Searching Scale Space for Activation in PET Images

K.J. Worsley,   S. Marrett,   P. Neelin,   and   A.C. Evans

Department of Mathematics and Statistics, McGill University, 805 Sherbrooke St. West, Montreal, Québec, Canada H3A 2K6 (K.J.W.), and McConnell Brain Imaging Centre, Montreal Neurological Institute, 3801 University St., Montreal, Québec, Canada H3A 2B4 (S.M., P.N., A.C.E.)

◇───◇

Abstract: PET images of cerebral blood flow (CBF) in an activation study are usually smoothed to a resolution much poorer than the intrinsic resolution of the PET camera. This is done to reduce noise and to overcome problems caused by neuroanatomic variability among different subjects undertaking the same experimental task. In many studies the choice of this smoothing is arbitrarily fixed at about 20mm FWHM, and the resulting statistical field or parametric map is searched for local maxima. Poline and Mazoyer (1994ab) have proposed a 4-D search over smoothing kernel widths as well as the usual three spatial dimensions. If the peaks are well separated then this makes it possible to estimate the size of regions of activation as well as their location. One of the main problems identified by Poline and Mazoyer (1994ab) is how to assess the significance of scale space peaks. In this paper we provide a solution. Our main result is a unified P-value for the 4-D local maxima that is accurate for searches over regions of any shape or size. Our results apply equally well to any Gaussian statistical field, such as those resulting from fMRI.

Key words: PET, fMRI, adaptive filter, Euler characteristic, random fields.

◇───◇

Addresses of corresponding author:
e-mail: keith@zaphod.math.mcgill.ca
ph: 1-514-398-3842
fax: 1-514-398-3899.

Short title: Searching Scale Space

1   Introduction

Many studies of brain function with positron emission tomography (PET) involve the interpretation of a subtracted PET image, usually the difference between two images of cerebral blood flow (CBF) under baseline and activation conditions. In many cognitive studies, the activation is so slight (4-8%) that the experiment must be repeated on several subjects. The images are then mapped into a standardised or stereotaxic coordinate space (Talairach *et al.*, 1988) to account for differences in brain size and orientation, and the subtracted images averaged to improve the signal to noise ratio (Fox *et al.*, 1985; Friston *et al.*, 1990, 1991; Evans *et al.*, 1992; Worsley *et al.*, 1992). The averaged ΔCBF image is then normalised by dividing by an estimate of the standard deviation and the resulting statistical field or parametric map is searched for local maxima.

In many studies the images are first smoothed to a FWHM of 20mm before the analysis is performed (Friston *et al.*, 1990, 1991; Worsley *et al.*, 1992). Some of the reasons for this are: to reduce the impact of miss-registration of PET data into stereotaxic space, to allow for anatomical variability among subjects, and to improve the signal to noise ratio. The value of 20mm is somewhat arbitrary and no thorough justification has been given. The matched filter theorem (Rosenfeld and Kak, 1982) states that the optimal choice of smoothing kernel should should match the signal to be detected, in order to best detect the signal. Thus 20mm objects are best detected by a 20mm filter. Miss-registration and anatomical variability increase the effective signal width when the subtracted image is averaged over subjects, so a wider filter should be used to allow for this. Rather than try to determine the signal width in advance, Poline and Mazoyer (1994ab) have proposed a search over a range of filter widths to find 4-D local maxima in location and scale space. This has the added advantage of estimating the signal width as well as its location. In Section 3 we show how to generate these images and search them in a computationally efficient way.

One of the main problems identified by Poline and Mazoyer (1994ab) is how to set the threshold of the 4-D local maxima in order to control the false positive rate of finding local maxima above the threshold when in fact no signal is present in the image. They note that the threshold must obviously be larger than that at any fixed scale because searching over an extra dimension (scale) must increase the false positive rate. As the filter width decreases, the field becomes 'rougher' and there is a greater chance of a local maximum exceeding a fixed threshold by chance alone (Friston *et al.*, 1991; Worsley *et al.*, 1992). Thus the threshold should be at least as high as that for the smallest filter width. However no theoretical solution to this problem is given.

In Section 4 we solve this problem by presenting a unified *P*-value for the 4-D local maxima. The result is based on recent theoretical work on 4-D Gaussian random fields (Siegmund and Worsley, 1995). It builds on work in the related paper of Worsley *et al.* (1995a) which gives a unified *P*-value for spatial (3-D) local maxima. The *P*-value is *unified* in the sense that it combines previous 2-D results (Friston et al, 1991) and 3-D results (Worsley *et al.*, 1992) to give a good estimated *P*-value for search regions of any shape or size. This makes it possible to restrict the search to small regions such as the cingulate gyrus, or two dimensional regions such as a slice or the cortical surface, or even single voxels, while at the same time searching over filter widths. Of course the additional flexibility of searching in an extra dimension does not come without some cost. We show that the

price paid for searching in scale space from 6.8mm FWHM (the intrinsic resolution of our scanner) to 34mm FWHM (five times the intrinsic resolution), as opposed to searching at 20mm FWHM, is an increase in the critical threshold of about 0.8 on the Gaussian scale, which is roughly equivalent to reducing the number of subjects by 30%. Hence about 30% of the sample is sacrificed to permit a search in scale space.

The method is not without its drawbacks. If two peaks are separated by less than twice their FWHM then they are optimally detected as a single wide peak rather than two narrow ones. Thus multiple peaks in a small region tend to be blurred together rather than separated, which adds to their detectability but not to their resolvability. We recommend that any peak detected at a wide filter width should be re-examined at a narrower filter width to look for multiple peaks. In Section 5 the unified P-value is validated on null PET data then applied to activation data, followed by a discussion in Section 6.

2 The effect of smoothing on the signal to noise ratio

In this Section we shall apply elementary results from signal processing (see for example Rosenfeld and Kak, 1982) to investigate the effect of signal width, anatomical variability and filter width on the detectability of the signal.

In a typical CBF activation study, PET image data are collected from n subjects under a set of m conditions or tasks. Each image is smoothed to an effective FWHM of w in each dimension either at reconstruction (Evans *et al.*, 1992) or using a Gaussian kernel smoother (Friston *et al.*, 1990). In the simplest approach, the images are normalised by dividing by the global mean of all intracranial voxels, then the image under one condition is subtracted from the image under another condition, and the resulting difference images are averaged over subjects to produce a ΔCBF image. If there is no strong evidence for fluctuations in the voxel variance, then Worsley *et al.* (1992) pool the subject variance over voxels to obtain an effectively constant estimator of the standard deviation. Dividing the ΔCBF image by this standard deviation produces a Gaussian statistical field with unit standard deviation, which is then searched for local maxima. An alternative is to divide by the voxel standard deviation, and then transform the resulting T statistic to a Gaussian random variable (Friston *et al.*, 1991), although the resulting statistical field is only approximately Gaussian provided the degrees of freedom is high enough (see Worsley *et al.* (1995a) for a discussion of this issue).

2.1 Model

We shall develop a very simple approximate model for PET activation which will allow us to get some idea of the factors that influence signal detectability. Let D be the number of dimensions of the spatial data; $D = 3$ for volumetric data and $D = 2$ for slices. Denote the D-dimensional Gaussian function with standard deviation σ by

$$\phi(\sigma) = (2\pi\sigma^2)^{-D/2} \exp[-||\mathbf{x}||^2/(2\sigma^2)]. \tag{2.1}$$

where $w = \sigma\sqrt{8\log_e 2}$ is the FWHM and $\mathbf{x}$ is a D-vector. The Gaussian function has the attractive property that the convolution of two Gaussians is still Gaussian, that is,

$$\phi(\sigma_1) \star \phi(\sigma_2) = \phi(\sqrt{\sigma_1^2 + \sigma_2^2}), \tag{2.2}$$

where $\star$ denotes convolution.

We shall assume that the D-dimensional ΔCBF image can be modelled as follows (see Figure 1 for an illustration in $D = 1$ dimensions); similar models were used by Friston *et al.* (1991) and Worsley *et al.* (1992):

- the signal consists of just one peak that can be modelled as the Gaussian function $h(2\pi\sigma_S^2)^{D/2}\phi(\sigma_S)$ where h is the signal height and $\sigma_S\sqrt{8\log_e 2}$ is the signal FWHM;

- the noise component is modelled as stationary white noise ϵ, which is added to the signal;

- the effect of image reconstruction and additional smoothing is to convolve the signal plus noise with a Gaussian point spread function $\phi(\sigma_R)$ with FWHM $\sigma_R\sqrt{8\log_e 2}$ to produce the ΔCBF image;

- anatomical variability produces registration error that perturbs the location of the ΔCBF image on each subject about a common location, which without loss of generality we shall take as the origin. Functional variability, due to differences in the anatomical location of the activation across subjects, is also included in this term. We shall model these perturbations as a 3-D Gaussian random variable with standard deviation σ_A.

The ΔCBF image, averaged over a large number n of subjects, can then be written as

$$
\begin{aligned}
\Delta\text{CBF} &= [\text{signal} \star \text{anatomical variability} + \text{white noise}] \star \text{reconstruction} \\
&= \left[h(2\pi\sigma_S^2)^{D/2}\phi(\sigma_S) \star \phi(\sigma_A) + \epsilon/\sqrt{n}\right] \star \phi(\sigma_R).
\end{aligned}
\tag{2.3}
$$

2.2 Signal to noise ratio

It is straightforward to check that the signal component of ΔCBF is the Gaussian function

$$
h(2\pi\sigma_S^2)^{D/2}\phi\left(\sqrt{\sigma_S^2 + \sigma_A^2 + \sigma_R^2}\right).
\tag{2.4}
$$

The noise component of ΔCBF has a Gaussian spatial covariance function proportional to

$$
[\phi(\sigma_R)/\sqrt{n}] \star [\phi(\sigma_R)/\sqrt{n}] = \phi(\sqrt{2}\sigma_R)/n.
\tag{2.5}
$$

The signal to noise ratio at the origin, SNR, is then proportional to the ratio of (2.4) to the square root of (2.5) evaluated at the origin:

$$
\text{SNR} \propto h\sqrt{n}\left(\frac{\sigma_S^2\sigma_R}{\sigma_S^2 + \sigma_A^2 + \sigma_R^2}\right)^{D/2}.
\tag{2.6}
$$

The constant of proportionality depends only on the amplitude of the white noise.

A few remarks are obvious: SNR increases with signal height h, number of subjects n and signal width σ_S, and decreases with registration error σ_A. However the SNR reaches a maximum when the filter width matches the perturbed signal width, that is when

$$
\sigma_R = \sqrt{\sigma_S^2 + \sigma_A^2}.
\tag{2.7}
$$

This is essentially the same result as the matched filter theorem (Rosenfeld and Kak, 1982) which states that the signal is best detected by smoothing with a filter that matches the signal. Figure 2(a) illustrates the SNR in $D = 3$ dimensions. If the filter is twice as wide or half as wide as the perturbed signal then the signal to noise ratio is reduced by a factor of 0.28, which is equivalent to the loss of almost half the subjects. Thus wider signals are best detected with a wide filter, and narrow signals with a narrow filter. This provides some justification for searching over filter widths.

Finally, we shall investigate the effect of anatomical variability. If we use the optimum filter width (2.7) then the SNR is

$$\text{SNR} \propto h\sqrt{n}\frac{(\sigma_\text{S}/2)^{D/2}}{(1+\sigma_\text{A}^2/\sigma_\text{S}^2)^{D/4}}. \tag{2.8}$$

For $D = 3$ dimensions, a plot of this in Figure 2(b) shows that if the anatomical variability equals the signal width then the signal to noise ratio drops by 0.41 which is equivalent to losing 65% of the subjects. Thus reducing the registration error component of anatomical variability by non-linear image warping (Evans *et al.*, 1991a; Collins *et al.*, 1992, 1994, 1995) could substantially improve the sensitivity of signal detection in averaged PET and MRI images.

2.3   Examples

Figure 3 gives an example of (a) $D = 1$ white noise with no signal plus (b) a FWHM=15mm signal. The horizontal axis is chosen to represent a lateral line through the brain centre of a ΔCBF image, divided into 100 1.34mm pixels. To simulate reconstruction at 6.8mm FWHM (the intrinsic resolution of our PET camera, see Section 5) and smoothing to 15mm and 34mm (5 times the intrinsic resolution), the white noise is smoothed with a 6.8, 15 and 34mm FWHM filter and corrected to give unit standard deviation. Note that the 15mm filter gives the largest peak, as the theory predicts. Figures 3(c) and 3(d) present the same data in the form of a two dimensional statistical field. As described later in Section 3.3, the FWHM is sampled at only 13 filter pixels (or 'fixels') equally spaced on the log(FWHM) axis (the field is very smooth in this direction); the bottom row is the 6.8mm FWHM reconstruction and the top row is the 34mm FWHM smoothed data. Notice that the statistical field maximum is very close to the location and scale of the 15mm FWHM signal.

Figure 3(e) shows the same noise component but with two signals. The first image contains a 26mm and a 9mm FWHM signal; note that both are accurately located by the local maxima of the statistical field. Figure 3(f) contains two 9mm FWHM signals 18mm apart. Here the local maximum is at 34mm FWHM between the two signals, so that they are estimated as one wide signal rather than two separate narrow signals. It can be shown that this behaviour happens, on average, whenever the two signals are less than 2.8 signal FWHMs apart. In practice, then, our procedure tends to blur together signals which are too close; this increases their detectability at the expense of their resolvability. However such 'bifurcation' is readily apparent in the scale space images so we recommend inspection of these slices for hidden smaller foci (see Figure 7).

3   Efficient smoothing of the data

A major obstacle to searching in scale space is the computational effort of repeating the smoothing and statistical analysis on all $n \times m$ images. However we shall show that this effort can be reduced appreciably in several ways. An obvious reduction comes from the fact that kernel smoothing is a *linear* operation on the data, and that the ΔCBF image is also a *linear* function of the data; linear operators can be interchanged, so that the ΔCBF for the smoothed images is equivalent to smoothing the ΔCBF image for the unsmoothed data. Thus the statistical field of the smoothed data is proportional to the smooth of the statistical field at the highest resolution.

The remaining task is to determine the constant of proportionality, equal to the ratio of the standard deviations, without resorting to smoothing all the data again. We shall do this in two ways. The first and quickest method is based on assuming that the correlation function of the data is Gaussian. Although this assumption is accurate enough for many purposes, we have found that it is not accurate enough to determine the standard deviation of our PET data (see Figure 5). For this reason we have developed a second method that works for any correlation function; it takes longer than the first method but not as long as smoothing all the data.

3.1   Gaussian correlation function

In this subsection we derive a very simple method for correcting the standard deviation which requires that the spatial correlation function of the highest resolution statistical field, Z_1 say, can be well approximated by a Gaussian function as in (2.5). This will be the case if the noise component of Z_1 can be modelled as stationary white noise smoothed by a stationary Gaussian point response function with FWHM$= w_1$ say. If Z_1 is then smoothed by a Gaussian kernel with FWHM$= \sqrt{w^2 - w_1^2}$, then by (2.2) the resulting image is a Gaussian statistical field Z with an effective FWHM$= w$. Siegmund and Worsley (1995) show that the standard deviation is reduced by a factor of $c = (w_1/w)^{D/2}$, so that dividing by c then restores the standard deviation of Z to unity. The operations on the data can be summarised as follows:

1. Reconstruct the data at the highest resolution with FWHM$= w_1$;

2. Calculate the Gaussian statistical field, Z_1, with a standard deviation of 1 at each voxel, for the highest resolution data;

3. For a desired FWHM$= w$, smooth Z_1 with the Gaussian kernel (2.1) with FWHM$= \sqrt{w^2 - w_1^2}$;

4. Correct the standard deviation by dividing by $c = (w_1/w)^{D/2}$;

5. Repeat steps 3 and 4 for each desired FWHM.

3.2   Non-Gaussian correlation function

Antoine *et al.* (1994) have reported that the correlation function of high resolution PET data is non-Gaussian, with negative side-lobes in the xy plane. This was attributed to a low apodization cutoff frequency for PET images reconstructed by filtered backprojection. For our PET data, we have found that the Gaussian correlation function is a good first approximation but not accurate enough to determine the standard deviation by the method presented in the previous section. In this section we present an accurate method for correcting the standard deviation that does not require the assumption of a Gaussian correlation function; the only requirement is stationarity of the pure noise component.

The first step is to estimate the correlation function $r(\mathbf{x})$ of the unsmoothed (highest resolution) data, that is, the correlation between voxels at a separation vector of $\mathbf{x}$ voxels. Fourier transform methods are the most efficient way of doing this, provided the number of voxels on each axis is a power of 2. Residual (pure noise) images are created by subtracting all estimated effects such as an overall mean and the effect of any regressors from the individual scans. A mask is defined that takes the value 1 for all intercranial voxels, and 0 otherwise. To avoid wrap-around effects, the mask should be at least one FWHM from the boundary of the image on all sides. The masked residual image, formed by multiplying the residual by the mask, is then passed through a D-dimensional Fourier transform, and the squared modulus of each transform is cumulated over all images. To speed up calculation, two residuals can be transformed simultaneously by putting one into the real part and one into the imaginary part of the complex array to be transformed. No sorting is required afterwards; the squared modulus of the transformed complex array is the sum of the squared moduli of the two Fourier transformed real arrays, plus an odd function whose inverse Fourier transform is purely imaginary. This convenient trick halves the computation time.

The sum of the squared moduli is then passed through an inverse D-dimensional Fourier transform. The real part is the sum of cross-products $C(\mathbf{x})$ of the residual images at lag $\mathbf{x}$; the imaginary part is discarded. The mask is convolved with itself in the same way, to produce an image of counts $N(\mathbf{x})$ of the number of pairs of voxels in the mask that are separated by $\mathbf{x}$. Note that $C(\mathbf{x})$ and $N(\mathbf{x})$ are periodic, so that values for negative indices of $\mathbf{x}$ can be found by adding the size of the image to such indices. The corrected unbiased estimate of the spatial correlation $r(\mathbf{x})$ at a lag of $\mathbf{x}$ is:

$$r(\mathbf{x}) = [C(\mathbf{x})/N(\mathbf{x})]/[C(\mathbf{0})/N(\mathbf{0})]. \tag{3.1}$$

This estimate is valid only for $\mathbf{x}$ smaller than the gap between the mask and the boundary of the image on each side, since larger lags will contain cross-products of the image wrapped around on itself. However we expect $r(\mathbf{x})$ to be zero for large lags so this restriction is not important.

Once the spatial correlation function is obtained, the correction factor $\tilde{c}$ is easily calculated for any smoothing filter $f(\mathbf{x})$:

$$\tilde{c} = \sqrt{\sum_{\mathbf{x}}(f \star f)(\mathbf{x})r(\mathbf{x})}, \tag{3.2}$$

where $\star$ is the convolution operator. If the filter is a product of one-dimensional filters in each of the D dimensions, such as the Gaussian filter (2.1), then $f \star f$ can be rapidly calculated as the product of the convolution of the one-dimensional filters.

As a check on the calculations, the standard deviation pooled over the voxels in the mask is

$$\sqrt{C(\mathbf{0})/[N(\mathbf{0})\nu]}, \tag{3.3}$$

where ν is the degrees of freedom of the residuals, equal to the number of images minus the number of effects removed (ν is one less than the number of subtractions in a simple subtraction experiment). As a further check, the effective FWHM of the point spread function of the smoothed image, calculated from the standard deviation of the derivative of the residuals assuming a Gaussian point spread function (Friston *et al.*, 1991; Worsley *et al.*, 1992), can also be calculated directly from the spatial correlation function. Let $\dot{f}_j$ be the one-step difference of f in dimension j, equal to f minus f lagged by one voxel in dimension j, and let δ_j be the voxel size in dimension j. Then the covariance matrix $\mathbf{\Lambda}$ of the derivatives of the smoothed standardised image has (j, k) element

$$\lambda_{jk} = \sum_{\mathbf{x}} (\dot{f}_j \star \dot{f}_k)(\mathbf{x}) r(\mathbf{x})/(\delta_j \delta_k \tilde{c}). \tag{3.4}$$

$j, k = 1, \ldots, D$. The effective FWHM in dimension j is

$$\sqrt{(4 \log_e 2)/\lambda_{jj}}, \tag{3.5}$$

$j = 1, \ldots, D$, and the overall effective FWHM is

$$\tilde{w} = \sqrt{(4 \log_e 2)/\det(\mathbf{\Lambda})^{1/D}}. \tag{3.6}$$

The effective FWHM for the unsmoothed data can be obtained directly from (3.6) by replacing f by a 'spike' at the origin, that is $f(\mathbf{0}) = 1$ and zero elsewhere.

The operations on the data can be summarised as follows:

1. Reconstruct the data at the highest resolution with FWHM= w_1;

2. Find the spatial correlation function $r(\mathbf{x})$ of the highest resolution data;

3. Check that the pooled standard deviation from (3.3) is the same as that from a direct calculation;

4. Check that the effective FWHM from (3.6) calculated for a spike function is close to the nominal FWHM w_1;

5. Calculate the Gaussian statistical field, Z_1, with a standard deviation of 1 at each voxel, for the highest resolution data;

6. For a desired FWHM= w, smooth Z_1 with the Gaussian kernel (2.1) with FWHM= $\sqrt{w^2 - w_1^2}$;

7. Correct the standard deviation by dividing by $\tilde{c}$ from (3.2);

8. Check that the effective FWHM $\tilde{w}$ from (3.6) is close to the desired FWHM w;

9. Repeat steps 6, 7 and 8 for each desired FWHM.

In summary, the methods of this subsection allow us to correct the standard deviation of a smoothed image without smoothing all the data and without assuming a Gaussian point spread function. The computational cost is about the same as half the cost of one smoothing of all the data. For searches in scale space, which require many smoothes, this is a considerable saving.

3.3 Sampling scale space

Another reduction in computations arises because in practice only a small number, typically 12, of different smoothes are required. The reason is that the scale space statistical field is very smooth in the scale dimension, so that very few 'fixels', or filter pixels, need to be used to adequately sample the scale axis. Siegmund and Worsley (1995) show that the scale space statistical field is stationary in $\log_e w$, where the effective FWHM is $\sqrt{(8/D)\log_e 2}$. For $D = 3$, this gives an effective FWHM of 1.36, so that a fixel size of 1/10 of the FWHM, or $1.36/10 = 0.136$, would seem appropriate to give adequate coverage (our highest resolution data has a voxel size of approximately 1/4 to 1/6 of the FWHM). A five-fold range of FWHM from w_1 to $w_2 = 5w_1$ seems large enough to cover most signals of interest. This then becomes a range of $\log_e 5 = 1.61$ on the log scale, which covers $1.61/0.136 = 12$ equal-sized fixel intervals. Adding one for the end point gives a total of 13 fixels, whose 'coordinates' are $w = w_1 5^{i/12}$ FWHM, $i = 0, 1, \ldots, 12$. For our data the intrinsic FWHM of the scanner is $w_1 =$ 6.8mm, so this gives fixel coordinates at $w =$ 6.8, 7.8, 8.9, 10.2, 11.6, 13.3, 15.2, 17.4, 19.9, 22.7, 26.0, 29.7 and 34.0mm FWHM. Note that the spacing increases as the FWHM increases, and only 12 kernel smoothing operations need to be performed.

Further reductions in both time and storage can be achieved by image compression (see Appendix A.1).

4 The unified P-value

Throughout this Section we will be concerned with volumetric data ($D = 3$) and the P-value of the maximum M of the 4-D location and scale space statistical field searched over all voxels inside a search region V, and searched over all fixels from w_1 to w_2 FWHM. Equivalent results for slice data ($D = 2$) and line data ($D = 1$) are given in Appendix A.2. We shall assume that the correlation function is Gaussian; for the non-Gaussian case, the results of this Section can serve as a useful approximation provided the standard deviation is corrected as in Section 3.2.

As in the case of searching spatial statistical fields, there is no known exact result for the P-value of M. An added complication is that the 4-D statistical field is not stationary in both location and scale simultaneously. This is apparent from Figure 3 where the image becomes smoother in the spatial direction as the scale increases. Thus we cannot use the results of Worsley *et al.* (1995a) since they only work for stationary fields. Accordingly special methods must be used.

We report a recent result of Siegmund and Worsley (1995), which gives exact results for a quantity closely related to the P-value, the expected *Euler characteristic* (EC) of the *excursion set* of a scale space statistical field. The excursion set is simply the set of voxels

where the statistical field exceeds a fixed threshold t and the EC counts the number of connected components of the excursion set, minus the number of 'holes' (see Worsley *et al.*, 1992). For high thresholds the holes disappear and the EC counts the number of local maxima of the statistical field. For even higher thresholds, near the global maximum M, the EC counts one if $M \geq t$ and zero otherwise, so that the expected EC approximates the P-value of M. The expected EC is therefore our proposed unified formula for the P-value of the 4-D maximum M. It has exactly the same form as the unified P-value in Worsley *et al.* (1995a):

$$P(M \geq t) \approx \sum_{d=0}^{3} R_d(V)\rho_d(t). \tag{4.1}$$

$R_d(V)$ is the d-dimensional *resel count* calculated for the highest resolution statistical field with FWHM$= w_1$; exact definitions, some examples and a method of calculating the resel counts from voxel data are given by Worsley *et al.* (1995a). $\rho_d(t)$ is the d-dimensional *EC density* which is given by the following:

$$
\begin{aligned}
\rho_0(t) &= \frac{1}{(2\pi)^{1/2}} \left\{ \sqrt{\frac{3}{4\pi}} (-\log_e u) e^{-t^2/2} + \int_t^{\infty} e^{-z^2/2} dz \right\} \\
\rho_1(t) &= \frac{(4\log_e 2)^{1/2}}{(2\pi)} \left\{ \sqrt{\frac{3}{4\pi}} (1-u)t + \frac{1+u}{2} \right\} e^{-t^2/2} \\
\rho_2(t) &= \frac{(4\log_e 2)}{(2\pi)^{3/2}} \left\{ \sqrt{\frac{3}{4\pi}} \frac{1-u^2}{2} \left(t^2 - \frac{1}{3}\right) + \frac{1+u^2}{2} t \right\} e^{-t^2/2} \\
\rho_3(t) &= \frac{(4\log_e 2)^{3/2}}{(2\pi)^2} \left\{ \sqrt{\frac{3}{4\pi}} \frac{1-u^3}{3} (t^3 - t) + \frac{1+u^3}{2} (t^2 - 1) \right\} e^{-t^2/2} \tag{4.2}
\end{aligned}
$$

where $u = w_1/w_2$. EC densities for smoothed lower dimensional data, such as slices or single lines through an image, are given in Appendix A.2.

The above P-value (4.1) is *unified* because we shall impose essentially no restrictions on the search region; it can be a 3-D volume of arbitrary shape, a 2-D slice, a 2-D surface (such as the cortical surface), a 1-D line or a 0-D point; it can be connected or disconnected. Note that if $u = 1$, so that the search is in location space only at the lowest FWHM, then (4.1) is identical to the unified P-value of Worsley *et al.* (1995a).

For large search regions the first three terms can be neglected and just the last term ($d = 3$) of (4.1) is accurate enough for the P-value. This means that we need only calculate $R_3(V)$, which equals the number of RESELS or resolution elements of the highest resolution image, defined as the volume of V divided by w_1^3 (Worsley *et al.*, 1992).

The P-value (4.1) is most accurate when the search region is convex, and it appears to be satisfactory only for high thresholds t whenever the P-value is less than 0.2. For lower thresholds, (4.1) should be interpreted as the expected number of false positive regions of activation above the threshold.

Figure 4 plots critical thresholds of M for a spherical search region with w_1=6.8mm and $w_2 = 34$mm, so that $u = 5$. Critical values for fixed $w = 6.8$, 20 and 34mm FWHM are also shown. These can be obtained from (4.1) by setting $u = 1$ and dividing $R_d(V)$ by $(w/w_1)^d$,

$d = 1, 2, 3$. It can be seen that the cost of searching over scale adds about 0.8 to the critical thresholds at w =20mm FWHM. This is a 20% increase, which is equivalent to reducing the number of subjects by about 30%. Hence about 30% of the sample is sacrificed to permit a search in scale space. Compared to a fixed 6.8mm FWHM the cost is negligible; a 0.06 increase in critical value or a 2.5% reduction in sample size.

5   Validation and application

5.1   Methods

PET scans were obtained using the Scanditronix PC-2048 system which produces 15 image slices 6.5mm apart with a transverse image resolution of 4.6-6.4mm FWHM and an axial resolution of 5.4-7.1mm FWHM (Evans *et al.*, 1991b). Using the bolus $H_2^{15}O$ methodology without blood sampling (Herscovitch *et al.*, 1983; Raichle *et al.*, 1983; Fox *et al.*, 1984), the relative distribution of CBF was measured in baseline and activation conditions. All subjects also had an MRI scan containing 64 2mm-thick T_1-weighted multi-slice spin-echo images ($T_R = 550$msec ; $T_E = 30$msec). Using a volumetric image registration procedure described previously (Evans *et al.*, 1989, 1991a), the MRI volume from each subject was aligned with the corresponding PET volume. An orthogonal coordinate frame was then established based on the anterior commissure - posterior commissural (AC-PC) line as identified in the MRI volume (Evans *et al.*, 1992). These anatomical frame coordinates were used to apply a trilinear re-sampling of each matched pair of MRI and PET datasets into a standardized stereotactic coordinate system (Fox *et al.*, 1985; Talairach *et al.*, 1988). Individual images were then sampled on a $128 \times 128 \times 80$ lattice of voxels, separated at approximately 1.34mm, 1.72mm and 1.50mm on x, y and z axes, respectively, then PET images were normalised by dividing each voxel by the mean value for all intra-cerebral voxels. The intra-cerebral voxels were defined as all voxels with a value greater than 150% of the mean value of the entire image. The ΔCBF image was obtained by averaging the subtracted images across subjects, then converted to a Gaussian statistical field by dividing each voxel by the mean standard deviation in normalised ΔCBF for all intra-cerebral voxels (Worsley *et al.*, 1992). Individual MRI images were subjected to the same averaging procedure.

5.2   Validation on null data

Coghill *et al.* (1994) carried out an experiment in which PET cerebral blood flow images were obtained for $n = 9$ subjects while a thermistor was applied to the left forearm at both warm (34^oC) and hot (48^oC) temperatures, each condition being studied twice on each subject. The purpose of the experiment was to find regions of the brain that were activated by the hot stimulus, compared to the warm stimulus. For validation, we analysed the difference images of the two warm conditions as a null dataset which should have an expectation of zero throughout.

For the null dataset, the variance of derivatives of the individual subtraction images was used to estimate the effective FWHM in each direction (Friston *et al.*, 1991; Worsley *et al.*, 1992). The effective FWHMs (3.5) were 6.2, 6.9 and 7.3mm in the x, y and z directions, and

so 6.8mm was taken as a common value. The images were smoothed over a five fold range from $w_1 = 6.8$mm to $w_2 = 34$mm FWHM at fixel coordinates 6.8, 7.8, 8.9, 10.2, 11.6, 13.3, 15.2, 17.4, 19.9, 22.7, 26.0, 29.7 and 34.0mm FWHM, equally spaced on the log(FWHM) scale. Image compression, as advocated in Section A.1, was not used, in order to maintain compatibility with existing software.

The correlation function was found using the methods of section 3.2 (Figure 5). The Gaussian assumption holds well in the y and z directions but not in the x direction, where two negative side-lobes are evident. Similar behaviour was noted in other studies. Figures 6(a) and 6(b) compare the Gaussian correction factor c with the non-Gaussian correction factor $\tilde{c}$ and the correction factor found by smoothing all 9 subtraction images and pooling the standard deviation; the latter took several hours of computation whereas $\tilde{c}$ took several minutes. It can be seen that $\tilde{c}$ is much more accurate than c for this data. Discrepancies for larger FWHM may be explained by a smaller number of resels and hence greater variability in the estimator (Worsley, 1995b). Figures 6(c) and 6(d) compare the effective FWHM of the smoothed data $\tilde{w}$ with the desired FWHM w; the agreement is reasonably close.

The search region V was a 1227cc whole-brain region. To apply the unified P-value, the resel counts were taken from Table 3 of Worsley $et\ al.$ (1995a), which gives values of $R_d(V)$ for 20mm FWHM. The resel counts at 6.8mm FWHM, obtained by multiplying $R_d(V)$ by $(20/6.8)^d$, are $R_3(V) = 3903$, $R_2(V) = 926$, $R_1(V) = 60$, and $R_0(V) = 1$. The approximate level 0.05 critical value for M is $t = 5.04$, found by equating (4.1) to 0.05 and solving for t. The 4-D maximum was $M = 4.45$ ($P = 0.43$) for the null dataset, so that no significant signal is detected.

5.3 Application to activation data

We also analysed the difference between the average of the two hot conditions and the average of the two warm conditions to search for activation due to the painful heat stimulus. For the same search region as above, the 4-D maximum was $M = 6.56$ ($P < 0.0001$); the maximum was located at 22.7mm FWHM in the anterior insula/supplementary motor area close to where activation was reported by Coghill $et\ al.$ (1994). Figure 7 shows xy, xw and yw slices through this peak.

Table 1 gives a list of all significant 4-D peaks, together with equivalent 3-D peaks reported in the original study of Coghill $et\ al.$ (1994), where these can be identified. Strict comparisons with the original study are not possible because the original data was reconstructed at 20mm FWHM in the x and y directions with no axial smoothing. This does not correspond to any of the images in our scale space because they are all smoothed to equal resolutions in all dimensions. Nevertheless, all the peaks in the original study also appear in Table 1. The peak heights are generally higher in the scale space search, and the P-values show that the peaks are more significant. There are two exceptions: the left putamen and the right posterior cingulate are more significant in the original analysis, presumably because these peaks are 'flatter' in the z direction and are better detected by a 'flat' filter. There is strong evidence for this in the case of the right posterior cingulate, where the original analysis found a single peak mid-way between the two sharp foci detected by the scale space search. It appears as if these two foci were merged by the 20mm xy filtering to produce one peak in the original analysis.

One interesting feature is that the reverse has happened to the SII/anterior insula peak, which was reported as two separate peaks, one in SII and and one in the anterior insula, in the original study. This merge can in fact be seen by looking at the yw slice of Figure 7, where the main peak appears to 'bifurcate' at lower scales, indicating two peaks about 2cm apart along the y axis, which are identifiable as the SII and anterior insula peaks of the original study. In addition, the scale space search revealed a small focus of 6.8mm in the caudate that did not appear in the original study, presumably because it was over smoothed. This may also explain why the small focus in the left anterior insula had a much smaller ($M = 3.72$) significance in the original study.

6 Discussion

6.1 Assumptions

All the above theoretical work does rely on certain assumptions about the distribution of the data, which we shall discuss below. These are:

- **Gaussian distribution for the averaged ΔCBF image.** This is not too problematic because averages approach a Gaussian distribution as the number of subjects (or scans) increases, thanks to the Central Limit Theorem.

- **Gaussian correlation function.** This is required for three things:

  - choosing the width of the smoothing filter to achieve a given effective FWHM. This is not important; errors in the resulting effective FWHM will produce errors in our estimates of signal extent, but since scale space is already very smooth then these estimates of extent are already quite variable.

  - correcting the pooled standard deviation. This is important because small errors in the standard deviation produce large changes in the P-value. To get an accurate estimate, we were forced to compute the correlation function (Section 3.2).

  - validity of the unified P-value formula. This is not important; the P-value actually depends not on the whole form of the correlation function but only on its curvature (second derivative) at zero lag, which equals the variance of the derivative of the process. Inaccuracies can occur if this curvature is not accurately estimated, and the most important part is the curvature of the highest resolution image, not the curvature of the lowest resolution image. This issue has been investigated by Poline *et al.* (1995) who shows that it affects sensitivity but not specificity.

- **Stationary noise correlation structure.** This implies that the point spread function is equal at all voxels, which appears to be valid for intra-cerebral voxels. The assumption is required for calculating the correlation function by pooling across voxels, and for the validity of the unified P-value.

- **Equal variance across voxels.** Worsley *et al.* (1995a) discuss the impact of this assumption on the fixed FWHM analysis, and the same arguments apply here to scale

space. Although this assumption is strictly incorrect, the main question is to what extent it affects the analysis. Worsley *et al.* (1995a) show that specificity is not seriously affected by fluctuations of less than 8% in the voxel standard deviation, but the sensitivity, particularly in cortical areas near the skull, could be reduced. This reduction must be offset against the inevitable reduction in sensitivity caused by replacing the pooled standard deviation by a voxel standard deviation with low degrees of freedom and higher variability.

The solution might be to accept this reduction and extend our results to images of *t*-statistics, as in Worsley *et al.* (1995a). However there is a subtle theoretical obstacle to such a proposal. It can be shown that the model used to justify a voxel standard deviation at one scale cannot hold at any other scale. In other words, the *t*-image at one scale may be free of an unequal voxel standard deviation (by dividing by the voxel standard deviation) but at another scale the unequal voxel standard deviations affect the stationarity of the *t*-image; in particular, the effective FWHM is non-stationary and depends on the unequal voxel standard deviations. There seems to be no way of avoiding this. The best course of action might be to work with the Gaussianized *t*-image (*à la* Friston *et al.*, 1991) at the highest resolution and check that the theory presented here is a reasonable approximation. This is a suitable topic for future work.

6.2 fMRI data

This latter suggestion could be the best approach for fMRI data, where the voxel standard deviation is markedly unstable. Our recommendation here is to divide the unsmoothed data by the voxel standard deviation, then spatially smooth to the lowest reasonable FWHM, such as 4-6mm, making sure the FWHM is sufficiently large relative to the voxel size; greater than 5 voxels seems adequate. The operation of smoothing will improve the Gaussian approximation to the *t*-distribution without the need to explicitly Gaussianize the *t*-statistics, by the Central Limit Theorem. The smoothed images should now have equal voxel standard deviations, and so they can be treated in exactly the same way as PET data, but allowing for time correlations between images (Friston *et al.*, 1994a,1995; Worsley *et al.* 1995c). The fact that the fMRI voxels are almost uncorrelated suggests that the spatial correlation function of the smoothed data might be well approximated by a Gaussian function, as in Forman *et al.* (1995). In this case the pooled standard deviation can be found by the quick method of Section 3.1 without computing the spatial correlation function.

6.3 Region size tests

Friston *et al.* (1994b) and Forman *et al.* (1995) have proposed a test for activation based on region size above a fixed threshold, rather than peak height. It might at first seem that this achieves the same goal as scale space searches, but in a much simpler way. This is not strictly true. First of all, the region size is measuring the extent above a *fixed* threshold; the extent estimated by scale space searches is the extent above a *variable* threshold set at half the maximum height. The former will increase or decrease as the signal strength varies; the latter remains fixed.

Secondly, low but extensive signals might fall below a high threshold and never be detected by a region size test, but a scale space search might detect such a signal at a large FWHM. Conversely, sharp foci would give small region sizes above a low threshold and also escape detection, whereas a scale space search might detect it at a small FWHM. In other words, the *sensitivity* of the region size test depends on the level of the threshold. Friston *et al.* (1994b) show more formally that the sensitivity of the region size test is best with a low threshold for extensive signals, and with a high threshold for narrow signals. Thus the *threshold* plays the same role for region size tests as the *FWHM* does for peak height tests. In this paper we have proposed searching over FWHM; the analogous concept for region size tests would be to search over thresholds. This could be a suitable topic for future work.

6.4 Conclusion

We have presented a method for scale space searches of PET data from a typical activation study, and a rigorous method for correcting P-values to allow for this. We have demonstrated the advantages to be gained from a scale space search: an increase in sensitivity at detecting activation when the extent is unknown; and an estimate of the unknown extent. There is a small price to pay for this: a loss in sensitivity when the extent is known and the FWHM is chosen to match it; and blurring of small foci that are too close together, which can be detected by examining scale space slices through the local maxima to look for evidence of 'bifurcations'. These small costs seem to be offset by real gains in the practical application to real PET data sets. For the application chosen here, a scale space search picked up the same activations as a fixed 20mm FWHM search, but in addition it found some small foci that were previously overlooked. In general, the benefits appear to outweigh the drawbacks.

We have also gone to some trouble to find efficient ways of smoothing the large amount of data in a typical PET experiment. Reductions in calculations come from: smoothing only the high resolution averaged ΔCBF image, not all the data; estimating the pooled standard deviation from the correlation function (the most costly item to compute); only performing a small number of smoothes on the log(FWHM) scale to give adequate, uniform coverage in scale space; compressing the smoothes by increasing the voxel size (see Appendix A.1). This makes scale space searches both feasible and attractive as an alternative to other methods.

A Appendix

A.1 Image compression

Although not implemented in this study due to software incompatibilities, considerable savings in time and storage space can be achieved by compressing the images as they are smoothed. Without compression, direct kernel smoothing, as opposed to Fourier transform methods of smoothing, takes longer as the kernel becomes wider. Fortunately the equal sampling on a log(FWHM) scale combined with the convolution property (2.2) of the Gaussian kernel can be used to actually reduce the time taken for direct kernel smoothing with wider kernels, by compressing the data as it is smoothed.

We first note that as the FWHM increases, the voxel size can be proportionately increased

to maintain the same FWHM per voxel. Thus the size of the image in voxels can be decreased in proportion to 1/FWHM, while still maintaining the same resolution per voxel. The width of the kernel, in voxels, needed to smooth from one fixel to the next is $v\sqrt{b^2 - 1}$, where v is the FWHM in voxels at the first fixel, and b is the ratio of FWHM's between the two fixels. What is noteworthy about this is that b and hence v remain constant when scale space is sampled uniformly on the log(FWHM) scale. This implies that the FWHM of the smoothing kernel (in voxels) also remains constant.

Thus the whole smoothing operation can be done sequentially by performing the same operation of simultaneously smoothing and increasing the voxel size (or compressing the image) on successive smoothes, starting with the highest resolution data. Because the images are shrinking (in voxels) at each step the operations go faster. Moreover no matter how many smoothes are performed, the total amount of storage is limited by the sum of a geometric progression with ratio $1/b^D$. For the example in Section 5, this amounts to a maximum of $1/(1 - 1/b^D) \approx 3$ times as much storage as required for one image.

A.2 Smooths in 2-D and 1-D

If the whole process of smoothing and searching is restricted to 2-D slice data then the EC densities are

$$
\rho_0(t) = \frac{1}{(2\pi)^{1/2}} \left\{ \sqrt{\frac{1}{2\pi}} (-\log_e u) e^{-t^2/2} + \int_t^\infty e^{-z^2/2} dz \right\}
$$

$$
\rho_1(t) = \frac{(4\log_e 2)^{1/2}}{(2\pi)} \left\{ \sqrt{\frac{1}{2\pi}} (1-u)t + \frac{1+u}{2} \right\} e^{-t^2/2}
$$

$$
\rho_2(t) = \frac{(4\log_e 2)}{(2\pi)^{3/2}} \left\{ \sqrt{\frac{1}{2\pi}} \frac{1-u^2}{2} t^2 + \frac{1+u^2}{2} t \right\} e^{-t^2/2}. \tag{A.1}
$$

Note that the above EC densities are not the same as (4.2); the reason is that for (A.1) the smoothing filter is restricted to the slice, whereas for (4.2) the smoothing filter filters data from outside the slice. For completeness, we give the EC densities for smoothing and searching over 1-D data:

$$
\rho_0(t) = \frac{1}{(2\pi)^{1/2}} \left\{ \sqrt{\frac{1}{4\pi}} (-\log_e u) e^{-t^2/2} + \int_t^\infty e^{-z^2/2} dz \right\}
$$

$$
\rho_1(t) = \frac{(4\log_e 2)^{1/2}}{(2\pi)} \left\{ \sqrt{\frac{1}{4\pi}} (1-u)t + \frac{1+u}{2} \right\} e^{-t^2/2}. \tag{A.2}
$$

Acknowledgement

References

Antoine MJ, Travère JM, Bloyet D (1994): *Proceedings of the IEEE Medical Imaging Conference, Norfolk, Virginia.*

Coghill, RC, Talbot JD, Evans AC, Meyer E, Gjedde A, Bushnell MC, Duncan GH (1994): Distributed processing of pain and vibration by the human brain. *Journal of Neuroscience* **14**:4095-4108.

Collins DL, Dai W, Peters TM, Evans AC (1992): Model-based segmentation of individual brain structures from MRI data. *Visualization in Biomedical Computing 1992* **Proc SPIE 1808**:10-23.

Collins DL, Peters TM, Evans AC (1994): An automated 3D non-linear deformation procedure for determination of gross morphometric variability in human brain. *Visualization in Biomedical Computing 1994* **Proc SPIE 2359**:180-190.

Collins DL, Holmes C, Peters TM, Evans AC (1995): Automatic 3D segmentation of neuroanatomical structures from MRI *Human Brain Mapping*, in press.

Evans AC, Marrett S, Neelin P, Collins L, Worsley KJ, Dai W, Milot S, Meyer E, Bub, D (1992): Anatomical mapping of functional activation in stereotactic coordinate space. *NeuroImage* **1**:43-53.

Evans AC, Marrett S, Peters TM (1989): Anatomical-functional correlative analysis of the human brain using three-dimensional imaging systems. *Proceedings of the International Society of Optical Engineering (SPIE): Medical Imaging III*, 264-274.

Evans AC, Marrett S, Torrescorzo J, Ku S, Collins L. (1991a): MRI-PET correlative analysis using a volume of interest (VOI) atlas. *Journal of Cerebral Blood Flow and Metabolism* **11(2)**:A69-A78.

Evans AC, Thompson, C.J., Marrett, S., Meyer, E. and Mazza, M. (1991b). Performance characteristics of the PC-2048: A new 15-slice encoded-crystal PET scanner for neurological studies. *IEEE Transactions on Medical Imaging*,**10(1)**: 90-98.

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995): Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magnetic Resonance in Medicine* **33**:636-647.

Fox PT, Perlmutter JS, Raichle ME (1985): A stereotactic method of anatomical localization for positron emission tomography. *Journal of Computer Assisted Tomography* **9(1)**:141-153.

Fox PT, Raichle ME (1984): Stimulus rate dependence of regional cerebral blood flow in human striate cortex, demonstrated with positron emission tomography. *Journal of Neurophysiology* **51**:1109-1121.

Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ (1991): Comparing functional (PET) images: the assessment of significant change. *Journal of Cerebral Blood Flow and Metabolism* **11**:690-699.

Friston KJ, Frith CD, Liddle PF, Lammertsma AA, Dolan RD, Frackowiak RSJ, (1990): The relationship between local and global changes in PET scans. *Journal of Cerebral Blood Flow and Metabolism* **10**:458-466.

Friston KJ, Holmes AP, Poline J-B, Grasby BJ, Williams CR, Frackowiak RSJ, Turner R. (1995): Analysis of fMRI time-series revisited. *NeuroImage* **2**:45-53.

Friston KJ, Jezzard P, Turner R (1994a): The analysis of functional MRI time series. *Human Brain Mapping* **1**:153-171.

Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta JC, Evans AC (1994b): Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping* **1**:214-220.

Herscovitch P, Markham J, Raichle ME (1983): Brain blood flow measured with intravenous $H_2^{15}O$ I. Theory and error analysis. *Journal of Nuclear Medicine* **24**:782-789.

Poline J-B, Mazoyer BM (1994a): Enhanced detection in activation maps using a multifiltering approach. *Journal of Cerebral Blood Flow and Metabolism* **14**:690-699.

Poline J-B, Mazoyer BM (1994b): Analysis of individual brain activation maps using hierarchical description and multiscale detection. IEEE Transactions on Medical Imaging **13(4)**:702-710.

Poline J-B, Worsley KJ, Holmes AP, Frackowiak RSJ, Friston KJ (1995): Estimating smoothness in statistical parametric maps: variability of P-values. *Journal of Computer Assisted Tomography*, submitted.

Raichle ME, Martin WRW, Herscovitch P, Mintun MA, Markham J (1983): Brain blood flow measured with intravenous $H_2^{15}O$ II. Implementation and validation. *Journal of Nuclear Medicine* **24**:790-798.

Rosenfeld A, Kak AC (1982): *Digital Picture Processing, Volume 2.* Orlando FA: Academic Press.

Siegmund DO, Worsley KJ (1995): Testing for a signal with unknown location and scale in a stationary Gaussian random field. *Annals of Statistics*, in press.

Talairach P, Tournoux J (1988): *A Stereotactic Coplanar Atlas of the Human Brain.* Stuttgart: Thieme.

ter Haar Romeny BM, Florak LMJ, Koenderink JJ, Viergever MA (1991): Scale space: its natural operators and differential invariants. *In: Colchester ACF, Hawkes DJ (eds): Information Processing in Medical Imaging*, Berlin: Springer-Verlag, pp. 239-255.

Worsley KJ, Evans AC, Marrett S, Neelin P. (1992): A three dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism* **12**:900-918.

Worsley KJ, Friston KJ (1995c): Analysis of fMRI time-series revisited - again. *NeuroImage*, submitted.

Worsley KJ, Marrett S, Neelin P, Friston KJ, Evans, AC (1995a): A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, submitted.

Worsley KJ, Poline J-B, Vandal AC, Friston KJ (1995b): Tests for distributed, non-focal brain activations. *NeuroImage*, submitted.

TABLE 1: SIGNIFICANT PEAKS IN THE PAIN STUDY

Region	Scale space searches location			scale			Fixed scale (original study)	
	x	y	z	w	M	P	M	P
Increases:								
R SII/anterior insula	35	-2	10	22.7	6.56	0.0000	5.65	0.0006
R SMA (Inferior)	4	-2	54	22.7	6.22	0.0001	5.00	0.014
R Thalamus	11	-19	-3	17.4	5.73	0.0012	5.36	0.0025
L Anterior insula	-33	13	7	6.8	4.86	0.0824	3.72	1.96
R SI	24	-28	57	19.9	4.77	0.12	4.02	0.72
L Putamen	-25	5	0	19.9	4.76	0.12	4.57	0.089
R Caudate	-12	-2	15	6.8	4.73	0.14	-	-
Decreases:								
R Posterior cingulate	5	-59	25	15.2	4.86	0.082	5.14	0.0072
M Posterior cingulate	1	-42	30	11.6	4.54	0.31	-	-

R = right, L = left, M = midline; stereotaxic coordinates are derived from Tailarach and Tournoux (1988), and refer to medial-lateral position x (mm) relative to mid-line (positive = right), anterior-posterior position y (mm) relative to the anterior commisure (positive = anterior), and superior-inferior position z (mm) relative to the commisural line (positive = superior). w is the FWHM (mm) of the filter. M is the peak height of local maxima, and P is its P-value; those that are greater than 0.2 should be interpreted as the expected number of false positive regions of activation. The last two columns are for the equivalent peak in the original study, found at a fixed 20mm FWHM scale in x and y, and no axial smoothing. A dash indicates that no matched peak could be found.

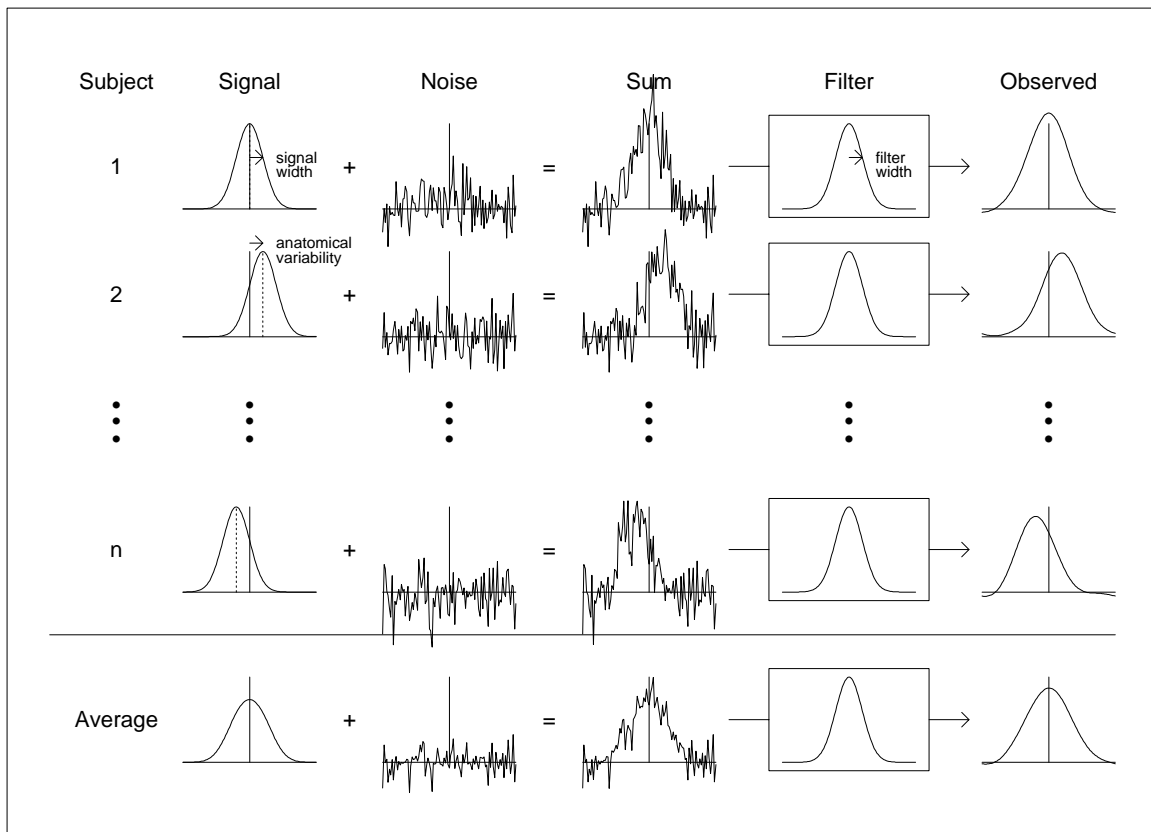Figure 1. Model for the activation data.

Figure 1: Model for the activation data. A Gaussian signal is added to white noise, then filtered by the point response function of the PET camera and further smoothed to improve signal to noise ratio (here shown as one combined operation). Anatomical variability perturbs the resulting individual images, which are then averaged over subjects to produce the final ΔCBF image.
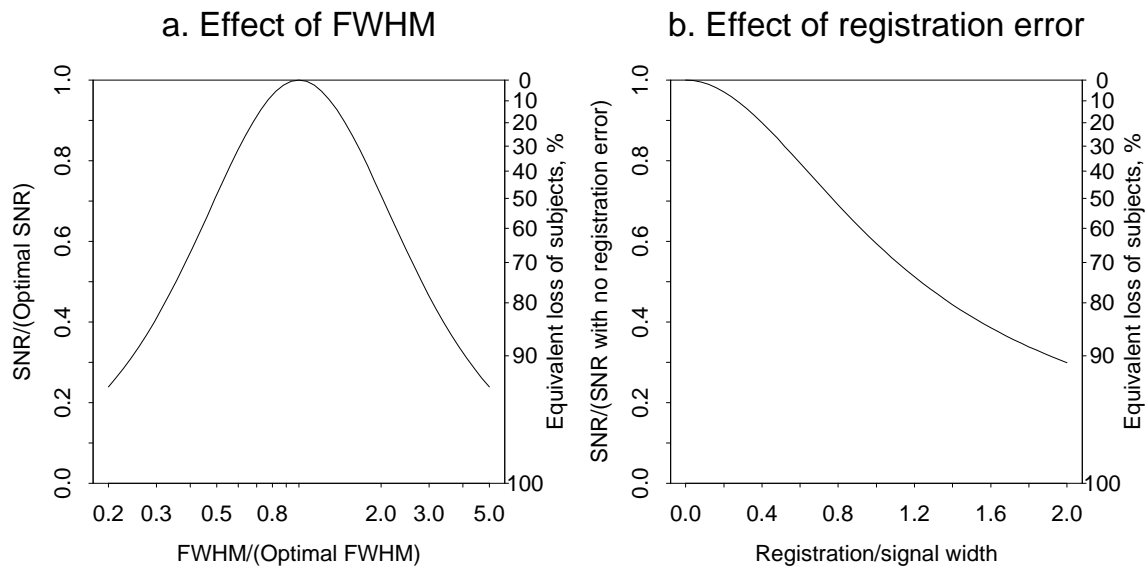
Figure 2. Signal to noise ratio (SNR)

a. Effect of FWHM

b. Effect of registration error

Figure 2: The effect of filter width (a) and anatomical variability (b) on signal to noise ratio (SNR).
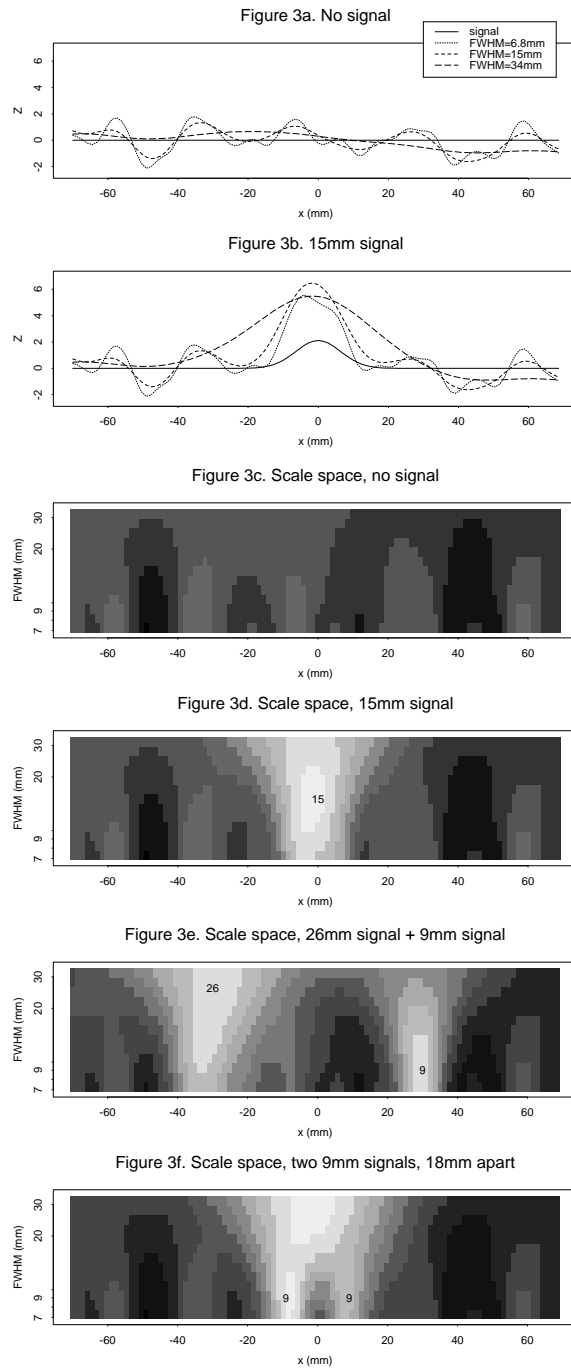
Figure 3: (a) 1-D White noise, plus (b) a 15mm signal smoothed by a 6.8, 15 and 34mm filter. All statistical fields are adjusted to have unit standard deviation. Note that the 15mm filter best identifies the 15mm signal; (c) and (d) are the same as (a) and (b) but shown as a 2-D location and scale statistical field with the centre of the 15mm signal marked; (e) the same noise component as in (a), but with a 26mm and a 9mm signal far apart, and (f) two 9mm signals 18mm apart. Note how the two well separated signals are correctly identified, but the two close signals are incorrectly identified as one wide signal.
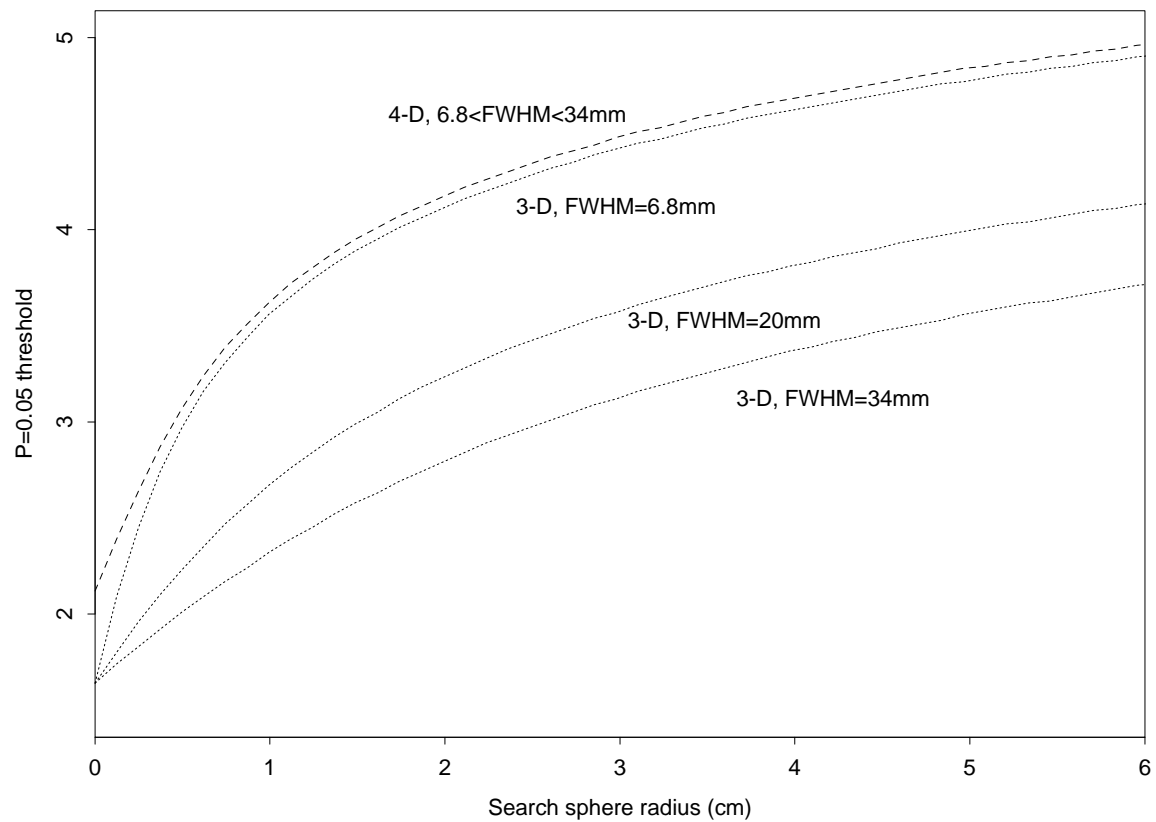
Figure 4: Critical thresholds of the maximum over a spherical search region with 6.8mm < FWHM < 34mm. Also shown are the critical thresholds when the FWHM is fixed at 6.8, 20 and 34mm FWHM.
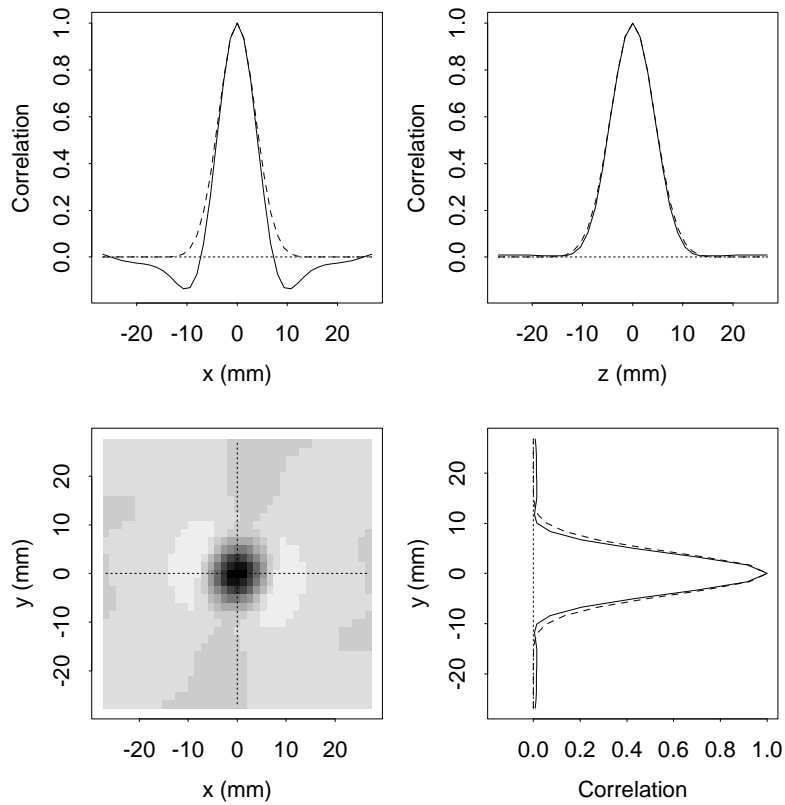
Figure 5. Correlation function

Figure 5: Correlation function for the null data set in the x, z, xy and y directions, respectively (solid lines). Also shown are the Gaussian correlation functions chosen to have the same effective FWHM (dotted lines). The fit appears to be reasonable in the y and z directions but not in the x direction, where large negative side lobes are evident.
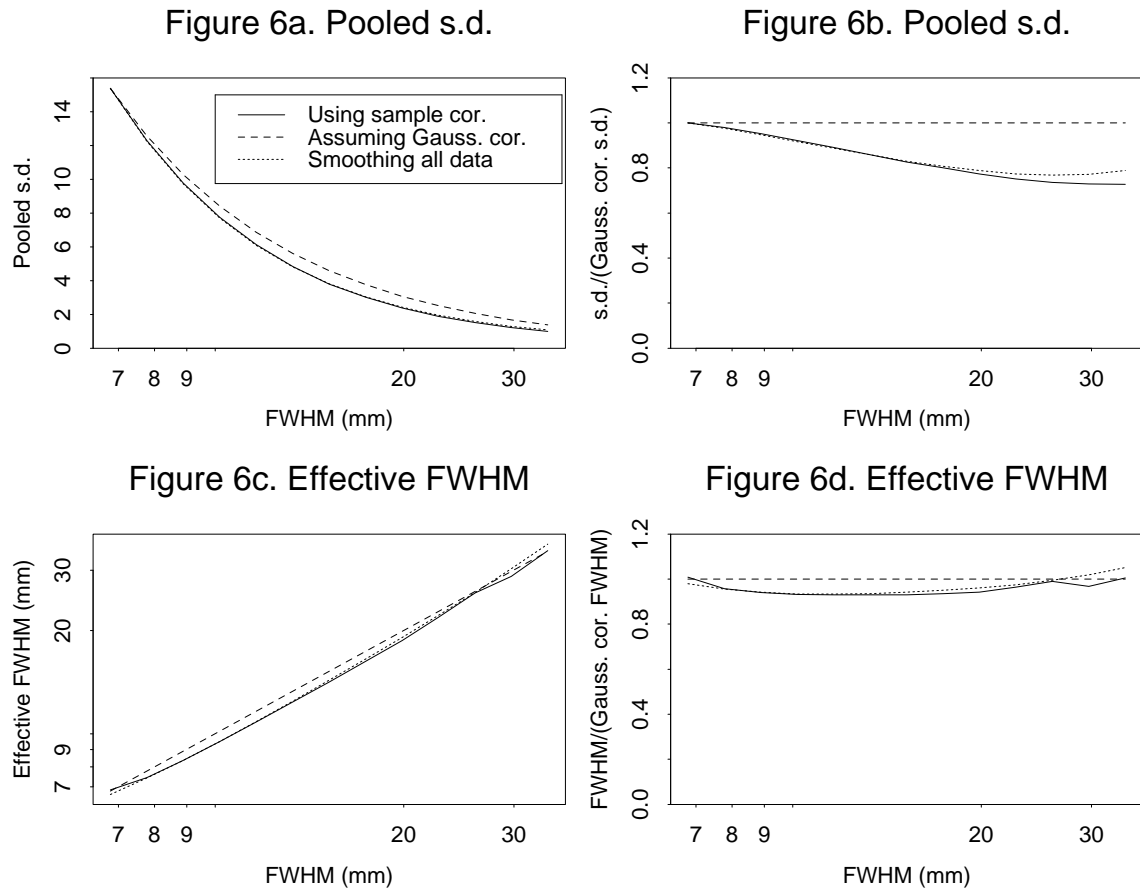
Figure 6: Validation of the pooled standard deviation (a,b) and effective FWHM (c,d) calculated by smoothing all the data (dashed line); assuming a Gaussian correlation function (dotted line); and using the sample correlation function of Figure 5 (solid line). The Gaussian assumption overestimates the pooled standard deviation, but the sample correlation estimate is quite accurate. The same applies to the effective FWHM, but the Gaussian assumption is reasonably accurate in this case.
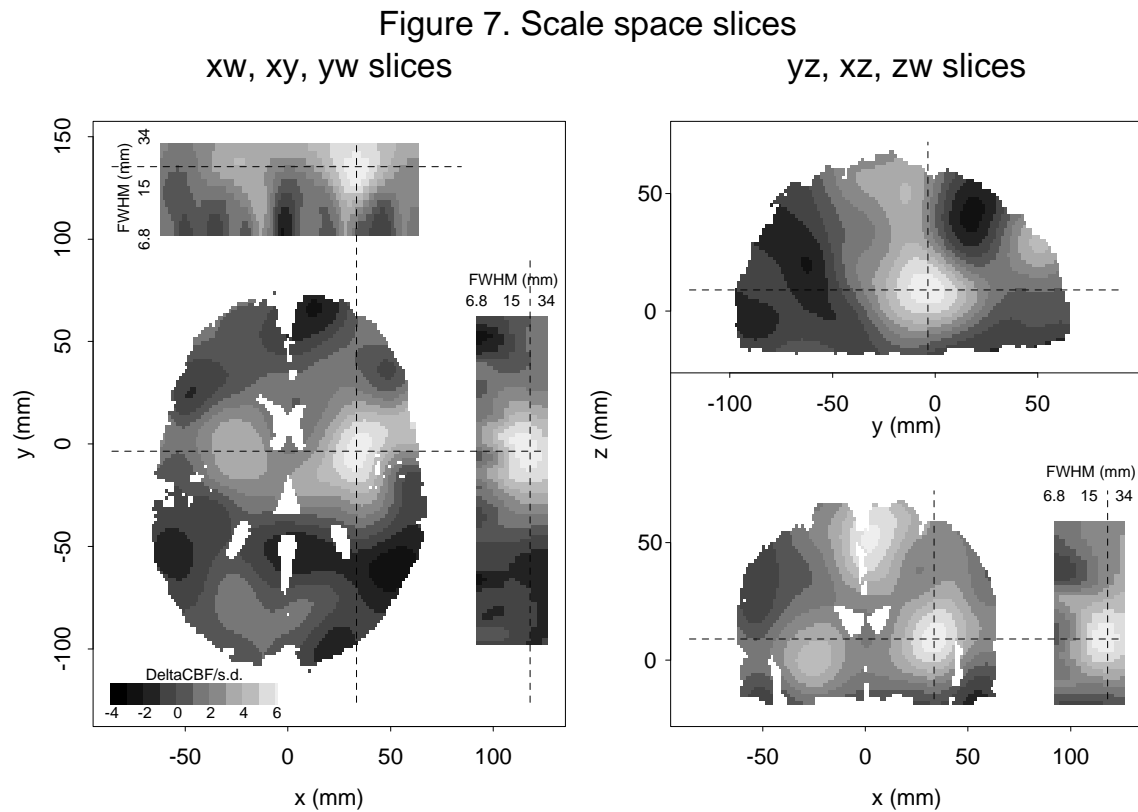
Figure 7. Scale space slices

Figure 7: Scale space slices of the pain data through the 4-D maximum at $x = 35$, $y = -2$, $z = 10$, $w = 22.7$mm. Lines through the maximum are marked on each slice. Images are masked by thresholding the average MRI image. The bifurcation in the yw slice suggests that the peak splits into two smaller foci, separated by 2cm, at a lower scale.