increased, lesion bounders became less sharp.

These effects could be anticipated. Smoothing serves to blur an image. Voxels fuse into the background and the borders of any lesion become less distinct.

Using Patient 1 (figure 6) as an example, analysis of an unsmoothed image produced results where the lesion identified was much smaller that the actual anatomical lesion. Referring to table 2 three effects can be identified as the FWHM of smoothing increases:

1. The t score of a voxel involved in the lesion increases
2. The t score of a voxel not involved in the lesion decreases and
3. The maximum t score (T) decreases.

These effects also are illustrated in figure 13. This would indicate that as smoothing increased the lesion increased in size with the maximum T value dropping off as the borders of the lesion merged with the surrounding scan. Referring again to figure 13, the rate of this decrease can be seen to reduce as the FWHM of smoothing increases. The final shape of the lesion detected matches that of the anatomical lesion. This suggests that the voxels actually

making up the lesion were close to the necessary p value to make them statistically significant. Smoothing was however necessary for this to occur.

Looking closely at the unsmoothed and 4mm smoothed images of Patient 1 in figure 6 you can also notice several small areas of lightly coloured abnormality over the cerebellum. These are likely to be artefacts because they are not visible on the anatomical scans. Artefacts could arise as a result of the imperfect nature of the earlier steps of pre-processing (e.g. normalisation). With smoothing however these small differences merge into the background and are no longer identified as abnormal. In effect, the smoothing has 'cleaned up' the images. Unlike the voxels discussed above, the p value of these voxels can be considered as being close to non-significance. Smoothing was however necessary them to be fully rendered as non-significant. The findings suggest that it would be inappropriate to use images that are unsmoothed or smoothed with a 4mm FWHM.

Lastly, 8mm and 12mm smoothed images are considered. On visually comparing the images in figure 6, the magnitude of the change between these images is not as great as that seen between 4mm and 8mm smoothing. The effect of smoothing thus becomes less pronounced when dealing with higher values of FWHM. Again looking at figure 13 we can see that the t value of Voxel 2 (which is involved in the lesion) actually decreased as we moved from an FWHM of 8mm to 12mm. This would suggest that the maximum t value for

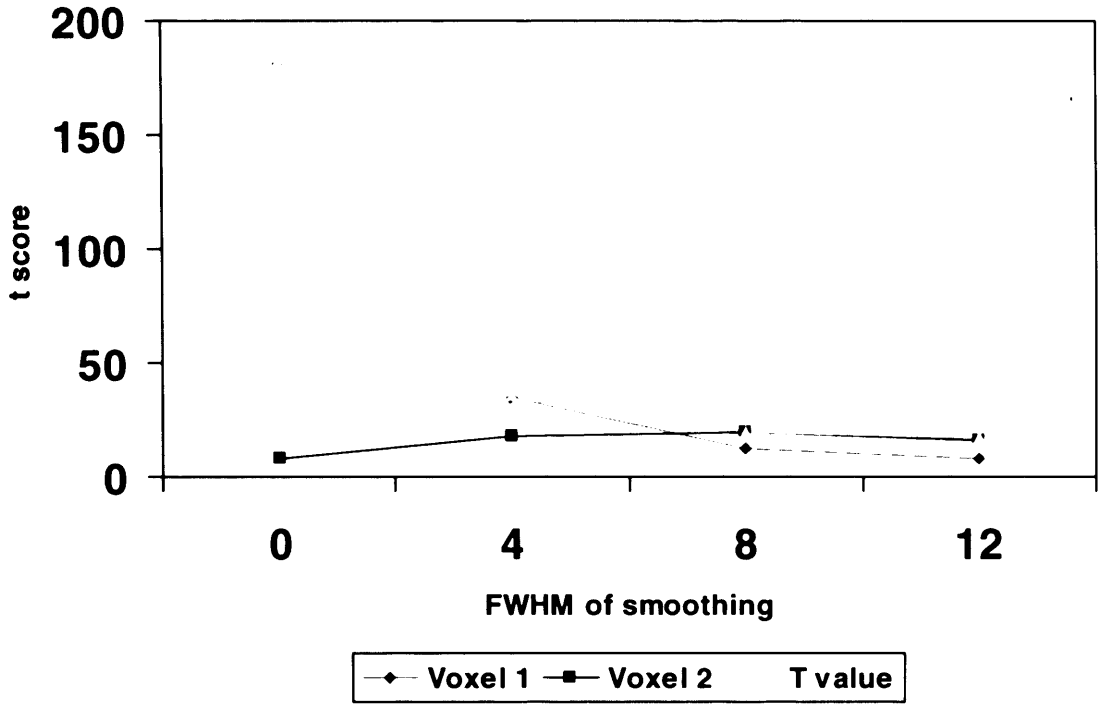
this 'lesioned' voxel was already attended at 8mm. Further smoothing had the effect of actually decreasing its significance.

As smoothing increases and the boarders of the lesion become blurred, it is important that we do not start encroaching on areas of the scan that are actually normal. It is possible that this will occur at very high smoothing values as the entire image is blurred into one and the abnormal lesion is spread throughout. It is thus best to use the lowest possible value of smoothing that will give us useful results. This gradual increase in lesion size has been described in other studies where smoothing was varied (Jones et al. 546-54).

For these reasons, in this study the FWHM found to be most useful was 8x8x8mm. At this value the lesion detected was similar in size and shape to the anatomical lesion. While a slightly larger lesion was outlined with 12x12x12mm smoothing, there is the risk of including voxels that are not actually involved in the lesion. These results are in line with previous studies using spatial smoothing during lesion identification (Stamatakis and Tyler 167-77).

Smoothing may also depend on the voxel size. Here the voxels were acquired at 1x1x1 mm and this was increased to 2x2x2mm during the spatial normalisation process. While 8x8x8 mm may be best for these images, the ideal smoothing may vary with the size of the voxel. This was not investigated in this study.

Figure 13 – Effect of smoothing on the t score of Voxel 1, Voxel 2 and the maximum t score, T.



Having completed pre-processing, attention is now turned to the lesion identification methods that were compared in this study.

5.2 Lesion identification methods

The methods compared in this study include Voxel Based Morphometry (VBM), Posterior Probability Mapping (PPM) and Fuzzy Clustering with fixed Prototypes (FCP). All methods were dependent on images being pre-processed prior to analysis. Provided that pre-processing was adequate, all methods produced meaningful results in so far as the lesions demarcated corresponded to the anatomical lesion seen in figure 5. Before further comparisons are made between methods, each will be discussed independently.

5.2.1 VBM

This is one of the earliest developed methods for image analysis. Results from VBM analysis relay on statistical testing. The quality of the results is thus likely to be related directly to the quality of pre-processing and the controls used in the study (for more details see Mehta et al). This made VBM the simplest method to use amongst the three.

5.2.2 PPM

As noted PPM relies on Bayesian statistics and analysis is done in two stages, the results of the first stage being used as the priors in the second analysis. It is this second stage of analysis that distinguished PPM from VBM. The second analysis requires a probability threshold to be applied to help decide if a voxel is abnormal or not. Defining this probability threshold is vital for the efficient functioning of PPM. As seen in the analysis, using too low a value can result in too many voxels being identified as abnormal. The opposite effect is also true, with far fewer voxels being identified as abnormal when the probability threshold is too high. In the case of Patient 5, this actually led to the patient being assessed as lesion free when a probability threshold of 0.9 was used. It was therefore decided that 0.5 would be the most adequate compromise between these two extremes.

5.2.3 FCP

This method is the most recent to be developed. Analysis is fundamentally different from both VBM and PPM analysis with neither t nor F tests being preformed. Rather FCP groups all voxels in different clusters. Within these

clusters abnormal voxels are identified as outliers that differ markedly from the mean effect of all voxels. As stated α acts as tuning parameter in the analysis. An increase in α will result in a decrease in the sensitivity of FCP. Thus it is likely that 3.0σ and 2.5σ both produced α values that were higher than acceptable. With an α value of 0.5, however, the area assessed as abnormal increased to match the lesion as seen on the anatomical scan. Considering this, an α value of 0.5 is most likely to be useful for analysis.

5.3 Comparison between methods

All three methods do have similarities. They all required adequate pre-processing to work. The important aspects of smoothing were discussed above. All methods were also successful at producing meaningful results. All lesions identified were of a similar size and shape to the anatomical images. Thus grossly, all methods behaved in a similar fashion.

VBM proved to be the easiest to use. Analysis does however require adequate pre-processing of images. Further analysis depends solely the results of statistical tests. In the past this has proven to be one of the strengths of the method, while others have viewed it as a weakness. In this study the unified segmentation algorithm was used for normalisation and segmentation and an FWHM of 8mm was found to be an optimal level of smoothing. Using these parameters for pre-processing VBM gave meaningful results. VBM can thus be seen as a robust and dependable method that is relatively easy to use. Mention should be made of Patient 5 however. Here

the lesion was small and VBM only identified part of the lesion. It thus functions better when the lesion is large.

PPM goes one step further than VBM. A probability threshold is built into the calculation. This extra information comes from two stages of calculations typical of the Bayesian approach. In this study we decided on a value of 0.5 or 50%. The results generated are thus more meaningful than those of VBM. When visually assessed, PPM calculation also produced results with relatively sharper border delineation. PPM can thus be most useful in situations where sharp delineation is important and the lesion must be outlined accurately. The inclusion of the probability threshold however makes PPM more difficult to work with. The use of too low a value will result in areas beyond the lesion being assessed as abnormal. However when the value is too high, results will produce a situation where small lesions are missed. This is exemplified when Patient 5 was analyzed using a probability threshold of 0.9. Caution is thus advised when using a high probability threshold.

FCP analysis is the most recently developed of the three methods. Analysis is dependent on the choice of the tuning parameter α. In this study we chose a value of 0.5 to give the necessary sensitively for analysis. Of the three, this unique characteristic of FCP may make it more useful in the analysis of scans with small lesions. Again using Patient 5 as an example, VBM and PPM both identified small lesions. FCP was however sensitive enough to identify a larger area of abnormality. It is important however that too small an α value is not used. This will result in increasing the false positive rate making the analysis invalid.

Thus none of the methods can be considered ideal. They each have strengths making them more suited under different circumstances. VBM is easy to use and carries the benefit of being in use for a long time. PPM however provides the most information from analysis and produced sharp lesion delineation. FCP is best suited when the lesion is small. This analysis is very sensitive in identifying lesions that other analyses miss.

Mention should also be made of computing time. A VBM Grey Matter Scan took an average of 3 minute for analysis. With the two levels of analysis required in PPM, analysis of a single Grey Matter scan took well over 10 minutes. This may have implications if computing power is a limiting factor and as voxel sizes decrease with higher resolution. FCP however was the fastest of the three with analysis being completed within 1 minute for a single Grey Matter T1 weighted MRI scan. It is thus likely that this maybe a more useful technique if computing power is a limiting factor.

6 Limitations

The large number of parameters varied during this study means that limitations are inevitable. On discussing these limitations, possible means of improving on the study will also be addressed.

The analyses in this study were conducted on only 5 patients. Such a small number makes it likely that the power of the study is low. At the beginning of the study it would have been useful to do a power calculation to determine exactly how many images need to be analysed to make the results significant. Time constraints however also need to be considered as further patient analyses would have been difficult in the time period allocated.

The patients analysed were all from a bank of images held by the Wellcome Foundation fMRI Unit. It is thus unlikely that this cohort represents the entire stroke population. All patients had middle cerebral artery strokes. It thus cannot be said for certain if the results can be applied to patients with posterior circulation strokes or lacuna strokes. It should be recalled that although the physiological changes are similar, the pathological conditions leading to these events could differ. Similar analysis should thus be conducted on other patients to confirm the results.

All scans analysed were conducted of stroke patients who has their primary event over six months ago. The scans thus showed marked atrophy over the area of infraction. It is actually this atrophy that the methods used in the

identification of lesions. It is thus uncertain how the methods would cope with acute strokes. The oedema that can accompany these insults may in fact be difficult for these methods to identify, as the MRI signal is higher than in areas of atrophy.

The results of the analyses where compared visually. Visual comparison, while easily performed, is highly subjective, user dependent and prone to error. Robust statistical test should instead be used in the comparison of results. Receiver Operating Characteristics (ROC) curves could have been used for this comparison. Ideally all results could be compared using ROC curves. They would provide a means of graphically representing how close to ideal the results are.

Analyses were also conducted using the scans of actual stroke patients. While this has the benefit of assessing how the methods work under 'real' circumstances, scans using manually inserted lesions could have been used. This would make comparison easier as the co-ordinates of the lesion would be known. The results obtained and the conclusions made would thus be more reliable.

This study also lacked a clear Gold Standard for comparison. In the discussion we used VBM analysis as a standard for comparison of PPM and FCP analysis. However the Gold Standard for reporting is actually the manually identified mapping of the lesion. As indicated in the introduction this manual mapping is very time consuming and requires in depth knowledge of

neuro-anatomy. It thus was not possible to carry out manual mapping of the lesions. If available however they would provide another possible standard for the generation of ROC curves even thought manual mapping has its own limitations.

As noted many parameters were varied in this study. These parameters were key to the conclusions generated. It should be noted however that further analyses could be conducted to ensure that the same conclusions are arrived at as in this study. Considering initially smoothing, the trend identified was a general drop in maximum T value with an evening off after an FWHM of 8mm (see figure 13). Other studies however have looked at smoothing values beyond 12mm (Jones et al. 546-54). It would have thus been useful to extend the range of FWHM. We could thus be certain of the trend identified and making our conclusions stronger. The other parameters varied included the probability threshold (for PPM analysis) and α (for FCP analysis). Considering first the probability threshold, we tested three values that identified 'high', 'medium' and 'low' probability. It is with the 'high' value of 0.9 that we ran into some difficulty (of example, with the analysis of Patient 5). We have thus suggested that the 'medium' value of 0.5 should be used in analysis. From our study this is the correct suggestion but further analyses should be conducted to look at values between 0.5 and 0.9, for example 0.7 may actually provide useful results at a higher probability. The same can be said for α in FCP analysis. From this study we identified 0.5 as a value making FCP most useful. However, further analyses should be conducted using α values smaller than 0.5. It would be even more useful if analyses were conducted using

scans with small lesions that other methods found difficult to detect. Using this approach the benefits of FCP analysis can be emphasised.

The suggestions can form the basis of further studies in the area of neuro-imaging. With the increasing importance of automated lesion identification methods, conducting these studies is becoming increasingly essential to furthering our knowledge into Lesion – Symptom mapping.

7 Conclusion

The science of neuro-imaging has increased significantly over the last 50 years. Using MRI, we are now able to produce images with great detail and resolution. The manual reporting of these images has however become increasingly complex. This has made the science of automated lesion identification important. There are many issues that still need to be resolved when using this relatively new technology. New techniques are constantly being developed without fully testing of old techniques. This has created a situation where the full potential of various techniques has not been completely resolved. Indeed there are no studies in the literature attempting to compare different techniques. This study aimed to achieve just that. Three commonly used lesion identification methods have been compared and the strengths and weakness of the methods identified. The study has also looked at the contentious issue of smoothing. A rational for the need of smoothing has been made and an optimal FWHM of smoothing has been identified. The main conclusions of the study are dealt with below.

The first part of the study investigated the impact of spatial smoothing on the results of VBM analyses. This highlighted the importance of smoothing as unsmoothed data produced incorrect results. The most valid results were identified when data were smoothed with a FWHM of 8mm. The use of a FHWM value above this runs the risk of over smoothing and increasing the false positive rate.

The second part of the study looked at the different types of lesion identification methods. VBM analysis is simplest to use and can produce useful results especially in patients with large lesions and if pre-processing is adequate.

PPM analysis however gave the most informative results. The probability threshold built into the analysis tells us the probability that the lesion identified is statistically significant. From this study we can say that a threshold of 0.5 produces meaningful results. Using too high a probability threshold may lead to false negative errors. PPM also produced results with the sharpest borders. It is thus most useful where the lesion needs to be clearly defined.

FCP analysis is the newest technique of the methods assessed. It has proven to be more sensitive then VBM and PPM. It employs the use of the tuning parameter α. In this study we have suggested a value of 0.5 for this parameter. The strength of FCP lies in its ability to detect small lesions. These lesions actually appear larger than with VBM or PPM analysis. This method would thus be most useful for the analysis of scans with small lesions.
Of all the methods FCP was least demanding on computer processing time while PPM took the longest to generate results. FCP may thus be most suited where computing power is a limiting issue.

In conclusion, it cannot be said that there is an ideal method. They all have different strengths and should be used for the task best suited to them.

This study has also raised several issues that should be addressed in future research. The ability of these methods to identify other lesions (e.g. oedema and tumours) should be addressed. This however was not one of the main aims of this study. More in depth analyses should also be conducted to investigate the ideal probability threshold and tuning parameter α. For both these parameters 0.5 has been suggested as a usable value. However to ensure that this is the best possible advice, further values should be tested. The results of such trials will further improve our knowledge and familiarity with these methods.

8 Reference List

Ashburner, J and KJ Friston. "Voxel-Based Morphology - The Methods." NeuroImage 11 (1999): 805-21.

---. "Why Voxel-Based Morphometry Should Be Used." NeuroImage 14 (2001): 1238-43.

Bates, E, et al. "Voxel-based lesion-symptom mapping." Nature Neuroscience 6.5 (2003): 448-50.

Bezdek, JC, LO Hall, and LP Clarke. "Review of MR image segmentation techniques using pattern recognition." Medical Physics 20.4 (993): 1033-48.

Clark, MC, et al. "MRI segmentation using Fuzzy Clustering Techniques.":730-742. 1994.

Crinion, J, et al. "Spatial normalization of lesioned brains: Performance evaluation and impact on fMRI analyses." NeuroImage (2007).

Davatzikos, C, et al. "Voxel-Based Morphometry Using the RAVENS Maps: Methods and Validation Using Simulated Longitudinal Atrophy.":1361-1369. 2001.

Fisher, R, et al. "Gaussian Smoothing.". May 29, 2007.

Foong, J, et al. "Investigating regional white matter in schizophrenia using diffusion tensor imaging." Brain Imaging 13.3 (2007): 333-36.

Friston, K and WD Penny. "Posterior probability maps and SPMs." NeuroImage 19 (2003): 1240-49.

Friston, KJ, et al. "Classical and Bayesian Inference in Neuroimaging: Theory." NeuroImage 16 (2001): 465-83.

Jones, DK, et al. "The effect of filter size on VBM analyses of DT-MRI data." NeuroImage 26 (2005): 546-54.

Mechelli, A, et al. "Voxel-Based Morphometry fo the Human Brain: Methdos and Applications." Current Medical Imaging Reviews 1.1 (2004): 1-9.

Mehta, S, et al. "Evaluation of voxel-based morphometry for focal lesion detection in individuals." NeuroImage 20 (2003): 1438-54.

Seghier, ML, KJ Friston, and CJ Price. "Detecting subject-specific activations using fuzzy clustering." NeuroImage (2007).

Stamatakis, EA and LK Tyler. "Identifying lesions on structural brain images--validation of the method and application to neuropsychological patients." Brain and Language 94.2 (2005): 167-77.

Wu, Y, et al. "Automated segmentstion of multiple sclerosis lesion subtypes with multichannel MRI." NeuroImage 32 (2006): 1205-15.

9 Appendix 1 – List of Figures

10 Appendix 2 – List of Tables