NEUROIMAGE 2, 183–194 (1995)

Tests for Distributed, Nonfocal Brain Activations

K. J. WORSLEY, J-B. POLINE,* A. C. VANDAL, AND K. J. FRISTON*

*Department of Mathematics and Statistics, McGill University, 805 Sherbrooke Street West, Montreal, Québec, Canada H3A 2K6; and *The Welcome Department of Cognitive Neurology, Queen Square, London WC1N 3BG, United Kingdom*

Received May 18, 1995

Most approaches to detecting changes in functional brain images assume that activations are focal or very localized. However, the brain's response to cognitive of sensorimotor challenge may be spatially or anatomically distributed. In this paper we consider the usefulness of a test based on the mean sum of squares of statistical parametric maps. The performance of this test is evaluated using simulated and real data and is compared to the γ_2 test, a test of the size of the activated region, and a focal activation test based on the intensity of local maxima. We demonstrate that the mean sum of squares test is more sensitive to nonfocal signals and propose that it could be used to complement approaches that are more sensitive to focal activations. © 1995 Academic Press, Inc.

1. INTRODUCTION

The main aim of this paper is to describe a simple test that can be applied to data obtained from activation studies. In particular, we emphasize the potential importance of this test in assessing distributed or nonfocal task-dependent brain changes.

In past years, a number of statistical methods have been proposed for the analysis of brain activation studies (see McColl *et al.,* 1994, for a review). Their purpose is to characterize, in a statistical way, the difference between two (or more) brain states, as measured by Positron Emission Tomography (PET) or functional Magnetic Resonance Imaging (fMRI). A number of methods have focused on assessing the chance probability of detecting focal differences using either the intensity of local maxima in statistical parametric maps (Friston *et al.,* 1991; Worsley *et al.,* 1992) or the spatial extent of foci above an arbitrary threshold (Poline *et al.,* 1993; Roland *et al.,* 1993; Friston *et al.,* 1994b).

These approaches are based on distributional approximations that hold for high thresholds and assume (implicitly) that the underlying physiological effect is focal. It is possible, however, that in some circumstances, nonfocal or anatomically distributed differences may occur, with or without focal changes. This possibility is suggested by the topography of eigenimages based on activation study time-series. These eigenimages or spatial modes reveal widespread brain systems with coherent physiological activity. This activity reflects systematic changes that often depend on the experimental conditions used (e.g., see Friston *et al.,* 1993a,b). A second observation, which is suggestive of nonfocal changes, is that many cognitive activation studies result in large distributed activations (and deactivations), particularly in multimodal and paralimbic cortex, for example, the dorsolateral prefrontal activations, and extensive bitemporal deactivation due to verbal fluency (Frith *et al.,* 1991). The topography of functional activations as measured with fMRI does not resolve the focal vs nonfocal issue. Eigenimage analysis of fMRI time-series (Friston *et al.,* 1994a) suggests a widespread focal and nonfocal functional organization and yet when thresholded, fMRI activations appear very localized.

The concept of nonfocal differences is important in describing the pathophysiology of certain clinical conditions. An example is hypofrontality in schizophrenia (Ingvar, 1983). Hypofrontality is a diffuse nonfocal reduction in prefrontal cortical physiology that is characteristic of "psychomotor poverty" syndromes (e.g., Liddle *et al.,* 1992).

In summary, the possibility of nonfocal response to cognitive or sensorimotor challenge points to the importance of (i) acknowledging that the assumption of focal change is implicit in many current approaches to data analysis and (ii) assessing the ability of new tests to reveal or discount the presence of nonfocal changes.

In 1989, Fox and Mintun proposed an omnibus test (the γ_2 statistic) using the kurtosis of the distribution of the local maxima in difference volumes. The proposal was based on the idea that activations should increase the number of local maxima outliers and therefore the fourth moment (kurtosis) of their distribution. In contrast, our proposed mean sum of squares test is based on the second moment of all the voxel

1053-8119/95 $12.00

values of the statistical parametric map, and not just the local maxima. Friston *et al.* (1990) proposed an omnibus test based on the activation proportion of thresholded statistical parametric maps, and in Section 3 of this paper, we report a correction to the specificity of this test, derived in Worsley and Vandal (1994). These tests can be described as "omnibus tests" in the sense that they allow rejection of a weak null hypothesis that no activation has occurred in the brain volume, without indicating where the activation is localized (a stronger null hypothesis is that an activation has not occurred in a specific brain region).

Most of the tests discussed above are based on thresholds or maxima and implicitly assume the presence of focal activations. In this paper we highlight the potential usefulness of an omnibus test based on the mean sum of squares of an SPM. It is simple, sensitive, and particularly suited for the detection of nonfocal activations. The second section of the paper presents the test and the underlying theory. In Section 4 we assess its specificity and sensitivity using Monte Carlo simulations (that include distributed foci) and apply the test to an experimental data set (a verbal fluency activation study of normal subjects) in Section 5.

2. THE MEAN SUM OF SQUARES STATISTIC

The result of many statistical analyses of PET or fMRI data is a statistical parametric map (SPM) $Z(\mathbf{x})$ at voxel locations $\mathbf{x}$ in a region of interest R (Friston *et al.*, 1991, 1994a; Worsley *et al.*, 1992). The map is a measure of activation standardized to have a Gaussian distribution with zero mean and unit variance at all locations when no activation is present. Our test for activation is based on the mean sum of squares S of $Z(\mathbf{x})$ in the search region R, defined as

$$S = \Sigma Z(\mathbf{x})^2/N, \tag{2.1}$$

where summation is over all voxels with coordinates $\mathbf{x}$ in R, and N is the number of voxels in R. Obviously the usual χ^2 distribution cannot be used to set the specificity of S because the voxels are highly correlated due to the so-called "partial volume" effect. Instead, we show how to adjust the degrees of freedom of the χ^2 distribution to take this into account; specifically, we show that the degrees of freedom depends on the RESELS in the search region, which equals the volume of the search region divided by the product of the full width at half maxima of the point response function of the PET camera (Worsley *et al.*, 1992).

2.1. Distribution

We suppose that under the null hypothesis of no activation the SPM can be represented as a Gaussian random field $Z(\mathbf{x})$ sampled on a uniform lattice of voxels. We assume that $Z(\mathbf{x})$ is stationary with mean zero and unit variance at any point, and correlation function $\rho(\mathbf{x})$. Then the expectation of S is

$$\begin{aligned} \mathrm{E}(S) &= \Sigma \mathrm{E}\{Z(\mathbf{x})^2\}/N \\ &= 1, \end{aligned} \tag{2.2}$$

and the variance of S is

$$\begin{aligned} \mathrm{Var}(S) &= \Sigma\Sigma \mathrm{Cov}\{Z(\mathbf{x}_1)^2, Z(\mathbf{x}_2)^2\}/N^2 \\ &= \Sigma\Sigma 2\mathrm{Cov}\{Z(\mathbf{x}_1), Z(\mathbf{x}_2)\}^2/N^2 \\ &= \Sigma\Sigma 2\rho(\mathbf{x}_1 - \mathbf{x}_2)^2/N^2, \end{aligned} \tag{2.3}$$

where summations are over $\mathbf{x}$, $\mathbf{x}_1$, and $\mathbf{x}_2$ in R. Letting $\mathbf{h} = \mathbf{x}_1 - \mathbf{x}_2$, we get, from (2.3),

$$\begin{aligned} \mathrm{Var}(S) &\leq \Sigma\Sigma 2\rho(\mathbf{h})^2/N^2 \\ &= 2\Sigma\rho(\mathbf{h})^2/N, \end{aligned} \tag{2.4}$$

where summations are over $\mathbf{x}$ in R and all $\mathbf{h}$. Worsley and Vandal (1994) show that the inequality (2.4) is a very accurate approximation provided that the search region is large.

If voxels were independent, then from (2.3) we see that $\mathrm{Var}(S) = 2/N$, and the distribution of NS is χ^2 with N degrees of freedom. This suggests that in the correlated case, we can find a value ν for the "effective" degrees of freedom of S so that νS is approximately χ^2 with ν degrees of freedom (Satterthwaite, 1946). The appropriate value of ν, obtained by solving $\mathrm{Var}(S) = 2/\nu$, and approximating the summation in (2.4) by an integral, is

$$\nu = V/\int \rho(\mathbf{h})^2 d\mathbf{h}, \tag{2.5}$$

where V is the volume of R. A similar method has been used by Friston *et al.* (1995) and Worsley and Friston (1995) for fMRI time-series.

Friston *et al.* (1991) and Worsley *et al.* (1992) show that the correlation function can be well approximated by a Gaussian function with full width at half maxima (FWHM) equal to those of the point spread function multiplied by $\sqrt{2}$. Substituting this in (2.5) gives

$$\nu = \mathrm{RESELS}(4\log_e 2/\pi)^{D/2}, \tag{2.6}$$

where D is the number of dimensions.

2.2. Fourier Analysis Approach

Fourier analysis can be used to help find the exact distribution of S in a rectangular search region. For simplicity we shall assume that $D = 1$, although the results are straightforward to generalize to higher dimensions. The search region is then an interval of

length V sampled at N points. An important constraint is that the correlation function is periodic; that is,

$$\rho(x) = \rho(V - x). \tag{2.7}$$

This assumption implies that voxels on one boundary of the search region are highly correlated with those on the opposite boundary. Admittedly this is an unrealistic assumption for PET or fMRI data but it will allow us to find the exact distribution of S in this artificial case.

Let $k(x)$ be the kernel associated with $\rho(x)$, where, in neuroimaging, this kernel is the point spread function. The spectral density function $g(\omega)$ of the process at frequency $\omega = 2\pi j/N$, $j = 0, \ldots, (N-1)$ is given by

$$g(\omega) = \mathrm{FT}\{\rho(x)\} = \mathrm{FT}\{k(x)\}\mathrm{FT}\{k(x)\}^*, \tag{2.8}$$

where FT denotes Fourier transform and * the complex conjugate.

In the frequency domain the spectral density $g(\omega)$ represents the amount of variance (or energy) at frequency ω. A property of the process in Fourier space is that the components $z(\omega)$ of the discrete Fourier transform of the SPM $Z(x)$ are independent Gaussian random variables with variance $g(\omega)$, $j = 0, \ldots, (N/2 - 1)$. For $j = N/2, \ldots, (N-1)$, $z(\omega) = z(2\pi - \omega)^*$. Because the total variance or energy of the process is the same in both representations (real and Fourier space), the mean sum of squares of the process in real space is also the mean sum of squares of the process in Fourier space (see, e.g., Cox and Miller, 1980, Chap. 7). Therefore we can write

$$S = \Sigma z(\omega)z(\omega)^*/N, \tag{2.9}$$

where the summation is over the $N/2$ distinct values of $z(\omega)z(\omega)^*$, $j = 0, \ldots, (N/2 - 1)$. Now $z(\omega)z(\omega)^*/g(\omega)$ has a χ^2 distribution with 2 degrees of freedom, and so the characteristic function (Fourier transform of the probability density function) of $z(\omega)z(\omega)^*/N$ is

$$\phi(t;\omega) = [1 - 2itg(\omega)/N]^{-1}. \tag{2.10}$$

The characteristic function of S is then the product of $\phi(t;\omega)$ over $j = 0, \ldots, (N/2 - 1)$. The density function of S at s can then be recovered by the standard inversion formula:

$$p(s) = (1/2\pi)\int\prod\phi(t;\omega)\exp(-its)dt. \tag{2.11}$$

A method of obtaining the upper tail probabilities of S directly from $g(\omega)$ without inverting the Fourier transform (2.11) is given by Imhoff (1961), although we found it was easier to use (2.11). We can also rederive the same results (2.2) and (2.4) from the Fourier representation (2.9):

$$\mathrm{E}(S) = \Sigma\mathrm{E}\{z(\omega)z(\omega)^*\}/N = \Sigma 2g(\omega)/N = 1, \tag{2.12}$$

$$\mathrm{Var}(S) = \Sigma\mathrm{Var}\{z(\omega)z(\omega)^*\}/N^2 = \Sigma 4g(\omega)^2/N^2 = 2\Sigma\rho(h)^2/N. \tag{2.13}$$

2.3. *Under What Conditions Does the χ^2 Approximation Fail?*

In the case of a periodic correlation in a rectangular region, we can use the exact expression for $p(s)$ (2.11) and the χ^2 approximation (2.2)–(2.5) to examine how "good" the approximation is under different conditions. Figure 1 shows the correspondence between the actual density $p(s)$ from (2.11) and the χ^2 approximation (2.2)–(2.6) for two autocorrelation functions, a Gaussian function and a sinc function, for a $D = 1$ dimensional process. Our analyses suggest that the χ^2 approximation behaves very well for any autocorrelation function irrespective of the dimension of the space, for large degrees of freedom. However, the χ^2 approximation fails when the effective degrees of freedom of the process is small, although if the search region is shrunk to a single point, so that S has a χ^2 distribution with 1 degree of freedom, then the approximation given by (2.2) and (2.4) is exact. Figure 2 shows the actual (2.11) and approximated (2.2–(2.6) distribution of S for small ν for a 1D and 2D process. Figure 2 suggests that the agreement is poor for small degrees of freedom, although the periodic assumption under which these results were obtained is not realistic for practical applications.

For some point spread functions, again assuming a periodic correlation structure, the χ^2 approximation is exact. If the kernel is a product of sinc functions $k(x) = \sin(x)/x$, then it can be shown that the power spectrum $g(\omega)$ is π for $|\omega| \leq 1$ and zero otherwise, that is, for $0 \leq j \leq N/(2\pi)$. The distribution of S is now the same as the sum of $N/(2\pi)$ identically distributed χ^2 random variables with 2 degrees of freedom, which gives an exact χ^2 distribution with $\nu = N/\pi$ degrees of freedom. The correlation function $\rho(x)$ is identical to the kernel $k(x)$, the integral of $\rho(x)^2$ is π, so that the degrees of freedom given by (2.5) is the same result $\nu = N/\pi$. Unfortunately the relationship between ν and the RESELS is not quite the same as (2.6); the exact relation for the sinc kernel, found by solving $k(x) = 1/2$ numerically to find the FWHM, is $\nu = \mathrm{RESELS}\ (1.207)^D$.

3. THE ACTIVATION PROPORTION STATISTIC

Friston *et al.* (1990) proposed a global test based on the activation proportion, measured by the proportion A of voxels in a search region R where the SPM exceeds a fixed threshold t. A focal version of this test has recently been proposed by Friston *et al.* (1994b) based on the size of the largest connected region where the SPM exceeds a fixed threshold. The null distribution of A given in Friston *et al.* (1990) assumes that voxels are independent and leads to far more false positives. The

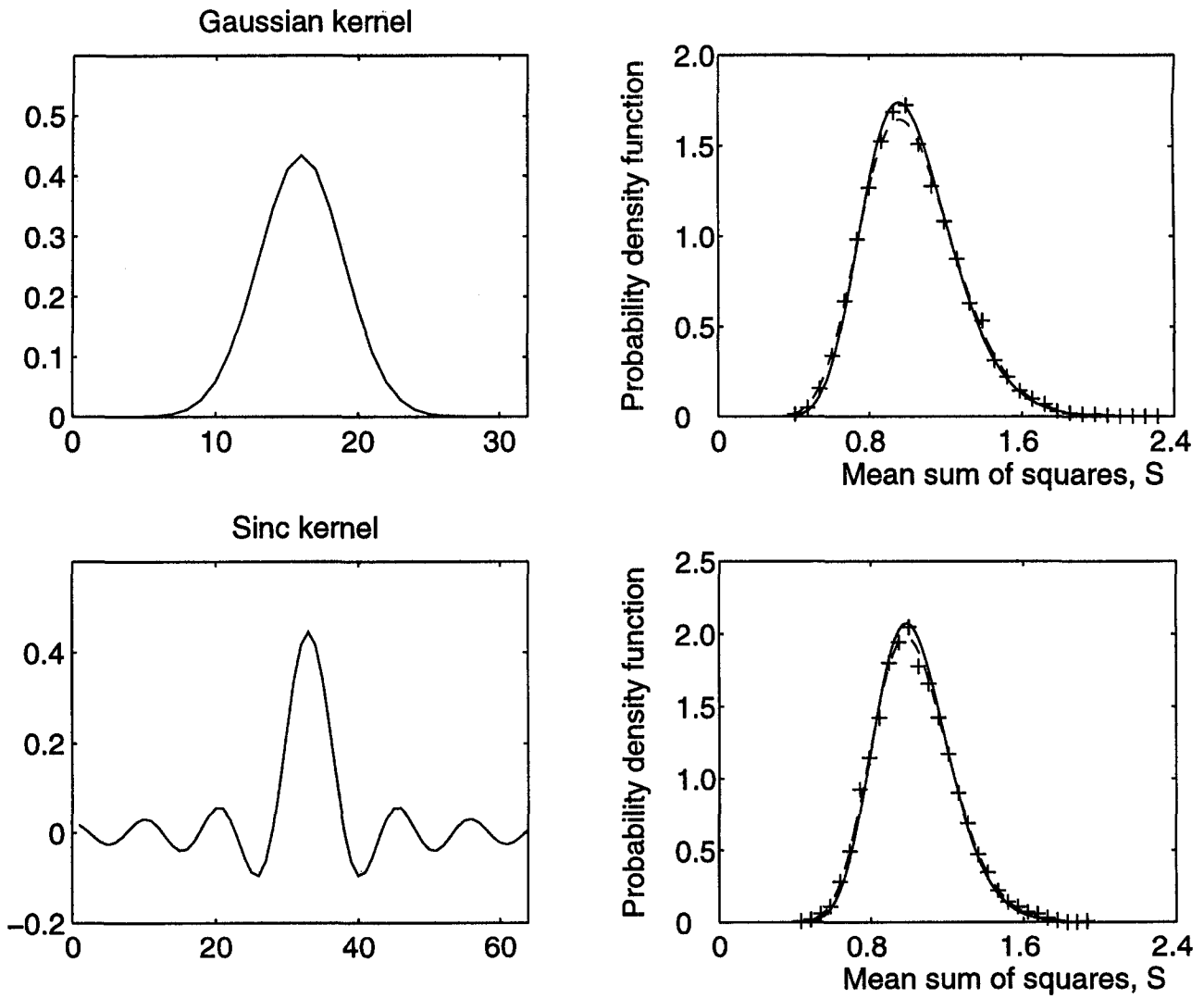


FIG. 1. The distribution of S with large degrees of freedom in $D = 1$ dimensions. A comparison of the χ^2 approximation and the exact distribution of the mean sum of squares statistic S for two different periodic point spread functions (upper left: Gaussian; lower left: sinc). The right panels show good agreement between the distribution of S using the χ^2 approximation (dashed line), the exact distribution (solid line), and simulations (crosses). The processes have $N = 256$ points and their effective degrees of freedom are $\nu = 34.0$ and $\nu = 50.2$ for the Gaussian and sinc functions, respectively.

correct limiting distribution, derived in Worsley and Vandal (1994), is as follows. The expectation of A is

$$\mathrm{E}(A) = \int_t^\infty (2\pi)^{-1/2}\exp(-z^2/2)dz. \tag{3.1}$$

If the process is isotropic (the correlation function $\rho(r)$ depends only on the distance r from the origin), then the variance is well approximated, for large R, by

$$\mathrm{Var}(A) = \frac{1}{V}\int_0^\infty \frac{\pi^{D/2-1}r^D}{D\Gamma(D/2)(1-\rho^2)^{1/2}} \exp\left(-\frac{t^2}{1+\rho}\right)\left(-\frac{d\rho}{dr}\right)dr. \tag{3.2}$$

For a Gaussian correlation function this becomes

$$\mathrm{Var}(A) = \frac{1}{\mathrm{RESELS}}\int_0^\infty \frac{(4\log_e 2)^{-D/2}\pi^{D/2-1}r^{D+1}}{D\Gamma(D/2)(1-\exp(-r^2))^{1/2}} \exp\left(-\frac{t^2}{1+\exp(-r^2/2)} - \frac{r^2}{2}\right)dr. \tag{3.3}$$

The distribution of A is then approximately normal with expectation (3.1) and variance (3.3).

Friston *et al.* (1990) suggest three choices of $t = 1.64$, 2.33, and 2.58, corresponding to an expected proportion of $\mathrm{E}(A) = 0.05$, 0.01, and 0.005. In $D = 3$ dimensions the corresponding variances are $\mathrm{Var}(A) = 0.0591/\mathrm{RESELS}$,

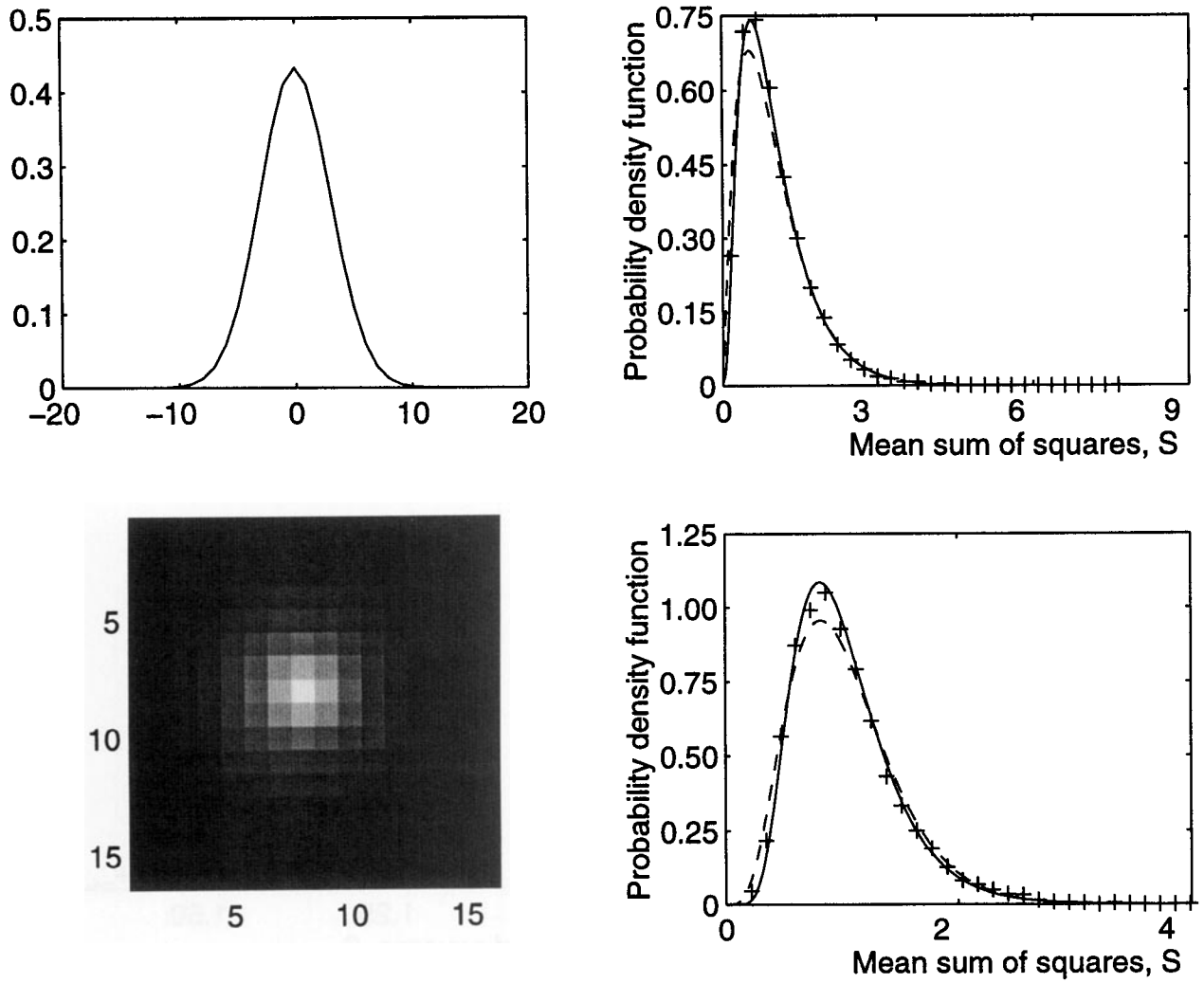


FIG. 2. The distribution of S with small degrees of freedom for $D = 1$ dimensions (upper right) and $D = 2$ dimensions (lower right). The χ^2 approximation (dashed line) is less accurate than the exact distribution (solid line) and simulations (crosses). The left panels show the corresponding Gaussian point spread functions. The size of the processes were $N = 32$ in $D = 1$ dimensions, and $N = 16^2 = 256$ in $D = 2$ dimensions, to give $\nu = 4.2$ and $\nu = 10.2$ degrees of freedom, respectively.

0.00698/RESELS, and 0.00278/RESELS. It is interesting to express Var(A) in terms of the "effective" number of independent voxels that would produce the same variance of A, defined as $E(A)\{1 - E(A)\}/\text{Var}(A)$. For the above thresholds these are 0.80 × RESELS, 1.42 × RESELS, and 1.79 × RESELS, respectively. The result in Friston *et al.* (1990) is based on assuming that the effective number of independent voxels equals the actual number of voxels N in the search region, which leads to an underestimate of the variance Var(A).

4. VALIDATION

4.1. Two Dimensions

We simulated 5000 SPMs in $D = 2$ dimensions using uncorrelated Gaussian fields of 64 × 64 voxels, convolved with a Gaussian kernel with standard deviation of 3 pixels, $k(\mathbf{x}) = \exp(-\|\mathbf{x}\|^2/18)$. The FWHM is $3\sqrt{8 \log_e 2} = 7.06$ pixels and the number of resels in the field is $(64/7.06)^2 = 82.1$. The specificity of the test was assessed by comparing the observed and expected distribution of S. Figure 3 shows the empirical distribution of S based on the simulated noise-only SPMs and the χ^2 approximation with $\nu = 82.1(4 \log_e 2/\pi) = 72.4$ from (2.6). The agreement is evident. Note that the χ^2 approximation gets better as the effective degrees of freedom increases.

Sensitivity was assessed by adding focal and nonfocal signals of different spatial extent. The signals were step functions (convolved with $k(\mathbf{x})$), with a deliberately small positive signal of 0.15 in half the signal area and the same size decrease of −0.15 in the other half. Three sorts of signal were used, ranging from focal to spatially distributed (representing 7.8, 14.6, 25.4, and 58.6% of the region tested). Figure 4 shows

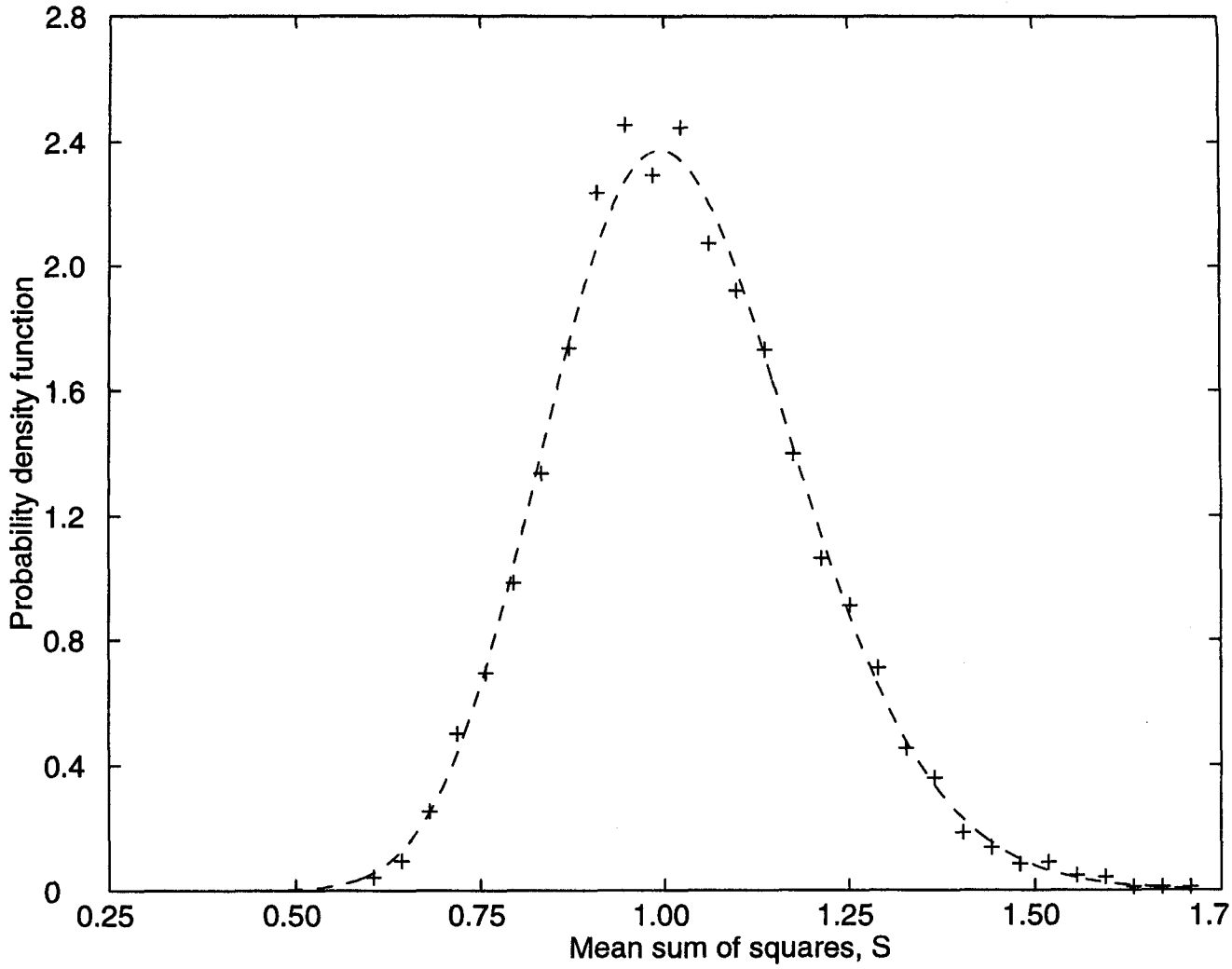


FIG. 3. The distribution of S with large degrees of freedom in $D = 2$ dimensions. A comparison of the χ^2 approximation and the simulated distribution (5000 realizations) of S for a Gaussian point spread function (FWHM = 7.06 pixels). The χ^2 approximation (dashed line) is in good agreement with simulations (crosses). The process has $N = 64^2 = 4096$ points and the effective degrees of freedom is $\nu = 72.4$.

the signals used in the simulation and an example of a simulated SPM with no signal. The sensitivity of the sums of squares was compared to the sensitivity of the γ_2 statistic (Fox and Mintun, 1989), a test based on the kurtosis of the distribution of SPM local maxima. Sensitivity was measured as the percentage of correctly detected signals at a fixed false positive rate ($P = 0.05$). Table 1 presents the results for the comparison of sensitivity between the test based on S and that on the γ_2 statistic for the different signal sizes. It shows that for all signals the S test is more or equally powerful than a test based on the distribution of local maxima. The S test was up to nine times more sensitive in the case of very nonfocal signals and retains a comparable sensitivity for focal signals, therefore proving itself to be (to some extent) versatile. Of course, in the case of very focal signals, the S test is much less sensitive than tests based explicitly on maxima (Friston *et al.,* 1991; Worsley *et al.,* 1992), as we shall see in the next section.

4.2. Three Dimensions

We simulated 200 Gaussian SPMs in $D = 3$ dimensions with zero mean and unit variance, sampled on a $128 \times 128 \times 64$ lattice of 1.5-mm voxels. The search region R was a hemisphere of radius 75 mm and volume $V = 884$ cm^3, which roughly approximated the brain region. The Gaussian random fields were created by convolving Gaussian white noise with a Gaussian smoothing kernel of resolution 18-, 18- and 7.5-mm FWHM, chosen to represent smoothing in the x and y directions but no axial (z) smoothing, giving RESELS = $884/(1.8 \times 1.8 \times 0.75) = 364$. Convolution was achieved via the Fast Fourier Transform.

Three tests were compared on the 3D data: the mean sum of squares test S, the activation proportion test A, and the maximum SPM test M. Critical thresholds for each test were calculated at the 5% nominal false positive rate. The degrees of freedom for the χ^2 distribution of S was $\nu = 301$ from (2.6). The mean and variance

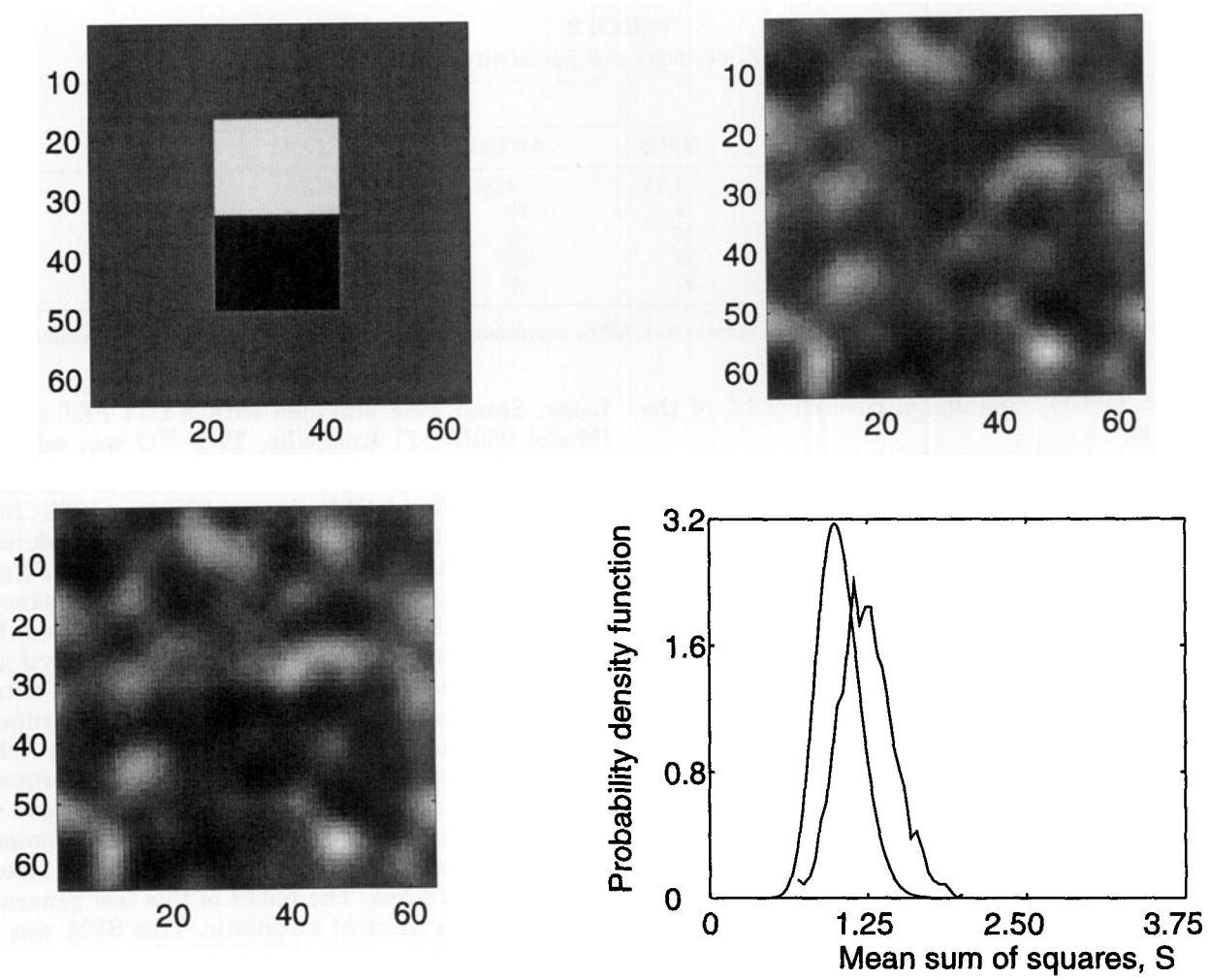


FIG. 4. The same as in Fig. 3, but with a simulated signal (upper left) added to the noise (upper right) to give signal plus noise (lower left). The lower right panel shows the χ^2 approximation to the distribution of S for noise only from Fig. 3 (smooth line), and the distribution of S from 1000 realizations of signal plus noise (broken line). The evident increase in S shows the sensitivity of the test to the addition of this signal.

of A were calculated from (3.1) and (3.3) using three values of the threshold t. The P value of M was calculated from Worsley *et al.* (1992). The observed specificities are shown in Table 2, and it can be seen that they are in good agreement with the nominal levels.

TABLE 1
2D Specificity and Sensitivity

	Signal		Detection (%)	
	Area (%)	SNR (%)	MSOS	γ_2
None	0	0	3.6	1.0
Signal 1	7.8	28	7.9	7.3
Signal 2	14.6	45	31.7	12.3
Signal 3	25.4	65	67.7	11.0
Signal 4	58.6	106	99.2	13.0

Note. SNR, signal to noise ratio; MSOS, mean sum of squares test; γ_2, γ_2 test.

Phantom signals were created and added to each simulated image and the tests were repeated. The signals were:

- a focal signal, created by convolving the point spread function with itself to produce a Gaussian shaped signal with FWHM equal to $\sqrt{2}$ times that of the point spread function. This was shown by Worsley and Vandal (1994) to be the signal shape best detected by S and M. The location was chosen to lie in the anterior cingulate close to where activation was in fact detected in a study of pain perception by Talbot *et al.* (1991). The peak height was chosen to be 4. In comparison with the 2D simulations, the focal signal occupied only a small part (0.8%) of the search region.
- a signal with three peaks, identical in shape and height to the first but centered in the anterior cingulate, the primary and secondary somatosensory regions of the brain, close to where activation was detected in

TABLE 2
3D Specificity and Sensitivity

	Signal		Detection (%)				
	Volume (%)	SNR (%)	MSOS	AP(1.64)	AP(2.33)	AP(2.58)	Max
None	0	0	4.25	4.25	5.25	5.75	4.50
One focal	0.8	23	16	20	31	38	56
Three focal	2.3	38	55	60	85	90	89
One flat	4.6	64	89	98	100	100	97
Diffuse	100	50	87	49	59	59	27

Note. SNR, signal to noise ratio; MSOS, mean sum of squares test; AP(t), activation proportion test with threshold t; Max, maximum SPM test.

Talbot *et al.* (1991). This signal covered 2.3% of the search region.

- a "flat" signal, created by convolving a 3 × 4.5 × 3-cm region of uniform height with the Gaussian point spread function to create a broader region of activation which covered 4.6% of the search region. The maximum height of this region was chosen to be the same as the peak heights above, and the region was located in the right hemisphere in roughly the same place where activation was detected by Talbot *et al.* (1991).
- a random "diffuse" signal, created as the realization of a Gaussian random field with the same correlation structure as the noise. For this signal, there is a straightforward way of assessing the sensitivity of all tests without using simulations. The sensitivity, averaged over all such random realizations, is equivalent to just the P value for the same statistic but with the image standard deviation multiplied by $(1 + \mathrm{SNR}^2)^{1/2}$, where SNR is the signal to noise ratio, as measured by the root mean square (RMS) amplitude of the signal relative to that of the noise. In our case we chose SNR = 0.5.

The sensitivity of the tests was then estimated by the proportions of simulated values exceeding the nominal critical values, and the results are given in Table 2, along with the SNR of all signals.

We note that, as expected, the maximum SPM test M is by far the most sensitive at finding a single peak, closely followed by the activation proportion test A at high threshold; all tests are better at finding three peaks or a flat peak. For the diffuse signal the S test is the most sensitive, followed by the activation proportion test A for highest threshold; the maximum test M has low sensitivity. The conclusion is that focal signals are best detected by the maximum test, diffuse signals are best detected by the mean sum of squares test S, and the activation proportion test A performing reasonably well for all types of signal, particularly at the highest threshold.

5. APPLICATION

Data were obtained from three subjects scanned 12 times (every 8 min) while performing one of two verbal tasks. Scans were obtained with a CTI PET camera (Model 953B CTI Knoxville, TN). ^{15}O was administered intravenously as radiolabeled water infused over 2 min. Total counts per voxel during the buildup phase of radioactivity served as an estimate of regional cerebral blood flow (rCBF) (Fox and Mintun 1989). Subjects performed two tasks in alternation. One task involved repeating a letter presented aurally at 1 per 2 s (*word shadowing*). The other was a paced verbal fluency task, for which the subjects responded with a word that began with the letter presented (*intrinsic word generation*). To facilitate intersubject pooling, the data were stereotactically normalized (Friston *et al.*, submitted) and analyzed as a randomized block design ANCOVA at every voxel (with global activity as covariate). The data were analyzed using a contrast designed to test for time-dependent effects. This contrast compared a linear effect during the first with a similar effect during last six scans. The result of this test generated an SPM of the Student t statistic. This SPM was transformed to the unit Gaussian distribution.

The test based on the mean sum of squares was very significant. Figure 5 presents the observed and expected distribution of the voxel values in the SPM volume, the design matrix, and the contrast used in the example. The mean sum of squares statistic was $S = 79378/62025 = 1.28$, corresponding to $P < 0.02$ with $\nu = 142$ degrees of freedom (171 RESELS). No localized activation was assessed as significant using the maximum M (Friston *et al.*, 1991; Worsley *et al.*, 1992) or the spatial extent (Friston *et al.*, 1994b). The γ_2 statistic was not significant. In this example, the spatially distributed time-dependent effects involved extensive regions in the posterior cingulate and fusiform gyrus (Fig. 6).

A null analysis was performed by randomizing the order of the scans and the elements of the contrast used to compute the SPM. The P value of the observed $S = 1.04$ did not reach significance ($P = 0.34$).

6. DISCUSSION

We have presented a simple and easily implemented test based on the distribution of the mean sum of

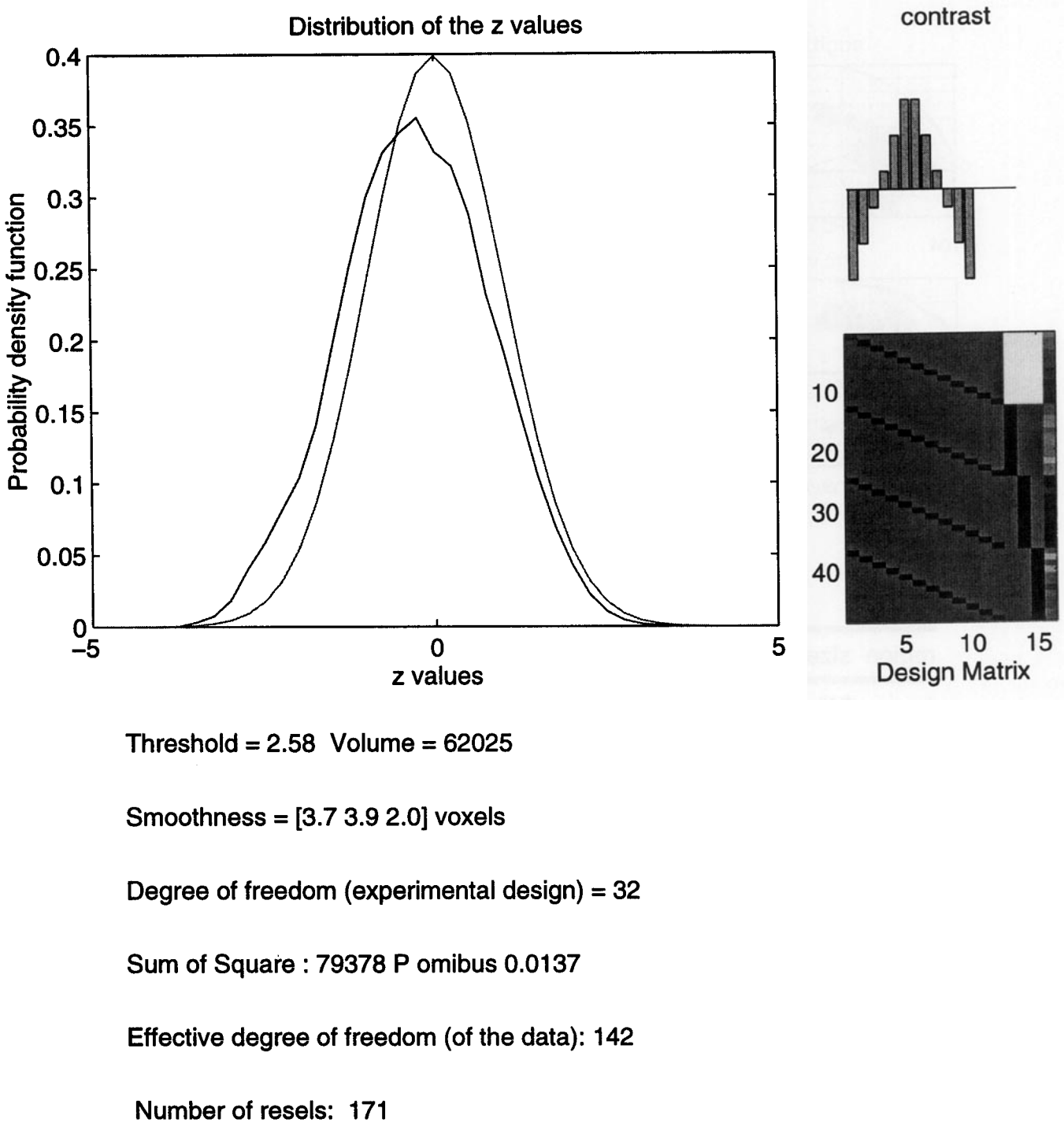
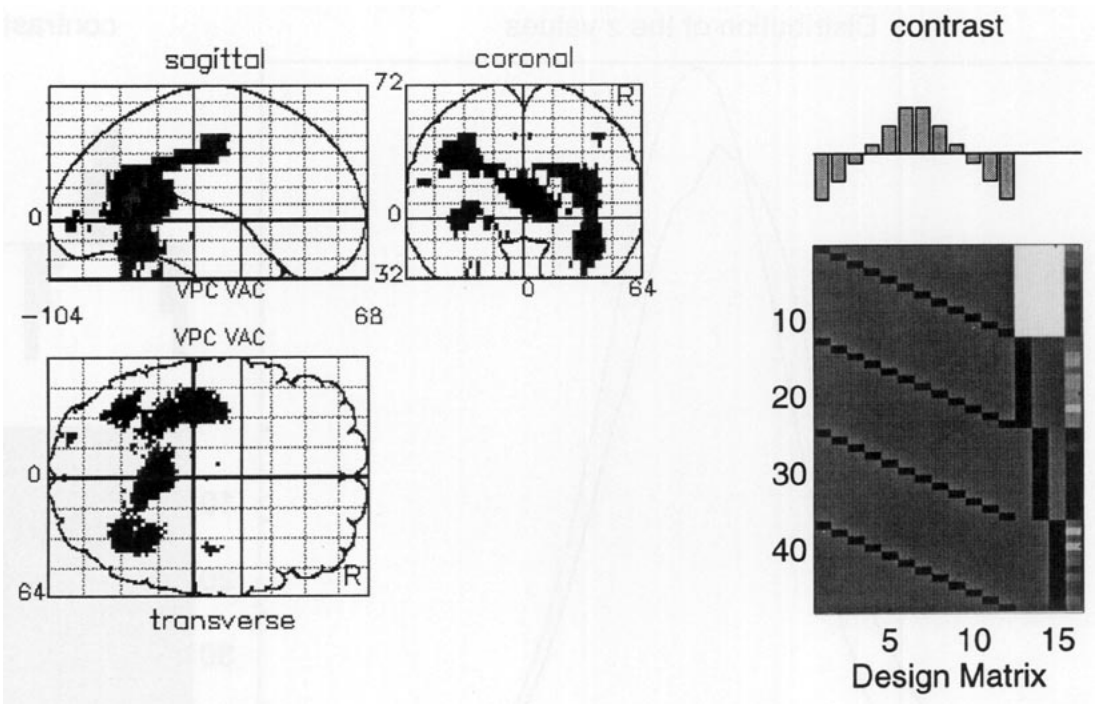


FIG. 5. Experimental dataset tested with S. The distribution of the experimental SPM Z values (broken line) and the Gaussian distribution (smooth line) are shown in the left panel. The design matrix and the contrast used in this application are shown in the right panel. With an effective degrees of freedom of $\nu = 142$ (171 RESELS), the mean sum of squares statistic is $S = 1.28$, leading to a significant omnibus test ($P < 0.02$). No other test (activation proportion, γ_2, peak height or spatial extent) gave significant results (see Fig. 6).

squares of a smooth stationary process. The test is applicable in any number of dimensions and is sensitive to diffuse nonfocal cerebral signals. The mean sum of squares can be significant in the absence of local activations. This raises the issue of whether it is appropriate to model the brain's physiological response in terms of highly focal activations. Although this model has proved itself extremely successful, it is a constrained model that may not reflect the distributed and varied spatial scales of neurophysiological changes.



region	size {k}	P(n max > k)	Z	P(Z max > u)	(Uncorrected)	{x,y,z mm}
1	**749**	**0.601**	**3.55**	**0.462**	(0.000)	**2 −46 12**
			2.84	2.865	(0.002)	−40 −20 40
			2.82	2.986	(0.002)	−30 −10 44
2	**388**	**0.911**	**2.76**	**3.368**	(0.003)	**38 −50 −8**
			2.66	4.097	(0.004)	32 −64 28
			2.65	4.166	(0.004)	38 −68 0
3	**9**	**1.000**	**2.46**	**5.879**	(0.007)	**−22 −90 −4**
4	**99**	**0.999**	**2.44**	**6.063**	(0.007)	**−32 −58 −4**
			2.40	6.441	(0.008)	−40 −68 0
			2.11	9.616	(0.017)	−32 −76 4
5	**3**	**1.000**	**2.33**	**7.186**	(0.010)	**−56 −56 20**
6	**13**	**1.000**	**2.19**	**8.783**	(0.014)	**40 −20 44**
7	**18**	**1.000**	**2.16**	**9.028**	(0.015)	**−46 −56 24**
8	**1**	**1.000**	**2.04**	**10.438**	(0.021)	**26 −42 −16**
9	**1**	**1.000**	**2.01**	**10.689**	(0.022)	**40 −68 −8**
10	**4**	**1.000**	**1.99**	**11.009**	(0.024)	**−16 −58 28**
11	**3**	**1.000**	**1.97**	**11.129**	(0.024)	**−26 −44 −24**
12	**1**	**1.000**	**1.94**	**11.493**	(0.026)	**−6 −14 44**
13	**3**	**1.000**	**1.89**	**11.970**	(0.029)	**4 −58 28**
14	**3**	**1.000**	**1.88**	**12.088**	(0.030)	**−34 −90 4**

Threshold = 1.64; Volume [S] = 62025 voxels; df = 32
FWHM = [17.2 18.2 18.4] mm (i.e. 171 RESELS)

FIG. 6. Statistical parametric map of Z reflecting the significance of a compound of effects. The SPM is displayed in a standard format as a maximum intensity projection viewed from the back, the right-hand side, and the top of the brain. The anatomical space corresponds to the atlas of Talairach and Tournoux (1988). The SPM has been thresholded at 1.64 and the color scale is arbitrary. (Upper right panel) Top, the contrast used for this SPM. The contrast is displayed above the appropriate effects (columns of the design matrix). (Lower panel) Table of regional effects (activations or regional differences) characterized by the volume of each region (k), its significance based on patial extent $P(n_{max} > k)$, the highest Z value (Z), its significance based on $P(Z_{max} > u)$, and the location of this primary maximum. We have also included up to three secondary maxima for each region and their associated significance based on the corrected and uncorrected P value.

Because of its explicit dependence on maxima and kurtosis, the γ_2 test is not very sensitive to a volume containing spatially distributed signals. Moreover, its specificity has only been estimated empirically and there is no theoretical basis for the validity of the test. Tests based on the activation proportion (the proportion of voxels exceeding a given threshold) behave more like the mean sum of squares test, since the number of suprathreshold pixels also increases with the prevalence of distributed signals. These tests, however, depend on the threshold chosen; a high threshold seems to be best. Among all these tests, our results show that the mean sum of squares test is the most sensitive at detecting diffuse nonfocal activation.

As stated in the introduction, the test reviewed in this paper is omnibus in the sense that it has no localizing power. However, it can be applied to parts of the brain as opposed to the entire brain volume, provided the choice of this subset is not based on the voxel values themselves. For example, a subvolume could be chosen on some anatomical basis (e.g., a specific gyrus). In such applications the volume and effective degrees of freedom may be quite small and the χ^2 approximation may not be accurate.

Note generally that if a focal activation can be detected, the "omnibusness" of any test becomes redundant. This is because if we can reject the strong hypothesis that no activation occurred at some point, then one is implicitly rejecting the null hypothesis that an activation has not occurred anywhere.

The degree of smoothing, or equivalently the scale at which the data are observed, has already been acknowledged as a key parameter in the analysis of functional neuroimaging data (Poline and Mazoyer, 1994a,b). This continues to be the case with the mean sum of squares test. Varying the width (from 1.5 to 5 pixels) of the Gaussian kernel (used in creating the simulated processes containing the second smallest 2D signal) showed that the sensitivity of the test is reduced when the filter is too small. In general, Worsley and Vandal (1994) have shown that the mean sum of squares test, like the maximum test, is most sensitive when the kernel associated with the SPM has the same spatial extent as the underlying physiological signals. Thus a wide filter should be used to detect a wide signal, and a narrow filter for a narrow signal.

7. CONCLUSION

We have described a simple test based on the mean sum of squares of the SPM. The test can be applied to any volume that approximates (under the null hypothesis) a stationary smooth Gaussian process. This test has proven powerful in the case of distributed, nonfocal activations and could be used as a prelude to variance partitioning procedures that do not allow for statistical inferences, such as singular value decomposition and eigenimage analysis (Friston *et al.,* 1993a,b).

ACKNOWLEDGMENTS

K.J.W. was supported by the Natural Sciences and Engineering Research Council of Canada, and the Fonds pour la Formation des Chercheurs et l'Aide à la Recherche de Québec. J-B.P. was funded by a European Union grant, Human Capital and Mobility, No. ERB4001GT932036. K.J.F. was funded by the Wellcome Trust.

REFERENCES

Cox, D. R., and Miller, H. D. 1980. *The Theory of Stochastic Processes.* Chapman & Hall, New York.

Fox, P. T., and Mintun, M. A. 1989. Non-invasive functional brain mapping by change distribution analysis of averaged PET images of $H^{15}O_2$ tissue activity. *J. Nuclear Med.* **30:** 141–149.

Friston, K. J., Frith, C. D., and Frackowiak, R. S. J. 1993a. Time-dependent changes in effective connectivity measured with PET. *Human Brain Mapping* **1:** 69–79.

Friston, K. J., Frith, C. D., Liddle, P. F., and Frackowiak, R. S. J. 1991. Comparing functional (PET) images: The assessment of significant change. *J. Cerebral Blood Flow Metab.* **11:** 690–699.

Friston, K. J., Frith, C. D., Liddle, P. F., and Frackowiak, R. S. J. 1993b. Functional connectivity: The principal component analysis of large (PET) data sets. *J. Cerebral Blood Flow Metab.* **13:** 5–14.

Friston, K. J., Frith, C. D., Liddle, P. F., Lammertsma, A. A., Dolan, R. D., and Frackowiak, R. S. J. 1990. The relationship between local and global changes in PET scans. *J. Cerebral Blood Flow Metabol.* **10:** 458–466.

Friston, K. J., Holmes, A. P., Poline, J-B., Grasby, B. J., Williams, C. R., Frackowiak, R. S. J., and Turner, R. 1995. Analysis of fMRI time-series revisited. *NeuroImage* **2:** 45–53.

Friston, K. J., Jezzard, P., and Turner, R. (1994a). Analysis of functional MRI time series. *Human Brain Mapping* **1:** 153–171.

Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C., and Evans, A. C. 1994b. Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping* **1:** 214–220.

Frith, C. D., Friston, K. J., Liddle, P. F., and Frackowiak, R. S. J. 1991. A PET study of word finding. *Neuropsychologia* **29:** 1137–1148.

Imhoff, J. P. 1961. Computing the distribution of quadratic forms in normal variables. *Biometrika* **48:** 419–426.

Ingvar, D. H. 1983. Serial aspects of language and speech related to prefrontal cortical activity. A selective review. *Human Neurobiol.* **2:** 177–189.

Liddle, P. F., Friston, K. J., Frith, C. D., and Frackowiak, R. S. J. 1992. Cerebral blood flow and mental processes in schizophrenia. *J. R. Soc. Med.* **85:** 224–227.

McColl, J. H., Holmes, A. O., and Ford, I. 1994. Statistical methods in neuroimaging with particular application to emission tomography. *Statistical Methods Med. Res.* **3:** 63–86.

Poline, J. B., and Mazoyer, B. M. 1994a. Enhanced detection in activation maps using a multifiltering approach. *J. Cerebral Blood Flow Metab.* **14:** 690–699.

Poline, J. B., and Mazoyer, B. M. 1994b. Analysis of individual brain activation maps using hierarchical description and multiscale detection. *IEEE Trans. Med. Imaging* **13**(4): 702–710.

Raichle, M. E., Martin, W. R. W., Herscovitch, P., Mintun, M. A., and Markham, J. 1983. Brain blood flow measured with intravenous $H_2{}^{15}O$ II. Implementation and validation. *J. Nuclear Med.* **24:** 790–798.

Roland, P. E., Levin, B., Kawashima, R., and Ackerman, S. 1993. Three dimensional analysis of clustered voxels in ^{15}O-butanol brain activation images. *Human Brain Mapping* **1:** 3–19.

Satterthwaite, F. E., 1946. An approximate distribution of estimates of variance components. *Biometrics* **2:** 110–114.

Talairach, J., and Tournoux, P. 1988. *Co-planar Stereotactic Atlas of the Human Brain: 3-Dimensional Proportional System: An Approach to Cerebral Imaging.* Thieme Verlag, Stuttgart/New York.

Talbot, J. D., Marrett, S., Evans, A. C., Meyer, E., Bushnell, M. C., and Duncan, G. H. (1991). Multiple representations of pain in human cerebral cortex. *Science* **251:** 1355–1358.

Worsley, K. J., Evans, A. C., Marrett, S., and Neelin, P. 1992. A three dimensional statistical analysis for CBF activation studies in human brain. *J. Cerebral Blood Flow Metab.* **12:** 900–918.

Worsley, K. J., and Friston, K. J. 1995. Analysis of fMRI time-series revisited—Again. *NeuroImage* **2:** 173–181.

Worsley, K. J., and Vandal, A. C. 1994. *Quadratic Tests for Local Changes in Random Fields with Applications to Medical Images.* Technical Report 94-08, Department of Mathematics and Statistics, McGill University.