Improved Assessment of Significant Activation in Functional Magnetic Resonance Imaging (fMRI): Use of a Cluster-Size Threshold

Steven D. Forman, Jonathan D. Cohen, Mark Fitzgerald, William F. Eddy, Mark A. Mintun, Douglas C. Noll

The typical functional magnetic resonance (fMRI) study presents a formidable problem of multiple statistical comparisons (i.e., >10,000 in a 128 x 128 image). To protect against false positives, investigators have typically relied on decreasing the per pixel false positive probability. This approach incurs an inevitable loss of power to detect statistically significant activity. An alternative approach, which relies on the assumption that areas of true neural activity will tend to stimulate signal changes over contiguous pixels, is presented. If one knows the probability distribution of such cluster sizes as a function of per pixel false positive probability, one can use cluster-size thresholds independently to reject false positives. Both Monte Carlo simulations and fMRI studies of human subjects have been used to verify that this approach can improve statistical power by as much as fivefold over techniques that rely solely on adjusting per pixel false positive probabilities.

Key words: spatial extent; significance; NMR; spatial correlation.

INTRODUCTION

Reliable identification of areas of brain activation is the sine qua non of functional neuroimaging. In functional magnetic resonance imaging (fMRI), we assume that condition-dependent changes in neural activity may be inferred from changes in image intensity (signal) at each pixel sampled across experimental conditions. Generally, exploratory analysis of changes in functional neural activity is performed using statistical parametric maps (SPMs). As described by Friston and colleagues (1), an SPM is a two- (or higher-) dimensional image of a test statistic determined at each pixel by the value of the signal and its variance across experimental conditions. As long as the distribution function of the test statistic is known, one can determine the false positive probability (or alpha probability) for each pixel in the SPM. A variety of approaches have been suggested for setting the appropriate alpha-level threshold with which to assess statistical significance of change across experimental conditions (2–4). While the optimum method for determining threshold has not been established, two points are incontestible: 1) because any SPM consists of a large number of pixels, some adjustment for multiple comparisons must be made in the analysis; 2) we want to choose a method that maximizes statistical power to detect events of interest.

Most previous approaches to the problem of multiple comparisons have attempted to protect against false positives (Type I errors) by adjusting (decreasing) the alpha level of the test statistic by methods analogous to Bonferroni corrections (e.g., see refs. 2 and 3). (In a study containing N multiple comparisons, a Bonferroni correction establishes that the overall probability of any false positive is ≤P, by setting the individual comparison false positive probability to P/N.) Such approaches inevitably incur a substantial loss of statistical power. We present an alternative approach, which relies on the assumption that, in general, interesting areas of neural activity will tend to be larger than individual pixel dimensions (in fMRI) and, as such, will tend to stimulate signal changes over contiguous groups of pixels. That is, in an SPM image containing a true signal, not only will individual pixels more likely exceed a particular statistical threshold, but such pixels will also tend to form clusters. The probability is low that a given number of pixels exceeding threshold due to chance will be contiguous, but it is higher if they are activated by a true source of signal. It should be possible to capitalize on this difference, by incorporating contiguity as a factor in our statistical analysis. To do so, we must first: 1) determine the probability with which clusters of various sizes occur by chance, and 2) determine the likelihood of detecting such clusters when a signal is actually present. With this information, we can then pick a cluster-size threshold that will help distinguish between pixels detected by chance and those activated by a true source of signal. More precisely, if one knows the probability distribution of cluster size as a function of alpha level, then one can use a cluster-size threshold, R, (of contiguous pixels) to establish a Type I error level. (In this paper contiguous is defined as any of the eight immediate neighbors to the index pixel.) We will show that an appropriate combination of R and alpha level can be chosen to provide equivalent Type I protection to a Bonferroni-type adjustment without the equivalent loss of power.

An alternative “approximate analysis” of this problem has been presented (5), but the approximations used...
appear largely inapplicable to fMRI data sets (see Discussion). A major goal of this paper was to extend cluster threshold analysis to fMRI. We first present Monte Carlo estimates of false positive probability distributions for pixel clusters as a function of cluster size and individual pixel alpha level. Then, we proceed to power calculations for detecting clusters that arise in response to a true source of signal, which are shown to depend on the sample size, effect size, alpha level, and the size and topographic configuration of the underlying signal source. The cluster size probability distributions (and the power calculations) change if the pixels in the SPM are spatially correlated. Spatial correlations in fMRI-derived SPMs may arise from a variety of sources including, but not limited to, filtering during image reconstruction, nonuniform sampling during acquisition, or physiological effects. Such sources of spatial correlation can be modeled as resulting from the application of Gaussian filters of various widths to the SPM. Using this Gaussian correlation model, we have simulated cluster-size probability distributions for a range of spatial correlations seen in fMRI experiments. We then present experimental data (from an fMRI study of frontal cortex activation in a working memory task) establishing that an appropriate combination of cluster-size threshold and alpha can detect more pixels as significantly active than an equivalent Bonferroni procedure. We also provide corroborative evidence that the additional pixels identified localize to areas of increased blood flow. This is shown in blood flow maps obtained in a PET study of one subject who repeated the working memory task. Finally, we present tables allowing retrospective and prospective power analysis to be performed for a variety of experimental conditions.

THEORY AND METHODS

For a given sample, we wish to maximize the probability of detecting true change and minimize the probability of detecting false positives due to noise or multiple comparisons. Generally, we convert images of signal intensity into images of a test statistic, dependent on the condition-related change and the variance at each pixel. A threshold value, $T(a)$, may be applied to the value of the test statistic at each pixel. For a given probability distribution of the test statistic, $T(a)$ determines the probability of false positive detection per pixel, $a$. Because any image consists of a large number of pixels, the probability of any false positives per image can become quite large, unless some adjustment is made for multiple comparisons. Most frequently, Bonferroni adjustments to $T(a)$ are made, although this has been called overly conservative [2]. Note that the event of interest is the individual pixel and $T(a)$ is adjusted to detect changes in individual pixels.

Consider, instead, the event of interest to be the detection of a contiguous cluster of $R$ (or more) pixels, all with values above $T(a)$. The probability of accepting a false positive pixel now depends on two separate thresholds, $T(a)$ and $R$. Regardless of $T(a)$, the probability of detecting false positive pixels decreases as $R$ increases. This is because the likelihood of two or more pixels exceeding threshold and being contiguous is lower than their simple probability of exceeding threshold. As such, we can adjust the expectation of false positives (per pixel, per image or per experiment) by modifying $R$. By definition, this precludes identifying regions of neural activity whose area is less than $R$. However, this is compensated by the increased statistical power obtained by decreasing $T(a)$. In essence, we trade off ability to detect very small areas of activity for increased power to detect larger areas. If activity tends to extend over several pixels, then this tradeoff can work in our favor.

Determinant of Detection Probabilities As a Function of Cluster Size

Much of this investigation consists of generating probability distributions for detecting “active” pixels as a function of cluster size. Generally, there are three steps in producing such probability distributions: 1) generate an appropriate image process (SPM) over the desired image space; 2) repeatedly sample and tabulate the frequency versus size distribution of active clusters over that space; 3) convert the frequency distribution into a probability distribution (by dividing the area under the curve from $R$ to $\infty$ by the total number of pixels sampled). Unless otherwise noted, probability distributions are expressed in the following units: probability of detecting an individual pixel (not a cluster) per pixel. These units are used to facilitate comparison with Bonferroni-type procedures which adjust the false positive probability per pixel. In the first experimental simulation section, we explicitly outline the conversion from cluster frequencies to pixel frequencies; subsequently the conversion is assumed.

Note that “active” as used above refers merely to that pixel’s value exceeding some arbitrary statistical threshold, $T(a)$. Whether that pixel is considered a false positive or a true positive depends on the context. In all cases of spatially correlated SPMs, the pixel values in the SPM are generated according to explicitly defined noise and signal distributions, and are filtered using a Gaussian filter function (as a model of the correlational structure) and then the pixel values are compared with $T(a)$. According to standard statistical convention, we define $\alpha$ as the one-tailed probability of a false positive event under a noise distribution. In this work $\alpha$ applies to the values at individual pixels. Conversely, the probability of detecting a true positive event at an individual pixel whose value is determined by a signal distribution, is defined as $1 - \beta$, where $\beta$ is the probability of Type II (false negative) error. All random numbers were generated using the ran2 routine from Press et al. [6], modified as needed to produce random normal deviates.

Uncorrelated Case (Noise). In this section we wish to determine the probability of false positive clusters of various sizes that occur by chance over a spatially uncorrelated SPM. A $128 \times 128$ image was generated using a uniform random number generator (0, 1) set to define a (false) positive pixel with probability, $\alpha$ (one-tailed). Ten thousand simulations were performed for each 0.005 $\leq \alpha \leq 0.2$, in 0.005 increments. The frequency versus size distribution of false positive clusters was tabulated at each $\alpha$. This was converted to the frequency distribution
of false positive pixels by multiplying frequency \( \times \) cluster size. We estimate false positive probability per pixel by numerically integrating the frequency distribution from \( R \) to \( \infty \), and dividing by the total sample of pixels (10,000 SPMs/estimate \( \times \) 128 \( \times \) 128) (see Table 1). The accuracy of the estimates in Tables 1 and 2 (see below) is reflected in the number of significant digits reported. The standard error for each estimate is less than one unit of the least significant digit for that estimate. Note that we also truncated trailing zeros for legibility. This latter convention can lead to underestimating a particular estimate's accuracy (e.g., in rounding 0.1999 to 0.2). However, any ambiguities can be resolved by examining the number of significant figures of the nearest neighboring estimates in the tables.

**Uncorrelated Case (Power).** We now determine the likelihood of detecting true positive clusters of various sizes over a series of \( n \times m \) regions of task-related neural activity. The \( n \times m \) configurations were chosen to represent a variety of possible topographic configurations of underlying signal. Again, we assume that there is no underlying spatial correlation among pixels. Using a cluster-size threshold, power to detect true positive pixels in clusters consists of two components: the probability of detecting individual true positive pixels, \( 1 - \beta \), and the conditional probability of detecting these pixels in clusters \( \geq R \) within the \( n \times m \) region. The former probability, \( 1 - \beta \), is a function of \( \alpha \) and of the separation (in standard z scores) between the means of the signal and the noise distributions. The separation between the signal and noise distributions is determined by the number of samples per experimental condition, \( N \), and by the effect size, \( \delta \). \( \delta \) is the difference between the mean signal in each condition divided by the coefficient of variation of the signal. Power (given \( 1 - \beta \), \( R \), and \( n \times m \)) is the integral of the frequency distribution of true positive pixels from \( R \) to \( \infty \) divided by the total sample of pixels (10,000 SPMs/estimate \( \times \) \( n \times m \)). We produced frequency distributions in a manner exactly analogous to the previous simulation with \( 0.1 \leq 1 - \beta \leq 0.9 \) replacing \( \alpha \) and \( n \times m \) areas from \( 3 \times 3 \) to \( 10 \times 10 \) replacing the 128 \( \times \) 128 matrix. To further explore the effect of asymmetric spatial configurations on power, we also generated the probability distributions for the cases of 1 \( \times \) 36 and 2 \( \times \) 18 pixels. Calculation of \( 1 - \beta \) (7, Eq. 12.2.1, pg. 544) for various \( \alpha \) was based on 100 samples/condition and \( \delta = 0.3 \). This was based on our previous work (8), where we found that the percentage mean difference between the activation and control conditions was 1.5-2.0% and the coefficient of variation of the signal in the functional scans is approximately 5%. Therefore, the \( \delta \) range is 0.3-0.4.

We plotted power (Fig. 1a) for a number of pairs of \( R \) and \( \alpha \), whose probability of false positives was \( \approx 0.0006 \). This latter value is a standard in the PET literature as providing appropriate Type I protection (9). We also determined relative power (i.e., the ratio of the above power divided by the power obtained by a Bonferroni adjustment to \( T_{0.0006} \)).

**Spatially Correlated Case (Noise).** The calculations above assume that the values of each pixel in the SPM are spatially uncorrelated (e.g., statistically independent events). In PET, this assumption is clearly invalid because significant degrees of data filtering are performed prior to obtaining the SPM image. With fMRI, although we suspect the degree of spatial correlation is much less than in PET, some spatial correlation can arise from a variety of causes. Any such spatial correlation will tend to shift the probability distributions for pixel detection toward larger clusters. Therefore, we now determine the detection probability distributions under correlation conditions typical of our fMRI studies. Using the general approach of Friston et al. (2), we modeled a Gaussian correlational structure in our SPMs by generating a 256 \( \times \) 256 image process of uncorrelated Gaussian deviates (mean = 0, standard deviation = 1) convolved with a 2D Gaussian filter of width 2 s (pixels). Formally, the convolution process should include weighted contributions from every point in the SPM at every point in the SPM. In practice, the decay of the Gaussian filter function with distance is such that points more than two or three filter widths from the index position contribute negligibly to the postfiltered value. Thus, for computational expediency, we truncated the size of the filter weight matrix for the convolution to 4 s \( \times \) 4 s. However, when this led to fractional values, we rounded up to the next larger odd integer.

To decrease the effect of pixelation, prior to further processing, the convolved 256 \( \times \) 256 image process is reduced to 128 \( \times \) 128 by taking the mean value of each four-pixel block. (This returns the effective filter width for the 128 \( \times \) 128 SPM back to s pixels, whose full-width, half maximum (FWHM) equivalent is approximately 2.35 \( \times \) s.) We then scale the 128 \( \times \) 128 image into standard normal deviates and apply z-score thresholds to determine whether a given pixel is considered active. The z-score thresholds are chosen so that detection frequency corresponds to \( \alpha \). Subsequent probability distribution calculations proceed exactly as previously outlined above. Eight alpha levels were simulated at 13 filter width values, 0.3 \( \leq \) s \( \leq \) 1.5, using 1000 SPMs/simulation. Detection probabilities for s = 0.4, 0.5, 0.6, 0.8, 1.0 are shown in Table 2.

**Spatially Correlated Case (Power).** In the uncorrelated case one could ignore the influence of surrounding pixels on the detection probabilities for the active \( n \times m \) region. Because of the spatial correlations and because the \( n \times m \) regions are fairly small, we must account for the edge effect of surrounding regions (at the noise level) on the probability of detection over an active region. To simulate this situation, we entered a 2n \( \times \) 2m active region consisting of uncorrelated Gaussian deviates (mean = \( \bar{z} \), \( \sigma = 1 \)) in the center of a 256 \( \times \) 256 matrix of uncorrelated Gaussian deviates (mean = 0, standard deviation = 1). Convolution and fourfold area reduction were performed as above over the entire 256 \( \times \) 256 matrix. Subsequent normalization to standard scale and tabulation of frequency and probability distributions were performed only over the \( n \times m \) region. Other calculations were as previously described (except to compensate for the smaller areas sampled, 5000 SPMs were averaged for each point in Fig. 1b).
Retrospective/Prospective Power Analyses

For the convenience of investigators desiring to utilize these methods, we have provided some additional information. While calculation of \( 1 - \beta \) (and power) requires knowledge of \( \delta \), samples per condition and \( \alpha \), the separation between signal and noise distributions is fully specified by \( \delta \) and samples per condition. For example, our experimental conditions (\( \delta = 0.3 \), samples per condition = 100) correspond to a separation between the means of the signal and noise distributions of 2.11 standard deviation units. An identical separation would be achieved with only 50 samples per condition if the effect size, \( \delta \), were increased to 0.427. In both cases the \( 1 - \beta \) (and subsequent power determination) for a given \( \alpha \) would be identical. In Fig. 5, we plot a series of \( \delta \) and samples per condition pairings that produce equivalent mean separations between signal and noise distributions. Given the signal-to-noise separation (discriminability), Table 3 provides power determinations for a series of cluster-size threshold and alpha combinations. The combinations in the table were chosen to produce expectations of false positive clusters (size \( \geq 5 \), per 128 x 128 SPM) in a 20% range around the values of 1.0, 0.1, and 0.05, respectively.

Estimation of the Spatial Correlation (Filter Width) in fMRI

As stated, we are modeling all spatial correlational structure of the SPM as if a Gaussian filter of width, \( s \), had been applied to form the observed SPM. Friston et al. (2) derived the following equation for estimating \( s \) (in pixel units) on an SPM:

\[
s = \frac{1}{\sqrt{2 \times \sigma^2}}
\]

where \( \sigma^2 \) is the variance of the difference between each pixel and its edgewise neighbors over the entire SPM. This equation assumes that pixel size is small relative to \( s \). In their PET study, the measured \( s \) was slightly greater than three pixel widths, so the assumption was upheld. It is clear, however, that the above equation becomes increasingly inaccurate as \( s < 1 \) pixel width. For example, in the limit of no filtering, true \( s = 0 \) but the equation produces an estimated \( s = 0.5 \). Our fMRI SPMs appeared to show a filter width of less than one pixel. To improve the accuracy of estimation of filter widths <1, we derived the following formula (see Appendix for derivation) where \( S^2 \) is the individual pixel variance across the SPM:

\[
s = \frac{1}{\sqrt{4 \times \ln \left( \frac{1 - \frac{S^2}{2S^2}}{2S^2} \right)}}
\]

Using simulated filtered SPMs (as above) at each of various filter widths, \( 0.3 \leq s \leq 1.5 \), we calculated \( S^2 \) and \( S^2 \) for each SPM and the mean value (8000 simulations/ filter width) was used to estimate \( s \) by both methods (see Fig. 2).

Activation Paradigm

To test whether these methods would improve detection of functional neural activity in actual subjects, we replicated a previously reported nonspatial working memory design (8). Both control and memory conditions consist of random sequences of letters appearing one at a time in the center of a visual display. In the control condition, subjects respond whenever the letter "H" appears (by pressing a button on a hand-held response box). In the memory condition, subjects respond only when a letter is repeated with exactly one nonidentical letter intervening (e.g., A-F-A, but not A-A or A-Q-G-A). All human subjects were volunteers and all studies were performed according to an approved experimental protocol (University of Pittsburgh IRB).

Functional Neuroimaging

Scanning was performed on a series of subjects in a standard 1.5T GE Signa whole-body magnet, using two 5-inch surface coils mounted parasagittally. One hundred functional images per condition (7.0 mm thick, 1.9 x 1.9 mm in-plane) were acquired at six to eight contiguous locations in an oblique coronal plane through the frontal cortex, using a spiral scan pulse sequence (10) with 10 interleaves. TR 600 ms, TE 35 ms, flip 45°, FOV 24 cm. In all studies a bite bar was used to control head movement. High resolution T1-weighted scans were used for intermodality alignment of the fMRI and the PET studies (11). PET scans were obtained from one subject who repeated the same tasks in a Siemens ECAT 951r/31 (Knoxville, TN). Reconstructed images (128 x 128) had an in-plane resolution of 5 mm FWHM. H215O was administered by intravenous bolus and counts were obtained during the first minute of brain uptake, during which the subject performed the experimental task in an alternating basis (x3). Full details of the PET to fMRI comparison will be presented elsewhere.

Data Analysis

All fMRI images were scaled to the same global intensity prior to generating SPMs. To verify that the chance cluster-size probability distributions (i.e., those due to noise) in the experimentally obtained SPMs correspond to those in the simulated SPMs, we plotted (Fig. 3) the cumulative probability of detecting clusters \( \geq 5 \) from 16 experimental SPMs in two subjects (\( \alpha = 0.1 \)). This was done on t-maps generated using random pairings within experimental conditions (i.e., activation - activation and control - control), to eliminate any potential effect of the experimental manipulation on our estimates. We then used Eq. [2] to estimate the degree of the spatial correlation (as indicated by measured filter width) affecting these SPMs. To these we overlaid the theoretical cumulative probability function generated from the Monte Carlo simulations at the appropriate filter widths.

In contrast, SPM t-maps for the experimental manipulation (Fig. 4) were obtained by random pairing across the experimental conditions (i.e., activation - control). Note that although we are using SPM t-estimates here, the cluster-size method generalizes to other statistics.

In the PET study (Fig. 4), because there were too few samples to reliably generate a t-map, we used the average of the mean difference in cerebral blood flow between condition pairs. The PET difference scores were then
given percentile rankings and contour maps were generated with each contour representing a 10 percentile range.

RESULTS
Type 1 Levels
Table 1 shows the per-pixel probabilities of detecting false positive pixels as a function of a level and cluster size in the absence of filtering. Table 2 shows similar probability distributions modified by filter width. It can be seen that increasing smoothing tends to shift the distributions to the right (i.e., toward larger clusters).

Power and Relative Power
Figure 1 shows the Power as a function of cluster-size threshold for a series of R, a combinations chosen for equivalent Type 1 protection. Across nearly all combinations the relative power advantage of the cluster-size threshold over a Bonferroni adjustment is substantial, exceeding fivefold in some cases. In Fig. 1b the same calculations are shown with a filter width = 0.6 pixels. The qualitative picture is the same as in the uncorrelated case in Fig. 1a. As expected, the power advantage relative to a Bonferroni adjustment has decreased. In most cases the relative power advantage is still substantial; however, this advantage appears lost in the limiting case (1 × 36).

Smoothing Estimates
Figure 2 plots the filter width estimated from simulated data using Eq. [1] and our new method Eq. [2]. The fit with Eq. [2] is clearly superior.

Validation of Smoothing Estimates
Figure 3 shows a plot of distribution of cluster sizes at alpha level = 0.1 using randomly paired, within-condition SPMs. It can be seen that these values closely correspond to the Monte-Carlo derived estimates obtained at that smoothing level.

Experimental Verification of Relative Power Advantage
Figure 4 shows a series of structural images upon which we have overlaid the significantly active pixels identified by the cluster-size threshold and the equivalent Bonferroni procedure. There are (4.6-fold) more pixels identified by the cluster-size threshold method. This was in close accord with theoretical power estimates indicating an expected advantage of about 4.3, assuming an effect size of 0.3, 100 samples per condition.

Validation of the Identified Pixels
Figure 4 also shows an aligned series of contour maps of cerebral blood flow changes obtained with PET. Areas where the percentile change in cerebral blood flow are the highest are shown in white and light gray. Black and dark gray indicate areas of decreased blood flow. Intermediate gray indicates approximately no change. The additional pixels identified as active by the cluster-size threshold method predominantly localize to areas of increased blood-flow change in PET.

DISCUSSION
Recognition that functional neural activity has both a spatial as well as an effect size component allows one to enhance detection sensitivity. We use this improvement to mitigate the multiple comparisons problem inherent in exploratory fMRI studies. By taking account of both spatial extent and effect size, we can now reliably reject false positives without incurring unacceptable losses of power due to decreased alpha levels. Our data demonstrate that use of cluster-size thresholds provides sub-

Table 1
Probability of False Positive Pixel Per Pixel As a Function of Cluster Size Threshold: Spatially Uncorrelated Pixels

<table>
<thead>
<tr>
<th>Cluster Size</th>
<th>0.005</th>
<th>0.01</th>
<th>0.02</th>
<th>0.025</th>
<th>0.03</th>
<th>0.05</th>
<th>0.1</th>
<th>0.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.005</td>
<td>0.01</td>
<td>0.02</td>
<td>0.025</td>
<td>0.03</td>
<td>0.05</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>2</td>
<td>0.00019</td>
<td>0.0008</td>
<td>0.0029</td>
<td>0.0045</td>
<td>0.0064</td>
<td>0.017</td>
<td>0.056</td>
<td>0.165</td>
</tr>
<tr>
<td>3</td>
<td>0.00006</td>
<td>0.000064</td>
<td>0.00041</td>
<td>0.0008</td>
<td>0.0013</td>
<td>0.0053</td>
<td>0.031</td>
<td>0.137</td>
</tr>
<tr>
<td>4</td>
<td>0.00001</td>
<td>0.000014</td>
<td>0.00004</td>
<td>0.00014</td>
<td>0.00027</td>
<td>0.0017</td>
<td>0.017</td>
<td>0.113</td>
</tr>
<tr>
<td>5</td>
<td>0.00001</td>
<td>0.00002</td>
<td>0.00005</td>
<td>0.00005</td>
<td>0.0005</td>
<td>0.0094</td>
<td>0.094</td>
<td>0.078</td>
</tr>
<tr>
<td>6</td>
<td>0.000001</td>
<td>0.000004</td>
<td>0.00004</td>
<td>0.00004</td>
<td>0.00017</td>
<td>0.00051</td>
<td>0.078</td>
<td>0.065</td>
</tr>
<tr>
<td>7</td>
<td>0.00002</td>
<td>0.00015</td>
<td>0.00034</td>
<td>0.00034</td>
<td>0.00051</td>
<td>0.0028</td>
<td>0.065</td>
<td>0.054</td>
</tr>
<tr>
<td>8</td>
<td>0.00001</td>
<td>0.00008</td>
<td>0.00045</td>
<td>0.00045</td>
<td>0.00045</td>
<td>0.00051</td>
<td>0.045</td>
<td>0.038</td>
</tr>
<tr>
<td>9</td>
<td>0.000002</td>
<td>0.000046</td>
<td>0.000046</td>
<td>0.000046</td>
<td>0.000046</td>
<td>0.000051</td>
<td>0.038</td>
<td>0.032</td>
</tr>
<tr>
<td>10</td>
<td>0.00026</td>
<td>0.00014</td>
<td>0.00027</td>
<td>0.00027</td>
<td>0.00027</td>
<td>0.00027</td>
<td>0.00027</td>
<td>0.00027</td>
</tr>
<tr>
<td>11</td>
<td>0.00008</td>
<td>0.00008</td>
<td>0.00008</td>
<td>0.00008</td>
<td>0.00008</td>
<td>0.00008</td>
<td>0.00008</td>
<td>0.00008</td>
</tr>
<tr>
<td>12</td>
<td>0.00004</td>
<td>0.00004</td>
<td>0.00004</td>
<td>0.00004</td>
<td>0.00004</td>
<td>0.00004</td>
<td>0.00004</td>
<td>0.00004</td>
</tr>
<tr>
<td>13</td>
<td>0.00002</td>
<td>0.00002</td>
<td>0.00002</td>
<td>0.00002</td>
<td>0.00002</td>
<td>0.00002</td>
<td>0.00002</td>
<td>0.00002</td>
</tr>
<tr>
<td>14</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
</tr>
<tr>
<td>15</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
</tr>
<tr>
<td>20</td>
<td>0.0012</td>
<td>0.0012</td>
<td>0.0012</td>
<td>0.0012</td>
<td>0.0012</td>
<td>0.0012</td>
<td>0.0012</td>
<td>0.0012</td>
</tr>
<tr>
<td>30</td>
<td>0.00022</td>
<td>0.00022</td>
<td>0.00022</td>
<td>0.00022</td>
<td>0.00022</td>
<td>0.00022</td>
<td>0.00022</td>
<td>0.00022</td>
</tr>
<tr>
<td>40</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
</tr>
<tr>
<td>60</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00001</td>
</tr>
</tbody>
</table>
Significance Assessment in fMRI By Cluster Size

### Table 2
Probability of False Positive Pixel Per Pixel As a Function of Cluster Size Threshold: Spatially Correlated Pixels

<table>
<thead>
<tr>
<th>Cluster Size</th>
<th>Alpha</th>
<th>0.005</th>
<th>0.01</th>
<th>0.02</th>
<th>0.025</th>
<th>0.03</th>
<th>0.05</th>
<th>0.1</th>
<th>0.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Filter width = 0.4 pixels</td>
<td></td>
<td>0.0016</td>
<td>0.0041</td>
<td>0.0014</td>
<td>0.0047</td>
<td>0.0069</td>
<td>0.0094</td>
<td>0.022</td>
<td>0.064</td>
</tr>
<tr>
<td>1</td>
<td>0.0004</td>
<td>0.0009</td>
<td>0.0022</td>
<td>0.0034</td>
<td>0.005</td>
<td>0.014</td>
<td>0.05</td>
<td>0.155</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.0011</td>
<td>0.0016</td>
<td>0.001</td>
<td>0.0016</td>
<td>0.0025</td>
<td>0.0085</td>
<td>0.039</td>
<td>0.141</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.0003</td>
<td>0.0005</td>
<td>0.0044</td>
<td>0.0008</td>
<td>0.0013</td>
<td>0.0053</td>
<td>0.03</td>
<td>0.128</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0019</td>
<td>0.0038</td>
<td>0.0007</td>
<td>0.0032</td>
<td>0.023</td>
<td>0.117</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.00001</td>
<td>0.00008</td>
<td>0.00019</td>
<td>0.00003</td>
<td>0.002</td>
<td>0.018</td>
<td>0.107</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.00004</td>
<td>0.00003</td>
<td>0.00008</td>
<td>0.00016</td>
<td>0.0012</td>
<td>0.014</td>
<td>0.097</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.00001</td>
<td>0.00002</td>
<td>0.00003</td>
<td>0.00008</td>
<td>0.0003</td>
<td>0.011</td>
<td>0.089</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.00001</td>
<td>0.00002</td>
<td>0.00004</td>
<td>0.00002</td>
<td>0.003</td>
<td>0.065</td>
<td>0.074</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.00003</td>
<td>0.000001</td>
<td>0.00001</td>
<td>0.00019</td>
<td>0.0048</td>
<td>0.068</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.00001</td>
<td>0.00002</td>
<td>0.00007</td>
<td>0.00012</td>
<td>0.0037</td>
<td>0.062</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.00006</td>
<td>0.00001</td>
<td>0.00007</td>
<td>0.00002</td>
<td>0.009</td>
<td>0.036</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.00006</td>
<td>0.00015</td>
<td>0.00001</td>
<td>0.00002</td>
<td>0.015</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.000017</td>
<td>0.000017</td>
<td>0.00003</td>
<td>0.00006</td>
<td>0.016</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00003</td>
<td>0.000004</td>
<td>0.000001</td>
<td>0.012</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B. Filter width = 0.5 pixels

<table>
<thead>
<tr>
<th>Cluster Size</th>
<th>Alpha</th>
<th>0.005</th>
<th>0.01</th>
<th>0.02</th>
<th>0.025</th>
<th>0.03</th>
<th>0.05</th>
<th>0.1</th>
<th>0.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0005</td>
<td>0.01</td>
<td>0.02</td>
<td>0.025</td>
<td>0.03</td>
<td>0.05</td>
<td>0.1</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.0021</td>
<td>0.0051</td>
<td>0.012</td>
<td>0.016</td>
<td>0.02</td>
<td>0.038</td>
<td>0.086</td>
<td>0.189</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.0007</td>
<td>0.0022</td>
<td>0.0066</td>
<td>0.0094</td>
<td>0.012</td>
<td>0.027</td>
<td>0.071</td>
<td>0.176</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.00025</td>
<td>0.001</td>
<td>0.0035</td>
<td>0.0053</td>
<td>0.0075</td>
<td>0.019</td>
<td>0.059</td>
<td>0.164</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.0007</td>
<td>0.00037</td>
<td>0.0018</td>
<td>0.0029</td>
<td>0.0043</td>
<td>0.012</td>
<td>0.047</td>
<td>0.151</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.0002</td>
<td>0.0015</td>
<td>0.0009</td>
<td>0.0016</td>
<td>0.0025</td>
<td>0.0083</td>
<td>0.038</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.00001</td>
<td>0.00006</td>
<td>0.00045</td>
<td>0.0008</td>
<td>0.0014</td>
<td>0.0056</td>
<td>0.031</td>
<td>0.129</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.00003</td>
<td>0.00002</td>
<td>0.00023</td>
<td>0.00046</td>
<td>0.0008</td>
<td>0.0038</td>
<td>0.025</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.00012</td>
<td>0.00024</td>
<td>0.00046</td>
<td>0.0025</td>
<td>0.02</td>
<td>0.111</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.00006</td>
<td>0.00014</td>
<td>0.00025</td>
<td>0.00017</td>
<td>0.016</td>
<td>0.102</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.00008</td>
<td>0.00008</td>
<td>0.00008</td>
<td>0.00013</td>
<td>0.011</td>
<td>0.095</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.00002</td>
<td>0.00004</td>
<td>0.00008</td>
<td>0.00008</td>
<td>0.01</td>
<td>0.088</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.00001</td>
<td>0.00003</td>
<td>0.00004</td>
<td>0.00005</td>
<td>0.0083</td>
<td>0.081</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.000004</td>
<td>0.00001</td>
<td>0.00002</td>
<td>0.000032</td>
<td>0.0066</td>
<td>0.075</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 2—continued

**Probability of False Positive Pixel Per Pixel As a Function of Cluster Size Threshold: Spatially Correlated Pixels**

<table>
<thead>
<tr>
<th>Cluster Size</th>
<th>0.005</th>
<th>0.01</th>
<th>0.02</th>
<th>0.025</th>
<th>0.03</th>
<th>0.05</th>
<th>0.1</th>
<th>0.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>0.00001</td>
<td>0.00002</td>
<td>0.00021</td>
<td>0.0053</td>
<td>0.069</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.00003</td>
<td>0.018</td>
<td>0.047</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.00021</td>
<td>0.022</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>0.00003</td>
<td>0.011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.000003</td>
<td>0.0049</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>0.0023</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>0.0005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.00014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>0.00002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**D. Filter width = 0.8 pixels**

<table>
<thead>
<tr>
<th>1</th>
<th>0.005</th>
<th>0.01</th>
<th>0.02</th>
<th>0.025</th>
<th>0.03</th>
<th>0.05</th>
<th>0.1</th>
<th>0.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.0034</td>
<td>0.0073</td>
<td>0.016</td>
<td>0.021</td>
<td>0.025</td>
<td>0.044</td>
<td>0.083</td>
<td>0.195</td>
</tr>
<tr>
<td>3</td>
<td>0.0019</td>
<td>0.0047</td>
<td>0.012</td>
<td>0.016</td>
<td>0.019</td>
<td>0.037</td>
<td>0.085</td>
<td>0.188</td>
</tr>
<tr>
<td>4</td>
<td>0.01</td>
<td>0.0029</td>
<td>0.0082</td>
<td>0.011</td>
<td>0.015</td>
<td>0.03</td>
<td>0.076</td>
<td>0.181</td>
</tr>
<tr>
<td>5</td>
<td>0.00047</td>
<td>0.0016</td>
<td>0.0054</td>
<td>0.0079</td>
<td>0.011</td>
<td>0.024</td>
<td>0.067</td>
<td>0.172</td>
</tr>
<tr>
<td>6</td>
<td>0.00024</td>
<td>0.0009</td>
<td>0.0036</td>
<td>0.0055</td>
<td>0.0076</td>
<td>0.019</td>
<td>0.059</td>
<td>0.164</td>
</tr>
<tr>
<td>7</td>
<td>0.00013</td>
<td>0.0005</td>
<td>0.0024</td>
<td>0.0038</td>
<td>0.0055</td>
<td>0.015</td>
<td>0.052</td>
<td>0.156</td>
</tr>
<tr>
<td>8</td>
<td>0.00006</td>
<td>0.0003</td>
<td>0.0015</td>
<td>0.0027</td>
<td>0.0039</td>
<td>0.012</td>
<td>0.045</td>
<td>0.148</td>
</tr>
<tr>
<td>9</td>
<td>0.00002</td>
<td>0.00017</td>
<td>0.001</td>
<td>0.0018</td>
<td>0.0028</td>
<td>0.0091</td>
<td>0.04</td>
<td>0.141</td>
</tr>
<tr>
<td>10</td>
<td>0.00001</td>
<td>0.00011</td>
<td>0.0007</td>
<td>0.0012</td>
<td>0.002</td>
<td>0.0071</td>
<td>0.035</td>
<td>0.134</td>
</tr>
<tr>
<td>11</td>
<td>0.000001</td>
<td>0.000006</td>
<td>0.00043</td>
<td>0.0008</td>
<td>0.0014</td>
<td>0.0056</td>
<td>0.03</td>
<td>0.128</td>
</tr>
<tr>
<td>12</td>
<td>0.0000004</td>
<td>0.000004</td>
<td>0.00029</td>
<td>0.0006</td>
<td>0.001</td>
<td>0.0043</td>
<td>0.026</td>
<td>0.121</td>
</tr>
<tr>
<td>13</td>
<td>0.0000002</td>
<td>0.000002</td>
<td>0.00018</td>
<td>0.00039</td>
<td>0.0007</td>
<td>0.0033</td>
<td>0.023</td>
<td>0.115</td>
</tr>
<tr>
<td>14</td>
<td>0.000001</td>
<td>0.000012</td>
<td>0.00028</td>
<td>0.0005</td>
<td>0.0026</td>
<td>0.02</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.000001</td>
<td>0.000008</td>
<td>0.0002</td>
<td>0.00036</td>
<td>0.002</td>
<td>0.018</td>
<td>0.104</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.0000003</td>
<td>0.000001</td>
<td>0.00004</td>
<td>0.00007</td>
<td>0.0006</td>
<td>0.0089</td>
<td>0.081</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.0000002</td>
<td>0.000006</td>
<td>0.00023</td>
<td>0.00063</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>0.0000003</td>
<td>0.00006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.000017</td>
<td>0.018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>0.000003</td>
<td>0.012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>0.000001</td>
<td>0.0046</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.0019</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>140</td>
<td>0.00029</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>180</td>
<td>0.00006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**E. Filter width = 1.0 pixels**

<table>
<thead>
<tr>
<th>1</th>
<th>0.005</th>
<th>0.01</th>
<th>0.02</th>
<th>0.025</th>
<th>0.03</th>
<th>0.05</th>
<th>0.1</th>
<th>0.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.0042</td>
<td>0.0085</td>
<td>0.018</td>
<td>0.023</td>
<td>0.027</td>
<td>0.047</td>
<td>0.097</td>
<td>0.197</td>
</tr>
<tr>
<td>3</td>
<td>0.003</td>
<td>0.0066</td>
<td>0.015</td>
<td>0.019</td>
<td>0.024</td>
<td>0.043</td>
<td>0.092</td>
<td>0.193</td>
</tr>
<tr>
<td>4</td>
<td>0.0021</td>
<td>0.0051</td>
<td>0.012</td>
<td>0.016</td>
<td>0.02</td>
<td>0.038</td>
<td>0.087</td>
<td>0.188</td>
</tr>
<tr>
<td>5</td>
<td>0.0013</td>
<td>0.0035</td>
<td>0.0095</td>
<td>0.013</td>
<td>0.017</td>
<td>0.033</td>
<td>0.08</td>
<td>0.183</td>
</tr>
<tr>
<td>6</td>
<td>0.0008</td>
<td>0.0025</td>
<td>0.0074</td>
<td>0.01</td>
<td>0.014</td>
<td>0.029</td>
<td>0.074</td>
<td>0.177</td>
</tr>
<tr>
<td>7</td>
<td>0.0005</td>
<td>0.0018</td>
<td>0.0056</td>
<td>0.0082</td>
<td>0.011</td>
<td>0.025</td>
<td>0.068</td>
<td>0.172</td>
</tr>
<tr>
<td>8</td>
<td>0.00033</td>
<td>0.0012</td>
<td>0.0043</td>
<td>0.0064</td>
<td>0.0088</td>
<td>0.021</td>
<td>0.062</td>
<td>0.166</td>
</tr>
<tr>
<td>9</td>
<td>0.00021</td>
<td>0.0008</td>
<td>0.0033</td>
<td>0.005</td>
<td>0.0071</td>
<td>0.018</td>
<td>0.057</td>
<td>0.16</td>
</tr>
<tr>
<td>10</td>
<td>0.00013</td>
<td>0.0005</td>
<td>0.0025</td>
<td>0.0039</td>
<td>0.0057</td>
<td>0.015</td>
<td>0.052</td>
<td>0.155</td>
</tr>
<tr>
<td>11</td>
<td>0.00008</td>
<td>0.00037</td>
<td>0.0018</td>
<td>0.0031</td>
<td>0.0046</td>
<td>0.013</td>
<td>0.048</td>
<td>0.15</td>
</tr>
<tr>
<td>12</td>
<td>0.00004</td>
<td>0.00025</td>
<td>0.0014</td>
<td>0.0024</td>
<td>0.0036</td>
<td>0.011</td>
<td>0.044</td>
<td>0.145</td>
</tr>
<tr>
<td>13</td>
<td>0.00002</td>
<td>0.00019</td>
<td>0.001</td>
<td>0.0019</td>
<td>0.0028</td>
<td>0.0092</td>
<td>0.04</td>
<td>0.14</td>
</tr>
<tr>
<td>14</td>
<td>0.00001</td>
<td>0.00013</td>
<td>0.0008</td>
<td>0.0015</td>
<td>0.0022</td>
<td>0.0076</td>
<td>0.037</td>
<td>0.135</td>
</tr>
<tr>
<td>15</td>
<td>0.00001</td>
<td>0.00009</td>
<td>0.0006</td>
<td>0.0011</td>
<td>0.0018</td>
<td>0.0066</td>
<td>0.033</td>
<td>0.13</td>
</tr>
<tr>
<td>20</td>
<td>0.000001</td>
<td>0.000015</td>
<td>0.00034</td>
<td>0.0006</td>
<td>0.0029</td>
<td>0.021</td>
<td>0.109</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.000001</td>
<td>0.000003</td>
<td>0.0006</td>
<td>0.0008</td>
<td>0.0083</td>
<td>0.077</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>0.00001</td>
<td>0.000011</td>
<td>0.0032</td>
<td>0.054</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.00002</td>
<td>0.014</td>
<td>0.038</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>0.000004</td>
<td>0.0005</td>
<td>0.027</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>0.000006</td>
<td>0.014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.00001</td>
<td>0.0071</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0.00024</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.00002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
FIG. 1. Power as a function of cluster-size threshold and configuration of the underlying region of neural activity. Configurations from $1 \times 36$ to $6 \times 6$ pixels are shown. The alpha level - cluster-size threshold combinations were all chosen to produce approximately the same false positive probability/pixel (i.e., 0.0006). Power was calculated assuming 100 samples/condition and an effect size of 0.3. Relative power is with respect to an equivalent Bonferroni-adjusted alpha. (a) Spatially uncorrelated SPM ($0.05 < \alpha < 0.175$, labels over points). (b) Spatially correlated SPM with Gaussian filter width = 0.6 pixels ($0.005 < \alpha < 0.1$, labels over points).

Substantial enhancements in power over a wide range of parameters.

The choice of $R$, $\alpha$ combination depends on the size and topographic configuration of underlying neural signal. The larger the active area, the larger the $R$, $\alpha$ combination which can be safely used and the greater gain in power. However, this advantage is offset by the larger $R$-value, $\alpha$ level combination's inability to detect smaller areas of activation. As an empirical compromise, we tend to favor those values around the peak of the $4 \times 4$ curve as shown in Fig. 1b (i.e., $7 \leq R \leq 9, 0.02 \leq \alpha \leq 0.03$). We feel that such intermediate values capture much of the power benefits available, while also retaining the ability to detect fairly small active areas.

The data shown here (as well as preliminary investigations of other subjects and pulse sequences, not shown) appears to validate our hypothesis that the degree of spatial correlation in fMRI images is significantly less than that seen in typical PET studies. By developing

![Graph](image)

FIG. 2. Methods for estimating filter width from variance of partial derivatives of the SPM. Results using the Eq. [1] (x) (Friston, 1991). Results using Eq. [2] (+). Points appear thickened because there are actually eight replications plotted at each filter width. Points plotted at filter width $= 0$ assume the limiting value for the variance of the partial derivative (see text) and are plotted within square figures to emphasize this difference in calculation method.

![Graph](image)

FIG. 3. Comparison of cluster-size distributions from simulated SPMs with distributions from experimentally obtained SPMs. Cumulative proportion of clusters of a given size or less were calculated from the SPMs. SPMs were simulated at four filter widths (line plots). Eight SPMs per subject were obtained in two subjects and the mean and standard deviation of the cumulative proportions are shown. Using the modified model, the estimated filter width for Subject 1 (■) was 0.54; similarly the estimated filter width for Subject 2 (□) was 0.58. The individual pixel alpha level was 0.1 for all the determinations.

An improved expression for estimating filter width, a Gaussian model can now be applied to fMRI data. An unanticipated benefit of the new formula is that it reduces the computational burden of estimating filter width. Equation [1] requires that an SPM be converted into a standard z-scale prior to calculating variance of the interpixel differences. Equation [2] may be applied to a “raw” SPM.

On first glance, spatial correlations appear to decrease the benefits of the cluster-size threshold methods be-
Forman et al.

FIG. 4. Comparison of activity detection in a working memory task. Active pixels represent areas more active in the memory compared with the control task and are shown in white overlaid on the appropriate structural image. (a) Active pixels obtained by thresholding the fMRI-generated SPM t-maps so that the false positive expectation per pixel was 0.0006. (b) Active pixels obtained on the same t-maps using an alpha/cluster-size combination that produced an equivalent false positive probability per pixel (i.e., one-tailed alpha level of 0.03 and a size threshold of seven pixels at an estimated filter width of 0.5). (c) Maps of CBF mean differences measured using PET. White is the largest positive mean difference; black is the largest negative mean difference. Maps are of the same slice locations as in the fMRI study in the same subject performing exactly the same task.

cause filtering causes increased clustering of pixels. While this effect must be considered and compensated for if one is to provide a true measure of Type 1 error, one must also consider that this increased clustering effect will occur with actual signal as well as noise and that this will tend to partially compensate for the relative power loss. Thus, spatially correlated SPMs will require somewhat larger cluster-size thresholds to obtain the same Type 1 error level. This will produce some real power loss. However, in most cases the relative power ratio still substantially favors the cluster threshold method over an equivalent Bonferroni adjustment. For most of the points in Fig. 1b, relative power is greater than one. Although this was not true for the 1 x 36 configuration, this configuration is a worst case. Even the 2 x 18 configuration produces a substantial relative power advantage in the presence of spatial correlations (Fig. 1b). In any event, several sources of artifact, such as motion and the signal produced by large vessels, often assume such linear patterns. Therefore, reduced sensitivity to such configurations may be a fortunate side-effect of this technique. On the other hand, we emphasize that this technique does not distinguish "true" signal from systematic, artificial sources of signal. Independent methods (such as magnetic resonance angiography to identify large vessels) should be used whenever available to screen out clusters which meet cluster threshold criteria but are due to artificial sources of signal.

Some readers may wonder at the range of signal to noise separations presented in Fig. 5 and Table 3. For the cluster threshold method to produce a useful benefit over Bonferroni-like corrections, two conditions must be met.

FIG. 5. Effect size and samples/condition determine the separation between the signal and the noise distributions. Lines show equi-separation contours for the differences between the means of both distributions in standard z-score units (a = 2.11; b = 2.46; c = 2.81; d = 3.17; e = 3.52). Data are calculated assuming that two experimental conditions are compared using an independent-sample t-test, with equal sample sizes and standard deviations for each condition. Signal and noise are assumed to be normally distributed with unit standard deviation.

First, the probability of false positive clusters must be fairly low, e.g., by setting individual pixel alpha to less than 0.1. Second, the probability of true positive clusters must be fairly high (i.e., individual true positive detection probability, 1 - β, greater than 0.5 or so). Both conditions can be met only if the signal and noise distri-
Significance Assessment in fMRI By Cluster Size

Table 3
Power As a Function of Signal-to-Noise Separation, Filter Width, and False Positive Rate (Assuming a 4 x 4 Underlying Area of Task-Dependent Activation)

<table>
<thead>
<tr>
<th>Cluster-Size Threshold</th>
<th>Alpha (per pixel)</th>
<th>Expected False Positive Clusters per SPM</th>
<th>Separation of Signal &amp; Noise Distributions in Standard Normal Units (Z)</th>
<th>Power As Function of Signal Discrimination</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.11</td>
<td>2.46</td>
</tr>
<tr>
<td>A. Filter width = 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.02</td>
<td>0.93</td>
<td>0.462</td>
<td>0.641</td>
</tr>
<tr>
<td>5</td>
<td>0.03</td>
<td>0.85</td>
<td>0.540</td>
<td>0.705</td>
</tr>
<tr>
<td>6</td>
<td>0.04</td>
<td>0.81</td>
<td>0.598</td>
<td>0.750</td>
</tr>
<tr>
<td>7</td>
<td>0.05</td>
<td>0.84</td>
<td>0.644</td>
<td>0.783</td>
</tr>
<tr>
<td>8</td>
<td>0.06</td>
<td>0.94</td>
<td>0.672</td>
<td>0.815</td>
</tr>
<tr>
<td>9</td>
<td>0.07</td>
<td>1.11</td>
<td>0.704</td>
<td>0.835</td>
</tr>
<tr>
<td>10</td>
<td>0.085</td>
<td>1.07</td>
<td>0.674</td>
<td>0.845</td>
</tr>
<tr>
<td>11</td>
<td>0.09</td>
<td>0.83</td>
<td>0.604</td>
<td>0.836</td>
</tr>
<tr>
<td>B. Filter width = 0.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.005</td>
<td>0.96</td>
<td>0.236</td>
<td>0.405</td>
</tr>
<tr>
<td>5</td>
<td>0.01</td>
<td>1.09</td>
<td>0.324</td>
<td>0.510</td>
</tr>
<tr>
<td>6</td>
<td>0.02</td>
<td>0.95</td>
<td>0.408</td>
<td>0.617</td>
</tr>
<tr>
<td>7</td>
<td>0.025</td>
<td>0.84</td>
<td>0.421</td>
<td>0.650</td>
</tr>
<tr>
<td>8</td>
<td>0.05</td>
<td>0.91</td>
<td>0.185</td>
<td>0.383</td>
</tr>
<tr>
<td>10</td>
<td>0.02</td>
<td>0.09</td>
<td>0.111</td>
<td>0.486</td>
</tr>
<tr>
<td>11</td>
<td>0.025</td>
<td>0.11</td>
<td>0.074</td>
<td>0.444</td>
</tr>
<tr>
<td>12</td>
<td>0.03</td>
<td>0.10</td>
<td>0.038</td>
<td>0.338</td>
</tr>
<tr>
<td>6</td>
<td>0.005</td>
<td>0.05</td>
<td>0.111</td>
<td>0.343</td>
</tr>
<tr>
<td>8</td>
<td>0.01</td>
<td>0.04</td>
<td>0.114</td>
<td>0.415</td>
</tr>
<tr>
<td>11</td>
<td>0.02</td>
<td>0.05</td>
<td>0.035</td>
<td>0.318</td>
</tr>
<tr>
<td>12</td>
<td>0.025</td>
<td>0.05</td>
<td>0.016</td>
<td>0.234</td>
</tr>
<tr>
<td>13</td>
<td>0.03</td>
<td>0.04</td>
<td>0.005</td>
<td>0.120</td>
</tr>
</tbody>
</table>

Asterisks indicate that the exact calculation could not be performed because the 1 - β level required for that combination of signal discrimination and alpha exceeded the limits of the simulation. However, based on the pattern of results, the power is estimated to be greater than 0.90 for * and greater than 0.95 for **.

Significance levels are fairly well separated. For example, if the separation of the signal distribution mean from the noise distribution mean is 2.0 z-score units (assuming standard deviation of 1.0 for both signal and noise distributions), then a one-tailed alpha of 0.025 allows detection of true positives slightly more than 50% of the time. We consider this to be close to the minimum usable signal to noise separation for this method. The experimental separation achieved in our memory studies (2.11 z-score units) is therefore just sufficient to benefit from these methods.

Friston and colleagues (5) in their "approximate analysis" of the cluster-size threshold method state that their "approximations hold reasonably well for thresholds as low as 2.4." Unfortunately, thresholds below 2.4 are precisely the range where cluster-size thresholding is most beneficial in terms of enhancing power. (In other words, if your individual pixel α level is 0.008 or less (threshold ≥2.4), it will be difficult achieving sufficient signal to noise separation to produce reasonable power.) In addition, they demonstrated their method using a phantom study with a FWHM of 7.2 pixels. This is a degree of Gaussian filtering typical of PET studies. They provide no data validating their method with SPMs filtered in the range we have seen in fMRI (less than 1/3 this value). For these reasons, we feel that the data and method presented here is preferable for fMRI investigations.

Table 3 is limited to considering only a 4 x 4 underlying matrix. Ideally, one would like to present results for a wide variety of underlying configurations of signal, false positive expectations, and filter widths. However, space and computational limitations prevented such a presentation (additional information is available upon request). The 4 x 4 matrix seems to us a reasonable
compromise value. Active areas much smaller than this probably can only be explored using individual pixel (and not cluster threshold) methods. In such a situation, an investigator is well-advised to consider Fig. 5 carefully to achieve as high a signal to noise separation as possible. Alternatively, if the expected region of activity exceeds 4 × 4 dimensions, then the values in Table 3 provide lower limit estimates of power. Power to detect will be at least that shown in Table 3 and in many cases considerably greater.

Finally, one may wonder at our choice of a Gaussian spread function, instead of a “sinc” (sin(kt)/kt) function. Note that the concept of ‘Gibb’s Ringing’ and the associated “sinc” point spread function applies to space limited (deterministic) objects for which the acquired Fourier domain data have been truncated. Noise processes, however, do not have any truncation effects, unless Fourier locations have been windowed to reduce truncation artifacts in deterministic objects or have been zero-filled to interpolate to a higher resolution. Provided that the noise samples in Fourier space are independent and uncorrelated, the spatial correlations would then be expected to be the Fourier transform of the window function. If no windowing is performed, we would expect the samples to be uncorrelated—not correlated by a “sinc” function.

For the spiral scan data, the acquisition pattern is nearly circular and further, the data is windowed by using a circularly symmetric apodization filter. The noise variance is not constant due to unequal sample density near the origin in Fourier space. Predicting the correlation by the Fourier relationship yields positive correlations with all edgewise neighbors and is not too unlike a Gaussian function in shape. It was our preference to keep the analysis as simple as possible by assuming a Gaussian spread function.

Thus, we emphasize that the Gaussian model of spatial cross-correlations is a first approximation to the true situation. There is a fair degree of agreement between expected and empirical cluster-size distributions (see Fig. 3) using this model, although some systematic deviations are apparent. Work is ongoing to design and incorporate more accurate correlational models for the spiral, as well as other pulse sequences. In the future, the model may, of necessity, be tailored specifically for a given pulse sequence. Such improvements promise even more progress towards optimizing detection of functional neural activity.

Appendix: Derivation of Gaussian Filter Width Estimator

Assume an infinite 2D image process consisting of an uncorrelated random variable, \( X_{i,j} \), with a mean \( \mu_{i,j} \) and variance \( \sigma^2_{i,j} \). Application of a bivariate Gaussian filter of width \( s \), to the above image process leads to the following measured value,

\[
X_{i,j}^s = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[ \frac{1}{2\pi s^2} \exp\left(-\frac{(x-i)^2 + (y-j)^2}{2s^2}\right) \right] X_{i,j} \, dx \, dy \tag{A1}
\]

Given that \( X_{i,j} \) is spatially uncorrelated and the image is thus isotropic, the variance of the paired difference, \( X_{i,j}^s - X_{i,j} \), can be estimated from the set of all edgewise pairs across an SPM. To estimate \( s \) from the variance of the edgewise paired differences, we must develop an expression for this term.

\[
X_{i,j}^s - X_{i,j} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[ \frac{1}{2\pi s^2} \exp\left(-\frac{(x-i+1)^2 + (y-j)^2}{2s^2}\right) \right. \\
- \left. \exp\left(-\frac{(x-i)^2 + (y-j)^2}{2s^2}\right) \right] X_{i,j} \, dx \, dy \tag{A2}
\]

Because the image is isotropic, \( i \to 0, j \to 0 \) without any loss of generality. Therefore,

\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[ \frac{1}{2\pi s^2} \exp\left(-\frac{(x+1)^2 + y^2}{2s^2}\right) - \exp\left(-\frac{x^2 + y^2}{2s^2}\right) \right] X_{i,j} \, dx \, dy \tag{A3}
\]

Noting that the variance of a random variable, \( X_i \), times a constant, \( A \), is \( A^2 \times \sigma^2_{i,j} \), we find after squaring and expanding Eq. [A3], that the variance of the paired differences, \( \sigma^2_{X_{i,j} - X_{i,j}} \),

\[
\frac{\sigma^2_{X_{i,j}}}{4\pi^2 s^4} \int_{-\infty}^{\infty} \left[ \frac{1}{2\pi s^2} \exp\left(-\frac{(x+1)^2 + y^2}{s^2}\right) \right. \\
- \left. 2\exp\left(-\frac{(x+1)^2 + y^2}{2s^2}\right) \exp\left(-\frac{x^2 + y^2}{2s^2}\right) + \exp\left(-\frac{x^2 + y^2}{s^2}\right) \right] \, dx \, dy \tag{A4}
\]

The solution for the first and third term is:

\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left(-\frac{(x+1)^2 + y^2}{s^2}\right) \, dx \, dy = \pi \tag{A5}
\]

The solution for the second term is:

\[
2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left(-\frac{(x+1)^2 + y^2}{2s^2}\right) \exp\left(-\frac{x^2 + y^2}{2s^2}\right) \, dx \, dy \tag{A6}
\]

Combining and rearranging in Eq.[A4] results in this expression for the variance:

\[
\sigma^2_{X_{i,j} - X_{i,j}} = \frac{\sigma^2_{X_{i,j}}}{2\pi s^2} \left[ 1 - \exp\left(-\frac{1}{4s^2}\right) \right] \tag{A7}
\]
However, from Eq. [A1] and Eq. [A5],

\[
\sigma^2_{x_{ij}} = \frac{1}{4\pi^2 s^4} \cdot \pi s^2 \cdot \sigma^2_{x_{ij}} = \frac{\sigma^2_{x_{ij}}}{4\pi^2 s^2} \tag{A8}
\]

Substituting for \(\sigma^2_{x_{ij}}\) in Eq. [7] and solving for \(s\) leads to the final expression.

\[
s = \sqrt{-\frac{1}{4\ln\left(1 - \frac{\sigma^2_{x_{ij}} - X_{ij}^2}{2\sigma^2_{x_{ij}}}\right)}} \tag{A9}
\]

ACKNOWLEDGMENTS

The authors thank the following individuals for helpful discussion and input: Rob Orr, Tom Nichols, Dan Degrandpre, Walt Schneider, Prahlad Gupta, Todd Braver, BJ Casey, Alan Koretsky, Afonso Silva and Weiguo Zhang; and Dr. Craig Meyer for the spiral pulse sequences used in our fMRI studies.

REFERENCES


