

# A General Statistical Analysis for fMRI Data

K. J. Worsley,\*† C. H. Liao,\* J. Aston,\*†‡ V. Petre,† G. H. Duncan,§ F. Morales,|| and A. C. Evans†

\*Department of Mathematics and Statistics and †Montreal Neurological Institute, McGill University, Canada; ‡Imperial College, London, United Kingdom; §Centre de Recherche en Sciences Neurologiques, Université de Montréal, Canada; and ||Cuban Neuroscience Center, Havana, Cuba

Received December 6, 2000

**We propose a method for the statistical analysis of fMRI data that seeks a compromise between efficiency, generality, validity, simplicity, and execution speed. The main differences between this analysis and previous ones are: a simple bias reduction and regularization for voxel-wise autoregressive model parameters; the combination of effects and their estimated standard deviations across different runs/sessions/subjects via a hierarchical random effects analysis using the EM algorithm; overcoming the problem of a small number of runs/session/subjects using a regularized variance ratio to increase the degrees of freedom.**

© 2002 Elsevier Science

**Key Words:** fMRI; hemodynamic response function; linear regression; random effects; EM algorithm; bias reduction.

## 1. INTRODUCTION

Many methods are available for the statistical analysis of fMRI data that range from a simple linear model for the response and a global first-order autoregressive model for the temporal errors (SPM'99), to a more sophisticated nonlinear model for the response with a local state space model for the temporal errors (Purdon *et al.*, 2001). We have tried to strike a compromise that is fast, efficient, general and valid. The main novel features of our method are: (1) a method for reducing the bias in estimating the parameters of an autoregressive model for the errors; (2) combining the results from separate runs/sessions/subjects using a simple random effects model estimated by the EM algorithm; (3) modifications to the EM algorithm to reduce bias and increase execution speed; (4) spatial regularization of the random effect to increase the degrees of freedom; (5) careful use of caching to improve the overall execution speed of the computer code.

The result is a set of four MATLAB functions, collectively known as *fmristar*, which are freely available on the web at <http://www.math.mcgill.ca/keith/fmristar>. The first, *fmridesign*, sets up the design matrix for a linear model by convolving the

external stimulus with a pre-specified hemodynamic response function (HRF). The second, *fmrilm*, fits the linear model to a single run of fMRI data allowing for spatially varying autocorrelated errors. It can also estimate the voxel-wise delay of the HRF, but this will be described elsewhere (Liao *et al.*, 2000). The third, *multistat*, combines the output from separate analyses of *fmrilm* from different runs within a session, different sessions on the same subject, and across subjects within a population, using a type of random effects analysis. Finally, *tstat\_threshold* sets the threshold for peak and cluster size detection using random field theory (Worsley *et al.*, 1996; Cao, 1999). The analysis is presented schematically in Fig. 1.

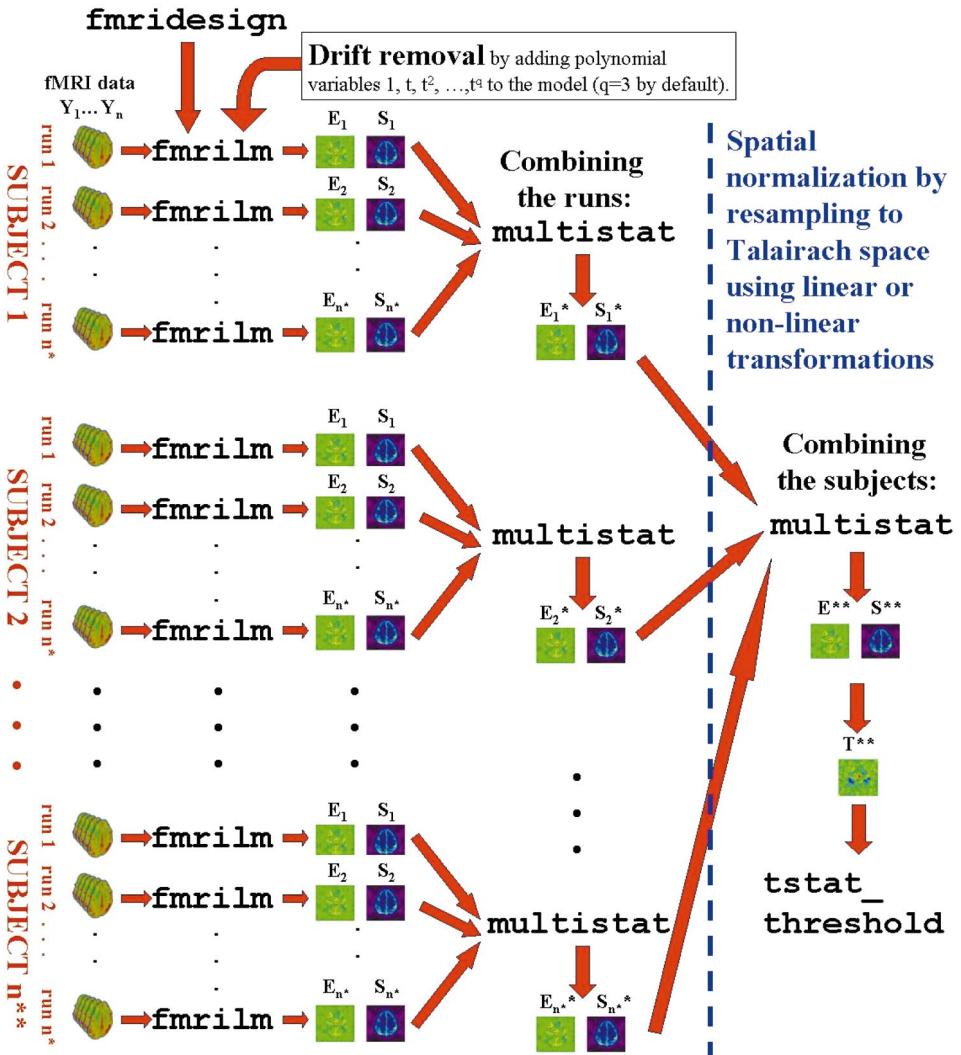
Note that spatial normalization to register the images to an atlas standard, such as Talairach coordinates, is only performed on the output from each session (since the subject will have moved), not on the original data nor the output from each run (since the subject is presumably still in the same position in the scanner). This reduces the amount of normalization considerably. Motion correction during the run is performed by different software in our lab.

This paper is organized as follows. Section 2 defines the linear model for fMRI data, then the statistical analysis of a single run is presented in Section 3. Results from multiple runs, sessions and subjects are analysed in Section 4. Peak and cluster size detection are discussed in Section 5. The methods are applied to a data set in Section 6, and some discussion and concluding comments are given in Sections 7 and 8.

## 2. DESIGNING THE STATISTICAL MODEL

### 2.1. Modeling the Signal

The first step in the statistical analysis of fMRI data is to build a model of how the data responds to an external stimulus. Suppose the (noise-free) fMRI response at a particular voxel at time  $t$  is given by  $x(t)$ , and the external stimulus is given by  $s(t)$ . The corresponding fMRI response is not instantaneous; there is a blurring and a delay of the peak response by about



**FIG. 1.** Fmristat flow chart for the analysis of several runs (only one session per subject);  $E$  = effect,  $S$  = standard deviation of effect,  $T = E/S = T$  statistic.

6 s. The simplest way of capturing this is to assume that the fMRI response depends on the external stimulus by convolution with a hemodynamic response function  $h(t)$  as follows:

$$x(t) = \int_0^\infty h(u)s(t-u)du. \quad (1)$$

Several models for  $h(t)$  have been proposed. A simple one is a gamma function (Lange and Zeger, 1997) or a difference of two gamma functions to model the slight intensity dip after the response has fallen back to zero (Friston *et al.*, 1998). An example is the HRF available in SPM'96:

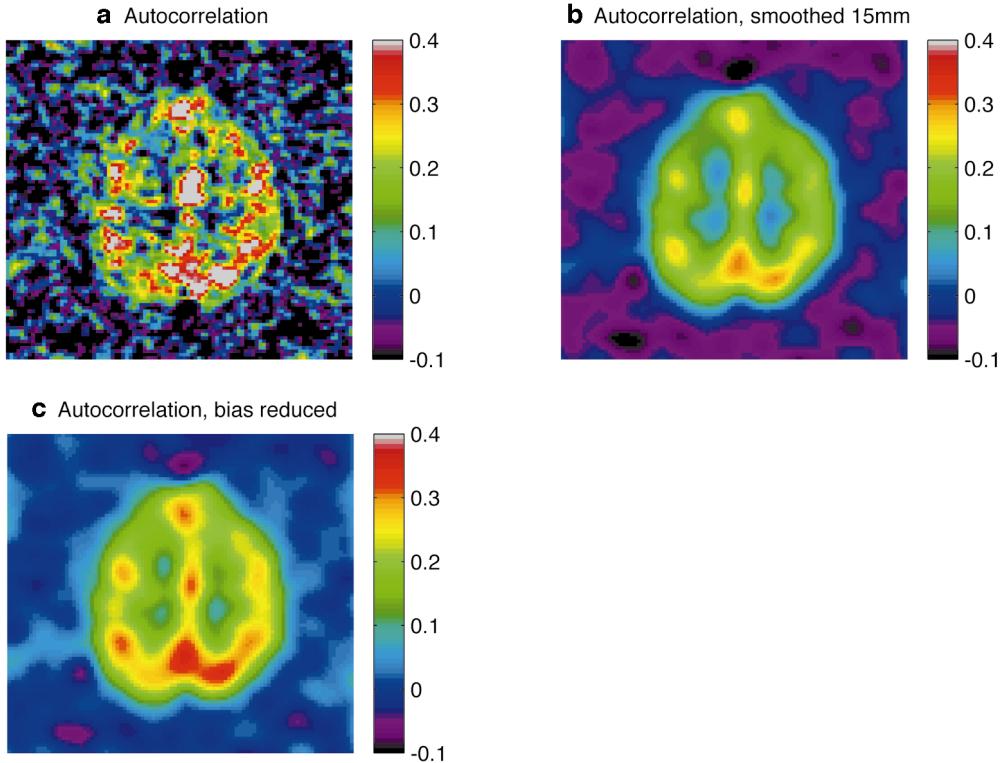
$$h(t) = (t/d_1)^{a_1} \exp(-(t-d_1)/b_1) - c(t/d_2)^{a_2} \exp(-(t-d_2)/b_2), \quad (2)$$

where  $t$  is time in seconds,  $d_j = a_j b_j$  is the time to the peak, and  $a_1 = 6$ ,  $a_2 = 12$ ,  $b_1 = b_2 = 0.9$  seconds, and  $c = 0.35$  (Glover, 1999). This is then subsampled at the  $n$  scan acquisition times  $t_1, \dots, t_n$  to give the response  $x_i = x(t_i)$  at scan  $i$ .

The combined effect of  $k$  different stimuli types on data in scan  $i$ , denoted by  $k$  different responses  $x_{i1}, \dots, x_{ik}$  is often assumed to be additive but with different multiplicative coefficients  $\beta_1, \dots, \beta_k$  that vary from voxel to voxel. The combined fMRI response is modeled as the linear model (Friston *et al.*, 1995)

$$x_{i1}\beta_1 + \dots + x_{ik}\beta_k.$$

Some voxels in fMRI time series data show considerable drift over time. Drift can be either linear, or a more general slow variation. If drift is not removed then it can either be confounded with the fMRI re-



**FIG. 2.** Estimation of the autocorrelation  $\rho$ : (a) unsmoothed, no bias reduction; (b) smoothed, no bias reduction; (c) smoothed, reduced bias. Note the biased autocorrelation of  $\approx -0.05$  outside the brain in (b), which goes to  $\approx 0$  after bias reduction in (c).

sponse, particularly if the stimuli vary slowly over time, or it can add to the estimate of the random noise. The first causes bias in estimates of the effect of the stimuli, the second causes bias in the estimate of the error of the estimated effect. Drift can be removed either by high-pass filtering or by introducing low frequency drift terms, such as cosines, polynomials, or splines, into the linear model. The cosine transform basis functions used in SPM'99 have zero slope at the ends of the sequence, which is not realistic, so we use a polynomial drift. All of these can be written as extra “responses”  $x_{i,k+1}, \dots, x_{i,m}$  at time  $i$ , and added to the linear model. For example, a polynomial drift of order  $q$  can be removed by adding to the linear model  $x_{i,k+1+j} = t_j^i$ ,  $j = 0, 1, \dots, q$ ,  $m = k + 1 + q$ . Finally a random error  $\epsilon_i$  is added to obtain the observed fMRI data,  $Y_i$ , at time index  $i$ :

$$\begin{aligned} Y_i &= \underbrace{x_{i1}\beta_1 + \dots + x_{ik}\beta_k}_{\text{fMRI response}} + \underbrace{x_{i,k+1}\beta_{k+1} + \dots + x_{im}\beta_m}_{\text{drift}} + \epsilon_i \\ &= \mathbf{x}'_i \boldsymbol{\beta} + \epsilon_i, \end{aligned} \quad (3)$$

where  $\mathbf{x}_i = (x_{i1}, \dots, x_{im})'$  and  $\boldsymbol{\beta} = (\beta_{i1}, \dots, \beta_{im})'$ .

## 2.2. Modeling the Noise

The errors are not independent in time; typically the correlation between scans 3 s apart can be as high as 0.4, particularly in cortical regions (Fig. 2c). Using least squares model fitting to estimate each  $\beta$ , and neglecting to take the correlation structure into account, or using the wrong correlation model, can cause biases in the estimated error of the estimates of  $\beta$ , but it does not bias the estimate of  $\beta$  itself, which will always be unbiased though perhaps less efficient (i.e., more variable). The simplest model of the temporal correlation structure is the first order autoregressive model (Bullmore *et al.*, 1996), in which the scans are equally spaced in time and we suppose that the error from the previous scan is combined with fresh noise to produce the error for the current scan:

$$\epsilon_i = \rho \epsilon_{i-1} + \xi_i,$$

where  $|\rho| < 1$  and  $\xi_i$  is a sequence of independent and identically distributed normal random variables with mean 0 and standard deviation  $\sigma_1$ , i.e.,  $\xi_i \sim N(0, \sigma_1^2)$  (known as “white noise”). The resulting autocorrelation at lag 1 is

$$\text{Cor}(\epsilon_i, \epsilon_{i-1}) = \rho^{|I|}.$$

This can be extended to autoregressive models of order  $p$ , denoted by AR( $p$ ), specified by

$$\epsilon_i = \alpha_1 \epsilon_{i-1} + \alpha_2 \epsilon_{i-2} + \dots + \alpha_p \epsilon_{i-p} + \xi_i,$$

in which the autocorrelation structure is more complex, including oscillatory terms as well as exponential decay. To take into account white noise from the scanner, Purdon *et al.* (2001) has extended the AR(1) model to add a second independent white noise term. This is a special type of state space model which are extremely powerful at capturing complex dynamic relationships, including drift.

### 2.3. *Fmrideresign*

The MATLAB function *fmrideresign* takes as input the stimuli  $s(t)$  (defined by onset time, duration and height) and the parameters of the HRF (2), calculates the responses  $x(t)$ , then samples them at the slice acquisition times to produce a different set of regressor variables  $\mathbf{x}_i$  for each slice. Our analysis uses regressors tailored to each slice whereas SPM simply augments the regressors to include extra covariates that model slight shifts in time induced by sequential slice acquisition. In other words, *fmrideresign* specifies a single but slice specific regressor, whereas SPM augments the design matrix to include an extra regressor that retains the same design matrix for all voxels. This is then used as input to *fmrilm* to be described next.

## 3. LINEAR MODELS ANALYSIS OF A SINGLE RUN

### 3.1. *Estimating the Signal Parameters*

Worsley and Friston (1995) proposed estimating  $\beta$  in (3) by least squares, then correcting the inference for correlated errors. A general unbiased estimator of  $\beta$  can be found by first multiplying the vector of observations  $\mathbf{Y} = (Y_1, \dots, Y_n)'$  and its model (3) by an  $n \times n$  matrix  $\mathbf{A}$  then estimating the parameters by least squares. The fully efficient (most accurate, i.e., minimum variance) estimator of  $\beta$  is obtained by choosing  $\mathbf{A} = \mathbf{V}^{-1/2}$ , where  $\mathbf{V}\sigma^2$  is the  $n \times n$  variance matrix of  $\mathbf{Y}$ , so that the variance of the resulting errors is proportional to the identity matrix, equivalent to “whitening” the errors. SPM'99 chooses  $\mathbf{A}$  to suppress both high and low frequencies. The removal of low frequencies implements an approximate whitening of the data and removal of high frequencies renders the resulting inference more robust to misspecification of the correlation structure (Friston *et al.*, 2000). SPM'99 adopts an AR(1) model for the original data, whose parameter  $\rho$  is assumed to be constant across the brain, rather than varying spatially as proposed in our method. Although there is considerable evidence that  $\rho$  is not constant (see Fig. 2 and Purdon *et al.*, 2001), the robustness

conferred by temporal smoothing appears to alleviate this problem.

Here we propose the fully efficient, pre-whitening strategy with a spatially varying  $\mathbf{A} = \mathbf{V}^{-1/2}$ . Doing this in practice can be very time consuming if it is repeated at every voxel. Fortunately there are computationally efficient ways of finding  $\mathbf{V}^{-1/2}$  if the errors are generated by an AR( $p$ ) process (see Appendix A.3) or a state space model (using the Kalman filter). For the AR(1) model, for example, pre-whitening the data and regressors is accomplished by

$$\begin{aligned}\tilde{Y}_1 &= Y_1, & \tilde{Y}_i &= (Y_i - \rho Y_{i-1})/\sqrt{1 - \rho^2}, \\ \tilde{\mathbf{x}}_1 &= \mathbf{x}_1, & \tilde{\mathbf{x}}_i &= (\mathbf{x}_i - \rho \mathbf{x}_{i-1})/\sqrt{1 - \rho^2},\end{aligned}\quad (4)$$

$i = 2, \dots, n$ , so that the model becomes

$$\tilde{Y}_i = \tilde{\mathbf{x}}_i'\beta + \xi_i \quad (5)$$

where  $\xi_i \sim N(0, \sigma^2)$  independently and  $\sigma^2 = \sigma_1^2/(1 - \rho^2)$ . Denoting the pseudoinverse by  $+$ , the least squares estimator of  $\beta$  in (5) is

$$\hat{\beta} = \tilde{\mathbf{X}}^+\tilde{\mathbf{Y}}, \quad (6)$$

where  $\tilde{\mathbf{X}} = (\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_n)'$  is the transformed design matrix, and  $\tilde{\mathbf{Y}} = (\tilde{Y}_1, \dots, \tilde{Y}_n)'$  is the transformed observation vector.

### 3.2. *Estimating the Noise Parameters*

The vector of residuals  $\mathbf{r} = (r_1, \dots, r_n)'$  is

$$\mathbf{r} = \tilde{\mathbf{Y}} - \tilde{\mathbf{X}}\hat{\beta} = \mathbf{R}\mathbf{Y}, \quad \mathbf{R} = \mathbf{I} - \tilde{\mathbf{X}}\tilde{\mathbf{X}}^+.$$

If  $\mathbf{V}$  is known, then  $\sigma^2$  is estimated unbiasedly by

$$\hat{\sigma}^2 = \mathbf{r}'\mathbf{r}/\nu, \quad (7)$$

where  $\nu = n - \text{rank}(\tilde{\mathbf{X}})$  is the degrees of freedom, equal to  $n - m$  if the model is not over parameterized.

We now turn to estimating the parameters in the correlation structure  $\mathbf{V}$ . We first estimate  $\beta$  by least squares for the original untransformed data (3), that is with  $\mathbf{A}$  equal to the identity matrix. While not fully efficient, this estimator is unbiased, so that the residuals have zero expectation and a correlation structure approximately equal to that of the errors  $\epsilon$ . The parameter of the AR(1) model can then be estimated by the sample autocorrelation of the residuals

$$\hat{\rho} = \sum_{i=2}^n r_i r_{i-1} / \sum_{i=1}^n r_i^2.$$

Replacing  $\rho$  in (4) by the reduced bias, regularized  $\hat{\rho}$  (see Section 3.3 below) produces approximately whitened data that are used to produce the final estimators of  $\beta$  and  $\sigma$  from (6) and (7). This could be iterated by reestimating  $\mathbf{r}$ ,  $\rho$ , then  $\beta$ , but in practice this does not seem to give much improvement.

### 3.3. Bias Reduction and Regularization of the Autocorrelations

A slight bias creeps into the estimator  $\hat{\rho}$  due to the correlation of the residuals induced by removing an estimated linear effect from the observations. This is because the variance matrix of the residuals is  $\mathbf{R}\mathbf{V}\mathbf{R}\sigma^2$ , which is not the same as the variance matrix of the errors  $\mathbf{V}\sigma^2$ . Typically  $\hat{\rho}$  is about 0.05 lower than expected.

Note that SPM'99 estimates a global value of  $\rho$  based on pooling the sample autocorrelations of the raw data after global normalization, so this problem is avoided. This is at the expense of introducing another potential source of bias due to the fact that signal and drift terms, not removed by global normalisation, have not been removed. However, only a small fraction of the voxels should show signal or drift, so when the autocorrelation is pooled across voxels this source of bias will be reduced. Bias due to using a global instead of a local estimate of  $\rho$  still remains, but this bias is reduced by the band-pass filtering strategy as noted in Section 3.1.

One method of obtaining less biased voxel-wise variance parameters is to use restricted or residual maximum likelihood (REML) estimates (Harville, 1974), but this involves costly iterations. Instead we propose a simple, fast, non-iterative method based on equating sample autocovariances to their approximate expectations. Details are given in Appendix A.1. Inserting the resulting autocorrelations into the Yule–Walker equations gives estimates of the AR( $p$ ) parameters, which were shown by simulation to have much reduced bias. However, before doing this, we propose regularizing the autocorrelations as described next.

In practice the estimated autocorrelation parameters vary considerably about their true values. Even when no correlation is present, it is well known that the standard deviation of the sample autocorrelation  $\hat{\rho}$  is about  $1/\sqrt{n}$ , or about 0.1 for typical runs of  $n = 100$  scans. This is quite large compared to the true correlations that range from 0 to 0.4. Some method of regularization that reduces this variability seems desirable. Purdon *et al.* (1998) achieve this by spatial smoothing of the likelihood before estimating the autocorrelation parameters. Instead we propose the simpler method of spatial smoothing of the (reduced bias) sample autocorrelations. This reduces the variability considerably, at the cost of a slight increase in bias.

The extension of these ideas to AR( $p$ ) models is treated in Appendix A.

### 3.4. Inference for Effects

An effect of interest, such as a difference between stimuli, can be specified by  $\mathbf{c}\beta$ , where  $\mathbf{c}$  is a row-vector of  $m$  contrasts. It is estimated by

$$E = \mathbf{c}\hat{\beta}, \quad (8)$$

with estimated standard deviation

$$S = \|\mathbf{c}\tilde{\mathbf{X}}^+\|^2\hat{\sigma}. \quad (9)$$

To detect the effect, we test the null hypothesis that the effect is zero. The test statistic is the  $T$  statistic

$$T = E/S,$$

which has an approximate  $t$  distribution with  $v$  degrees of freedom (exact if  $\mathbf{V}$  is known) when there is no effect ( $\mathbf{c}\beta = 0$ ). To detect more than one effect at the same time, that is, if  $\mathbf{c}$  is a matrix with rank  $k$ , the  $T$  statistic is replaced by an  $F$  statistic defined by

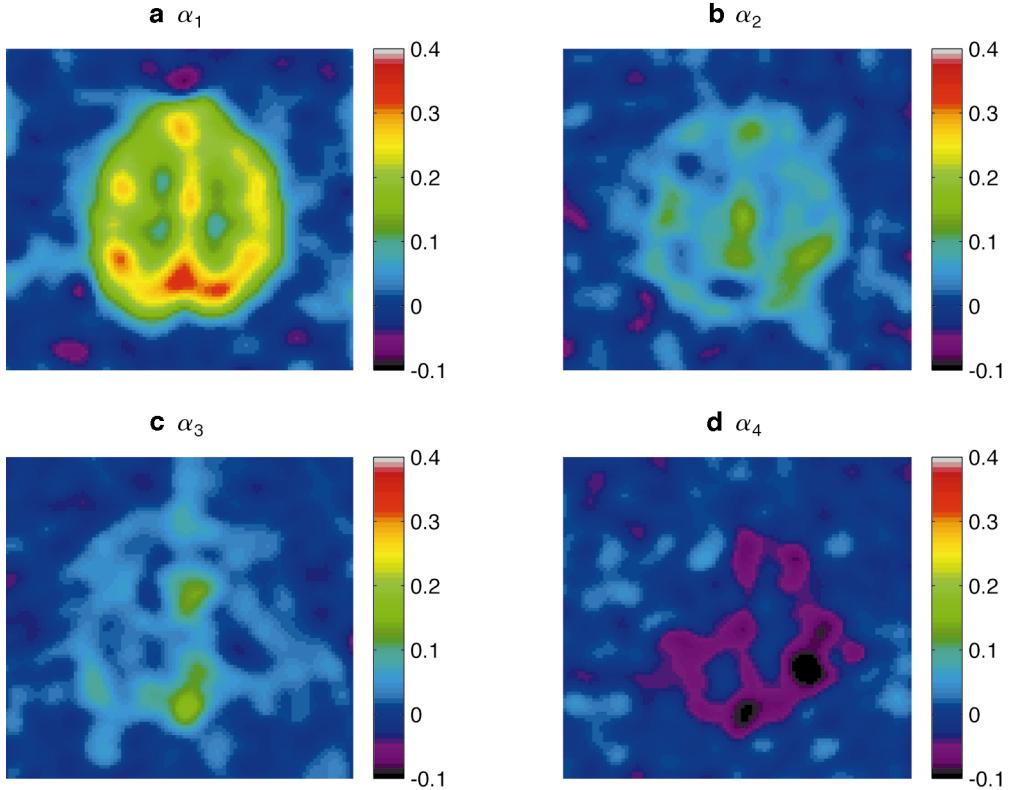
$$F = E'(\mathbf{c}\tilde{\mathbf{X}}^+(\mathbf{c}\tilde{\mathbf{X}}^+)^{-1})^{-1}E/(k\hat{\sigma}^2),$$

which has an approximate  $F$  distribution with  $k$  and  $v$  degrees of freedom (exact if  $\mathbf{V}$  is known) when there are no effects.

### 3.5. *fmrilm*

The methods described in this section have been implemented in the MATLAB function *fmrilm* (*fMRI* linear model). This takes as input the design matrices from *fmridesign* and adds polynomial trends. Note that the polynomial trends are not smoothed by the hemodynamic response function, so that the constant term codes for a rest or baseline condition that is supposed to exist before scanning started. Coding for a baseline “stimulus” is not necessary unless the baseline is a separate stimulus, such as a cross on a computer screen, that was first presented the moment that scanning commenced. If this is the case, the constant term (in the drift) codes for the small amount of background effect that is carried over by the HRF from before scanning started, and should be ignored.

In a first pass through the data, the reduced bias  $\hat{\rho}$  is calculated from the least-squares residuals of the untransformed model. This is regularized by spatial smoothing with a 15 mm (default) FWHM Gaussian kernel. In a second pass through the data, the fMRI data and design matrix are whitened using  $\hat{\rho}$



**FIG. 3.** Estimation of the AR(4) parameters, smoothed, reduced bias: (a)  $\hat{\alpha}_1$ , (b)  $\hat{\alpha}_2$ , (c)  $\hat{\alpha}_3$ , (d)  $\hat{\alpha}_4$ . The most important parameter is obviously the first, but the presence of some anatomical details in the higher order parameters suggests that they are not identically zero.

and the final estimates of  $\beta$  and  $\sigma$  are calculated, then effects  $E$ , their standard errors  $S$ , and  $T$  and  $F$  test statistics. Note that calculation of  $S$  by (9) avoids further matrix inversion. Note also that these steps give the right answers even if the design matrix has less than full rank (i.e., the model is over-parameterized).

Computations were reduced as follows. Note that in this second pass, a separate pseudoinverse must be calculated at each voxel, because the whitened design matrix depends on the autocorrelation parameter  $\hat{\rho}$ . To cut down computations,  $\hat{\rho}$  was rounded to the nearest 0.01 from  $-1$  to  $1$ , and a separate pseudo-inverse was calculated and cached for each rounded  $\hat{\rho}$  in the slice. This reduced the number of pseudo-inverse calculations by a factor of about 100. Unfortunately this trick does not work for AR(2) models or higher because two or more parameters are involved. However the whitening step was efficiently coded by a very small number of MATLAB commands (see Appendix A.3).

The execution time of `fmrilm` is about 6 min for  $118$  scans  $\times$   $13$  slices  $\times$   $128 \times 128$  pixels per slice on a  $450$  MHz Pentium, but  $22$  minutes for fitting an AR(2) model, increasing by  $4$  min per extra order of the AR( $p$ ) model, to over  $50$  min for an AR(8) model.

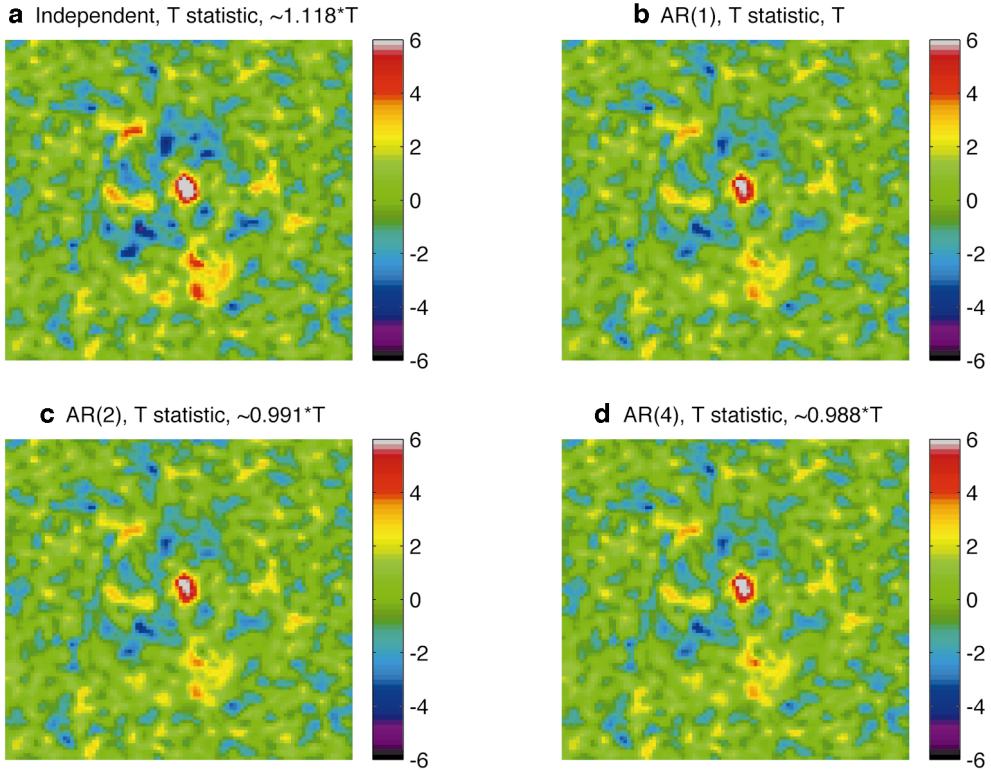
#### 4. COMBINING MULTIPLE RUNS, SESSIONS, OR SUBJECTS

##### 4.1. The Random Effects Model

The next step is to combine the effects and their standard errors from different runs in the same session, then different sessions on the same subject, then different subjects from a population. This makes our inference applicable to a population, rather than just a single case. Let us start by combining  $n^*$  runs; the combination of sessions and subjects is similar. Let  $E_j$  denote the effect on run  $j$  from (8), and let  $S_j$  denote its standard deviation from (9),  $j = 1, \dots, n^*$ .

Besides simply averaging the effects, we may also wish to compare one set of runs with another under different experimental conditions, or we may wish to remove a trend in the effects taken over time. This can be simply captured by another vector of regressor variables denoted by  $\mathbf{z}_j$  measured on run  $j$ . We then assume that  $E_j$  is a random variable whose expectation is a linear model  $\mathbf{z}'\gamma$  in the regressor variables.

In order to extend our inference to randomly sampled runs, rather than further scans within the same run, we assume that the effect has two components of variability. The first is the variance  $S_j^2$  from further scans within the same run, regarded as fixed. The



**FIG. 4.**  $T$  statistics for the hot–warm contrast by fitting different order AR( $p$ ) models: (a) independent errors ( $p = 0$ ), (b)  $p = 1$ , (c)  $p = 2$ , (d)  $p = 4$ . The  $T$  statistics for the AR(1), AR(2), and AR(4) models are almost identical, differing by less than 1%, suggesting that the AR(1) model is adequate for this data. In contrast, assuming independent errors gives a  $T$  statistic that is 12% too large.

second is a random effects variance  $\sigma_{\text{random}}^2$  from the variability of the expected  $E_j$  (found by infinitely long runs) from run to run. The random effects model (strictly speaking a mixed effects model) is then

$$E_j = \mathbf{z}'\gamma + \eta_j \quad (10)$$

where  $\eta_j$  is normally distributed with zero mean and variance  $S_j^2 + \sigma_{\text{random}}^2$  independently for  $j = 1, \dots, n^*$ . Note that setting  $\sigma_{\text{random}}^2 = 0$  produces a fixed effects analysis.

#### 4.2. Estimating the Parameters

If  $\sigma_{\text{random}}^2$  were known, then the best estimator of  $\gamma$  is obtained by weighted least squares with weights inversely proportional to  $S_j^2 + \sigma_{\text{random}}^2$ . The fixed effects estimates would use weights inversely proportional to  $S_j^2$ , which was the basis of our first version of multi-stat (Worsley *et al.*, 2000). If  $\sigma_{\text{random}}^2$  is unknown then we can estimate  $\gamma$  and  $\sigma_{\text{random}}^2$  by restricted maximum likelihood (REML) (Harville, 1974), which gives less biased variance estimates than maximum likelihood (ML). However, the fact that the error variances add means that there is no analytic expression for the

REML or ML estimators, so we must resort to iterative algorithms. Stability is more important than speed of convergence since it is not easy to monitor the algorithm at each voxel, so we have chosen the EM algorithm, which always produces positive estimates of variance components. The details are given in Appendix B.

When the random effect is small, the EM algorithm is known to be slow to converge and it produces severely biased estimates. This is because the EM algorithm always returns a positive estimate of  $\sigma_{\text{random}}^2$  even when the true value is zero. To overcome this, we subtract  $S_{\min}^2 = \min_j S_j^2$  from each  $S_j^2$ , which gives an equivalent model for  $E_j$  with variance

$$(S_j^2 - S_{\min}^2) + (\sigma_{\text{random}}^2 + S_{\min}^2).$$

We now use the EM algorithm to estimate the unknown parameter  $\sigma_{\text{random}}^{*2} = (\sigma_{\text{random}}^2 + S_{\min}^2)$ . The starting values are the least squares estimates assuming that  $S_j = 0$ . Convergence is faster because  $\sigma_{\text{random}}^{*2}$  is further from zero than  $\sigma_{\text{random}}^2$ ; 10 iterations appear to be enough. Finally we subtract  $S_{\min}^2$  from  $\sigma_{\text{random}}^{*2}$  to get an estimator of  $\sigma_{\text{random}}^2$  that appears to be almost unbiased for simulated data.

TABLE 1

The Effective Degrees of Freedom  $\tilde{\nu}^*$  for an Experiment with  $n^* = 4$  runs ( $\nu^* = 3$  degrees of freedom), of  $n = 118$  Scans ( $\nu = 112$  df), and  $w_{\text{effect}} = 6$  mm Effective FWHM

	Random effects analysis	Variability $\leftrightarrow$ Bias						Fixed effects analysis
FWHM of smoothing filter $w_{\text{ratio}}$ (mm)	0	5	10	15	20	25	$\infty$	
Effective degrees of freedom $\tilde{\nu}^*$	3	11	45	112	192	264	448	

#### 4.3. Regularizing the Random Effects Variance

A serious problem with the above analysis is that the number of runs, sessions or subjects is usually small and so the estimate of  $\sigma_{\text{random}}^2$  is highly variable, which results in a low effective degrees of freedom  $\nu^*$  for the estimated variance  $\widehat{\text{Var}}(\hat{\gamma})$ . For example, if all standard deviations  $S_j$  were equal, so that the variance of each effect  $E_j$  is an unknown but equal constant, then it can be shown that the degrees of freedom is  $\nu^* = n^* - p^*$ , where  $p^* = \text{rank}(\mathbf{Z})$  and  $\mathbf{Z} = (\mathbf{z}_1 \cdots \mathbf{z}_{n^*})'$ . In our application, the  $n^* = 4$  runs were averaged to give  $\nu^* = 3$  df. This is particularly serious when detecting local maxima by thresholding an image of  $T$  statistics, because the corrected threshold can become extremely high, making it very difficult to detect a signal that is really present (Worsley *et al.*, 1996).

Notice that the image of  $\hat{\sigma}_{\text{random}}^2$  (Fig. 6a) is very noisy (due to the low degrees of freedom) but that it does contain substantial anatomical structure. This same anatomical structure is also evident in the image of the fixed effects standard deviation,  $\sigma_{\text{fixed}}^2$  (Fig. 6b), obtained as

$$\sigma_{\text{fixed}}^2 = \sum_j \nu_j S_j^2 / \nu_{\text{fixed}},$$

where  $\nu_j$  is the degrees of freedom of  $S_j$  and  $\nu_{\text{fixed}} = \sum_j \nu_j$  is the degrees of freedom of the fixed effects analysis. Notice that the image of the fixed effects standard deviation has much less noise since it has high degrees of freedom coming from the large number of scans for each run. We can therefore use the fixed effects standard deviation as a template for a better estimate of the random effects standard deviation. To see this, consider the ratio of the two:

$$\tau^2 = \frac{\sigma_{\text{fixed}}^2 + \hat{\sigma}_{\text{random}}^2}{\sigma_{\text{fixed}}^2},$$

which has much less anatomical structure (Fig. 6d). We could therefore reduce the noise in  $\tau^2$  by spatial smoothing (Fig. 6e), then multiply back by the fixed effects standard deviation to get an improved estimator of the random effects standard deviation that has higher effective degrees of freedom (Fig. 6c):

$$\begin{aligned} \hat{\sigma}_{\text{random}}^2 &= \text{smooth}(\tau^2) \times \sigma_{\text{fixed}}^2 - \sigma_{\text{fixed}}^2 \\ &= \text{smooth}\left(\frac{\hat{\sigma}_{\text{random}}^2}{\sigma_{\text{fixed}}^2}\right) \times \sigma_{\text{fixed}}^2. \end{aligned} \quad (11)$$

This can then be used in place of  $\hat{\sigma}_{\text{random}}^2$  to get a better estimator of  $\gamma$  and its standard deviation.

Note that no smoothing gives a purely random effects analysis, whereas smoothing with an infinitely wide filter, so that the variance ratio is 1 everywhere, produces a purely fixed effects analysis. The desired amount of smoothing can be determined by the desired degrees of freedom, which we look at next.

#### 4.4. Estimating the Degrees of Freedom

By smoothing the variance ratio image  $\tau^2$  we decrease its variability at the expense of a possible slight increase in bias. The actual decrease in variability is measured by its effective degrees of freedom. To get a rough estimate of this, suppose that the FWHM of the filter is  $w_{\text{ratio}}$  and that the spatial correlation function of each  $E_j$  is a Gaussian function with FWHM  $w_{\text{effect}}$ . In Appendix C we show that the effective degrees of freedom of the regularized variance ratio is approximated by

$$\nu_{\text{ratio}} = \nu^* \left( 2 \left( \frac{w_{\text{ratio}}}{w_{\text{effect}}} \right)^2 + 1 \right)^{3/2}.$$

The final effective degrees of freedom of the standard deviation  $\tilde{\nu}^*$ , denoted by  $\tilde{\nu}^*$ , is estimated by

$$\frac{1}{\tilde{\nu}^*} = \frac{1}{\nu_{\text{ratio}}} + \frac{1}{\nu_{\text{fixed}}}. \quad (12)$$

Thus the  $w_{\text{ratio}}$  parameter acts as a convenient way of providing an analysis midway between a random effects and a fixed effects analysis. Table 1 shows the degrees of freedom for some typical values of the amount of smoothing. Setting  $w_{\text{ratio}} = 0$  (no smoothing) produces a random effects analysis; setting  $w_{\text{ratio}} = \infty$ , which smoothes the variance ratio to one everywhere, produces a fixed effects analysis. In practice, we choose  $w_{\text{ratio}}$  to produce a final  $\tilde{\nu}^*$ , which is at least 100, so that

errors in its estimation do not greatly affect the distribution of test statistics.

Finally, effects defined by contrasts  $\mathbf{b}$  in  $\gamma$  can be estimated by  $E^* = \mathbf{b}\hat{\gamma}$  with standard deviation  $S^* = \sqrt{\mathbf{b}\text{Var}(\hat{\gamma})\mathbf{b}'}$ . The  $T$  statistic  $T^* = E^*/S^*$ , with a nominal  $\tilde{v}^*$   $df$ , can then be used to detect such an effect.

#### 4.5. Multistat

The analysis of this section is done by `multistat`. It takes as input the output effect  $E_j$  and standard deviation  $S_j$  images from `fmrilm`, together with their degrees of freedom  $v_j$ . If the standard deviation images are not provided then `multistat` assumes they are 0 with  $v_{\text{fixed}} = \infty$ . The FWHM of the input effects  $w_{\text{effect}}$  is supplied, usually estimated by the FWHM of the raw data. Finally the FWHM for smoothing the variance ratio  $w_{\text{ratio}}$  should be chosen to give at least 100  $df$  (the default is 15 mm).

The variance ratio  $\tau$  is smoothed in 3-D, then the final  $\hat{\gamma}$  and its estimated variance is obtained by weighted least squares using weights  $1/(S_j^2 + \hat{\sigma}_{\text{random}}^2)$  (see Appendix B). On rare occasions near scanner artifacts outside the brain, the weights can be negative; see Appendix D for the reason and a fix. Output is the effective degrees of freedom  $\tilde{v}^*$ , the effect  $E^*$  and its standard deviation  $S^*$ , which can be used as input to further runs of `multistat` as illustrated in Fig. 1. Run time is very quick, less than 1 min for averaging four runs on a 450 MHz Pentium.

## 5. DETECTING ACTIVATED REGIONS

The final step in the analysis is to detect the regions activated by the stimulus, or contrast in the stimuli. This part is almost identical to that employed by SPM. `Tstat_threshold` finds the threshold for the maximum of a  $T$  statistic image for a given  $P$  value, 0.05 by default. Input is the volume of the search region, the volume of one voxel, the FWHM of the data, and the degrees of freedom.

This calculation is fairly robust to the shape of the region so it assumes the search region is spherical, which gives a threshold that is a lower bound for regions with the same volume (Worsley *et al.*, 1996). Sometimes if the FWHM is small relative to the voxel volume, a Bonferroni reduction gives a lower threshold than the random field theory, so `tstat_threshold` takes the smallest of the two. The threshold is also robust against non-stationarity of the data (Worsley *et al.*, 1999), so the fact that a different design matrix is used for each slice, and different autocorrelations are used to pre-whiten the data, does not greatly affect the threshold.

Finally a critical cluster size for the  $T$  statistics greater than an input threshold is calculated using new results from random field theory (Cao, 1999). The

critical cluster size is sensitive to nonstationarity, so it must be used with caution. A better method, based on cluster resels (Worsley *et al.*, 1999), has not yet been implemented.

## 6. AN EMPIRICAL ILLUSTRATION

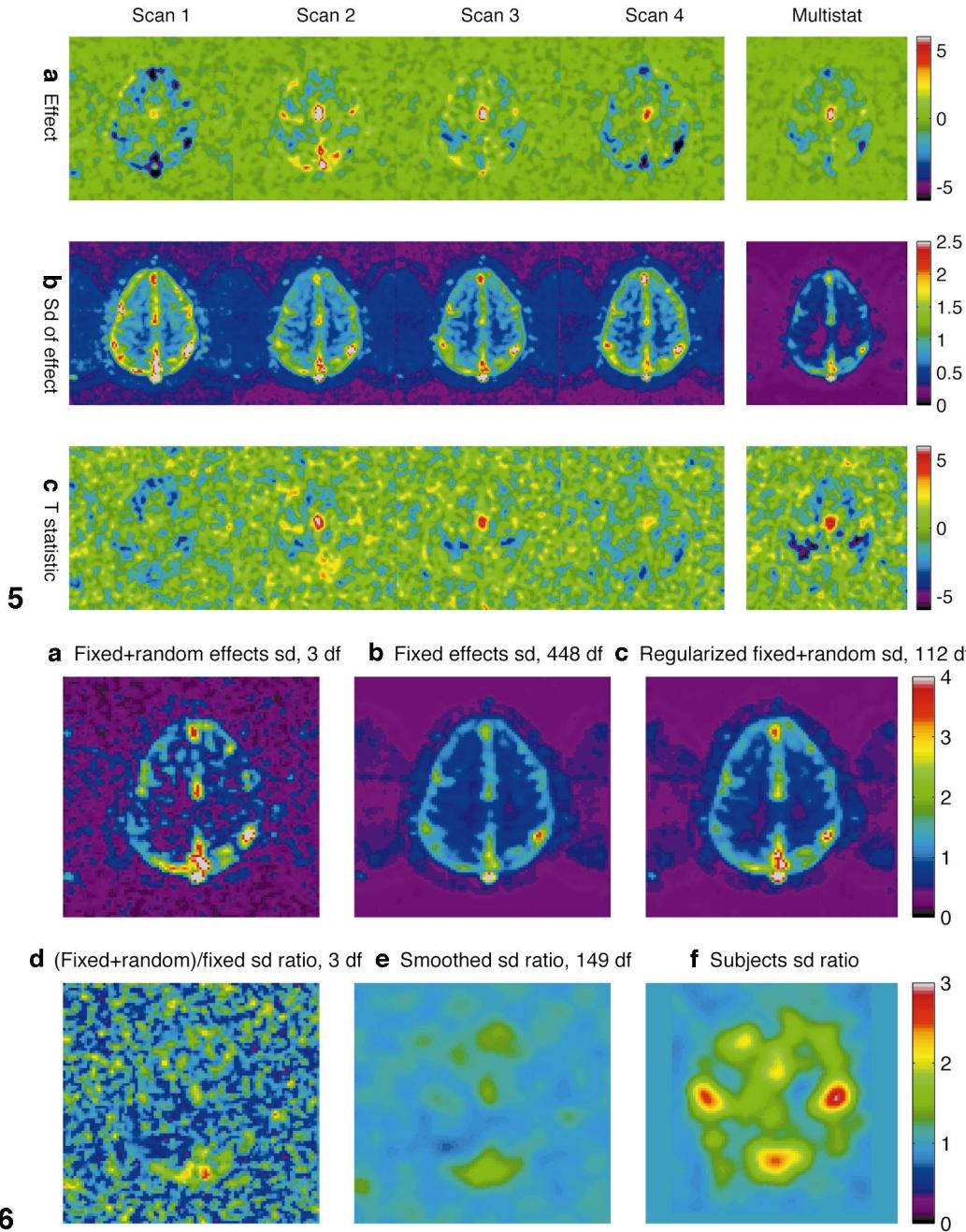
The methods described in this paper were applied to an fMRI experiment on pain perception (Chen *et al.*, 2000). After 9 s of rest, a subject was given a painful heat stimulus (49°C) to the left forearm for 9 s, followed by 9 s of rest, then a warm stimulus (35°C) for 9 s, repeated 10 times for 6 min in total. During this time the subject was scanned every TR = 3 s (120 scans) using a Siemens 1.5 T machine and 13 slices of 128 × 128 pixel BOLD images were obtained. The first two scans were discarded, leaving  $n = 118$  scans as data.

`Fmrider` was used to set up 13 different design matrices for the 13 slices, offset by the 0.12 s interleaved slice acquisition times. Two box functions were created for the hot and warm stimuli as described above, then convolved with the HRF (2) and sampled at the slice acquisition times.

`Fmrilm` was then used to analyze a single run with cubic drift removal ( $q = 3$ ) and 15 mm FWHM smoothing of the reduced bias autocorrelation parameter from the fitted AR(1) model. Execution time was 6 min on a 450 MHz Pentium. We are interested in a pain effect, so a contrast of  $\mathbf{c} = (1 -1 0 0 0 0)$  was used to compare the hot stimulus (1) with the warm stimulus (-1), ignoring the 4 coefficients of the cubic drift (0 values).

The two stages of smoothing and bias reduction for the autocorrelation are shown on one slice in Fig. 2. Without smoothing and bias reduction, the image is quite noisy ( $\pm 0.1$ ) but with clear evidence of higher correlations in cortical areas than in white matter or outside the brain (Fig. 2a). After smoothing (Fig. 2b) the noise is much reduced and there is now clear evidence of a negative bias of  $\approx -0.05$  outside the brain where the autocorrelation should be zero. This is successfully removed by bias reduction (Fig. 2c) to reveal correlations ranging from 0.2 to 0.4 in cortical areas.

`Fmrilm` was used to fit an AR(4) error model (30 min execution time), and the resulting 15 mm smoothed reduced bias parameter estimates of  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$  are shown in Fig. 3. The first coefficient  $\alpha_1$  clearly dominates, but the presence of nonzero values in the higher order parameters cannot be entirely ruled out, since some anatomical structure is clearly discernible particularly in  $\alpha_2$ . To see if this has any effect on the resulting  $T$  statistics for the pain effect, we calculated  $T$  statistics for AR( $p$ ) models with  $p = 1, 2, 4$ . As is evident in Fig. 4, there is almost no difference in the  $T$  statistics. Even for an AR(16) model, the  $T$  statistics were almost identical (not shown). However, if the correlation is ignored and the data are assumed to be independent then the  $T$  statistic is about 11% higher



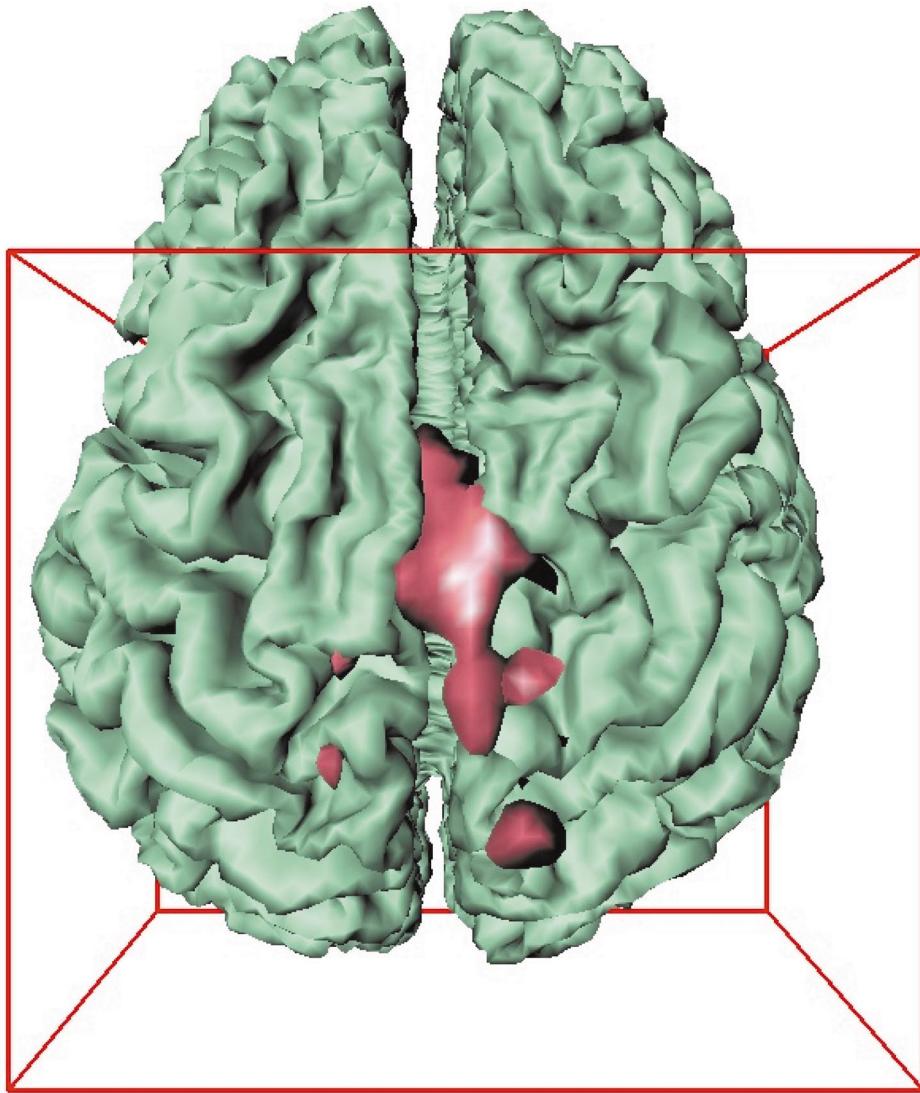
**FIG. 5.** Statistics from *fmrilm* for four different consecutive runs and their combination using *multistat* (left to right). (a) the estimated effect of the hot-warm stimulus  $E$ , (b) its estimated standard deviation  $S$ , (c) the  $T$  statistic  $T = E/S$ . The second run (second column) was used for Figs. 2–4.

**FIG. 6.** Regularizing the variance ratio: (a) the fixed + random effects standard deviation,  $\hat{\sigma}_{\text{fixed} + \text{random}} = \sqrt{\hat{\sigma}_{\text{fixed}}^2 + \hat{\sigma}_{\text{random}}^2}$  ( $3 \text{ df}$ ), (b) the fixed effects standard deviation  $\sigma_{\text{fixed}}$  ( $448 \text{ df}$ ), (c) the regularized fixed + random standard deviation  $\tilde{\tau}\sigma_{\text{fixed}}$  ( $\approx 112 \text{ df}$ ). (d) The ratio of the (fixed + random)/fixed standard deviations  $\tau = \hat{\sigma}_{\text{fixed} + \text{random}}/\sigma_{\text{fixed}}$  ( $\approx 3 \text{ df}$ ), (e) the smoothed standard deviation ratio  $\tilde{\tau} = \sqrt{\text{smooth}(\tau^2)}$  ( $\approx 149 \text{ df}$ ), (f)  $\tilde{\tau}$  for different subjects. Note that the standard deviation ratio (e) is roughly equal to one, indicating little evidence for a random effect over runs, except in frontal and anterior regions. However, there is strong evidence for a random effect over subjects (f).

than it should be (Fig. 4a). Thus the AR(1) model seems to be adequate for this experiment.

The experiment was repeated in  $n^* = 4$  consecutive runs separated by 18, 7, and 4 min, respectively. The run just analysed was in fact the second run, chosen to

illustrate a high pain effect. Using an AR(1) model, the resulting effects  $E_1-E_4$ , standard errors  $S_1-S_4$  and  $T$  statistics  $T_j = E_j/S_j$  with  $\nu = 118 - 6 = 112 \text{ df}$  are given in Fig. 5. There is some evidence for a pain effect in the central region, but the most striking feature is the



**FIG. 7.** Final thresholded  $T$ -statistic  $\tilde{T}^*$  for detecting the hot–warm stimulus by combining the four runs and thresholding at the  $P < 0.05$  threshold of 4.86 found using `tstat_threshold`.

considerable variability from run to run; the first run in particular seems to show very little evidence for a pain effect in this slice.

Multistat was used to average the effects using covariates  $\mathbf{z}_j = 1$  and the final effect  $E^*$ , standard deviation  $S^*$  and  $T$  statistic  $T^*$  are shown in the last column of Fig. 5. The method of regularizing the variance ratio is illustrated in Fig. 6. The random effects standard deviation (Fig. 6a) is very noisy due to its low degrees of freedom ( $\nu^* = n^* - p^* = 4 - 1 = 3$ ). It shows similar spatial structure to the fixed effects standard deviation (Fig. 6b) that has very high degrees of freedom ( $\nu_{\text{fixed}} = n^* \nu = 448$ ). The ratio of the two (Fig. 6d) is much more stationary, though still very noisy. After smoothing (Fig. 6e), it is apparent that the ratio is close to one outside the brain and in many brain regions as

well, indicating little evidence for a random effect, except in frontal and anterior regions where it attains values as high as 1.6. This means that in these regions the variability of the effects is up to 60% higher than their variability due to the scans. If we had just done a fixed effects analysis, our  $T$  statistics would be up to 60% higher in these regions than they should be. To correct for this, we multiply the smoothed ratio by the fixed effects standard deviation to obtain a roughly unbiased estimate of the variance (Fig. 6c). Note that the standard deviation is up to 60% higher in frontal and anterior regions to compensate for the random effects. The random effects due to repetitions over subjects is much higher (Fig. 6f), up to 3 times the fixed effects. The smoothing increases the degrees of freedom of the variance ratio from 3 to 149, then the

multiplication back by the fixed effects standard deviation decreases it to approximately  $\tilde{v}^* = 112 \text{ df}$  (middle column, Table 1).

`Tstat_threshold` with a search volume of 1000 cc, a voxel volume of 38.4 cc, 6 mm FWHM smoothing, and  $\tilde{v}^* = 112 \text{ df}$  gives a  $P = 0.05$  threshold of 4.86. Even if the degrees of freedom is in error by a factor of 2, it does not greatly affect the critical threshold. The final thresholded  $T$  statistic is shown in Fig. 7.

## 7. DISCUSSION

Many of the methods presented in this paper are refinements on existing ones: bias reduction and spatial regularization of temporal autocorrelation parameters; model fitting by prewhitening; careful coding to reduce computations. The main novel contribution is a simple random effects analysis for combining runs/sessions/subjects, including spatial regularization to increase the degrees of freedom. When we talk about an under-estimation of the standard error this is from the perspective of a random or "mixed" effects analysis. In our experience the sorts of questions addressed by neuroimaging experiments call for a mixed effects analysis so that the inferences can be generalised to the population from which the subjects were drawn. Our novel approach embodies some sensible constraints that allow for a much better estimation of the error variance required for mixed effects inferences. This better estimation improves sensitivity. However, it should be noted that this does not preclude an important role for fixed effects analyses, which are very sensitive but can only be generalised to the subjects studied. Fixed effects analyses have to be used in the context of single case studies (e.g., patients with specific brain lesions) or when the functional anatomy of interest does replicate from subject to subject (e.g., ocular dominance columns in visual cortex). Ignoring the random effect and treating the effects as fixed could lead to standard errors that are three times too small (see Fig. 6f) resulting in  $T$  statistics that are three times too large. On the other hand, using a purely random effects analysis gives very small degrees of freedom resulting in less sensitivity. For example, the  $P = 0.001$  point of the  $t$  distribution with 3  $df$  is  $T = 10.2$ , more than three times larger than with 100  $df$  ( $T = 3.2$ ). The method proposed here is a compromise between the two that maintains high degrees of freedom while still capturing most of the random effect.

There are many approximations leading to the result for the effective degrees of freedom  $\tilde{v}^*$ , so the true effective degrees of freedom could be quite different. By choosing the amount of smoothing to give a calculated effective degrees of freedom of at least 100, we hope to cover up for any errors in this approximation. Even if the calculated  $\tilde{v}^*$  is off by a factor of two, it will not

greatly affect our inference, since  $T$  statistics with more than 50  $df$  are very close to a standard Gaussian.

Regularization by spatial smoothing is used first to stabilize the autocorrelation, and second to increase the degrees of freedom of the random effects analysis. Obviously this introduces some bias due for instance to smoothing data outside the brain with data inside the brain. This could be alleviated by smoothing along the cortical surface, either by spatially informed basis functions (Kiebel *et al.*, 2000) or by cortical surface mapping (Andrade *et al.*, 2001).

It is the last stage when `multistat` is used to combine the subjects that has the greatest influence in determining the final standard deviation of the effect. This is because the random subject effect is much larger than the random effects from previous stages. This means that if the cost of the experiment is proportional to total scanner time, then it is better to enroll more subjects and scan them less, than to increase the number sessions, runs, or scans per run on fewer subjects. In reality there is additional cost associated with recruiting and training new subjects, so some compromise between the two is better.

## 8. CONCLUSION

We have presented a simple method for the statistical analysis of fMRI data that does a reasonable job of giving valid results. We have paid particular attention to the combination of results from separate runs, sessions and subjects, because it is the latter in particular that has the greatest affect on our final inference about the nature of the fMRI response in the population. There are many places where this could be improved, at the expense of more computation time. For example, Volterra kernels for nonlinear hemodynamics are not yet available in `fmrifstat`. However `fmrilm` can estimate the delay of the HRF at every voxel and for every stimulus (Liao *et al.*, 2000). This makes it possible to produce separate images of the delay of the HRF and its standard deviation for each stimulus, that themselves can be combined over runs, sessions and subjects using `multistat`.

The user will naturally ask how this method differs from SPM'99. The differences are given in Table 2. They are all small but in each case we think the choices we have made may be slightly better. The similarities are many: both methods use linear models with a simple autocorrelation structure, drift removal by additive trends, and the same random field theory to detect activated regions. Both methods are written in MATLAB and are freely available on the web. The differences are perhaps more of taste than substance, so we expect both methods to give roughly similar results.

`Fmrifstat` is only in the early stages of development. It only does a simple statistical analysis of fMRI data, and other functions such as motion correction and spa-

**TABLE 2**  
Comparison of SPM'99 and fmristat

Feature	SPM'99	fmristat
Different slice acquisition times	Augments the model with a temporal derivative	Shifts the model
Trend removal	Low frequency cosines (flat at the ends)	Polynomials (free at the ends)
Temporal correlation	AR(1), global parameter, bias reduction not necessary	AR( $p$ ), local parameters with bias reduction and spatial regularization
Estimation of effects	Band pass filter, then least squares, then correction for temporal correlation	Prewhiten the data, then least squares (no further corrections needed)
Rationale	More robust, but lower degrees of freedom	More accurate (if model is correct), higher degrees of freedom
Random effects	No regularization, low degrees of freedom	Spatial regularization, high degrees of freedom
Maps of estimates and standarderrors of the delay of the HRF	No	Yes

tential normalization are done by other software in our lab. It now has a GUI, and it reads and writes both `analyze` and `minc` formatted image files (the latter accessed through the `emma` toolbox available for Windows, Linux and SGI IRIX (Morales *et al.*, 2000)).

SPM'99 is a much more complete package with a GUI, a help line, and a course, that analyses much more than fMRI data. Future releases of SPM will remove temporal smoothing or low pass filtering as an option and replace the AR(1) model on the original data with AR(1) plus white noise estimated using the EM algorithm. SPM'01 will still retain a global parameter for the serial correlations and continue to use the least squares estimator as described in Worsley and Friston (1995) and in Section 3.1 for simplicity and flexibility (Friston, personal communication).

## APPENDIX

### A. Extensions to AR( $p$ ) Models

*A.1. Bias reduction of the autocovariances.* Let  $\mathbf{D}_l$  be a matrix of zeros with ones on the  $l$ th upper off-diagonal, so that the expectation of the lag  $l$  autocovariance,  $0 \leq l \leq p$ , is

$$\begin{aligned} E(\mathbf{r}'\mathbf{D}_l\mathbf{r}) &= \text{tr}(\mathbf{R}\mathbf{D}_l\mathbf{R}\mathbf{V})\sigma^2 \\ &\approx \text{tr}(\mathbf{R}\mathbf{D}_l)\nu_0 + \sum_{j=1}^p \text{tr}(\mathbf{R}\mathbf{D}_l\mathbf{R}(\mathbf{D}_j + \mathbf{D}'_j))\nu_j \end{aligned}$$

where  $\mathbf{V}\sigma^2$  is approximated by  $\nu_j$  on the  $j$ th off-diagonal,  $0 \leq j \leq p$ , and zero elsewhere. Equating  $\mathbf{r}'\mathbf{D}_l\mathbf{r}$  to its approximate expectation for  $0 \leq l \leq p$  gives  $p+1$  linear equations in  $p+1$  unknowns, which can be solved for  $\nu_j$ ,  $0 \leq j \leq p$ , giving approximate unbiased estimators of the autocovariances. To do this, let

$$a_l = \sum_{i=l+1}^n r_i r_{i-l}, \quad m_{lj} = \begin{cases} \text{tr}(\mathbf{R}\mathbf{D}_l) & j = 0, \\ \text{tr}(\mathbf{R}\mathbf{D}_l\mathbf{R}(\mathbf{D}_j + \mathbf{D}'_j)) & 1 \leq j \leq p, \end{cases}$$

$$\mathbf{a} = \begin{pmatrix} a_0 \\ \vdots \\ a_p \end{pmatrix}, \quad \mathbf{M} = \begin{pmatrix} m_{00} & \dots & m_{0p} \\ \vdots & \ddots & \vdots \\ m_{p0} & \dots & m_{pp} \end{pmatrix}, \quad \hat{\mathbf{v}} = \begin{pmatrix} \hat{\nu}_0 \\ \vdots \\ \hat{\nu}_p \end{pmatrix}.$$

The estimated covariances are

$$\hat{\mathbf{v}} = \mathbf{M}^{-1}\mathbf{a},$$

and the autocorrelations at lag  $l$  are estimated by

$$\hat{\rho}_l = \hat{\nu}_l / \hat{\nu}_0.$$

Note that  $\mathbf{M}^{-1}$  does not depend on the voxels, so it only needs to be calculated once per slice, making this bias reduction step very fast.

*A.2. Smoothing the autoregressive model.* As explained in Section 3.3, the autocorrelation coefficient for an AR(1) model is smoothed spatially to reduce its variability. For an AR( $p$ ) model there are two choices: smooth the autocorrelations, or smooth the elements of  $\mathbf{V}^{-1/2}$  (for AR(1) models they are identical). We have decided to take the first approach and smooth the autocorrelations. The main reason for this is that it requires less effort, since there are only  $p$  autocorrelations but there are  $(p+1)(p+2)/2$  different elements that define  $\mathbf{V}^{-1/2}$  (see Section A.3). Moreover it is easy to show that if the filter is positive, then smoothed autocorrelations still define a positive definite correlation matrix, hence a stationary AR( $p$ ) model.

The reason is this. Let  $f$  be the filter, normalized to integrate to 1, and let  $\mathbf{C}$  the Toeplitz matrix whose diagonals are  $\rho = (1 \rho_1 \dots \rho_p)$ , that is the reduced bias

correlations of the first  $p + 1$  scans. We shall assume that  $\mathbf{C}$  is still positive definite, even after bias reduction. Let  $f^* \mathbf{C}$  be the convolution of  $f$  with each element of  $\mathbf{C}$ . Consider an arbitrary quadratic form in  $f^* \mathbf{C}$ :

$$\mathbf{a}'(f^* \mathbf{C})\mathbf{a} = f^*(\mathbf{a}'\mathbf{C}\mathbf{a}) \geq 0$$

since  $f \geq 0$  and  $\mathbf{a}'\mathbf{C}\mathbf{a} \geq 0$ . Hence the smoothed correlation matrix  $f^* \mathbf{C}$  is a positive definite Toeplitz matrix. The diagonal elements are 1, so it is a correlation matrix and can be used to derive a stationary AR( $p$ ) model.

*A.3. A simple method of whitening.* The whitening matrix  $\mathbf{V}^{-1/2}$  can be found directly from the autocorrelations, without solving the Yule–Walker equations, by the following piece of MATLAB code. The row vector `rho` contains the regularized autocorrelations ( $1 - \hat{\rho}_1 \cdots \hat{\rho}_p$ ),  $n$  contains  $n$  and `Vmhalf` contains  $\mathbf{V}^{-1/2}$ :

```
[Ainv posdef]=chol(toeplitz(rho));
p1=size(Ainv,1);
A=inv(Ainv');
Vmhalf=toeplitz([A(p1,p1:-1:1)
zeros(1,n-p1)],zeros(1,n));
Vmhalf(1:p1,1:p1)=A;
```

Note that this code allows for the possibility that due to round-off error the estimated AR( $p$ ) model is nonstationary ( $p_1 < p + 1$ ), in which case the higher order AR coefficients are set to zero and an AR( $p_1 - 1$ ) model is fitted instead.

### B. The EM Algorithm

The EM algorithm for REML estimates is laid out in Laird, Lange, and Stram (1987), specifically Eq. (3.8). Let  $\mathbf{S} = \text{diag}(S_1, \dots, S_{n^*})$  and  $\mathbf{I}$  be the  $n^* \times n^*$  identity matrix. Then the variance matrix of the effects vector  $\mathbf{E} = (E_1, \dots, E_{n^*})'$  is, from (10),

$$\Sigma = \mathbf{S}^2 + \mathbf{I}\sigma_{\text{random}}^2. \quad (13)$$

Define the weighted residual matrix

$$\mathbf{R}_{\Sigma} = \Sigma^{-1} - \Sigma^{-1}\mathbf{Z}(\mathbf{Z}'\Sigma^{-1}\mathbf{Z})^+\mathbf{Z}'\Sigma^{-1}, \quad (14)$$

where  $\mathbf{Z} = (\mathbf{z}_1 \cdots \mathbf{z}_{n^*})'$ . We start with an initial value of  $\sigma_{\text{random}}^2 = \mathbf{E}'\mathbf{R}_{\mathbf{I}}\mathbf{E}/v^*$  based on assuming that the fixed effects variances are zero. The updated estimate is

$$\begin{aligned} \hat{\sigma}_{\text{random}}^2 &= (\sigma_{\text{random}}^2(p^* + \text{tr}(\mathbf{S}^2\mathbf{R}_{\Sigma})) \\ &\quad + \sigma_{\text{random}}^4 \mathbf{E}'\mathbf{R}_{\Sigma}^2\mathbf{E})n^*. \end{aligned} \quad (15)$$

Replace  $\sigma_{\text{random}}^2$  with  $\hat{\sigma}_{\text{random}}^2$  in (13) and iterate (13–15) to convergence. In practice 10 iterations appears to be enough.

After convergence and regularization as described in Section 4.3, step (13) is repeated with  $\sigma_{\text{random}}^2$  replaced by  $\tilde{\sigma}_{\text{random}}^2$ , then the estimate of  $\gamma$  is

$$\hat{\gamma} = (\mathbf{Z}'\Sigma^{-1}\mathbf{Z})^+\mathbf{Z}'\Sigma^{-1}\mathbf{E},$$

and its estimated variance matrix is

$$\widehat{\text{Var}}(\hat{\gamma}) = (\mathbf{Z}'\Sigma^{-1}\mathbf{Z})^+.$$

### C. The Effective Degrees of Freedom of the Smoothed Variance Ratio

We start by supposing that standard deviations  $S_j$  are all fixed and equal, and that  $\sigma_{\text{random}}^2 = 0$ . In this case the model (10) is equivalent to a model with constant unknown variance so that  $v^*\tau^2$  has a  $\chi^2$  distribution with  $v^*$  degrees of freedom at every point  $\mathbf{u}$  in  $D$  dimensional space. Suppose the error in the effect  $\eta_j$  is modeled as  $D$  dimensional white noise convolved with a Gaussian kernel  $f(\mathbf{u})$  with variance  $v_{\text{effect}} = W_{\text{effect}}^2/(8 \log 2)$ :

$$f(\mathbf{u}) = \phi(\mathbf{u}; v_{\text{effect}})$$

where

$$\phi(\mathbf{u}; v) = (2\pi v)^{-D/2} e^{-\|\mathbf{u}\|^2/(2v)}$$

is the  $D$  dimensional Gaussian density with mean zero and variance  $v$ . Then the spatial correlation function of  $\eta_j$  is

$$c(\mathbf{u}) = \frac{(f^* f)(\mathbf{u})}{(f^* f)(\mathbf{0})} = (2\pi 2 v_{\text{effect}})^{D/2} \phi(\mathbf{u}; 2v_{\text{effect}})$$

where  $*$  denotes convolution. Then it can be shown (Adler, 1981, Chapter 7) that the spatial covariance function of the squared error is

$$2c(\mathbf{u})^2 = 2(2\pi v_{\text{effect}})^{D/2} \phi(\mathbf{u}; v_{\text{effect}}).$$

Now  $\tau^2$  has the same distribution as a linear function of the average of  $v^*$  independent  $\eta_j^2$  and so its spatial covariance function is  $2c(\mathbf{u})^2/v^*$ . If  $\tau^2$  is convolved with a Gaussian filter

$$g(\mathbf{u}) = \phi(\mathbf{u}; v_{\text{ratio}})$$

with variance  $v_{\text{ratio}} = W_{\text{ratio}}^2/(8 \log 2)$ , to give  $\tilde{\tau}^2$ , then the covariance function of  $\tilde{\tau}^2$  is that of  $\tau^2$  convolved with  $(g * g)(\mathbf{u})$ , namely

$$2(c^2 * g * g)(\mathbf{u})/\nu^* = 2(2\pi\nu_{\text{effect}})^{D/2}\phi(\mathbf{u}; \nu_{\text{effect}} + 2\nu_{\text{ratio}}).$$

Thus the variance of  $\tilde{\tau}^2$  is the spatial covariance function at  $\mathbf{u} = \mathbf{0}$  namely

$$\begin{aligned}\text{Var}(\tilde{\tau}^2) &= 2(c^2 * g * g)(\mathbf{0})/\nu^* \\ &= 2(\nu_{\text{effect}}/(\nu_{\text{effect}} + 2\nu_{\text{ratio}}))^{D/2}/\nu^*.\end{aligned}$$

Finally the effective degrees of freedom of the variance ratio is, by the Satterthwaite approximation,

$$\tilde{\nu}^* = \frac{2E(\tilde{\tau}^2)^2}{\text{Var}(\tilde{\tau}^2)} = \nu^* \left( 2 \left( \frac{\nu_{\text{ratio}}}{\nu_{\text{effect}}} \right)^2 + 1 \right)^{D/2}.$$

Since this result does not depend on the value of  $S_j$ , it is also true if  $S_j$  is random. This result is expected to hold approximately when the  $S_j$ s are unequal and non-stationary, provided their expectations are fairly smooth.

The last question is the effective degrees of freedom of  $\sigma_{\text{fixed}}^2 + \hat{\sigma}_{\text{random}}^2$ , equal to the product of  $\tilde{\tau}^2$  and  $\sigma_{\text{fixed}}^2$ . These two factors are approximately independent. The degrees of freedom of  $\sigma_{\text{fixed}}^2$  is  $\nu_{\text{fixed}}$ . The coefficients of variation (percentage errors) of the product of independent quantities add (approximately, if they are small), so the reciprocals of the effective degrees of freedom add (by the Satterthwaite approximation), which leads to the final result (12).

#### D. Positivity of the Regularized Mixed Effects Variance

Even though the estimated effects variance  $S_j^2 + \hat{\sigma}_{\text{random}}^2$  is guaranteed to be positive at each voxel (even if the EM algorithm has not converged), we discovered that after spatial smoothing  $S_j^2 + \hat{\sigma}_{\text{random}}^2$  from (11) can sometimes be negative. This is because  $\hat{\sigma}_{\text{random}}^2$  can be negative, particularly when there is a small random effect; recall that this ensures that  $\hat{\sigma}_{\text{random}}^2$  is unbiased. The problem happens in scanner artifacts outside the brain near the border of regions of exact zeros, and we have never observed it in the brain itself. To overcome this, we cannot simply set negative variances to ‘missing’ values, since there are always some very small variances at neighboring voxels that give very large  $T$  statistics. Instead, multistat uses the maximum of  $S_j^2 + \hat{\sigma}_{\text{random}}^2$  and  $S_j^2/4$  as the variance of  $E_j$  whenever  $w_{\text{ratio}} > 0$ . This guarantees that the standard deviation of  $E_j$  is never less than half the fixed effects standard deviation. It does not alter any of the voxel values in any of the figures.

#### REFERENCES

- Adler, R. J. 1981. *The Geometry of Random Fields*. Wiley, New York.
- Andrade, A., Kherif, F., Mangin, J-F., Paradis, A-L., Worsley, K. J., Simon, O., Dehaene, S., Le Bihan, D., and Poline, J-B. 2001. Detection of fMRI activation using cortical surface mapping. *Human Brain Mapping* **12**: 79–93.
- Bullmore, E. T., Brammer, M. J., Williams, S. C. R., Rabe-Hesketh, S., Janot, N., David, A. S., Mellers, J. D. C., Howard, R., and Sham, P. 1996. Statistical methods of estimation and inference for functional MR image analysis. *Magn. Reson. Med.* **35**: 261–277.
- Cao, J. 1999. The size of the connected components of excursion sets of  $\chi^2$ ,  $t$  and  $F$  fields. *Adv. Appl. Prob.* **31**: 579–595.
- Chen, J-I., Ha, B., Bushnell, M. C., Pike, B., and Duncan, G. H. 2000. Differentiating pain- and innocuous mechanical-related activation of human S1 cortex using temporal analysis of fMRI. *In preparation*.
- Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J-B., Frith, C. D., and Frackowiak, R. S. J. 1995. Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Mapp.* **2**: 189–210.
- Friston, K. J., Fletcher, P., Josephs, O., Holmes, A. P., Rugg, M. D., and Turner, R. 1998. Event-related fMRI: Characterising differential responses. *NeuroImage* **7**: 30–40.
- Friston, K. J., Josephs, O., Zarahn, E., Holmes, A. P., Rouquette, S., and Poline, J-B. 2000. To smooth or not to smooth? Bias and efficiency in fMRI time series analysis. *NeuroImage* **12**: 196–208.
- Glover, G. H. 1999. Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage* **9**: 416–429.
- Harville, D. A. 1974. Bayesian inference for variance components using only error contrasts. *Biometrika* **61**: 383–385.
- Kiebel, S. J., Goebel, R., and Friston, K. J. 2000. Anatomically informed basis functions. *NeuroImage* **11**: 656–667.
- Laird, N. M., Lange, N., and Stram, D. 1987. Maximum likelihood computations with repeated measures: Application of the EM algorithm. *J. Am. Stat. Assoc.* **82**: 97–105.
- Lange, N., and Zeger, S. L. 1997. Non-linear Fourier time series analysis for human brain mapping by functional magnetic resonance imaging (with Discussion). *Appl. Statist.* **46**: 1–29.
- Liao, C. H., Worsley, K. J., Poline, J-B., Duncan, G. H., and Evans, A. C. 2000. Estimating the delay of the hemodynamic response in fMRI data. *NeuroImage*, submitted.
- Morales, F., MacDonald, D., Valdés, P., and Evans, A. C. 2000. Transferring the Montreal Neurological Institute image processing tools to a Windows platform. *NeuroImage* **11**: S922.
- Purdon, P. L., Solo, V., Weisskoff, R. M., and Brown, E. 2001. Locally regularized spatiotemporal modeling and model comparison for functional MRI. *NeuroImage* **14**: 912–923.
- Worsley, K. J., and Friston, K. J. 1995. Analysis of fMRI time-series revisited—Again. *NeuroImage* **2**: 173–181.
- Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., and Evans, A. C. 1996. A unified statistical approach for determining significant signals in images of cerebral activation. *Hum. Brain Mapp.* **4**: 58–73.
- Worsley, K. J., Andermann, M., Koulis, T., MacDonald, D., and Evans, A. C. 1999. Detecting changes in non-isotropic images. *Hum. Brain Mapp.* **8**: 98–101.
- Worsley, K. J., Liao, C., Grabcov, M., Petre, V., Ha, B., and Evans, A. C. 2000. A general statistical analysis for fMRI data. *NeuroImage* **11**: S648.