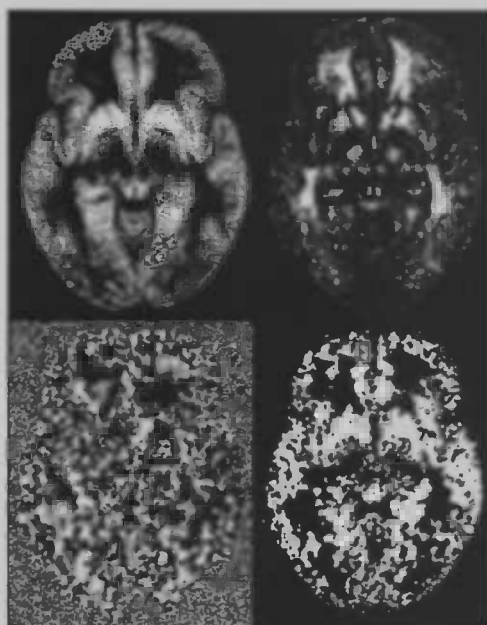




WORDT  
NIET UITGELEEND

# Voxel Based Morphometry, implementation and sensitivity

M.H.Kunst



Januari 2003

Rijksuniversiteit Groningen  
Bibliotheek Wiskunde & Informatica  
Postbus 800  
9700 AV Groningen  
Tel. 050 - 363 40 01



RuG

WORDT  
NIET UITGELEEND

**Supervisors:**

Dr. F.W. Cornelissen  
Laboratory of Experimental Ophthalmology  
School for Behavioral and Cognitive Neurosciences (BCN)  
University of Groningen

Dr. R.P. Maguire  
Department of Neurology  
Academic Hospital Groningen

Prof. dr. J.B.T.M. Roerdink  
Department of Mathematics and Computing Science  
University of Groningen

Rijksuniversiteit Groningen  
Bibliotheek Wiskunde & Informatica  
Postbus 800  
9700 AV Groningen  
Tel. 050 - 363 40 01

## ***Abstract***

Voxel-Based Morphometry (VBM) is a whole-brain, unbiased technique for characterising regional cerebral volume and tissue concentration differences in structural magnetic resonance images (MRI), between or within groups of subjects.

The goal of this project is to provide the University Hospital Groningen (AZG) with a standardised protocol and an automated procedure, that implements the VBM method for detecting morphological changes or differences in the human brain. Specific attention will be given to the sensitivity of the method (analytical and based on simulations).

# Contents

<b>1 INTRODUCTION.....</b>	<b>4</b>
<b>2 BACKGROUND .....</b>	<b>4</b>
<b>3 RELATED WORK .....</b>	<b>5</b>
<b>4 OVERVIEW OF THE VBM METHOD .....</b>	<b>6</b>
4.1 THE PURPOSE OF VOXEL BASED MORPHOMETRY .....	6
4.2 WHAT ACTUALLY IS MEASURED USING VBM .....	7
4.3 STATISTICS.....	7
4.3.1 A General linear approach for statistical parametric mapping in functional imaging .....	8
4.3.2 Statistical tests using the general linear approach, an example .....	9
4.3.3 Multiple comparison problem .....	10
4.3.4 Cluster Level tests .....	11
4.4 DATA PROCESSING .....	11
4.4.1 Normalisation.....	12
4.4.2 Segmentation.....	13
4.4.3 Preservation of tissue volume .....	14
4.4.4 Smoothing .....	14
4.5 CURRENT VBM PROTOCOLS .....	15
4.5.1 Standard VBM protocol .....	15
4.5.2 An optimised VBM protocol: .....	16
4.6 DATA ACQUISITION AND PROPERTIES .....	17
<b>5 IMPLEMENTATION OF THE VBM ROUTINES .....</b>	<b>18</b>
5.1 ENVIRONMENT .....	18
5.2 VBM PROCESS DESCRIPTION FILES .....	18
5.3 SEGMENTATION .....	19
5.4 SEGMENT CLEANING.....	20
5.4 ACCOUNT FOR VOLUME CHANGES.....	20
5.5 STATISTICS.....	20
<b>6 SENSITIVITY OF THE METHOD.....</b>	<b>21</b>
6.1 EXPERIMENT DESIGN.....	22
6.2 SIMULATED ATROPHY IN MRI-VOLUMES .....	22

6.2.1 Generation of an MRI dataset with controlled atrophy in a specified region. ....22

6.2.2 Simulating Segmented Data.....26

6.2.3 Modelling spatial coherence .....28

6.2.4 Evaluation of the simulated segments.....31

6.3 SENSITIVITY OF VBM WITH RESPECT TO LESION SEVERITY .....34

6.3.1 Grey matter reduction percentage .....34

6.3.2 Results from experiments using simulated groups with increasing atrophy.....35

6.4 ANALYTICAL POWER ESTIMATION.....38

6.4.1 Power .....38

6.4.2 Data model for the power estimation.....38

6.4.3 Calculations.....38

6.4.4 Results from analytical power estimation .....40

**7 EVALUATION.....46**

**8 FURTHER WORK.....46**

**9 APPENDICES .....48**

APPENDIX A. GREY-SCALE EROSION WITH FLAT AND NON-FLAT STRUCTURING ELEMENT.....48

APPENDIX B: VBM PROCESS FILES.....50

Normalisation:.....50

Segmentation.....53

Volume Preservation (optional).....54

Smoothing .....55

Automated Statistics .....55

**10 BIBLIOGRAPHY .....58**

## 1 Introduction

For many reasons the morphology of the human brain might change over time or differ between people. One can think of age-related changes, or sex related differences. Certain diseases will lead to local or global changes in the amount of the different brain tissues. Cortical morphology is known to vary in epilepsy, schizophrenia, anorexia nervosa and Alzheimer's disease. In most cases, a (local) decrease in the amount of grey matter, and thus a thinning of the cortical mantle, is observed.

The contemporary way of examining the structure of the human brain is by means of structural MRI (Magnetic Resonance Imaging) scans. From MRI scans the different brain tissues can be identified. Measurements can be taken without physically intruding the body. Certain techniques have been developed, or are currently being developed to detect differences between groups, or longitudinal changes, in data obtained from MRI scans.

The neurology department of the AZG has a large amount of experience with fMRI (functional MRI). For a research project on cortical reorganisation and degeneration following visual field defects by the ophthalmology department of the AZG, however, apart from the functional interest, there is an interest in the morphological changes in the visual cortex, as a consequence of the visual field defect. For this reasons a method that is able to detect local changes in the concentration of grey matter (GM), white matter (WM) and or Cerebro Spinal Fluid (CSF) is needed. Within the system used by the AZG for fMRI/PET analysis, there is no standard/automated method for performing such a morphological analysis.

The goal of this project is to provide the AZG with a standardised protocol and an automated procedure for detecting morphological changes or differences in the human brain, which implements this protocol. Specific attention will be given to the sensitivity of the method, and the practical implications of this sensitivity for research.

## 2 Background

A study for atrophic changes in the calcarine area resulting from retinal degeneration using MR image [Kitajima et al., 1997], suggest that there indeed are changes in the cortex following visual field defects. In this context, the examination of morphologic changes in the visual cortex because of visual field defects might prove interesting.

To perform morphological analysis different automated methods have been developed in the past few years. Before, measurements were taken manually at a number of points in a Region (or Volume) of Interest (ROI), either using anatomical MRI/CT scans, or real slices from the brain. For large data

sets, this is a tedious and cumbersome method, which yields results that largely depend on the person performing the measurement. Far more powerful automated techniques have been recently introduced. These techniques can be broadly divided in methods that deal with differences in brain shape, and methods that look at differences in the local composition of brain tissue after compensating for macroscopic differences in shape.

For the research project on cortical reorganisation and degeneration following visual field defects, a method called Voxel Based Morphometry (VBM) is chosen to examine the differences between test and control group. The VBM method is part of the second class of methods. In essence, it involves a voxel-wise comparison of the local concentration of grey matter (or white matter or CSF) between two groups of people.

VBM (as defined in [Ashburner et al, 2000]) consists of the following steps: non-linear-spatial normalisation, segmentation, (optionally) modulation, smoothing and analysis. Schemes that are more elaborate are possible. Choices have to be made for a large number of parameters and the precise scheme for the method that is most suitable for our purposes should be devised. Some of these parameter values will be invariant over different research projects. The user will be offered a way to easily provide the system with the other parameter values that are specific for his/her project.

### **3 Related Work**

A number of studies have used the technique of voxel-based morphometry to examine structural brain differences among different populations. For this thesis we will use the methods described by John Ashburner and Karl J. Friston in "Voxel-Based Morphometry – The methods", 1999, as a starting point on the development of a suitable protocol. They state: "The importance of the VBM approach is that it is not biased to one particular structure and gives an even-handed and comprehensive assessment of anatomical differences throughout the brain." This makes this method suitable for general research purposes. Different improvements on the methods described in this article have been proposed: [Good et Al., 2001] suggest an optimised protocol in which they try to enhance normalisation and segmentation with respect to a specific tissue of interest.

A related approach is Regional Analysis of Volumes Examined in Normalised Space (RAVENS), [Goldsal et al. 1998]. Voxel-Based morphometry using RAVENS maps is described in [Davatzikos et al. 2001], including a new method for elastic, volume-preserving spatial normalisation. Both VBM and RAVENS are concerned only with local volumetric measurements and do not measure any higher order shape characteristics. RAVENS and VBM differ in two ways. The first difference is in the spatial normalisation transformation. RAVENS uses a high dimensional elastic transformation, which is driven by point correspondences established on distinct anatomical surfaces. VBM relies on

relatively smoother parametric transformations. A second difference is in the way in which the spatial distributions of GW, WM and CSF are determined. RAVENS preserves the volumes of different tissue volumes, both local and at a global level. VBM uses relatively low parameter shape transformation to eliminate global shape differences, and the residual variability is examined using SPM analysis. However, a kind of volume-preservation is possible in VBM.

Other methods examine shape differences. Deformation-Based and Tensor-Based Morphometry (DBM and TBM) are methods that use deformation fields, obtained from non-linear registration of brain images. DBM is used to identify differences in the relative positions of structures within the subjects' brains. TBM refers to methods that localise differences in the local shape of brain structures.

Since functional changes are most commonly reflected in the cortex, methods for actually measuring the thickness of the cortex have been developed, and still are. ([Fischl et al. 2000], [Lerch 2001]). An important issue is the definition of cortical thickness. In earlier studies, this was commonly taken to be the shortest distance between the GM/WM boundary and the pial surface (GM/CSF boundary). However a better physiological assumption would be that the thickness should be measured along the trajectory of the cortical columns, which in general is not a straight line [Lerch 2001]. The thickness metric outlined by [Jones et al. 2000] uses a partial differential equation, Laplace's equation, which models the properties of the cortical columns more closely than the straight-line approach. Extraction of the required surfaces (in general needed for measurements of the cortical thickness) and volumes, and solving the initial value problem, however, is computationally very expensive.

## **4 Overview of the VBM method**

### ***4.1 The purpose of Voxel Based Morphometry***

VBM is used to localise significant structural brain differences among populations, using univariate data, typically by means of a statistical parametric map of regional differences. (For multivariate data, a method called tensor-based morphometry can be used.) Overall brain structure is analysed by other methods such as deformation-based morphometry.

VBM is less suitable for identifying focal brain lesions in individuals [Mehta et al. 2002]. Due to large anatomical variance, the lesion has to be very large to be found significant. Also, for single subject longitudinal research, especially when little data are available, there are more suitable methods to depict changes over time. (For example SIENA, which is part of the FMRIB Software Library, Oxford Centre for Functional Magnetic Resonance Imaging of the Brain). In general it can be stated that VBM should be used mainly for group analysis.



## **4.2 What actually is measured using VBM**

Typically the analysis is performed on segmented volumes. The intensity of each voxel in such a volume indicates the probability, ranging from 0 to 1, that the voxel belongs to the tissue of interest. When we assume that the output of the segmentation process is accurate, such a probability can be regarded as the concentration of the tissue in that voxel. A zero value indicates that no tissue from this type is present, and a value of one indicates that the voxel is completely filled with the tissue of interest. All values in the range 0-1 are possible.

In this simplest case, concentrations of tissue in each voxel are compared among the subjects. It is also possible to regard the volume of tissue present rather than the concentration (depending on the goals of the research). When the data are not spatially warped (voxel dimensions and place remain unaltered) concentration and volume correspond one to one. Normalising data introduces shape/volume changes in the data, it is possible to correct the concentration in each voxel, to account for the total volume change of the tissue that is represented by that voxel.

Before analysis, the volumes are smoothed with smoothing kernels of different sizes, to improve signal/noise ratio and pool voxels belonging to the same functional region. A wide range of statistical tests is available to test hypotheses about the data. More about the various post-processing steps will be said in paragraph 4.4 below.

## **4.3 Statistics**

Statistical tests are used to test hypotheses. Generally, they are used to disprove a null hypothesis with a certain level of confidence (alpha level). In morphometry this null hypothesis usually states that there are no significant structural differences among a number of populations, or related to particular covariates such as age [Ashburner 2000, Ch. 6].

SPMs of univariate data allow us to address simple questions like whether there is a change in a particular measure, in correlation with some effect of interest. Standard t- and F- tests can be used to test the hypothesis within the general linear model (see below). Gaussian Random Field (GRF) theory is used to correct for multiple dependent comparisons and produce the appropriate rate of false-positive results. Without this correction, the number of false-positive results would be proportional to the number of independent tests; a large number when comparisons for each voxel are made. The multiple comparison problem is discussed in [Worsley et al. 1996]. In paragraph 4.3.3 below, a short introduction to the problem is given.

The theory on Statistical Parametric Mapping usually deals with functional imaging, such as PET or fMRI. In general, the response of the subject to a controlled change in the environment is examined. The subject is offered stimuli whilst scans are made of the brain. Effects of the stimuli, the response,

can be detected by comparing the scans after or during the stimulus with scans in a different state, usually rest. In VBM an analogy to functional MRI is made by stating that each of the subject groups corresponds to a different stimulus or state of the mind e.g. action and rest. Instead of placing the scans of a single subject over time, and measuring the effect of the expected conditions, we place the scans of different individuals 'over time' and measure the effect of the group that the image belongs to.

For the purpose of this report, the descriptions are adjusted to be directly applicable to a structural MRI study.

#### 4.3.1 A General linear approach for statistical parametric mapping in functional imaging

This section is a brief extract from [Friston et al. 1995]. For more information and a thorough explanation on the subject we recommend this paper.

SPM99 uses a general linear model to construct the statistical images. The general linear model for a response variable  $x_{ij}$  at voxel  $j = 1, \dots, J$  is:

$$x_{ij} = g_{i1}\beta_{1j} + g_{i2}\beta_{2j} + \dots g_{iK}\beta_{Kj} + e_{ij} \quad (1)$$

where  $i = 1, \dots, I$  indexes the subject. Here the  $\beta_{kj}$  are  $K$  unknown parameters for voxel  $j$ . The coefficients  $g_{ik}$  are explanatory variables relating to the conditions under which the observation (e.g. scan)  $i$  was made. The errors  $e_{ij}$  are assumed to be independent and identically distributed normally  $[N(0, \sigma_j^2)]$ .

Equation(1) can be written in matrix form as:

$$\mathbf{X} = \mathbf{G}\boldsymbol{\beta} + \mathbf{e} \quad (2)$$

where data matrix  $\mathbf{X}$  has one column for each voxel  $j$ , and one row for each scan (subject). Matrix  $\mathbf{G}$  contains the coefficients  $g_{ik}$ , and is called the design matrix. This design matrix has one row for every scan and one column for each effect in the model.  $\boldsymbol{\beta} = [\beta_1 | \beta_2 | \dots | \beta_J]$  is the parameter matrix where  $\beta_j$  is a column vector of parameters for voxel  $j$ . A least squares estimate of  $\boldsymbol{\beta}$ , say  $\mathbf{b}$ , satisfies the equation:

$$\mathbf{G}^T \mathbf{G} \mathbf{b} = \mathbf{G}^T \mathbf{X} \quad (3)$$

If  $\mathbf{G}$  is of full rank then the least squares estimate is uniquely given by

$$\begin{aligned} \mathbf{b} &= (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{X} \\ \text{where } E\{\mathbf{b}_j\} &= \beta_j, \text{ and} \\ \text{Var}\{\mathbf{b}_j\} &= \sigma_j^2 (\mathbf{G}^T \mathbf{G})^{-1} \end{aligned} \quad (4)$$

The error  $\mathbf{R}(\Omega)$  is obtained from the difference between the actual and the estimated values of  $\mathbf{X}$  ( $\Omega$  is the alternative hypothesis to the null hypotheses and includes the effects of interest):

$$\mathbf{R}(\Omega) = (\mathbf{X} - \mathbf{G}\mathbf{b})^T (\mathbf{X} - \mathbf{G}\mathbf{b}) \quad (5)$$

The significance of specific effects can be assessed with the t-statistic using linear compounds or contrasts of the parameter estimates  $\mathbf{b}_j = [b_{1j}, b_{2j}, \dots]^T$ . For example if we wanted to test for activation effects between conditions one and two then we would use the contrast  $\mathbf{c} = [-1 \ 1 \ 0 \ 0 \dots]$ . The significance of a particular linear compound of effects at voxel  $j$  is tested with:

$$t_j = \frac{\mathbf{c} \mathbf{b}_j}{\varepsilon_j} \quad (6)$$

where the standard error  $\sigma_j^2$  for voxel  $j$  is estimated by  $\varepsilon_j^2$ :

$$\varepsilon_j^2 = (\mathbf{R}_j(\Omega)/r) \mathbf{c} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{c}^T \quad (7)$$

$t_j$  has the Student's t-distribution with  $r$  degrees of freedom. When gathered in a volume, the  $t_j$  constitute an  $\text{SPM}\{t\}$  and represent a spatially extended statistical process that directly reflects the significance of the profile of effects 'prescribed' by the contrast  $\mathbf{c}$ .

#### 4.3.2 Statistical tests using the general linear approach, an example

A wide range of different statistical parametric maps can be obtained using the described general linear approach, including t-tests, standard regression analysis and An(C)oVa. As a simple example, we will

show how a t-test is described using this model. When assuming that the data are mean-corrected, we can specify a t-test in the following way:

Suppose we have two groups of subjects, group one with  $N_1$  and group two with  $N_2$  subjects. The design matrix  $\mathbf{G}$  consists of two columns, one for each group. A value of 1 indicates that the scan belongs to that group, otherwise, a value of 0 is assigned.

The contrast vector  $\mathbf{c}$  is defined as  $[1 \ -1]$ , (i.e. second condition smaller than the first)

$\mathbf{b}_j$  from equation (4) becomes  $[\mu_{1j} \ \mu_{2j}]^T$ . The mean value of voxel  $j$  in group 1 ( $\mu_{1j}$ ) and in group 2 ( $\mu_{2j}$ ).

The sum of squared errors for voxel  $j$ , by equation (5):

$$\mathbf{R}_j(\Omega) = (\mathbf{X}_j - (\mathbf{G}\mathbf{b})_j)^T (\mathbf{X}_j - (\mathbf{G}\mathbf{b})_j) = \sum_{i=1}^{N_1} (x_{ij} - \mu_{1j})^2 + \sum_{i=N_1+1}^{N_1+N_2} (x_{ij} - \mu_{2j})^2$$

$\mathbf{c} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{c}^T = \frac{1}{N_1} + \frac{1}{N_2}$ , as a result of the specific design of  $\mathbf{c}$  and  $\mathbf{G}$ .

$r = N_1 + N_2 - 2$ , the degrees of freedom; here the number of subjects minus the number of estimates used to estimate the variances (2, because of the two estimates for the group means).

$$\text{The error is estimated by equation (7): } \varepsilon_j^2 = \frac{\sum_{i=1}^{N_1} (x_{ij} - \mu_{1j})^2 + \sum_{i=N_1+1}^{N_1+N_2} (x_{ij} - \mu_{2j})^2}{N_1 + N_2 - 2} \left( \frac{1}{N_1} + \frac{1}{N_2} \right) = s_j^2 \left( \frac{1}{N_1} + \frac{1}{N_2} \right)$$

Now from equation (6):  $t_j = \frac{\mathbf{c}\mathbf{b}_j}{\varepsilon_j} = \frac{\mu_{1j} - \mu_{2j}}{s_j \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}$ , which is the familiar t-statistic.

#### 4.3.3 Multiple comparison problem

For correcting the p-values derived from the SPM, the expected number of clusters above a certain threshold within a random field is used. This expectation is based on the Euler characteristic. For a full overview of the used models in SPM99, see [Worsley et al. 1996]. An uncorrected p-value for a t- or Z-score is the chance that, given the characteristics of the random field, at a certain voxel that or a

higher score occurs. The total number of independent observations in a common SPM is so large, that even values corresponding to very low p-values will occur regularly by chance. For this reason, a correction should be made for the large number of comparisons. A corrected p-value can then be regarded as the chance that one or more voxels in the total search volume have equal or larger values.

The motivation to use the expected Euler characteristic to correct p-values lies in the observation that for high thresholds, near the maximum in a gaussian field, the Euler characteristic takes the value 1, and 0 for thresholds higher than this maximum. The expected Euler characteristic can then be regarded as the chance of obtaining one or more values above this threshold, which is precisely needed for the corrected p-value.

The same theory is used to obtain a threshold for the SPM that produces the correct false positive rate. In short, the user gives a confidence level, or alpha level, usually 0.05, at which the statistical test is performed. Now, based on the Euler characteristic, a threshold is calculated for which the data is expected to exceed the calculated threshold by chance the number of times corresponding to the alpha level. For an alpha level of 0.05, this means that an SPM with the same spatial properties is expected to exceed the threshold value only 5 out of 100 times (the expected number of false positives).

Corrected p-values are calculated based on the statistical score, data-smoothness and volume of interest. The threshold corresponding to a certain alpha level is obtained by the inverse function, in which alpha is the corrected p-value.

#### 4.3.4 Cluster Level tests

SPM allows the user to perform tests based on the number of connected voxels in a cluster. The user sets the threshold for the number of voxels within a cluster. For this test to be valid, the smoothness of the residuals is required to be spatially invariant. For anatomical data is known not to be the case due to the non-stationary nature of the underlying neuro-anatomy. In [Ashburner et al. 1999] the extent statistic is tested on data for VBM analysis, and found that the voxel-based extent statistic should not be used with VBM. Three times more false positive clusters were found than the expected number. They also tested the height correction for use in VBM and in this case the number of false positives matched the theory. It seems that the latter correction is more robust with respect to the variances in the data smoothness.

### 4.4 Data processing

Before statistical tests are applied to the data, some processing is done, for example aligning scans and removing artefacts in the data. In this paragraph, the most important processing steps for VBM

(normalisation, segmentation, and modulation) are briefly covered before current VBM protocols are given.

#### 4.4.1 Normalisation

To be able to compare scans of different subjects, they are normalised into the same stereotactic space. This is done by applying a (non-)linear warp, that matches the data to some template in a standard space. The use of spatially normalised images also facilitates comparison of results from different studies. Apart from aligning different subjects, spatial normalisation is necessary when apriori information about tissue distributions is used in segmentation.

The most commonly used co-ordinate system within the brain imaging community is that described by the atlas of Talairach and Tournoux (Talairach space). However, SPM99 uses standard brains from the Montreal Neurological Institute (MNI). These templates are approximately matched to the Talairach brain using a linear transform, but they do not match completely.

To maximise the correspondence between images, a non-linear transformation should be applied. Different methods for non-linear normalisation exist, which are broadly divided into label-based and non-label based methods. Label-based methods identify corresponding features in image and template, and try to find a spatial warp that best superposes these features. The identification of these corresponding features is usually done manually (by indicating corresponding points or lines), and is a time-consuming process. One could also use corresponding surfaces. The extraction of these surfaces can be (semi-) automated. When using feature points, the transformation is known at these points, and between the points the transformation is forced to be as smooth as possible. One method is fitting splines through the points in order to minimise bending energy. For surfaces, methods that are more complex are often used (e.g. fluid model).

The non-label-based approaches try to minimise some index of difference between object image and a template of the same modality. Usually, this means minimising the sum of squared differences or maximising the correlation coefficient between warped image and template. For such an approach, the template should appear to be a warped version of the object image, this means that there should be a correspondence in the grey-levels of the different tissue types between the object and template image.

The non-label-based method adopted by SPM99 and described by John Ashburner in his PhD thesis, tries to minimise the residual squared difference by performing a non-linear warp defined by a linear combination of low spatial frequency basis functions. A maximum smoothness is maintained during the process using a maximum a posteriori (MAP) approach.

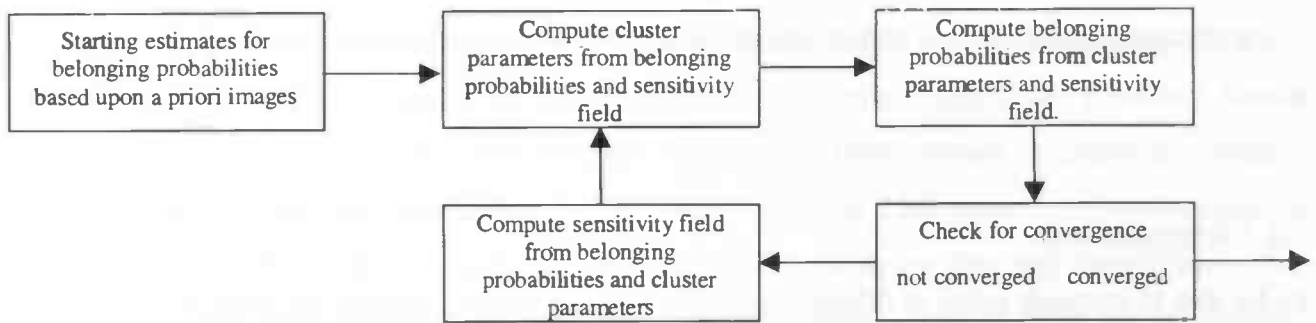


Figure 1: SPM99 segmentation algorithm

#### 4.4.2 Segmentation

From MR images, the human brain can be broadly divided in three tissue types. Grey matter, white matter and CSF. SPM99 uses a fully automated clustering algorithm for segmentation. The voxels are assumed to be drawn from three distinct tissue types. The intensities of the voxels within each cluster conform to a multivariate normal distribution. Such a distribution can be described by a mean vector, covariance matrix and number of voxels belonging to the cluster. Using prior probability images, information about the spatial distribution of the clusters is incorporated into the model. To be able to use this spatial information, the volume is assumed to be normalised into a standard stereotactic space.

The intensity of voxels within the same class can vary smoothly over the image due to the physics behind magnetic resonance imaging. These non-uniformities can greatly influence tissue classification. To perform a reliable segmentation some form of intensity inhomogeneity correction should be applied. The non-uniformity field is assumed to vary smoothly with the data, and is modelled as multiplicative noise. If the smoothness is not forced onto the inhomogeneity field, the field will fit to the higher frequencies in the data, which are due to different tissue types, rather than to the low frequency intensity inhomogeneity.

SPM99's segmentation procedure uses a modified mixture model cluster analysis, which incorporates this correction for intensity non-uniformity see [Ashburner, 2000, Ch. 5] for a full description). It consists of an iterative procedure that segments the data into probability maps. The MRI volume is segmented into 5 distinct clusters (probability maps): GM, WM, CSF, non-brain and background voxels. Each cycle of the algorithm consists of four steps:

1. estimate cluster parameters, using the non-uniformity corrected image
2. assign new belonging probabilities based on the cluster parameters
3. check for convergence. The algorithm is terminated when the change from the previous iteration

becomes negligible.

4. make a new estimation of the modulation (non-uniformity) field, and apply this to the original image.

The process begins by assigning starting estimates for each of the parameters. The modulation field typically has a constant value equal to 1. Initial belonging probabilities of WG, GM and CSF are based on the corresponding prior probability maps. No prior probability maps are available for background and non-brain voxels, for these clusters, the starting estimates are made by subtracting the maps for GW, WM and CSF from a map of all ones, and dividing the result equally between the remaining clusters.

To improve the algorithm, the additional information in a registered T2 weighted image could be incorporated into the analysis. When T1 and T2 relaxation times are combined for each voxel, the separation between tissue classes becomes more distinct. (This is not used in SPM99.)

#### 4.4.3 Preservation of tissue volume

Non-linear normalisation will change relative tissue volumes. In order to preserve tissue volumes (if desirable), the voxel intensities in the spatially normalised and segmented images should be modulated according to the relative volume change in each voxel. In SPM99 this is done by multiplying with the determinant of the Jacobian of the transformation matrix from template to original image. The transformation matrix is stored in a (.mat) file, which is created during normalisation. A script `spm_preserve_quantity` that implements the mathematics is written by John Ashburner as an additional step to the basic VBM method, and is available at the SPM web-site.

#### 4.4.4 Smoothing

Smoothing of the volumes before statistical analysis is done for two reasons. The first is to increase signal to noise ratio by pooling voxels which have the same function. For VBM this means that the accuracy of the registration will influence the amount of smoothing needed. Since the anatomical registration error in the cortex is about 9mm [Hellier et al. 2002], it is recommended to use a larger smoothing kernel to reduce the chance of detecting differences due to misregistration. The registration errors were estimated using corresponding structures from anatomical images, but the functional correspondence between regions cannot be deducted from these images, so the registration errors for functionally equivalent structures are likely to be even larger.

The second reason is that SPM99 relies on the theory of gaussian random fields for statistical



inference. In SPM99 the average smoothness over the entire search volume is used for the multiple comparison correction and cluster level inference (see 4.3.3 multiple comparison problem). For these corrections to be valid over the entire volume, the residual fields should conform to a stationary gaussian field, e.g. constant smoothness over the search volume. Less smooth regions contain more independent observations than average, so more independent t-tests are done and there is more chance of getting a false positive result (in these less smooth regions), thus if there are differences in smoothness, the single estimate is not valid. A way to overcome this problem (in case of the height correction) is by using a local smoothness estimate, but this is not done in SPM99. Smoothing the data with a gaussian smoothing kernel renders the residual fields more normally distributed. In anatomical data, the inherent spatial correlations vary over position, so even if the residual fields conform to a gaussian field, they are not stationary. Smoothing the data using larger kernel dimensions (with Full Width at Half Maximum (FWHM) of 8-12mm or more) renders the residual fields also more stationary.

The general rule of thumb to maximise sensitivity is to choose a smoothing level approximately equal to the expected size of the signal. However, as mentioned above smoothing at less than 9mm FWHM (for multi-subject VBM) is not recommended. Using multiple smoothing levels a multiresolution analysis can be performed, which may be worthwhile.

## **4.5 Current VBM Protocols**

### **4.5.1 Standard VBM protocol**

The basic VBM protocol (see figure 2), defined in [Ashburner et al. 2000] consists of five steps:

1. Template creation
2. Spatial normalisation
3. Segmentation
4. Smoothing
5. Statistics

An extra step after segmentation is suggested to be a volume change correction step (see 4.4.3).

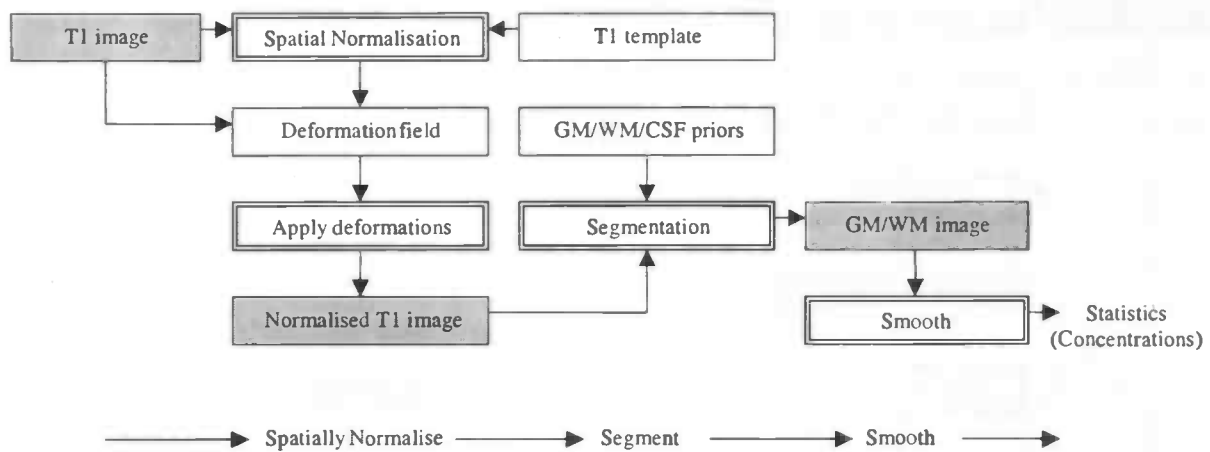


Figure 2: standard VBM protocol

#### 4.5.2 An optimised VBM protocol:

An optimised segmentation/normalisation procedure for VBM has been proposed by [Good et al. 2001] (see figure 3). Since VBM compares the segments of a single tissue type of interest, the segmentation can be optimised for this specific tissue. Normalisation of the original data is performed using the transformation that best maps the segment of interest to the general template for this segment. In this way the portions of the data that belong to this class are better aligned, and segmentation is likely to be more accurate. For grey matter this would be:

1. Segment non-normalised images, create a mask for the brain tissue and use this mask to cleanup the segments
2. Estimate spatial normalisation parameters from the cleaned up grey matter image by matching the grey matter segment to a grey matter template.
3. Write spatially normalised versions of the original images by applying the estimated warps to them.
4. Segment the spatially normalised images, and again clean up the segments.
5. Optional modulation (volume preservation procedure).
6. Smoothing
7. Statistics.

Of course, cleaning up of the segments can also be done in the standard protocol.

One comment on this protocol is that segmentation isn't meant to be performed on non-normalised

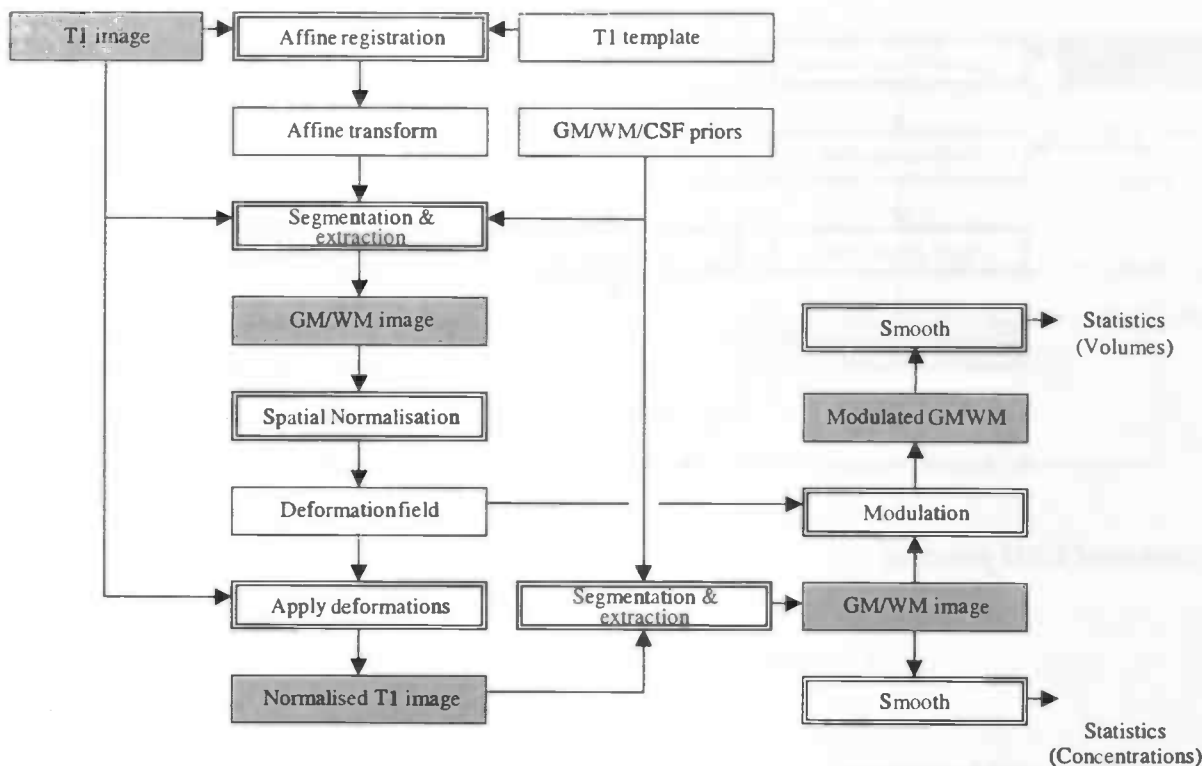


Figure 3: Optimised VBM protocol

images, since it relies on spatial a priori information. This might occasionally introduce errors.

#### 4.6 Data acquisition and properties

Current MRI scanning devices can produce images that have voxel dimensions down to  $0.5 \times 0.5 \times 0.5 \text{ mm}^3$ . To reduce the presence of partial volume effects, a very high-resolution image should be obtained. For segmentation it is also important that the signal-noise ratio in the images is high, which will produce more distinct clusters due to low variance in the intensity values within each cluster. In general, the best resolution available will be near  $1 \times 1 \times 1 \text{ mm}^3$ . The cerebral cortex (grey matter) has a thickness varying from 2 to 5 mm, thus is in some areas no more than two voxels wide. Due to the relatively low resolution of the scans partial volume effects occur. The segmentation algorithm should deal with these effects until higher resolution data with similar quality is available.

A large scanning time might produce motion artefacts due to movement of the subject. In practice, scanning time for anatomical data is short enough to prevent data-distortions due to movement. In general this is much less an issue for anatomical than for functional studies, where subject motion will cause consecutive scans to misalign.

A typical  $1 \times 1 \times 1 \text{ mm}^3$  scan has volume dimensions of  $256 \times 256 \times 168$  voxels. For 16 bit uncompressed data, this yields single volume data files of over 20MB. Each processing step will create more data, although these files will be smaller since normalisation will fix the region of interest (brain)

and a large portion of non-brain data can be excluded, reducing size by more than half. In addition, segmented images can be set to 8 bit, halving the file size again. Still, a lot of space is required to store data especially for a large number of subjects. When having limited storage space, special attention should be given to file management.

## 5 Implementation of the VBM routines

### 5.1 Environment

The system on which the data are processed consists of a SUN Enterprise 450 server, running UNIX Solaris 2.6, Matlab 6 and SPM99.

The software currently used is organised in a number of layers: on top are the user defined procedures and input data, which control the automation scripts, created by the K.U.L. (Royal University of Leuven, Belgium), and partly written on site. These automation scripts provide an environment for automatically processing and analysing fMRI data, using SPM99 functions. All scripts are written for Matlab.

Instead of K.U.L. functions. SPM99's batch mode could be used for automating, however, this requires much more expertise from the user.

### 5.2 VBM process description files

To be compatible with the current file naming and storage conventions used in the current PET/MRI analysis protocols, the directory structure and file naming for each project is assumed to follow some rules. Each VBM project is stored in it's own directory. Scans for a subject are written to a subfolder `mxxxx/`, `xxxx` is the subject identifier. In this folder, anatomical scans from one session are put in folder `antxx/`, where `xx` indicates session number.

For each group, the files (\*.img) with full pathname that are to be included in the VBM analysis are listed in a file called `groupx.txt` (`x` indicating group number and starting from 1), each file on its own line. These filenames are read into a data structure using the function `read_filenames`. Parameters for the pre-processing are described in a file called `descrip_preprocess.txt`. It contains the following fields:

```
TISSUE: grey | white | both
NORMALISATION: linear | standard | optimised
MODULATION: none | partial | full
CLEAN: yes | no
SMOOTH: {fwhm1 <,fwhm2> *} % array of integers
```

```

SCANS: n % integer, n is the total number of
          scans in the project
NGROUPS: ngroups % integer: number of groups >= 1
NUISANCE: [val_1;val_2; ... ;val_n] % optional, array of n values
NUISANCENAME: name % optional, string
MASK: maskfile % optional, file used to mask images in
                the final statistical analysis

```

A typical processing script for VBM would look like this:

```
<Process.m:>
```

```

spm_fmri;
descr = vbm_read_descrip('descrip_preprocess.txt');
files = '';
groups = [];
for i = 1:descr.ngroups;
    group = read_filenames(['group' mat2str(i) '.txt']);
    groups = [groups; {group}];
    files = strvcats(files,group);
end
global basedir fmriDIR VWD;
basedir = pwd;
fmriDIR = basedir; %required for K.U.L. functions
VWD = spm('dir');

normalise_process_VBM(files,descr);
segment_process_VBM(files,descr);
modulate_process_VBM(files,descr);
smooth_process_VBM(files,descr);
stat_process_vbm(descr,groups);

```

In case a process is for some reason interrupted, the process can be resumed by commenting the fully executed steps in the `process.m` file, and restarting this script.

### 5.3 Segmentation

To implement the segmentation step, the file `spm_segment.m` is used. For consistency with the K.U.L. scripts, a file `f99_segment.m` is created, which handles automation and logging. The function `segment_process_vbm` implements the standard segmentation protocol for VBM, whereas the

function `normalize_process_VBM`, apart for standard normalisation, implements the enhanced segmentation/normalisation protocol. Both processes are called from within the `process.m` file, based on the description file, the correct segmentation function is chosen.

For registration of the WM segment to the WM (a priori) template during the optimised protocol, the images should be masked with the extracted brain mask, otherwise the process frequently fails. Also for grey matter normalisation, this mask is applied. See [Ashburner et al. 1998] for the use of masks in normalisation.

## 5.4 Segment Cleaning

If the parameter `clean` is set to 'yes', then the segments are 'cleaned' with the `spm_imcalc` function using  $ix/(i1+i2+eps).*i3$ , where `ix` is current segment, `i1` is the grey-matter segment, `i2` white matter, and `i3` the extracted brain mask obtained by `spm_xbrain` function. In this way small regions of misclassified non-brain voxels are removed from the segments, and GM and WM segments are forced to add up to 1 at each voxel. Cleaned segments have a prefix '`_clean`' added to the filename.

## 5.4 Account for volume changes

Volume preservation is implemented in the function `spm_preserve_quantity`. This function is called for each volume in the `modulate_process_vbm` script. The resulting modulated files have an '`m`' prefixed to the filename. The modulation process does not do any processing if this is indicated in the description file. Thus, it is unnecessary to remove the function call from the process script.

## 5.5 Statistics

A simple automated multiple groups AnCoVa analysis with one optional nuisance variable is provided, in which the design parameters are chosen to be suitable for VBM. The parameters are:  $n$  groups; the files that are output by the processing and indicated by the description file and group files; the specified nuisance variable; no grand mean scaling; no threshold masking; not masking zero's; optional explicit mask for the data; no global calculation.

By calling `stat_process_vbm` this procedure is started. In the `descrip_preprocess.txt` file the nuisance variable, if any, should be provided. When the data for this variable are known prior to the whole data-processing procedure (for example sex or age), the example `process.m` script can be used without modifications. However, if the data is not known until after the pre-processing, for example total intra-cranial volume (usually obtained by adding the grey, white and CSF segments),

stat\_process\_vbm should not be executed until the correct data is added to descrip\_preprocess.txt. Of course, the user is free to do a manual analysis with a custom statistical design on the data, and in this case only use the scripts for automated post-processing, without statistical analysis.

## 6 Sensitivity of the method

For evaluation of the performance of VBM the process should be tested on data with a form of controlled atrophy. In absence of suitable data from previous research (in which we know the amount and location of atrophy, determined by an alternative method to VBM), the test data are to be constructed from available data.

Questions of interest are:

- How small is the amount of atrophy that is needed for VBM to detect differences? What is the sensitivity of VBM, with respect to lesion severity?
- How large should be the number of subjects in each data set be to make inferences about the difference between groups?
- How are these two properties related? (If differences are large, smaller groups are needed to get significant results.)
- Sensitivity of the segmentation with respect to the within-tissue noise level, given a fixed lesion.
- Sensitivity of VBM with respect to group sizes or lesion severity and positional variance of the lesion between subjects.

To keep the evaluation process close to a practical situation, the simulated data should closely represent such real data. The atrophy will in general be a loss in grey and/or white matter and an increase in CSF. This gives rise to the idea of performing some erosion on the segmented GM and WM probability maps, and combining the new segments to form a new data set. The reconstruction algorithm should be able to maintain the different tissue properties, noise levels and inhomogeneity of the data set.

Adjustable parameters for the simulation to be able to address the above questions include:

- severity (intensity and/or size) of the simulated lesion
- inter-subject positional variance of the lesion.
- within tissue noise level of the simulated MRI image (this affects the segmentation sensitivity, and might obscure the lesion).

- Group sizes

## 6.1 Experiment design

All experiments to assess the given questions are similar in design. An experiment basically consists of the simulation of two groups of subjects to be input to the SPM99 statistics. This process is repeated (a number of times) for each value of a gradually changing input parameter. The resulting SPM's are then processed to visualise the results of the experiment. For each experiment, our null-hypothesis is that there is no significant difference between the mean grey matter volumes of the two different groups.

General design:

- Intensity inhomogeneity normalised MRI images are simulated for each person based on the original data, one volume with lesion, and one volume without lesion (see next section). In this step, lesion severity and positional variance, and noise level can be adjusted.
- These images are segmented, and for the grey-segment a mean image for each group is constructed.
- For each group, based on the mean GM segment, a number of single individual images are constructed, which will be our simulated subject data (see 6.2 for details). These images are smoothed with a 10mm FWHM gaussian kernel, and input to SPM99.

Simulations are based on T1 MRI scans of 9 persons (obtained with a 1.5 Tesla GE scanner,  $T_R$  30 ms,  $T_E$  5 ms, flip angle  $30^\circ$ , field of view 26 cm,  $256 \times 256 \times 160$  matrix size at a resolution of  $1 \times 1 \times 1.5$  mm). For efficiency only one slice at  $z = -9$ mm (intersecting the visual cortex) is used.

## 6.2 Simulated atrophy in MRI-volumes

The experiments and results will be covered at the end of this chapter, first the simulation of the lesions in MRI images and the construction of simulated GM segments are explained.

### 6.2.1 Generation of an MRI data set with controlled atrophy in a specified region.

A first step in testing our VBM implementation is creating controlled lesions in existing MRI-volumes. These can be used directly as input to VBM, or we can use these for further simulation experiments. A region of the GM and WM segment volumes is eroded. For each voxel the removed amount is added to the CSF volume. The segmented components are reintegrated to form the new data set.



For the calculations the following data are used:

- original volume
- intensity inhomogeneity corrected volume (output by the `spm_segment` procedure).
- GM, WM and CSF volumes (output by the `spm_segment` procedure)

### **Create lesion in GM and WM.**

The grey and white matter atrophy that is simulated, is assumed to be local and near the boundaries between CSF and grey- or white matter regions, where the tissue is replaced by CSF. To accomplish this, grey-scale erosion in a certain region of the combined GM and WM segments is performed. The severity of the reduction is determined by the structuring element used for this operation. In the choice of shape, size and values for this kernel there is much freedom. More information about erosion and the typical structuring element that is used can be found in appendix A.

```
grey_white_lesioned = erode(white_segment + grey_segment)
```

### **Construction of the new segments.**

Assuming that the erosion first affects the grey-matter, the new grey segment can be obtained by subtracting the original white segment from the eroded (combined grey-white-matter) volume. In areas where mainly white matter is replaced, this yields negative values. Since there is no such thing as a negative concentration or probability, these negative values in the new grey segment are set to zero:

```
grey_lesioned = grey_white_lesioned - white_segment  
mask = grey_lesioned < 0; grey_negative = mask .* grey_lesioned;  
grey_lesioned = grey_lesioned - grey_negative;
```

After the erosion has fully cleared the grey matter portion in some voxel, the white matter is affected. The negative values from the previous step are passed to the white matter segment, resulting in a valid distribution of grey and white matter in this voxel.

```
white_lesioned = white_segment + grey_negative;
```

The eroded grey- and white matter is assumed to have changed into CSF, therefore the difference between original combined grey- and white matter segment with the eroded segment is added to the CSF segment:

```
csf_lesioned = csf_segment + ((white_segment + grey_segment) -
grey_white_lesioned);
```

The non-brain parts of the data remain unchanged. We need this data to fill-in the other parts of the reconstructed volume. The sum of the three segment volumes gives a mask for the brain data (values in the range of 0-1; 0 = non-brain, to 1 = 100% brain-tissue). The (inhomogeneity corrected) non-brain data are now calculated by:

```
non_brain = (1 - (grey_segment + white_segment + csf_segment)) .* inhomogeneity
corrected original
```

### Combining the constructed volumes.

The translation of the segmented volumes to a representative brain volume is ideally done by a simple multiplication of each segmented volume with a voxel intensity that is characteristic for the corresponding tissue type, and adding these three volumes to the non-brain volume. A 50% grey and 50% white-matter voxel thus gets an intensity of  $0.5 \times \text{'intensity for grey'} + 0.5 \times \text{'intensity for white'}$ . Partial volume effects will create a less distinct boundary between segments where voxels can't be exactly classified. Unfortunately there is no such thing as for example the 'intensity for grey', because the intensities for each tissue type fall within a certain range. Moreover there might be a distinct pattern in the voxel intensities within the voxels of one tissue class. To include some intensity variance, noise is added to the image, assuming that voxel-intensities for each cluster are characterised by a normal distribution. Mean and variance values were chosen by examination of a number of regions of interest for each tissue type in a number of scans.

Mean voxel intensities of 1230, 1700 and 470 for respectively GM, WM and CSF are used. For intensity variation a zero-mean gaussian noise with standard deviation of 85 is added. Because the noise is only applied to the combined segments (non-brain remains unchanged), the noise volume is multiplied with the brainmask.

```
reconstructed = noise + non_brain + grey_lesioned * grey_mean + white_lesioned *
white_mean + csf_slesioned * csf_mean
```

### Adding inhomogeneity field.

For more realism, the generated image so far can be multiplied with the inhomogeneity field for which SPM99 has corrected the original image during segmentation. Because original image = inhomogeneityfield .\* corrected image, the intensity inhomogeneity field is calculated by:

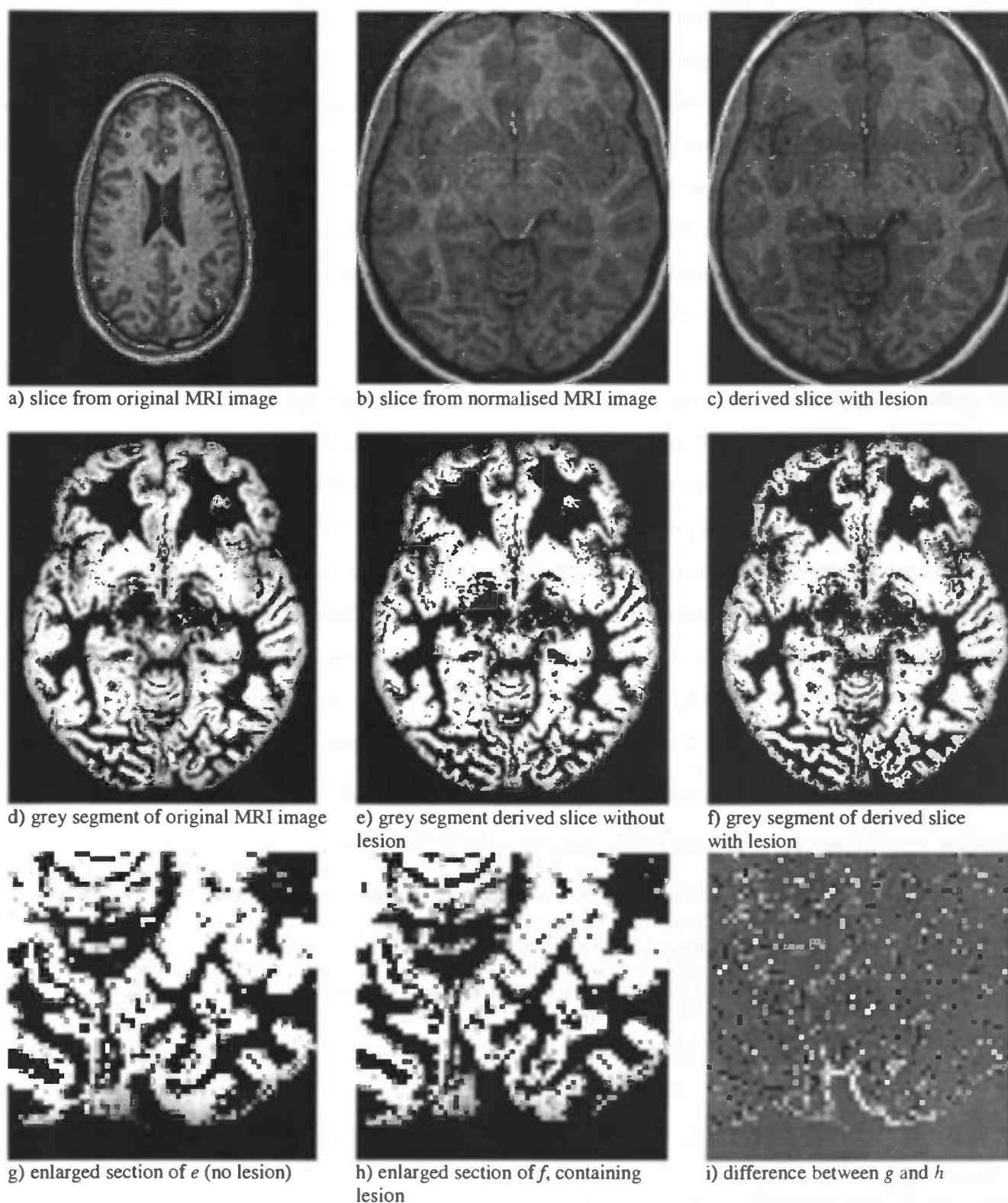
`inhomogeneityfield = original image ./ (corrected image + eps)`. In this formula `eps` is added to prevent division by zero.

`reconstructed = reconstructed .* inhomogeneityfield`

If the images were normalised, the inverse transformation can be applied to obtain images in native space. The standard bounding-box for normalisation should then be adjusted to preserve the total volume data (not just the brain area).

Means:	grey	white	CSF
sample size	44	219	42
intensity	1,23E+03	1,70E+03	0,47E+03
$\sigma$	76,0	85,0	105
noise level	6,2%	5,0%	22,3%

*Figure 4: Means and standard deviations from manual measurements on two intensity inhomogeneity corrected slices, obtained in Matlab, using a polygonal region of interest. In each image, three ROIs for each tissue type were selected.*



*Figure 5: Creating a simulated lesion in a T1 MRI image. (Original and resulting images.)*

## 6.2.2 Simulating Segmented Data

MRI volumes created by the described procedure are suitable for input to the entire VBM process. A drawback however is that we are limited to the original number of subjects. To overcome this problem, we will try to generate images with the same statistical properties as original data. The generation of a completely new morphologically probable volume (which is needed for normalisation and segmentation to be applicable) is beyond the scope of this research. Instead of truly simulating brain morphology, we will focus on devising a GM segment, based on the simulated atrophy images, that has realistic spatial distribution and data variance. These simulated segments will then be smoothed and input to the statistical analysis.

In order to simulate a grey matter belonging-probability map for a single subject, we have to devise a model that describes how the data is formed.

Our first assumption is, that at each point in the human brain there are two possible situations: 1) the tissue at this location is grey matter, 2) it is not grey matter. Let us now divide the human brain in a regular grid of small elements, so that in each element only one of these situations is true. A voxel contains a number of these elements. The value of a voxel is defined as the fraction of the volume of the voxel that consists of grey matter. Thus the voxel value is calculated by dividing the number of grey matter elements in that voxel by the total number of elements in a voxel.

At each point in the brain there is a certain probability  $p_j$  that there is grey matter present; where  $j$  indexes the element. If we have a group of subjects, an estimate for  $p_j$  is the number of subjects having grey matter at  $j$  divided by the total number of subjects. By the law of large numbers this converges to  $p_j$  for a large number of subjects. However, we do not know for each particular element whether it consists of grey matter or something else, since MRI scanning resolution is not near high enough. Let us consider voxel  $I$  containing a set of elements. The mean  $p$ -value  $\bar{p}_I$  within voxel  $I$  can be estimated as follows:

$$\bar{p}_I = \frac{1}{|I|} \sum_{j \in I} p_j \equiv \frac{1}{|I|} \sum_{j \in I} \frac{1}{S} \sum_{k=1}^S \text{brain}_j^k = \frac{1}{S} \sum_{k=1}^S \frac{1}{|I|} \sum_{j \in I} \text{brain}_j^k = \frac{1}{S} \sum_{k=1}^S v_I^k$$

Where,  $\text{brain}_j^k = \begin{cases} 1, & \text{if subject } k \text{ has grey matter at element } j \\ 0, & \text{otherwise} \end{cases}$ ,

and  $v_I^k$  = value of voxel  $I$  in the grey-matter segment of subject  $k$ . The mean  $\bar{p}_I$  within voxel  $I$  is thus estimated by the mean value of the grey matter volumes  $v_I^k$  over subjects  $k = 1, \dots, S$ . The  $p$  value at an arbitrary point in the brain can be estimated by interpolating the  $p$ -values at the nearest voxels.

A second assumption concerns the fact that the actual presence of grey matter at an element is dependent on its surrounding area. Given that a neighbour element contains grey matter, it is more likely than can be expected from the actual  $p_j$ , that the current element also contains grey matter. The influence of the surround can be expressed by a variance/covariance matrix. SPM estimates the smoothness of data from this variance/covariance matrix, and expresses it as the FWHM of a gaussian smoothing kernel, that produces a Gaussian Random Field (GRF) with similar smoothness. The smoothness estimation at each voxel will be used to model the spatial coherence in the data, see below how this estimation is obtained.

A simulation of an MRI segment can now be made by applying a 'spatially-coherent' Bernoulli process using the estimated p-values and smoothness for each point. The standard Bernoulli process can be implemented in the following way: take a sample  $s$  from a uniform distribution over the interval  $[0, \dots, 1)$ . If  $s < p$  then the result is 1, otherwise the result is 0.

We could create a simulated segment by performing a Bernoulli experiment for each voxel, and using the corresponding p-value as the threshold. However, if we perform this experiment for each voxel independently, the spatial coherence is lost. To resolve this, we create a random field that at each point has the desired smoothness, and in which all values between 0 and 1 have approximately an equal chance of occurring, for the details of implementation, see the next section. The value to be assigned to a voxel is now determined by comparing the value  $\tilde{s}_i$  in the adaptively smoothed field with the estimated p-value for that voxel as above: if  $\tilde{s}_i < \bar{p}_i$  then 1, else 0.

Directly assigning voxel-values in the final image, using the described process, produces binary images. To create partial volume effects, the process can be performed at a higher resolution than the original scan-data. Higher resolution random field and p-value estimates are used to generate a binary high-resolution segment. Then, sub-sampling and averaging creates grey-scale data at the original resolution. Adding partial volume effects will produce voxel variances closer to real data.

### 6.2.3 Modelling spatial coherence

Anatomical data has an inherent spatial coherence that is not the same throughout the brain. As a measure of spatial coherence in anatomical data, we will use the estimated smoothness of the data at each voxel. Of course, the smoothness of the data doesn't represent the true physical auto-correlation in the data, but it is the best thing we have.

Smoothness is estimated by SPM99 based on a subset of the residual images (the difference between the actual and estimated data values). If we take a simple design matrix, like we use in the

simplest form of AnCoVa analysis, these residual images are the original images minus the mean for each group. SPM99 estimates the smoothness of these images during parameter estimation for the general linear model.

The smoothness is described in terms of FWHM. During parameter estimation in SPM99, a file 'RPV.img' is generated which contains the number of RESolution ELEments (RESELS) Per Voxel (RPV). A RESEL is defined as  $\frac{S}{f_x f_y f_z}$ , where  $S$  is the search volume in  $\text{mm}^3$ , and  $f_x, f_y, f_z$  is the FWHM smoothness in mm in  $x, y$  and  $z$ . From this file, we can calculate the FWHM for each voxel. Since the data we use is two-dimensional, an infinite smoothness is forced in the  $z$ -direction and the smoothness in this direction is dropped from the equation. Because each voxel in our data has spatial dimensions of 1mm in  $x$  and  $y$  direction, the RESELS image is transformed to a FWHM per voxel image by

$FWHM_i = \sqrt{\frac{1}{RPV_i}}$ , where  $i$  indexes the voxels (see corresponding section in SPM99's `spm_spm.m` file).

To estimate the spatial smoothness in a grey matter segment, we input our nine grey matter segments and their inverse images to a one-group AnCoVa experiment. The inverse images are added because smoothness estimations are based on residual images. In the case of one-group An(C)oVa, this means that smoothness of an image is estimated from the difference between that image and the mean of the group. If the mean image is subtracted, the estimated smoothness will change, unless this mean is a constant value over the whole volume. When inverse images are added, a mean of 0.5 at each voxel is realised. Because the smoothness at each point of an inverse image is equal to the smoothness of the original image at that point, the estimated smoothness is not further affected. The latter is of course only true when all the images are indeed used for the estimation, since SPM99 uses the total number of images, with a maximum of 64 images (see `spm_spm.m` source code), we are safe with our 18 images.

The experiment yields an RPV image, which is transformed in Matlab to a FWHM image. Most of the voxels have a low smoothness ( $< 2\text{mm}$ ) the few brain voxels with infinite smoothness are truncated to 8 mm, in order to set an upper-bound to the smoothing kernels we will use in the simulations (figure 6a). The areas outside the brain are estimated at infinite smoothness, but because we know that this region does not contain any spatially correlated data, the smoothness in this region is set to a very low value. To produce a little less group-specific image, this FWHM image is then smoothed at 2mm FWHM (figure 6b).

The next step is to use this smoothness information in the generation of our random field. We will try

to create a field in which every voxel intensity has an equal chance of occurring, and the smoothness of the field is equal to the estimated smoothness of the original data at each voxel. The general idea is that first a random noise field is created in which the values at each voxel are samples from a uniform distribution. Then, for each voxel, a) smoothing at the correct smoothness is applied, and b) an intensity correction for the specific FWHM is applied to maintain a uniformly distributed intensity histogram.

In practice, to speed up calculations, the random noise field is smoothed at a number of smoothness-levels. The intensity at a voxel in the final noise field is determined by first looking at the estimated smoothness at that voxel. The field smoothed with the closest matching kernel is selected. The intensity in the final noise field is the intensity of the corresponding voxel in the matched field. Smoothing a grey-value image affects the image histogram. Low and high intensity values occur less often, and in general intensities are shifted to the mean of the image. To ensure that the intensities in the smoothed fields are evenly distributed over the interval [0,1], histogram equalisation is performed on each smoothed field separately. As a pre-equalisation step, the intensities in the image are transformed according to an intensity transformation that, based on the used smoothing width, approximately flattens the histogram. This transformation does not completely flatten the histogram for all FWHM, therefore it is used as a step before true histogram equalisation to increase performance and the number of distinct intensity output levels.

To perform the mapping from smoothness in the FWHM image and the actual width of the smoothing kernel to be used, the smoothness of random noise-fields, smoothed at different levels is estimated.

The expected FWHM is calculated by  $FWHM_{result} = \sqrt{FWHM_{applied}^2 + FWHM_{intrinsic}^2}$ . We do not know the intrinsic smoothness of our noise field, but that will be relatively small, so for larger smoothing kernels, the resulting smoothness should approach the applied smoothness.

Nine noise images, smoothed with a certain FWHM, and equalised intensity histogram, are input to SPM one-group AnCoVa analysis. The resulting smoothness estimations are compared to the applied smoothing levels (figure 5). The minimum smoothness of the data, when no smoothing is applied, is estimated 1.6 mm FWHM. This is the smallest smoothness we can achieve, so for smaller target smoothness we should not apply any smoothing. Also, the estimated smoothness for larger smoothing levels lies below the applied smoothing, either SPM underestimates the smoothness of the data, or the intensity correction renders the data less smooth. Either way, we should use the kernel that produces the estimated smoothness, since the target smoothness at each voxel was also estimated using SPM99. The inverse function is created, and used as a mapping between desired smoothness and applied FWHM. The intensity transformation is applied to the smoothness image from figure 6b to



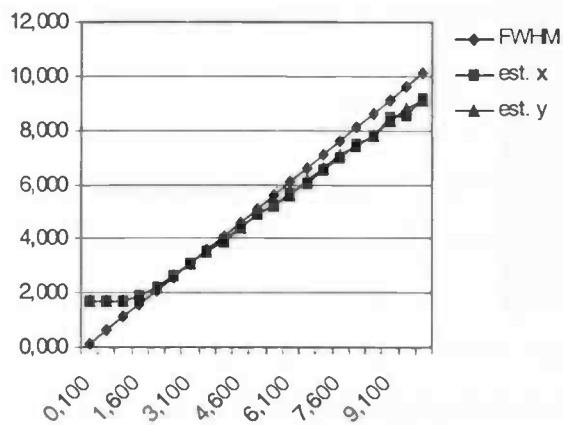


Figure 5: applied smoothness and estimated smoothness by SPM99. 'FWHM' is a reference line ( $y = x$ ). 'est. x' and 'est.y' are the estimated smoothnesses in x and y directions.

map the estimated smoothness values to smoothness values to use on our random field (figure 6c.).

#### 6.2.4 Evaluation of the simulated segments

Some results from the segment-simulation process are shown in figure 7. In the left column the noise fields are displayed that were used in combination with the mean value image to produce the simulated segments in the middle column. To show how the data is affected by smoothing and visualise the anatomical variation more clearly, in the right column the segments are smoothed. Given the fact that these images are the result of a simple noise model, they seem to get surprisingly close to the real thing.

In figure 8, an 'evolution' of the simulation process is displayed. In the left column, starting with a simple model in which segments are based on the mean image to which a smooth gaussian noise field with adaptive standard deviation based on the standard deviation image is added. Then, since non-smoothed segmentation data isn't normally distributed, but approximately takes values 1 or 0, a simulation based on mean image and a simple smoothed Bernoulli process, in the second column. The adaptive spatial coherence is added in the third column. Finally, the real thing is displayed rightmost.

For a simple comparison, standard deviation images per voxel of nine subjects are displayed in the third row, and standard deviation images of the smoothed data are shown in the bottom row. Ideally, the images should be close to the images in the last column, both in general behaviour and intensity. The improvements that each added step in the simulation process introduces are clear from these images.

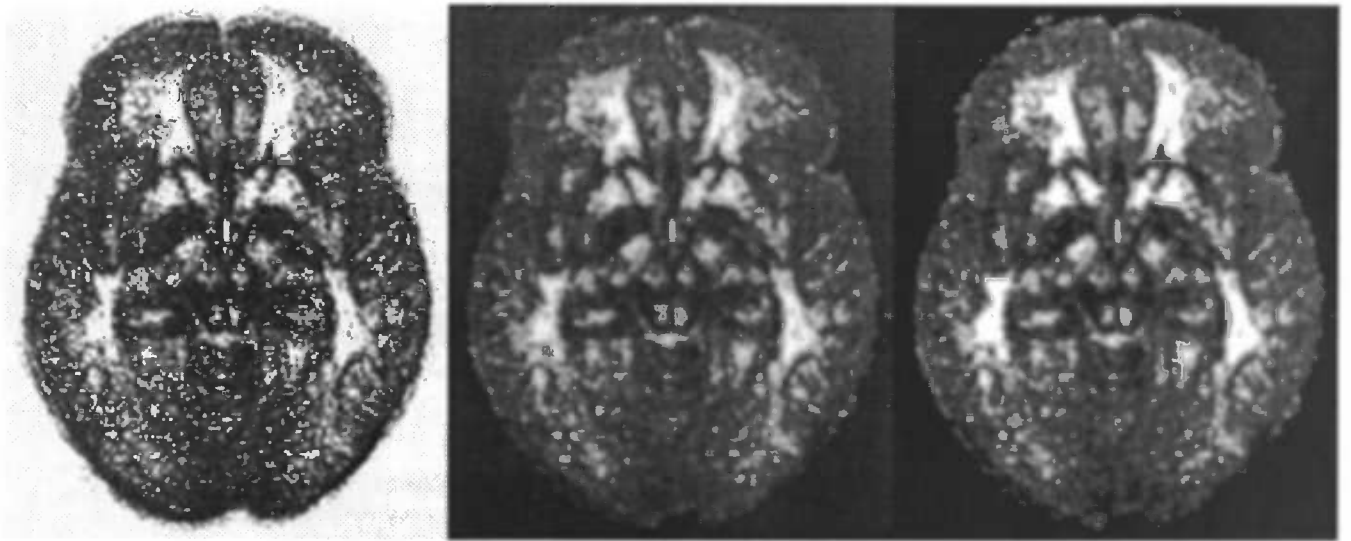


Figure 6: left: *smoothness per voxel*, middle: *masked smoothness image and smoothed 2mm FWHM*, right: *smoothness image after applying intensity transformation*

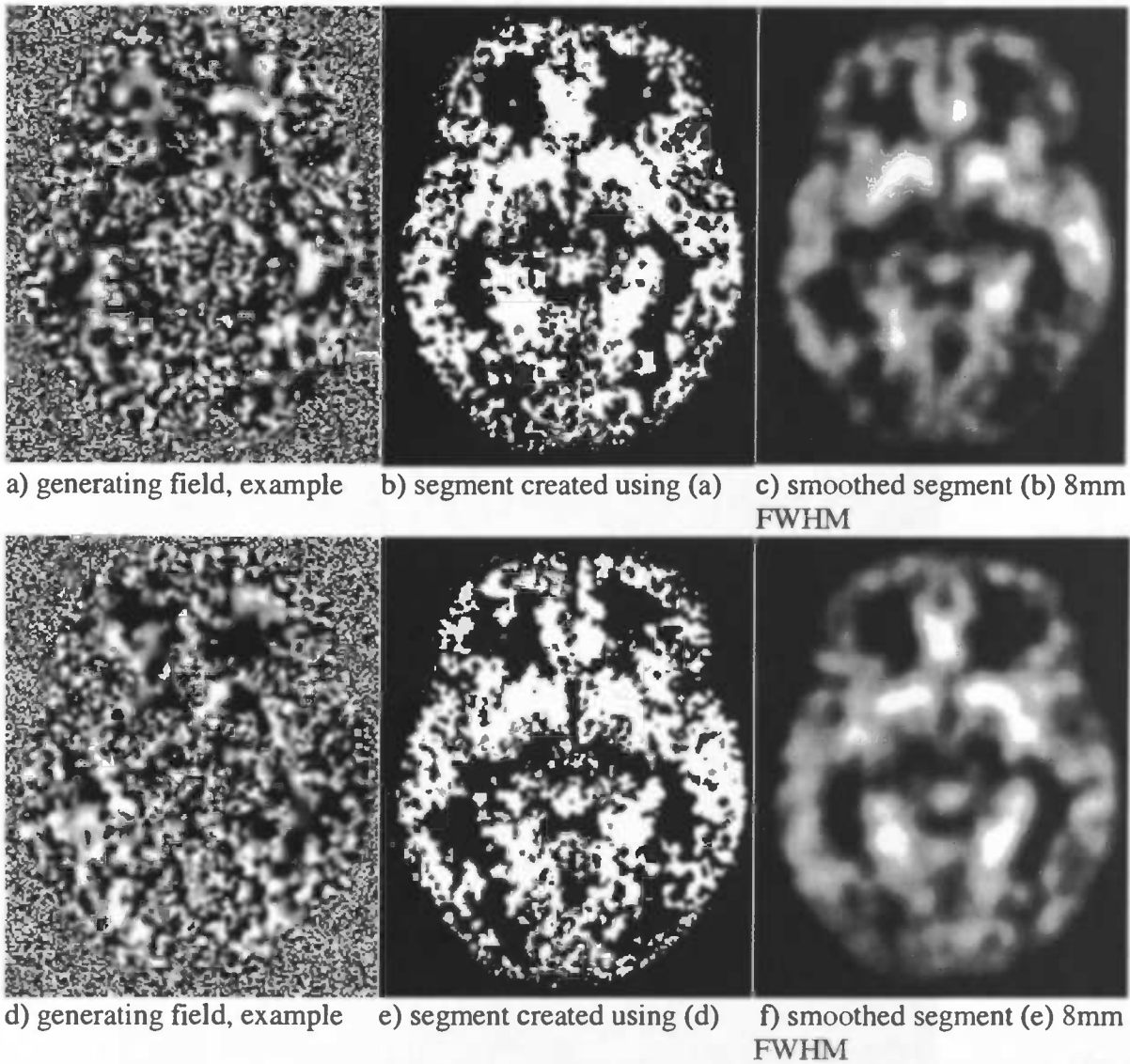
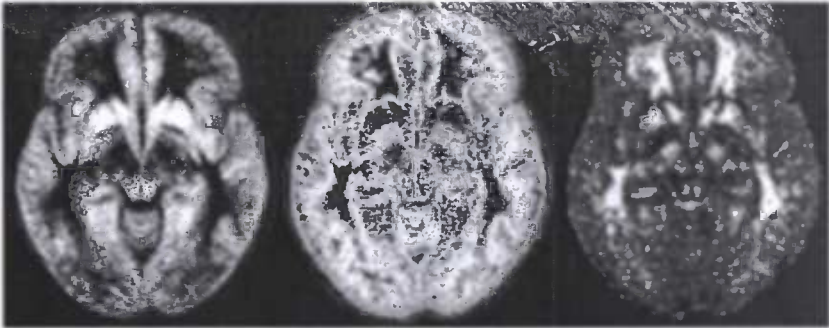


Figure 7: *Simulated segments, based on mean and smoothness group images.*

Figure 8:  
development of  
the simulation  
process

Images on which  
the simulations are  
based, derived  
from original data  
of 9 subjects



mean

standard deviation

Smoothness in FWHM  
(derived from RPV.img)

Simulation based on  
mean + deviation  
calculated from standard  
deviation image and  
gaussian field (2mm  
FWHM)

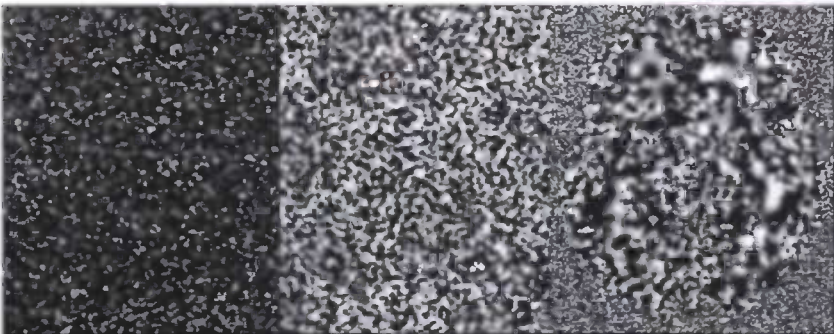
Simulation based on  
mean and uniformly  
distributed intensity field  
(2mm FWHM)

Simulation based on  
mean and uniformly  
distributed intensity  
field with adaptive  
smoothness.

Original

Examples of random  
fields used to generate a  
simulated image.

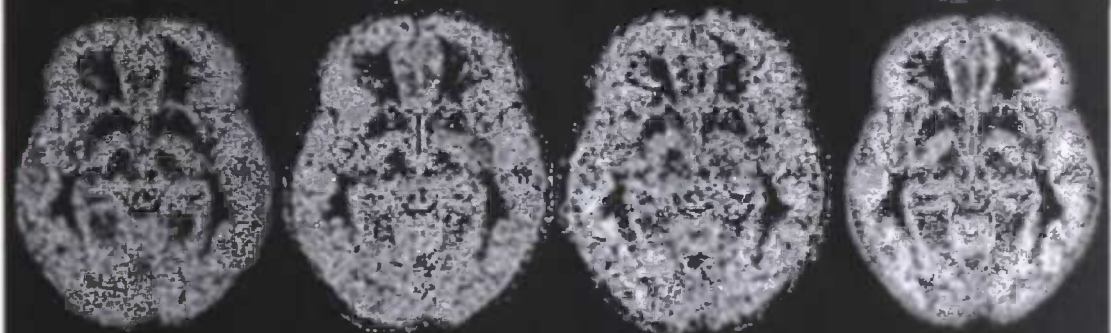
- a) gaussian random  
field (2mm FWHM)
- b) uniformly  
distributed noise  
(2mm FWHM)
- c) uniformly  
distributed noise  
(variable  
smoothness)



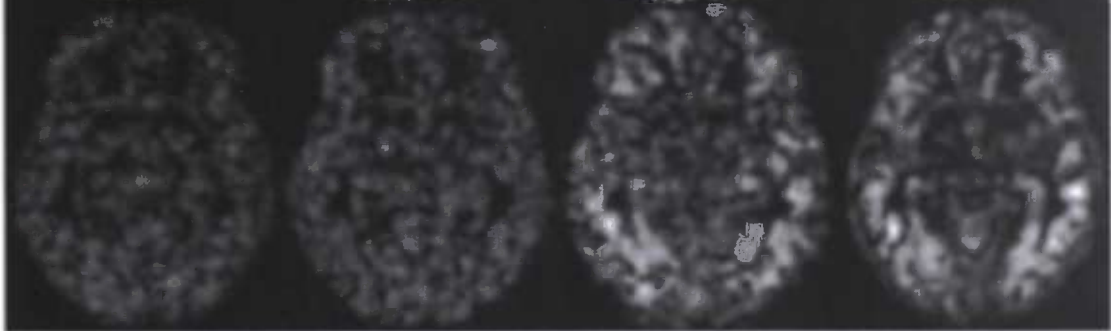
Examples of segments  
generated by each  
simulation method and  
one original segment.



Standard deviation  
images for each method,  
based on 9 segments.  
Each image is displayed  
in the same intensity-  
range.



Standard deviation  
images for each method,  
based on 9 smoothed  
segments. Each image is  
displayed in the same  
intensity-range.



6.3 Sensitivity of VBM with respect to lesion severity

We will look at how the SPM is affected by the severity of the lesion, given a fixed number of group subjects. For the project on visual field defects, the number of subjects is typically expected to be 30 to 40. We choose to base our tests on 40 subjects. The true difference in mean and variation between groups is only a consequence of the severity of the lesion. All other sources of variance are equal in both groups. By varying the severity of the lesion, we can see how sensitive SPM is, given a fixed number of subjects.

6.3.1 Grey matter reduction percentage

To determine how the shape of the structuring element for the erosion and the amount of atrophy created are related, the erosion process is performed for a number of different radii ( $r$ ). See appendix A for an explanation of how this radius defines the shape of the structuring element.

Figure 9 shows the relation between  $r$  and the percentage of removed grey matter tissue. The tissue was removed for nine subjects from a slice in the brain at  $z = -9.0\text{mm}$ , in the visual cortex. The region was equal for all subjects,  $23 \times 18 \text{ mm}^2$ . Plotted are the mean reduction of the grey matter (solid line), which is the actual reduction present in the later experiments, and dashed lines indicating the standard deviation from the mean, which illustrates the intra-subject differences of the operation. From this function, radii were calculated to obtain a mean GM reduction percentage in the range of 0 – 30%.

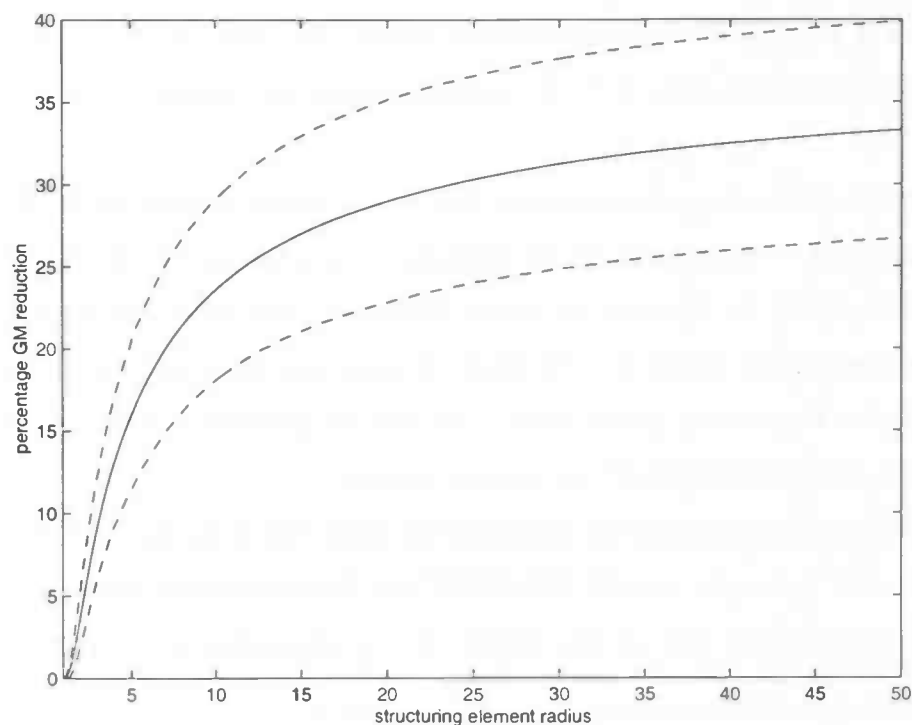


Figure 9: Relation between structuring element ‘radius’ and GM reduction after erosion. The solid line represents the mean GM reduction percentage. Dashed lines are offsets at one standard deviation from the mean.



### 6.3.2 Results from experiments using simulated groups with increasing atrophy.

Groups of 5, 10 and 20 subjects with increasing atrophy and equal sized control groups are created. Starting with an atrophy of 0% and increasing to about 30%, we will see at what point the VBM process starts to detect a significant difference between the groups in the correct place.

The simulated segments are smoothed with a 10mm kernel, and analysed using the standard AnCoVa model, with no covariates. The SPMs are thresholded using a height corrected threshold at  $\alpha = 0.05$ . The cluster extent is set to 0. This process is repeated 10 times (10 runs) for each lesion intensity and group size. For each run, the process outputs an SPM\_t map. The data are automatically processed after all analysis is finished. Data collected for each run are:

- the number of supra-threshold clusters
- the area of each cluster
- the maximum statistical value for each cluster
- the mean statistical value of each cluster
- the bounding box and voxel co-ordinates of each cluster, which gives us an indication of its position

The reduction percentages range from 0 to 30, with step-size 2.5. Figure 10 shows a box-plot over 10 runs of the maximum statistical value in the thresholded SPM\_t image. A supra-threshold value is only counted when it actually occurs in the lesioned region. The plots for 20 and 10 subjects in each group are displayed. For group sizes of 5, no voxels survived the calculated threshold in the correct position.

For 20 subjects in each group, we seem to be able to accurately identify the applied lesion starting at a reduction percentage of about 15%. For 10 subjects, we at least need to double the intensity of the lesion to get similar results. In this case we would need over 30% reduction before any changes are systematically detected, which might be a bit high. Groups sizes of 5 subjects are unsuitable in this case, and probably for most cases, unless there is an obvious difference, like a missing hemisphere. Smoothing the data at higher FWHM did not improve results.

The maximum area that is detected is plotted in the same way in figure 11. The actual signal was defined on a  $414\text{-mm}^2$  rectangle. In most cases only one supra-threshold area was found. From the figures it can be seen that the size of this cluster slowly approaches the true area with increasing reduction percentage.

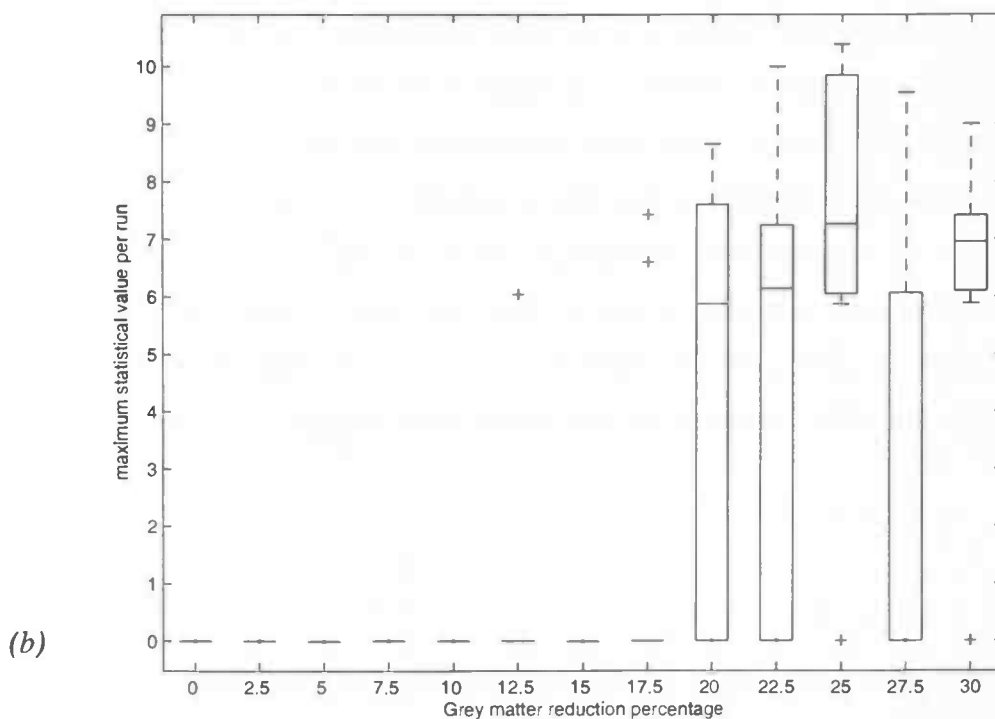
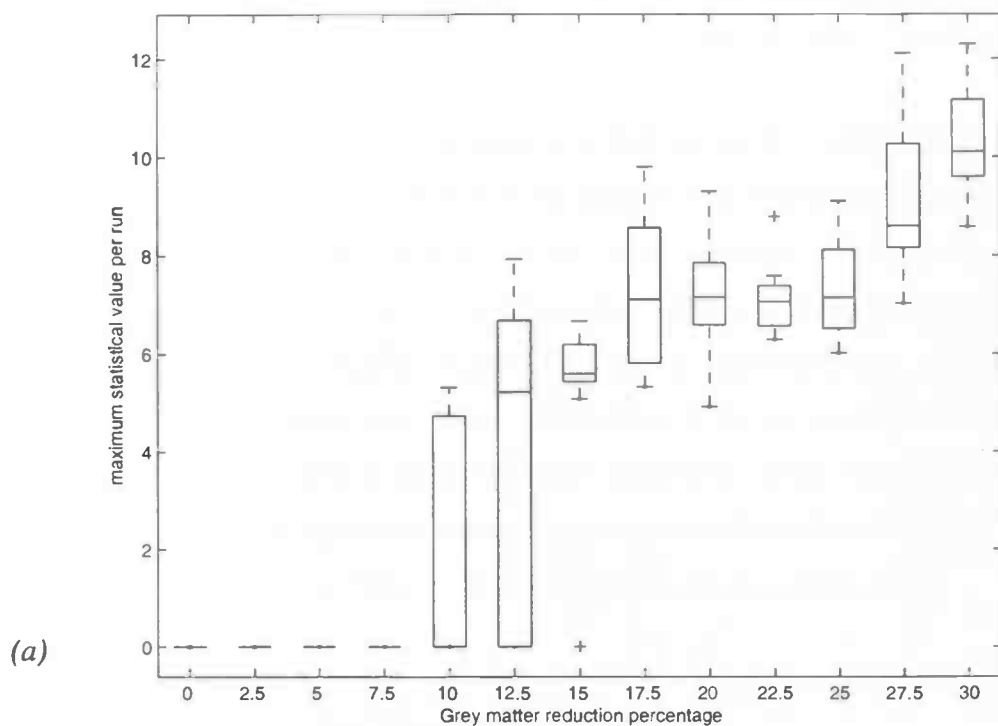


Figure 10: Maximum (supra-threshold) statistical value per run, for increasing GM reduction percentage.  $N = 20$  (a) and  $N = 10$  (b) subjects. For each reduction percentage a boxplot over 10 runs is displayed (median, upper, lower quartiles and data extent). Data was smoothed at 10-mm FWHM, alpha level 0.05, corrected critical height threshold.

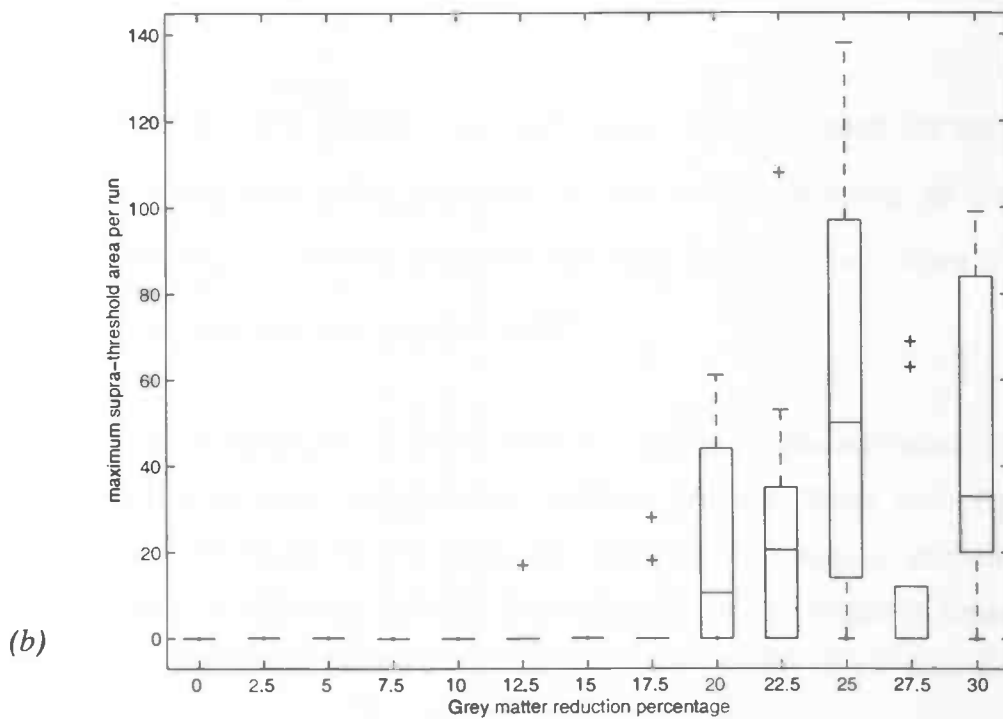
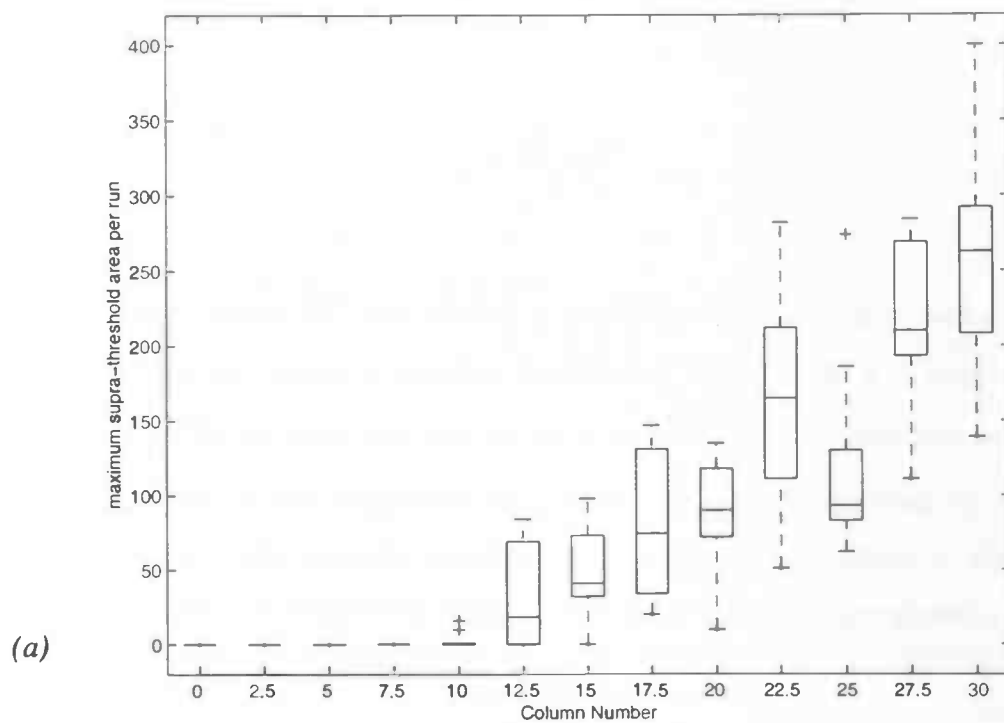


Figure 11: Maximum supra-threshold area per run, for increasing GM reduction percentage.  $N = 20$  (a) and  $N = 10$  (b) subjects. For each reduction percentage a boxplot over 10 runs is displayed (median, upper, lower quartiles and data extent). Data was smoothed at 10-mm FWHM, alpha level 0.05, corrected critical height threshold.

## 6.4 Analytical power estimation

Apart from the experiments will try to provide a more analytical measure of the power of VBM.

### 6.4.1 Power

The power (or sensitivity) of a statistical method is defined as the probability of rejecting a false null hypothesis. The power in general is related to the sample size, effect size (signal relative to noise) and alpha level (significance level, or the odds that the observed result is due to chance). In SPM we deal with a large number of spatially non-independent observations, and signal-noise ratios that depend on smoothing of the data, the original signal dimensions (size/intensity) and varying spatial autocorrelation and intrinsic smoothness. This means that it is not straightforward to give a simple expression for power, since it depends on so much properties of the input data. Nevertheless, we are curious to see how for some values of the input data an estimation for power can be obtained. For our calculations we will assume our SPM is calculated according to a simple t-test.

### 6.4.2 Data model for the power estimation

Calculations are based on two equal sized groups of  $N$  'subjects'. For each subject, the data consists of a single  $189 \times 157 \times 1$  mm slice with voxel sizes of  $1 \times 1 \times 1$  mm<sup>3</sup>. The signal is defined over the brain region defined by the SPM99 brain-mask image at  $z = -9$  mm. Apart from a single region with lowered mean signal in the second group, both groups have same constant mean, standard deviation and smoothness. The added signal has constant height, and is restricted to a single rectangular region of  $23 \times 18 = 414$  -mm<sup>2</sup>, the same area as in the experiments. Furthermore, we assume that this baseline signal conforms to a stationary gaussian field. Of course, this is an extremely simplified data-model, but since we are interested in the general behaviour of the power the model should not be too complicated. We will fix our alpha level at 0.05; e.g. the chance of obtaining a false positive result is 5 times out of 100.

### 6.4.3 Calculations

For a given alpha level, region, and smoothness of the data, we can obtain a minimum difference between the two groups that produces a supra-threshold value. SPM provides us with a function that calculates the corrected critical height threshold at a specified significance level (SPM\_uc). For a significance level of 0.05, that is the threshold at which in only 5 out of 100 times a supra-threshold region occurs by chance, given the smoothness and shape of the volume. The volume in this case is the total search volume of the brain region in our slice. The resel counts are calculated for the specific



search region and smoothness, using SPM99 functions according to [Worsley et al 1996].

The SPM-t map is calculated for each voxel by

$$t = \frac{\mu_1 - \mu_2}{\sigma \sqrt{\frac{2}{N}}} \quad (1)$$

Assuming equal variances ( $\sigma^2$ ) and number of subjects ( $N$ ) in each group, we can obtain the minimum difference between the means of samples from group 1 and group 2 to exceed threshold  $t$ , that is  $\mu_1 - \mu_2 = t \sigma \sqrt{\frac{2}{N}}$ . We can write this as a signal to noise ratio (SNR), which is usually expressed as 20 times the logarithm of the amplitude ratio, or 10 times the logarithm of the power ratio. For convenience we will take only the amplitude ratio. Here, the amplitude is defined as the standard deviation of the signal (or noise) with respect to the mean of the baseline signal:

$$snr_{\min} = \frac{\mu_1 - \mu_2}{\sigma} = t \sqrt{\frac{2}{N}}. \quad (2)$$

Since the data is normally distributed at each voxel, the group mean for that voxel is distributed according to a normal distribution with mean  $\mu_{1,2}$  and standard deviation  $\sigma N^{-\frac{1}{2}}$  (the standard error of the mean). Therefore, the difference between the mean images at each voxel is normally distributed with mean  $\mu_1 - \mu_2$  and standard deviation  $2\sigma N^{-\frac{1}{2}}$ .

Now, let us focus on the region in which there is a constant difference between the true mean of the two groups. Since all input images were stationary gaussian fields with equal smoothness, the difference between the means in this region also conforms to a stationary gaussian field with the same spatial smoothness. A difference between the two groups in this region is found significant when it exceeds the previously calculated minimum difference for the given smoothness, and number of subjects.

Take the mean of group 1 in this region to be  $\mu + \delta$ , and the mean of group 2 to be  $\mu$ , where  $\delta \geq 0$  i.e. group 2 has atrophy with respect to group 1. The difference is distributed normally,  $N(\delta, 2\sigma N^{-\frac{1}{2}})$ . We will find a significant result if this field exceeds a value of  $snr_{\min} \sigma$ .

The corresponding Z score ( $u$ ) is calculated by:

$$u = \frac{snr_{\min} \sigma - \delta}{2\sigma\sqrt{\frac{1}{N}}} = \frac{snr_{\min} - snr}{2\sqrt{\frac{1}{N}}} \quad (3)$$

The power estimation is now given by the chance of actually obtaining this or a larger difference anywhere in our region of interest (equivalently, the chance of obtaining this or a larger Z-score). For our small region this is the corrected p-value for the Z-score  $u$ :

$$p_{corr} = P(Z_{\max} > u) \quad (4)$$

Now we can use SPM99's function `spm_P`, that implements a number of functions described in [Worsley et al. 1996] to calculate this corrected p-value. The estimation for the corrected p-value is only accurate for high thresholds, or p-values below 0.2. For lower thresholds the estimate becomes less valid. One should bear this in mind when looking at the figures. Most results will be shown only for thresholds larger than 1.

#### 6.4.4 Results from analytical power estimation

##### **Power as function of SNR and number of subjects for 2D data**

Figure 13 shows the estimated power as a function of SNR and number of subjects used in the analysis. The signal was defined as explained above. Also estimations for lower thresholds than 1 are shown, to show that the function flattens as the power increases, and eventually even decreases (not shown). The latter obviously is incorrect, but explained by the nature of the Euler characteristic. One can expect the power to approach 1, without too much flattening. In figure 14 the power estimate is plotted against SNR, for  $n = 5, 10, 15, 20, 25$  and 30 subjects.

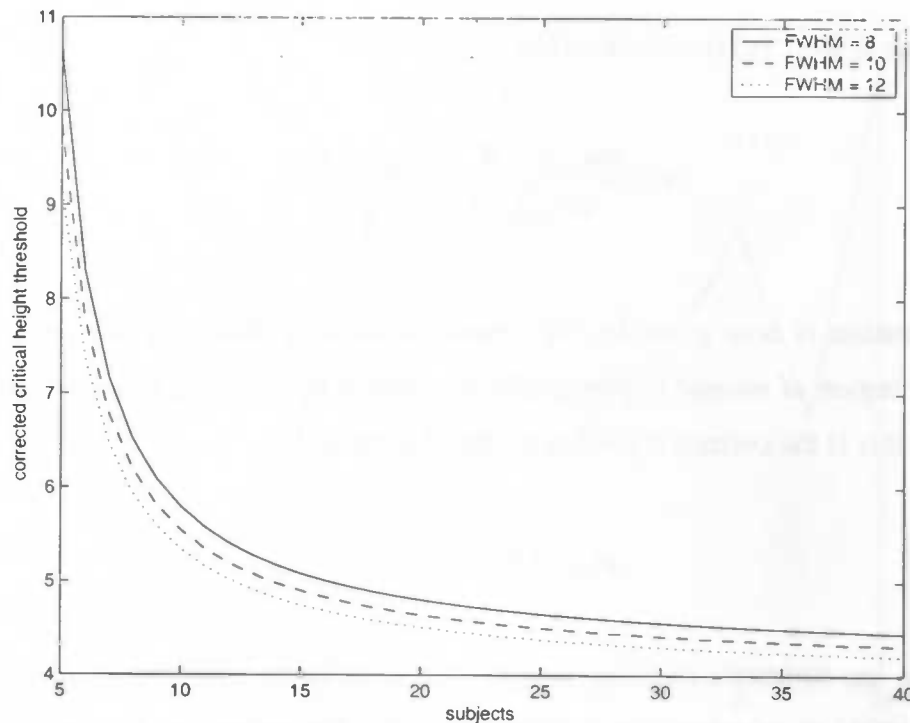


Figure 12: Corrected critical height threshold for single-slice ( $z = -9\text{mm}$ ) brain region at smoothness of 8, 10 and 12 mm FWHM, for increasing number of subjects. Alpha level = 0.05

To compare the computed power estimates with experimental results, a power estimation based on the experiments is plotted with analytical estimates based on approximately matched conditions (figure 15). To obtain a power estimate from the experiments for those conditions we simply divide the number of true positive results by the total number of results for each condition. This is the number of runs in which we found a supra-threshold area (which intersected the signal-region) divided by the total number of runs (in this case 20), for a given signal intensity and number of subjects. Due to the low number of runs, this is a very rough estimate.

For matching the analytical estimates to the experimental estimates, the data from our 9 original subjects were used, smoothed with 10mm FWHM. The smoothness was estimated by SPM99 at about 9mm FWHM. Based on 9mm smoothness analytical power estimates were calculated. The SRN corresponding to a reduction percentage of 30 was calculated by dividing 30 percent of the mean by an estimation of the standard deviation from the noise. The mean is estimated at 0.5 after observing that this was approximately the mean value throughout the cortex. The standard deviation was estimated by taking the mean of the standard deviation image where it had a value greater than 1/8 of the total image mean. The latter to prevent under estimation due to outside brain voxels. Signal to noise ratio at 30 percent reduction and 10mm FWHM smoothness was in this way estimated at approximately 2. The x-axis was scaled to match both functions.

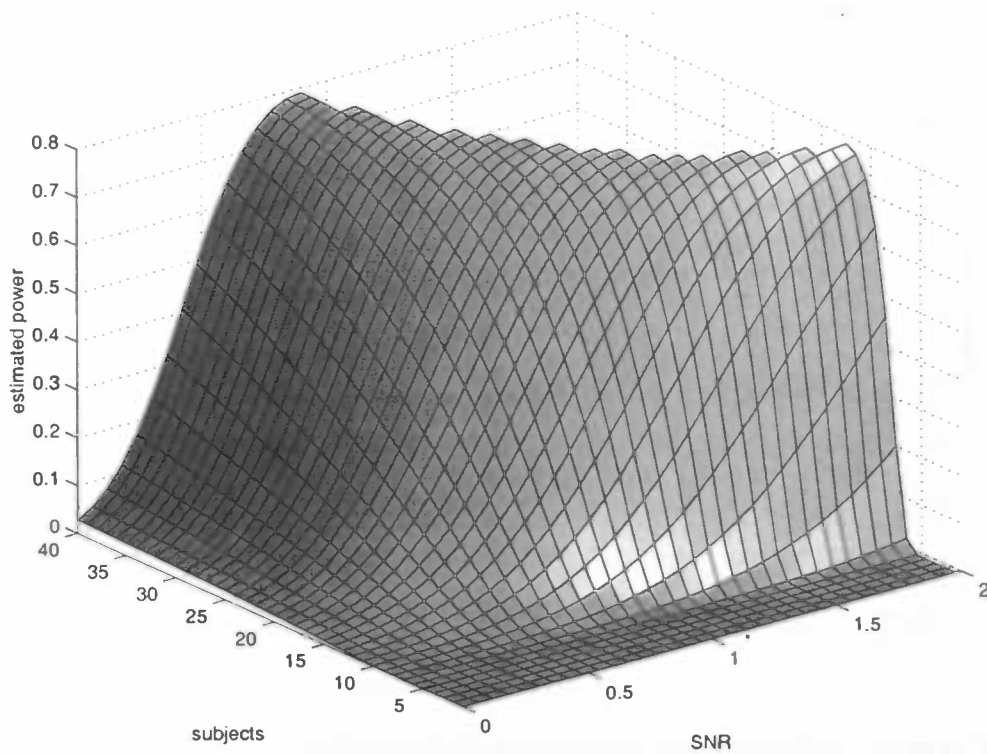


Figure 13: estimated power as function of SNR and number of subjects. Signal was defined on a 22x13-mm rectangle on a 2D slice, data smoothness was set to 10 mm FWHM. Alpha level 0.05.

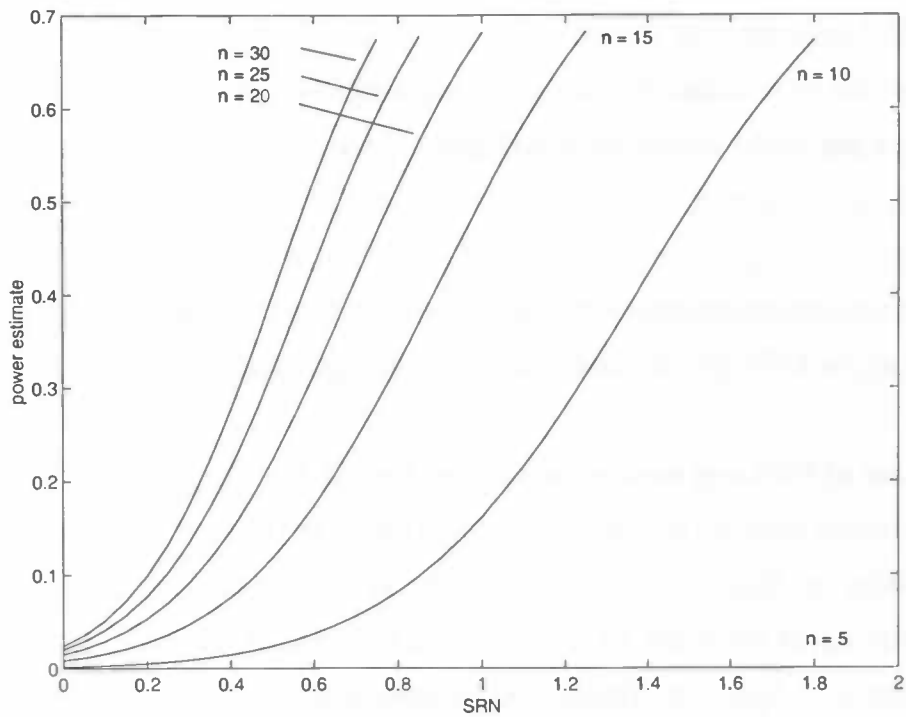


Figure 14: estimated power as function of SNR for different number of subjects. (Side view of figure n). Maximum threshold is 1.

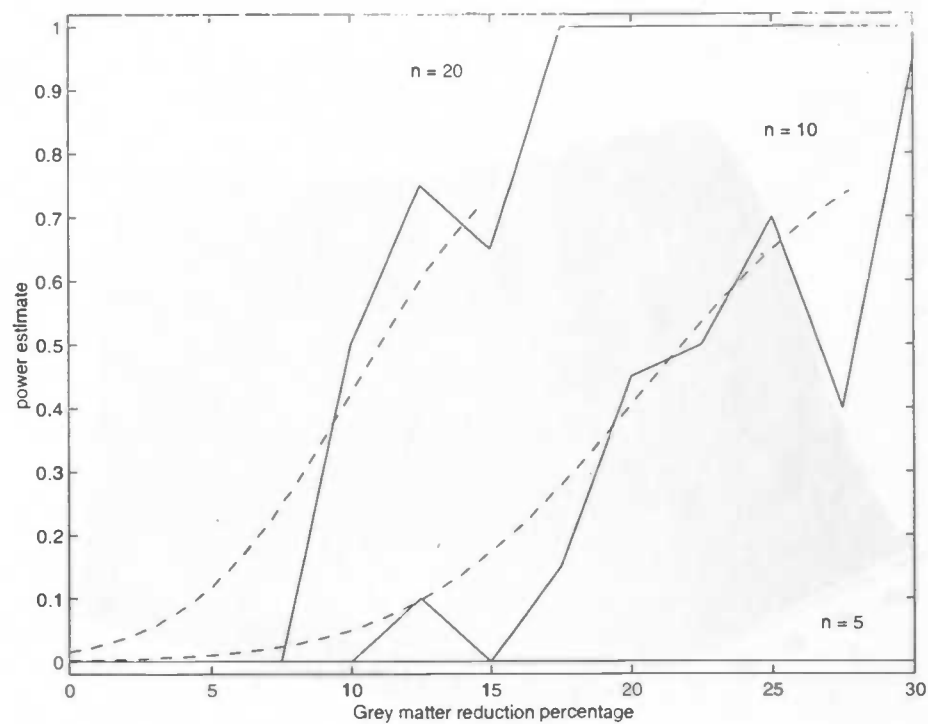


Figure 15: Power estimated from the experiments plotted with power estimations based on matched conditions.

Although the power estimates from the experimental results are very rough, the matching between the model and the experiments, and the model itself are highly approximated, the actual graphs show a surprising similarity for each number of subjects. Only at lower SNR levels it seems that the experiments under-perform with respect to theory, which might be attributed to a 'failure' of the segmentation process to translate the subtle intensity differences to concentration differences in these cases. Which was one of the reasons to include the segmentation step in the simulation process in the first place.

Both from analytical power estimates and estimates based on the experiments, it seems that for the main power-range, the SNR roughly needs to double when group size is halved to get the same power.

### Power as function of SNR and number of subjects for 3D data

To see whether results translate to three-dimensional space, the power calculation is also performed in the same way using the brain volume defined by the SPM99 brain-mask. Instead of taking a 23x18-mm rectangle, the signal for figure 17 is defined on a 23x18x23 mm<sup>3</sup> box. Corrected critical height thresholds are shown in figure 16. The power of course is highly influenced by the size of the signal, so the absolute values in the figures are hardly comparable. We do see however the same behaviour for increasing number of subjects as in the 2D case. Reassuringly, the power does not drop when looking

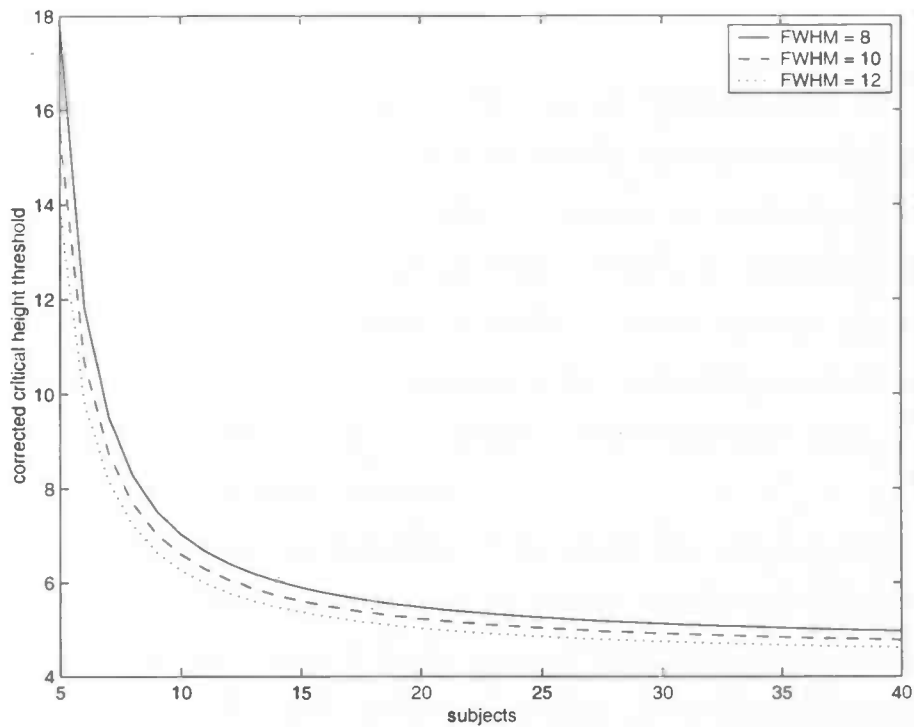


Figure 16: Corrected critical height threshold for whole brain region at smoothness of 8, 10 and 12 mm FWHM, for increasing number of subjects. Alpha level = 0.05

at three rather than two dimensions, which might be the case since the total fraction of the search volume occupied by the signal region will probably be less in three dimensions. (In this case for the box:  $9.5\text{cc} / 1203\text{cc} = 0.8\%$  and for the rectangle  $0.41\text{cc} / 25\text{cc} = 1.7\%$ )

### Power as function of signal size and number of subjects for 3D data

Also we take a look at the impact of changing the size of signal region. The region is defined as a cube in the same search volume as before. For increasing side length the power is estimated. We fix the SNR to 1, and take data smoothness to be 10mm FWHM. The resulting curves for  $n = 5, 10$  and  $20$  subjects are shown in figure 18. Power increases strongly with increasing signal width. Also from this figure it is clear that the degrees of freedom should be large enough to get substantial power.

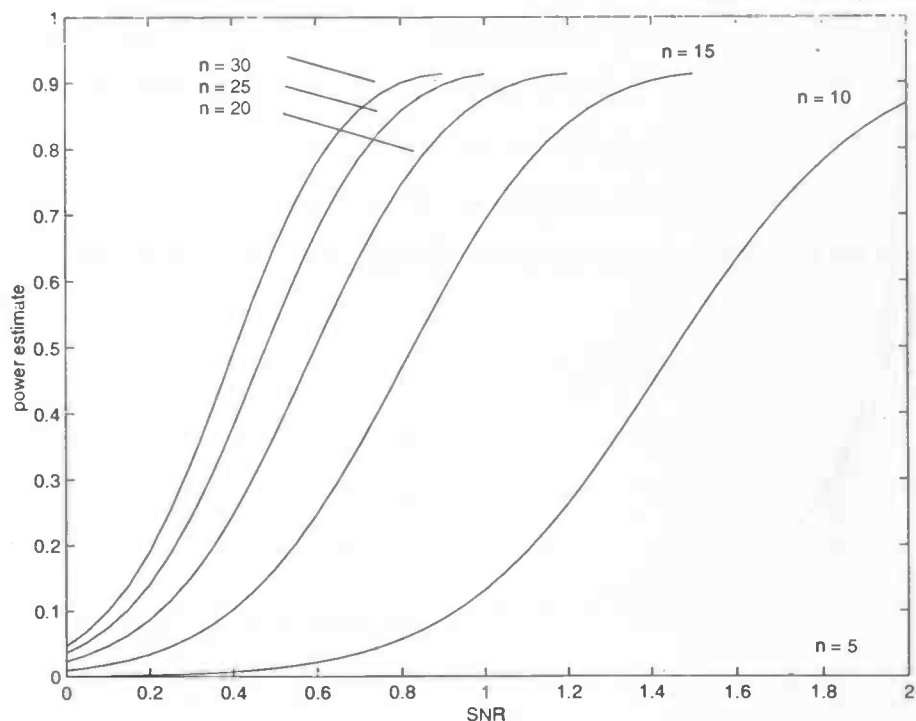


Figure 17: estimated power as function of SNR for different number of subjects. Signal is defined on a 23x18x23-mm box. Data smoothness is 10 mm FWHM, alpha level 0.05.

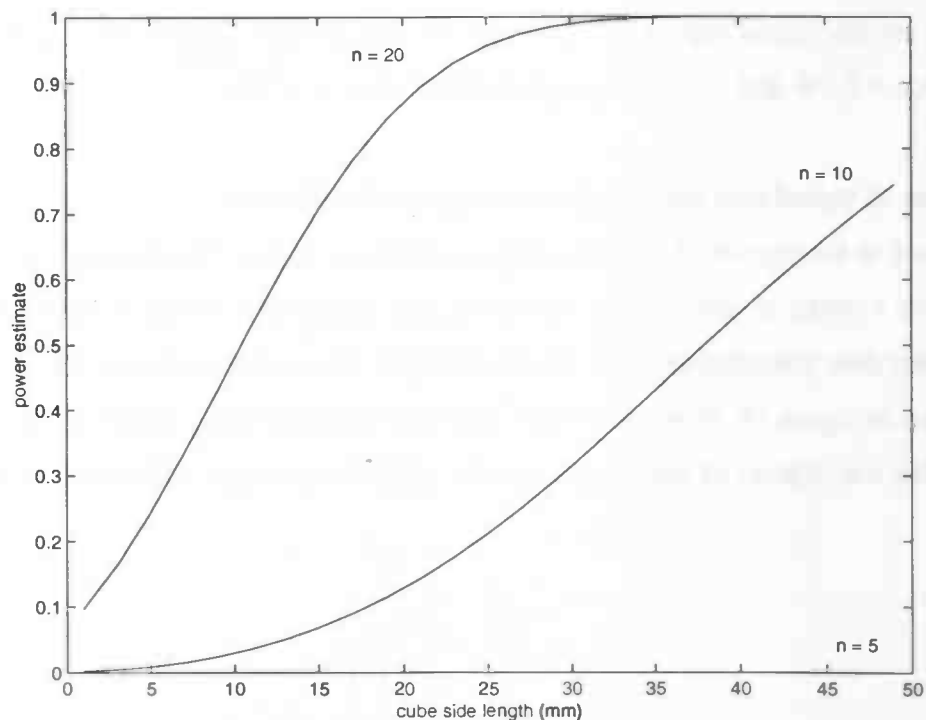


Figure 18: estimated power for cube with increasing edge length, for  $n = 5, 10$  and  $20$  subjects. SNR is fixed at 1, data smoothness is 10-mm FWHM and alpha level 0.05

## 7 Evaluation

With this document we hope to have given an insight into the VBM method, and how it can be used. Based on experiments and theory estimations of the sensitivity under various conditions have been made. Using the formula, and the scripts written, a researcher can make his own estimation for the performance of VBM, given the expectations of difference between groups, group-sizes or required power. For this, estimations have to be made of signal mean and variance (after smoothing) in the expected region of interest. The estimations provided in this document are rough estimates based on 9 normalised segments in a specific region, it is therefore recommended to make custom estimates for specific regions, preferably using more subjects.

In the various figures the general behaviour of the sensitivity with respect to different variables has been visualised. These can be used as tools to provide insight into their effect and from this, optimise experiment design and results. It should be noted that in each case the entire brain volume (or segment) was used as the search region. Better performance can be obtained when a priori the region of interest is known, since this will lower the corrected height threshold required for a given alpha level. However, the analysis will be only valid for that region, and not throughout the brain.

## 8 Further work

With the release of SPM2 in November 2002 it is probable that in the future the switch from SPM99 will be made. This of course has consequences for usability of the scripts, which may not be compatible with the new software. The improved segmentation algorithm of SPM2 might be interesting for VBM, and the facilitated batch-mode will possibly make the implementation of VBM more easily realised using a batch script and thus eliminate the need for external scripts. The VBM method itself will remain more or less the same. This document provides an overview of the different aspects of VBM, and together with the scripts gives a reference for implementing and utilising VBM in the future.

The models that have been devised for both simulating segment data and analytically estimating the power of VBM, are very rough models. For simulating segments, more complex behaviour can be incorporated, to eventually obtain a morphologically probable volume. For the power estimation, one could add the effect that the applied smoothing has to the shape and height of the signal. Variance estimates for the different brain regions at a number of smoothing levels will facilitate estimating the power for a specific research. [Worsley et al. 1996] list the resel counts and critical height thresholds for a number of brain regions, these counts are based on a smoothness of 20mm FWHM in each direction. It would be convenient to have this table for more smoothness levels (for example 10mm



FWHM), when one is interested in a specific search region. These tables might already exist.

Not all of the questions in chapter 6 have been investigated. Further work might take a look at the positional variance of lesions, or for example the influence of noise to the segmentation procedure. Methods for single subject analysis may be of interest, and probably new methods that will be developed in the field of morphometry and statistical parametric mapping in the future.

# 9 Appendices

## Appendix A. Grey-scale erosion with flat and non-flat structuring element.

The atrophy is simulated by means of grey-scale erosion of the combined grey-white matter probability image. The voxel values all lie between 0 and 1. Because of the small thickness of the cerebral cortex, the expected atrophy might be no more than one mm or less. With voxel dimensions of  $1 \times 1 \times 1 \text{ mm}^3$ , the effect of flat erosion in general will be too coarse to simulate realistic atrophy. Two possible solutions come to mind. The first solution is sub-sampling the volume, performing flat grey-scale erosion and then averaging of voxels to return to the original dimensions. This approach might be computationally expensive. The second solution is performing grey-scale erosion with a non-flat structuring element. The effect of both approaches on voxel-level (one-dimensional case) is illustrated in figure 20.

Although rather similar results can be obtained by both methods, the control over the real dimensions of the erosion is greater when using the sub-sampling process. The effect of non-flat erosion still depends greatly on the differences in intensity between neighbouring pixels. With this course representation of the (ideally) smooth structuring element, it is harder to make inferences about the amount of tissue that is eroded when using a certain structuring element. To obtain a comparable and reliable result, the resolution of the structuring element should be increased. In effect this is the counterpart of sub-sampling: increase the resolution of the structuring element instead of the volume, while still having control over the shape of the structuring element.

Without too much worrying about resolution issues, we will use simple grey-scale erosion that produces a noticeable reduction at the scale of one pixel. The structuring element used on single-slice data is a  $3 \times 3$  matrix, with user specified 'radius' (see figure 19). With matlabs `imerode` function, a portion of the volume can be eroded using this structuring element.

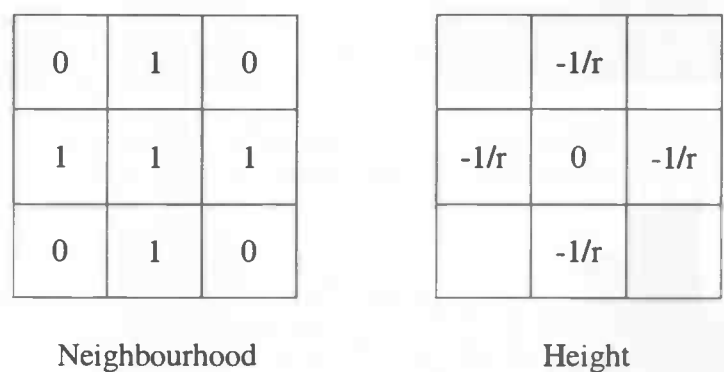
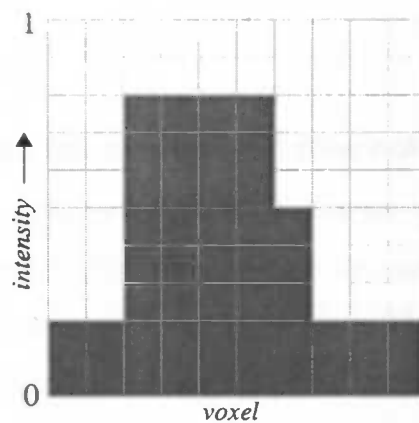


Figure 19: Definition of the structuring element, depending on a user specified radius

Figure 20: Erosion process illustrated for two methods

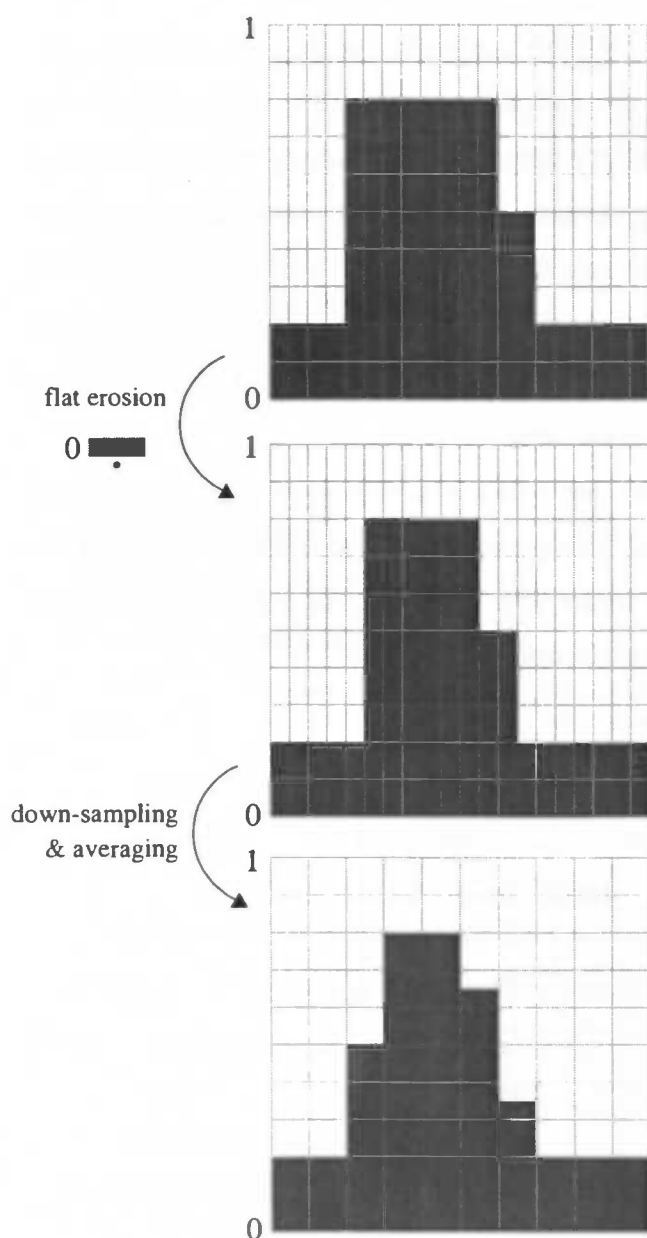
- Sub-sampling, then performing flat erosion using a 3 units wide flat structuring element.
- Grey scale erosion using different structuring elements.



Original image with intensity values in the range [0, 1].

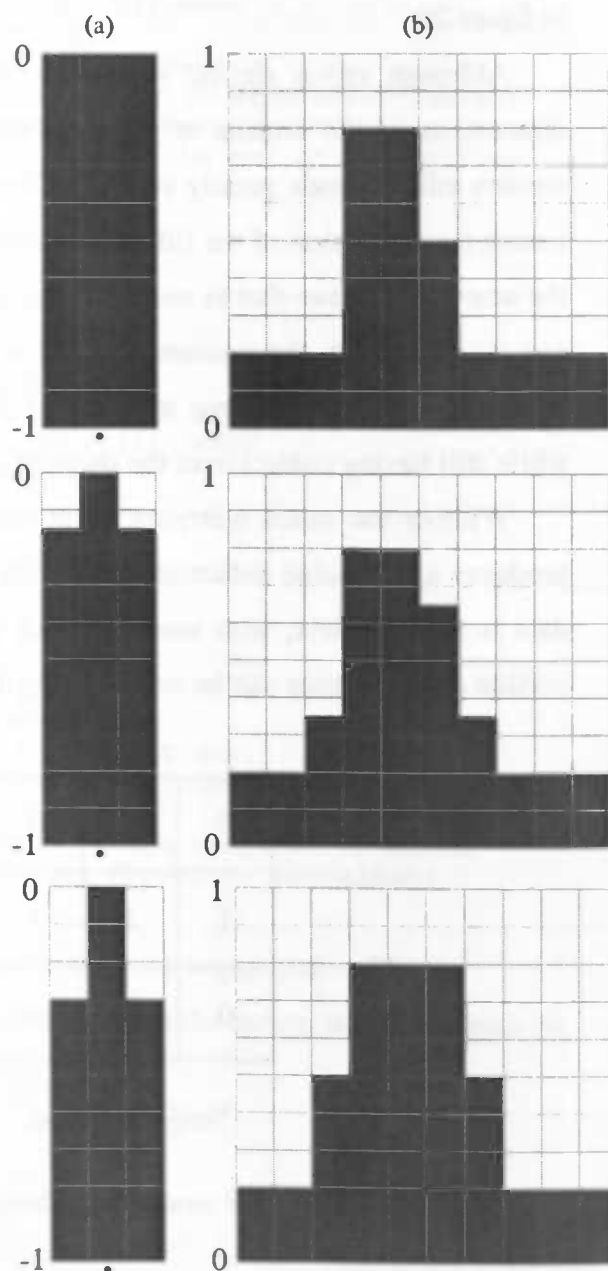
#### Sub-sampling to double resolution and flat erosion

The result is obtained by down-sampling and averaging, after the erosion is performed.



#### Grey-scale non-flat erosion

Three examples with (a) structuring element and (b) result.



## Appendix B: VBM process files

This appendix is an overview of the available process files for VBM. These are the scripts that implement the VBM protocol using the description given in the `descrip_preprocess.txt` file. The protocol is divided over 5 consecutive scripts: `normalise_process_VBM`, `segment_process_VBM`, `modulate_process_VBM`, `smooth_process_VBM` and `stat_process_VBM`. These scripts should be called from the `process.m` file with an array containing the source files and the analysis description structure.

For each script, the description is given (the heading in the matlab file), and an overview of the main function calls. This will give an impression of the way the scripts work, and the parameter values that are used. For further details, see the code in the scripts themselves.

For the main function calls, the parameters and their values are listed. Also, an explanation of the parameter and the meaning of the used value is given (separated by ';').

### Normalisation:

scriptfile: `normalise_process_VBM.m`

### description:

```
function normalise_process_VBM(files,descr)
% normalise_process A function to normalise T1 MRI data in SPM using
% K.U.L. functions and a standard protocol optimised for VBM
%
% normalise_process_VBM(files,descr) will normalise the files in string matrix
% files using the K.U.L. functions, according to options in descr.
%
% options used for normalisation in descr are:
% normalisation:
%   'linear':    perform linear normalisation
%   'standard':  non-linear normalisation, 7x8x7 basis functions, medium
%                 regularisation
%   'optimised': 'iterative' normalisation, see
%                 Good et al., 2000, A Voxel-Based Morphometric study of
%                 ageing in 465 normal adult human brains.
%                 NeuroImage 14: 21-36 (2001)
%
% ***** following options are used when optimised protocol is chosen *****
%
% tissue:
%   'grey':      normalisation & segmentation optimised for grey-matter
%   'white':     normalisation & segmentation optimised for white-matter
%   'both':
%
% clean:
%   'yes':       segments are 'cleaned' using ix/(i1+i2+eps).*i3, where
%                 ix is current segment, i1 is the grey-matter segment,
%                 i2 white matter, and i3 the extracted brain mask.
%   'no':        no cleaning on the segments
%
% A global variable VWD containing the directory in which the 'templates' and
% 'apriori' folders with the correct templates are stored, should be defined.
```

general actions:

### option 1) linear:

#### 1) affine normalisation

calls `vbm_sn3d` with parameters:

(name)	(value)	(explanation)
option	3	Do What; Determine parameters and write normalised
objima	files	Image to determine parameters from (objectimage)
objimages	files	Images to write normalised
mask_objb	0	Mask the object; no masking
maskima	' '	Image that contains the mask for the objectimage
templima	[VWD '/templates/T1.img']	[One or more templates]
mask_tmpl	0	Template masking; no masking'
spec_brainmask	[]	User specified brainmask for template
affinet	'affine3'	Affine Normalisation Type'
ase	[0 0 0 0 0 0 -1 1 1 0 0 0]	Affine Starting Estimates; Radiological Convention (L is R)?
nbfb	1	# Basis Functions (x y z); none
nli	1	# Nonlinear Iterations
nlr	0.01	Nonlinear Regularization; medium
intplt	-9	Interpolation Method; Sinc Interpolation (9x9x9)
bbt	[-78 78 -112 76 -50 85]	Bounding Box; Default SPM99
voxt	[1 1 1]	The voxel sizes (x, y & z, in mm) of the written normalised images

### option 2) standard:

#### 1) non-linear normalisation

calls `vbm_sn3d` with parameters (the same as for *linear* but):

nbfb	13	# Basis Functions (x y z); 7x8x7
nli	12	# Nonlinear Iterations

### option 3) optimised:

#### 1) segment non-normalised files

calls `f99_segment` with parameters

files	files	Matrix of filenames
opts	['cw']	Options; attempt to correct small intensity inhomogeneities and write inhomogeneity corrected image(s)
template	[VWD '/templates/T1.img']	Filename of the template to which the images should be normalised

`f99_segment` calls `vbm_segment` for each volume, which performs the segmentation like

`spm_segment`, but saves additional information to a `.mat` file and uses the apriori templates from the 'apriori' directory in the VWD directory.

#### 2) normalise segmented images

##### 2a) extract brain

calls **vbm\_xbrain** with parameters

P	[grey; white]	xxx_seg1.img & xxx_seg2.img - grey and white matter segments
mode	1	Mode; write volume, no rendering of surface extraction

2b) determine normalisation parameters for tissue of interest

calls **vbm\_sn3d** with parameters the as in *standard* but

option	1	Do What; Determine Parameters Only
objima	seg1(2)files	Image to determine parameters from (objectimage)
objimages	seg1(2)files	Images to write normalised
templima	[VWD '/apriori/gray(white).img'] [one or more templates]	

2c) apply estimated warps to original images

calls **vbm\_sn3d** with parameters the same as in *standard* but

option	2	Do What; Write Normalised Only
objima	seg1(2)files	Image image for which parameters exist
objimages	files	Images to write normalised'

3) segment normalised image

calls **f99\_segment** with parameters

files	normalisedfiles	Matrix of filenames
opts	['cw']	Options; attempt to correct small intensity inhomogeneities and write inhomogeneity corrected image(s)'

**f99\_segment** calls **vbm\_segment** for each volume, which performs the segmentation like

**spm\_segment**, but saves additional information to a .mat file and uses the apriori templates from the 'apriori' directory in the VWD directory.

4) clean segments (optional)

4a) extract brain

calls **vbm\_xbrain** with parameters

P	[grey; white]	xxx_seg1.img & xxx_seg2.img - grey and white matter segments
mode	1	Mode; write volume, no rendering of surface extraction

4b) do cleaning

calls **spm\_imcalc\_ui** with parameters

P	{grey;white;brain}	Matrix of input image filenames
Q	outfile	Name of output image
f	'i1./(i1+i2+eps).*i3'	Expression to be evaluated
dmtx	[]	Read images into data matrix; default, no
mask	[]	Implicit zero mask; default, no
type	[]	Data type for output image; default, 4 (16 bit

```

signed shorts)
hold      1      Interpolation hold; bilinear

```

## Segmentation

### (only for linear and standard normalisation)

scriptfile: segment\_process\_VBM.m

description:

```

function segment_process_VBM(files,descr)
% segment_process_VBM A function to segment normalised data in SPM using K.U.L.
% functions and a standard protocol for VBM
%
% segment_process_VBM(files,descr) will segment the files resulting from normalisation
% on the files in string matrix 'files' according to options in 'descr'.
%
% options in descr used by this process:
%
% normalisation:
%   'linear', 'standard': images already normalised, perform segmentation
%   'optimised':         images are already segmented by normalise_process_VBM
%
% tissue:
%   'grey', 'white', 'both': tissues of interest
%
% clean:
%   'yes':               segments are 'cleaned' using ix/(i1+i2+eps).*i3, where
%                       ix is current segment, i1 is the grey-matter segment,
%                       i2 white matter, and i3 the extracted brain mask.
%   'no':                no cleaning on the segments
%

```

general actions:

#### 1) segment the input files

calls `f99_segment` with parameters:

files	mrifiles	Matrix of filenames
opts	['cw']	Options; attempt to correct small intensity inhomogeneities and write inhomogeneity corrected image(s)

`f99_segment` calls `vbm_segment` for each volume, which performs the segmentation like `spm_segment`, but saves additional information to a `.mat` file and uses the apriori templates from the 'apriori' directory in the VWD directory.

#### 2) clean segments (optional)

##### 2a) extract brain

calls `vbm_xbrain` with parameters

P	[grey; white]	xxx_seg1.img & xxx_seg2.img - grey and white matter segments
mode	1	Mode; write volume, no rendering of surface

## 2b) do cleaning

calls `spm_imcalc_ui` with parameters

P	{grey;white;brain}	Matrix of input image filenames
Q	outfile	Name of output image
f	'i1./(i1+i2+eps).*i3'	Expression to be evaluated
dmtx	[]	Read images into data matrix; default, no
mask	[]	Implicit zero mask; default, no
type	[]	Data type for output image; default, 4 (16 bit signed shorts)
hold	1	Interpolation hold; bilinear

## Volume Preservation (optional)

scriptfile: `modulate_process_VBM.m`

## description:

```
function modulate_process_VBM(wildcards, descr)
% modulate_process_VBM A function to preserve tissue volumes in SPM
% using K.U.L. functions and a standard protocol for VBM.
%
% modulate_process_VBM(files,descr) will adjust the grey and/or white matter
% segments resulting from VBM normalisation and segmentation on the files in
% string matrix 'wildcards' according to options in 'descr'.
%
% options in descr used by this process:
%
% normalisation:
%   'linear', 'standard': single normalisation matrix
%   'optimised':         use different parameters for each segment
%
% modulation:
%   'none': no modulation -> return
%   'full': preserve tissue volumes, tissue volumes have changed due
%           to normalisation
%   'partial': ignore volume changes as a consequence of affine
%              normalisation, but correct for volume changes created by
%              non-linear normalisation
%
% tissue:
%   'grey', 'white', 'both': tissues of interest
%
% clean:
%   'yes': input segments are cleaned -> suffix '_clean'
%   'no':  input segments are not cleaned -> no additional suffix
%
% output: modulated files will have prefix 'm' added to the filename.
```

## general actions:

calls for each input file the procedure `spm_preserve_quantity` with parameters:

matname	matgrey(white)	The <code>`_sn3d.mat`</code> file containing the spatial normalization parameters
infile	grey(white)	The spatially normalized image to be modulated
mode	mode	Optional: 0 = account for all normalization transformations 1 = only regularization for the



Modulation mode is given in the description file: 'full'  $\rightarrow$  0, 'partial'  $\rightarrow$  1.

## Smoothing

scriptfile: smooth\_process\_VBM.m

description:

```
function smooth_process_VBM(files,descr)
% smooth_process_VBM A function to smooth data in SPM using K.U.L. functions
% and a standard protocol for VBM.
%
% smooth_process_VBM(files,descr) will smooth the files resulting from the
% VBM processing on the files in string matrix 'files' according to options
% in 'descr'. Smoothed files will have the prefix 's' added to the filename.
%
% options in descr used by this process:
%
% modulation:
%   'none':      input images are not modulated -> prefix 'n'
%   'full',
%   'partial':   input images are modulated -> prefix 'mn'
%
% tissue:
%   'grey', 'white', 'both': tissues of interest
%
% clean:
%   'yes':       input segments are cleaned -> suffix '_clean'
%   'no':        input segments are not cleaned -> no additional suffix
%
% smooth:
%   array with smoothing values
%
```

general actions:

calls for each smoothing value `vbm_smooth` with parameters:

SP	SP	FWHM of smoothing kernel
files	mrifiles	[matrix of files]

`vbm_smooth` calls `spm_smooth`, and stores each smoothed file in a `smoothSP` subdirectory in the directory of the unsmoothed file.

## Automated Statistics

scriptfile: stat\_process\_VBM.m

description:

```
function stat_process_vbm(descr, groups);
% stat_process_VBM A function to do AnCoVa analysis in SPM using K.U.L.
% functions and a standard protocol for VBM.
%
```

```
% stat_process_VBM(descr, groups) will do an AnCoVa analysis on the files
% resulting from the VBM processing on the files in the matrix of cell arrays
% of strings 'groups'. Each cell array in 'groups' contains the filenames
% of the original files belonging to that group.
%
% Data is assumed to be processed and smoothed according to options in 'descr'.
%
% options in descr used by this process:
%
% modulation:
%   'none':          input images are not modulated -> prefix 'sn'
%   'full',
%   'partial':       input images are modulated -> prefix 'smn'
%
% tissue:
%   'grey', 'white', 'both': tissues of interest
%
% clean:
%   'yes':           input segments are cleaned -> suffix '_clean'
%   'no':            input segments are not cleaned -> no additional suffix
%
% smooth:
%   array with smoothing values
%
% nuisance:
%   array of values for optional nuisance variable, the effect of this
%   variable is estimated and removed, before further analysis of the group
%   differences
%
% nuisancename:
%   name for nuisance variable
%
% Results from the analysis are stored in the experiment's basedir in
% subdirectory 'res_grey_smooth<smoothing level>' and
%               'res_white_smooth<smoothing level>'
```

#### general actions:

stat\_process\_vbm calls vbm\_spm\_ui for each tissue type and smoothing value. vbm\_spm\_ui is derived from spm\_spm\_ui, but fills in an VBM - AnCoVa analysis design matrix according to the specifications in the description file.

The parameters for this AnCoVa model are:

number of groups:	from description file
input files:	the grouped input files
nuisance:	the specified nuisance variable
grand mean scaling:	no grand mean scaling
threshold masking:	no threshold masking
implicit mask:	not masking zero's
explicit mask:	optional from description file

global calculation:      omit

vbm\_spm\_ui.m calls vbm\_spm.m, which is the same as spm\_spm.m, but saves some additional information. This can be safely replaced by a call to spm\_spm.m.

## 10 Bibliography

- Abell, F., Krams, M., Ashburner, J., Passingham, R., Friston, K., Frackowiak, R., Happé, H., Frith, C., Frith, U. 1999. **The neuroanatomy of autism: a voxel-based whole brain analysis of structural scans.** *NeuroReport* **10**: 1647-1651 (1999)
- Ashburner, J. 2000. **Computational Neuroanatomy.** *PhD thesis, University College London*, 2000.
- Ashburner, J., Friston, K.J. 1999. **Nonlinear spatial normalisation using basis functions.** *Human Brain Mapping* **7**: 254-266 (1999)
- Ashburner, J., Hutton, C., Frackowiak, R., Johnsrude, I., Proce, C., Friston, K. 1998. **Identifying Global Anatomical Differences: Deformation-Based Morphometry.** *Human Brain Mapping* **6**: 348-357 (1998)
- Ashburner, J., Friston, K.J. 1997. **Spatial Transformation of Images.** *SPM 1997 Course notes* Ch. 2.
- Ashburner, J., Friston, K.J. 1999. **Voxel-Based Morphometry – The methods.** *NeuroImage* **11**: 805-821 (2000)
- Ashburner, J., Friston, K.J. 2001. **Why Voxel-Based Morphometry should be used. (Comments and Controversies.)** *NeuroImage* **14**: 1238-1243 (2001)
- Bookstein, Fred L. 2000. **Voxel-Based Morphometry should not be used with imperfectly registered images.** *NeuroImage* **14**: 1454-1462 (2001)
- Bullmore, E.T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., Brammer, M.J. 1999. **Global, Voxel and Cluster Tests, by Theory and Permutation, for a Difference Between Two Groups of Structural MR Images of the Brain.** *IEEE Transactions on Medical Imaging* **18** no. 1: 32-42 (1999)
- Davatzikos, C., Genc, A., Xu, D., Resnick, S.M. 2001. **Voxel-Based Morphometry using the RAVENS maps: methods and validation using simulated longitudinal atrophy.** *NeuroImage* **14**: 1361-1369 (2001)

Fischl, B., Dale, A.M. 2000. **Measuring the thickness of the human cerebral cortex from magnetic resonance images.** *PNAS*, **97(20)**: 11050-11055 (2000)

Friston, K.J. 1997. **Data Analysis: Basic concepts and overview.** *SPM 1997 Course notes* Ch. 1.

Friston KJ , Holmes AP , Worsley KJ , Poline J-P, Frith CD and Frackowiak RSJ. 1995. **Statistical parametric maps in functional imaging: A general linear approach.** *Human Brain Mapping* **2**:189-210 (1995).

Friston, K.J., Holmes, A.P., Kiebel, S.J., Poline, J-B. and Worsley, K.J. 1999. **Robust smoothness estimation in statistical parametric maps using normalised residuals from the general linear model.** *NeuroImage* **10**, 756-766. (1999)

Friston, K.J., Holmes, A., Poline, J-B., Price, C.J., Frith, C.D. 1995. **Detecting Activations in PET and fMRI: Levels of Inference and Power.** *Neuroimage* **4**, 223-235 (1995)

Goldszal, A., Davatzikos, C., Pham, D. L., Yan, M. X. H., Bryan, R. N., Resnick, S. M. 1998. **An image processing protocol for qualitative and quantitative volumetric analysis of brain images.** *J. Comput. Assisted Tomogr.* **22**: 827-837 (1998)

Good, C.d., Johnsrude, I.S., Ashburner, J., Henson, R.N.A., Friston, K.J., Frackowiak, R.S.J. 2000, **A Voxel-Based Morphometric study of ageing in 465 normal adult human brains.** *NeuroImage* **14**: 21-36 (2001)

Hellier P., Ashburner J., Corouge I., Barillot C., Friston K. 2002. **Inter Subject Registration of Functional and Anatomical Data Using SPM.** *Proceedings of MICCAI* (2002).

Holmes, A., Friston. K.J. 1997. **Statistical models and Experimental Design.** *SPM 1997 Course notes* Ch. 3.

Howseman, A.M., Josephs, O., Rees, G., Friston, K.J., 1997. **Special Issues in Functional Magnetic Resonance Imaging.** *SPM 1997 Course notes* Ch. 9.

Jones, S.E., Buchbinder, B.R., Aharon, I., 2000. **Three-dimensional mapping of cortical thickness using laplace's equation.** *Human Brain Mapping*, **11**:12-32 (2000).

- Kitajima, M., Korogi, Y., Hirai, T., Hamatake, S., Ikushima, I., Sugahara, T., Shigematsu, Y., Takahashi, M., Mukuno, K. 1997. **MRChanges in the calcarine area resulting from retinal degeneration.** *American Journal of Neuroradiology* 18: 1291-1295 (1997)
- Lerch, J. 2001. **Measuring Cortical Thickness**, *MSc Thesis Proposal, Department of Neurology and Neurosurgery, McGill University Montreal, Canada.* (2001)
- May, A., Ashburner, J., Büchel, C., McGonigle, D.J., Friston, K.J., Frackowiak, R.S.J., Goadsby, P.J. 1999. **Correlation between structural and functional changes in brain in an idiopathic headache syndrome.** *Nature Medicine* vol. 5, n. 7: 836-838 (1999)
- Mehta, S., Grabowski, T.J., Trivedi, Y., Damasio, H., 2002. **Evaluation of Voxel-Based Morphometry for Delineating Focal Lesions in Individual Subjects** *NeuroImage Human Brain Mapping 2002 Meeting*
- Nijland, P. 1999. **Analysing brain activity using Statistical Parametric Mapping.** *University of Groningen, Department of Mathematics and Computing Science.* 1999
- Noe, A., Gee, J.C. 2001. **Partial Volume Segmentation of Cerebral MRI scans with Mixture Model Clustering.** *Proceedings of the XVIIth International Conference on Information Processing in Medical Imaging, Davis California, USA.* 2001
- Noe, A., Kovačič, S., Gee, J.C. 2001. **Segmentation of Cerebral MRI Scans Using a Partial Volume Model, Shading Correction, and an Anatomical Prior.** *The Proceedings of the SPIE Symposium on Medical Imaging 2001, San Diego, California, USA* 2001
- Poline, J-B., Holmes, A.P., Worsley, K.J., Friston, K.J., 1997. **Statistical Inference and the Theory of Random Fields.** *SPM 1997 course notes*, Ch. 4.
- Safran, A.B., Landis, T. 1996 **Plasticity in the adult visual cortex: implications for the diagnosis of visual field defects and visual rehabilitation.** *Current Opinion in Ophthalmology* 1996, 7 (VI): 53-64
- Shah, P.J., Ebmeier, K.P., Glabus, M.F., Goodwin, G.M. 1998. **Cortical grey matter reductions associated with treatment-resistant chronic unipolar depression.** *British Journal of Psychiatry* 172:

Watkins, K.E., Paus, T., Lerch, J.P., Zijdenbos, A., Collins, D.L., Neelin, P., Taylor, J. Worsley, K.J., Evans, A.C., 2001. **Structural Asymmetries in the Human Brain; a Voxel-based Statistical Analysis of 142 MRI Scans.** *Cerebral Cortex* **11**, 868-877 (2001)

Worsley, K.J., Marrett, S., Neelin, P., Vandal, A.C., Friston, K.J., Evans, A.C. 1996. **A unified statistical approach for determining significant signals in images of cerebral activation.** *Human Brain Mapping*, **4**:58-73. (1996)