Effects of large-scale nonstationarity on parametric maps. A study of rest perfusion CASL data

Roberto Viviani

Department of Psychiatry III
University of Ulm, Germany

Abstract

This study investigates the emergence of characteristic patterns in clusters thresholded at uncorrected significance levels, using as a case study rest perfusion images obtained with the continuous arterial spin labelling technique (CASL). The origin of these patterns is traced back to the existence of large-scale spatial covariance, a violation of the stationarity assumption on the spatial distribution of residual errors. It is shown that in the presence of large-scale covariance, several principles or intuitions common among experimenters when evaluating the inferential strength of their analyses are not applicable. Thresholded maps and clusters are confounded by the spatial patterns of large-scale covariance, irrespective of the existence of a true effect, as shown in t maps constructed by resampling groups at random from a large pool of volumes. Filtering clusters according to their size made the problem worse, and corrections on cluster size based on random field theory models of smoothness had only a minor impact on their tendency to appear in characteristic locations. A formal analysis shows that the large-scale covariance at the origin of these problems is retained in the parametric map irrespective of sample size. Therefore, neither sample independence nor sample size protect from replications of effects being confounded by replications of the spatial covariance of residual errors. In contrast, cluster peaks were not affected by large-scale covariance but only by local differences in smoothness levels, as predicted by random field theory for the distribution of maxima, highlighting the different inferential robustness of cluster-based and maxima-based statistics. A framework is provided to generalize these results to mixed effects models with nested random effects, applicable also to activation studies. Large-scale nonstationarity is most problematic when the variance source at the origin of the characteristic patterns is not specific to a function or variable involved in the inferential process, as typically in observational studies of individual differences. These results raise the question of the existence and impact of large-scale nonstationarity in studies with data obtained with other techniques.

Effects of large-scale nonstationarity on parametric maps. A study of rest perfusion CASL data

1. Introduction

This work originates from efforts to clarify the nature of spatial patterns encountered in baseline rest perfusion images obtained with the continuous arterial spin labelling technique (CASL, Alsop and Detre 1998; Detre et al. 1992; Wang et al. 2005; Williams et al. 1992). These patterns are visible when statistical parametric maps are thresholded at uncorrected levels to display the estimated effects. In our work, CASL was applied to study individual differences in baseline perfusion defined by psychological variables or genotypic variants (Abler et al. 2008; Beschoner et al. 2008; Kirchheiner et al. 2010; Viviani et al. 2010a). We noted that the same patterns tended to arise in t maps of linear estimates of individual differences, even if the explanatory variables used across studies had little relation to each other, sometimes confounding the interpretation of clusters of effects. Figure 1 exemplifies this phenomenon. The posterior part of the brain loads on the individual differences variable in a study of rest perfusion and working memory capacity (Beschoner et al. 2008), *SERT* promoter genotype (Viviani et al. 2010a) and emotion suppression (Abler et al. 2008).

FIGURE 1

Maps thresholded at uncorrected significance levels constitute an important aspect of the statistical analysis in neuroimaging practice (Poldrack et al. 2008). Even when formal evidence relies on corrected significance levels (Friston et al. 1994; Poline and Mazoyer 1993; Worsley et al. 1992), the clusters visible in the maps thresholded at uncorrected levels convey information on the shape of the effect. These overthreshold maps also provide information for meta-analytic studies (Wager et al. 2007). The reduced number of clusters,

relative to voxels, may be exploited to improve the precision of the effect estimate and reduce the size of the testing family (Chumbley et al. 2009; Heller et al. 2006; Penny and Friston 2003). In contrast, maps thresholded at strong control levels (voxel level), which are much more stringent, often display punctiform foci of activation, and may not reflect the real interest of experimenters concerned with mapping overall regional effects. Variables such as sex, age, genotype or diagnosis may be assumed to have a spatially extended effect on brain structure. In cases such as these, interest may center on regional effects, and the focus of inference may be the clusters of adjacent voxels reaching a cluster-defining threshold, rather than the sparse, punctiform activation identified by voxel-level statistics. In this context, the emergence of patterns as noted above constitutes an important methodological and practical issue.

A clue to the possible origin of the patterns we observed is given by their similarity to spatial patterns of covariance obtained in a principal component analysis of these data (Viviani et al. 2009; Figure 1D). The pattern in Figure 1, obtained after adjusting for global perfusion levels, opposes the posterior part of the volume, roughly in the territory of the vertebral artery, to the rest of the brain. This means that in each estimated cerebral blood flow (CBF) volume, a shape similar to the one shown in the figure may be staked out by signal that, at random, is either darker or brighter than its surroundings. The shape in the t map appears when the shapes in each volume in the sample happen not to cancel each other between the groups tested. However, this explanation of the existence of these patterns is, in its simplicity, misleadingly reassuring. To foreshadow the nature of the problem examined here, note that the map in Figure 1A was obtained from a sample of over 140 individuals, larger than that of many neuroimaging studies. At which point may one expect to have sampled enough subjects to avoid worrying about this problem?

Because of the large number of voxels in a volume, many voxels reach an uncorrected threshold by chance alone. The random tendency of these voxels to reach the threshold

together in the same experiment, i.e. the spatial pattern formed by them, is given by the covariance of the map under the null distribution. A common assumption made on this covariance is second-order stationarity, as when adopting a random field model for the errors of a general linear model at each voxel (Worsley et al. 1992, 1999). In a stationary random field, the joint distribution, or covariance, between any voxel and all others is the same for all voxels. Under this assumption, the pattern formed by overthreshold voxels is one of blobs of activation randomly scattered across the brain. However, if the form of spatial covariance is not constrained in some way such as by stationarity, the possibility arises that a characteristic cluster pattern may be observed resulting from spatial covariance binding together the signal in the voxels of the cluster. The existence of large-scale patterns in spatial covariance, as detected in a principal components analysis, and the tendency of the components to emerge in the thresholded maps as shown in Figure 1, suggests that the stationarity assumption is violated in these data.

It will be useful to distinguish between two ways in which stationarity may be violated. The first is due to local differences in spatial smoothness (or 'local isotropy', Taylor and Worsley 2007). The importance of smoothness is due to the fact that it determines the distribution of maxima in the random field (Taylor and Worsley 2007), which is used to obtain voxel-level (strong control) rejection thresholds. In the discrete datasets used in practice, smoothness and its local variation across the volume is estimated from first-order differences between the data at contiguous voxels, and is encoded in the 'RESEL density map' (where RESEL stands for 'resolution element', Worsley et al. 1992). Based on this aspect of spatial covariance, Hayasaka et al. (2004) have introduced a correction for cluster size that takes into account the local inhomogeneity of spatial smoothness.

However, because of its intrinsically local nature, this notion of smoothness homogeneity may not capture nonstationarity arising from large-scale patterns in the joint distribution of voxels. By 'large scale' covariance we mean substantial covariance beyond the

covariance in the neighbourhood of a voxel that makes data smooth, and that may introduce nonstationarity in the joint distribution of the voxels irrespective of smoothness (for example, differing degrees of left-right symmetry constitute a violation of the stationarity assumption, but it is easy to see that such nonstationarity may occur in the presence of uniform smoothness). Such large-scale nonstationarity may not only produce patterns in the data that may look more structured than scattered blobs; it also raises the issue of the adequacy of procedures that account only for local inhomogeneity of smoothness when correcting the spatial distribution of voxels (Hayasaka et al. 2004).

The distinction between two types of violations of the stationarity assumption may also be useful to highlight the possible existence of differences in the influence of nonstationarity on the distribution of maxima on the one hand and of clusters on the other. The theoretical result that the distribution of maxima in the random field depends on local isotropy only (Taylor and Worsley 2007) implies that large-scale nonstationarity has no effect on the distribution of maxima of the random field.

While originally motivated by a practical problem in the analysis of baseline perfusion data, where nonstationarity is well documented (Viviani et al. 2009, 2010b), this study contains a formal analysis of the covariance of parametric maps that is applicable more generally to studies of individual differences conducted with the statistical parametric mapping (SPM) approach (where the test statistic is estimated separately voxel by voxel). Going against the conventional view of this approach as 'massively univariate', the first part of this study is concerned with the formulation of a multivariate model in which the SPM approach and the Gaussian random field assumption on the spatial distribution of the errors are recovered as a special case. In the second part of the paper the effect of covariance patterns outlined by the model is investigated empirically in baseline perfusion images obtained with arterial spin labelling.

2. Theory

To clarify the behaviour of the patterns of overthreshold voxels, we seek here an explicit expression for the covariance of the parametric map (which differs from the covariance of the coefficient estimates) in a model that includes data in all voxels at once. We will first consider a simple design in which one volume is acquired per subject, and subsequently extend the treatment to more complex setting in which the two sources of variance of rest perfusion studies are modelled explicitly.

2.1 One volume per subject

Consider a dataset with n volumes, each containing v voxels, modelled by a design matrix $\mathbf{X}$ of n rows, as in a second level analysis where each volume corresponds to a subject. Let $\mathbf{y}$ be a single vector obtained by stacking the data obtained in each voxel on top of each other. Hence, the first n elements of $\mathbf{y}$ are the data of the first voxel of all volumes, the second n elements those of the second voxel, and so on for all voxels. The model equation is

$$\mathbf{y} = (\mathbf{I}_v \otimes \mathbf{X})\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \ \ \mathrm{Var}(\boldsymbol{\varepsilon}) = \mathbf{S} \otimes \mathbf{I}_n. \tag{1}$$

The Kronecker product '$\otimes$' between the $v \times v$ identity matrix $\mathbf{I}_v$ and the design matrix $\mathbf{X}$ is used here to formalize voxel-by-voxel modelling. In the expression $\mathbf{I}_v \otimes \mathbf{X}$, the design matrix $\mathbf{X}$ has been stacked across the diagonal to model the data in each voxel separately. In the expression for the covariance of the errors $\mathrm{Var}(\boldsymbol{\varepsilon})$, the element on the left of the Kronecker product is the spatial covariance $\mathbf{S}$ of the errors, while the $n \times n$ identity matrix $\mathbf{I}_n$ on the right expresses the independence of sampling from subject to subject in each voxel (Appendix A gives a more general expression). The covariance matrix $\mathbf{S}$ absorbs the voxelwise variances on the diagonal. Let σ_k^2 be the variance in each voxel k, $k = 1, 2, \ldots, v$, and collect all these voxelwise variances in a $v \times v$ diagonal matrix $\mathbf{K}$. If $\mathbf{G}$ is the $v \times v$ spatial correlation of the errors, the spatial covariance may be written $\mathbf{S} = \mathbf{K}^{\frac{1}{2}}\mathbf{G}\mathbf{K}^{\frac{1}{2}}$. In practice, $\mathbf{G}$ would incorporate the effect of smoothing usually applied in the preprocessing stages. At kernels sizes

commonly in use, smoothing removes high spatial frequencies from the data, but residual inhomogeneity of smoothness is commonly observed (Ashburner and Friston 2000). It would also not remove any large-scale covariance, whose spatial frequencies are low.

The least squares estimates $\hat{\boldsymbol{\beta}}$ of this model, computed voxel by voxel, are

$$\hat{\boldsymbol{\beta}} = \left[\mathbf{I}_v \otimes (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \right] \mathbf{y} . \tag{2}$$

Here, $\hat{\boldsymbol{\beta}}$ is a vector containing the voxels of the beta images stacked on top of each other. As before, the Kronecker product between the identity matrix $\mathbf{I}_v$ and the projection operator $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ formalizes the estimate of the coefficients of the model separately in each voxel. It may be shown (see Appendix A) that the sampling variance of these voxelwise estimates is

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}) = \mathbf{S} \otimes (\mathbf{X}'\mathbf{X})^{-1} . \tag{3}$$

Furthermore, given an operator $\mathbf{C} = \mathbf{I}_v \otimes \mathbf{c}$ that produces the contrast image $\mathbf{C}'\hat{\boldsymbol{\beta}}$ when applying the contrast vector $\mathbf{c}$ to the beta images in each voxel separately,

$$\mathrm{Var}(\mathbf{C}'\hat{\boldsymbol{\beta}}) = \mathbf{S} \otimes \mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c} , \tag{4}$$

where $\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}$ is a scalar. In these expressions, the element on the left of the Kronecker product is the spatial covariance the beta images, which is the same $\mathbf{S}$ of the errors, and the element on the right is the contribution of the structure of the design matrix to the sampling variance.

While the least squares estimator presented here was constructed by formalizing the procedure of estimating voxel by voxel, the beta images turn out to be identical to the coefficients estimates of the multivariate model (Appendix A). Model equation (1) is a notational variant of the standard multivariate model in which each voxel is a separate outcome variable (Mardia et al. 1979, pp. 173-174). The SPM and a standard multivariate approach differ in how the test statistic is derived from these estimates. In the multivariate approach, the test statistic is obtained by scaling the coefficient estimates by the whole

sampling variance $\hat{\mathbf{S}}$, estimated from the data at hand. In the SPM approach, in contrast, the elements of $\mathbf{C}'\hat{\boldsymbol{\beta}}$ are scaled at each voxel k separately by the voxelwise sample standard error $\sqrt{\hat{\sigma}_k^2 \mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}$ to obtain a parametric map that is subsequently thresholded. The SPM estimator, therefore, does not use the whole covariance $\hat{\mathbf{S}}$ (not estimable in entirety due to its dimensionality), but only its diagonal elements $\hat{\sigma}_k^2$. It may be shown (see Appendix B) that for known σ_k^2 the covariance of the parametric map $\mathbf{z}_{\text{SPM}}$ obtained in this way is the spatial correlation $\mathbf{G}$ of the original data:

$$\text{Var}(\mathbf{z}_{\text{SPM}}) = \mathbf{G}. \tag{5}$$

Note that the expression for the sampling covariance of the parametric map $\text{Var}(\mathbf{z}_{\text{SPM}})$ of equation (5) differs from the corresponding expression for the beta images $\text{Var}(\hat{\boldsymbol{\beta}})$ or $\text{Var}(\mathbf{C}'\hat{\boldsymbol{\beta}})$ of equations (3) and (4) in one important respect. In the sampling variance of the beta images, the term $(\mathbf{X}'\mathbf{X})^{-1}$ represents the influence of the number of observations in the sample, expressing the increased precision of the $\hat{\boldsymbol{\beta}}$ estimate with samples of increasing size. No such term is present in the sampling variance of the z map: the spatial covariance of the z map is the correlation of the original errors, irrespective of the size of the sample. For t maps computed in practice, where the variance at each voxel is estimated from the sample, these results are a limiting case.

In the Gaussian random field model, $\mathbf{S}$ is assumed to be the realization of a stationary Gaussian stochastic process with mean zero and variance σ_k^2 in each voxel k, which defines the form of $\mathbf{G}$ accordingly, and in terms of fewer parameters than a general positive definite correlation matrix. In this case, the result of equation (5) is the formalization of the standard neuroimaging assumption that the parametric map is a z or t random field with the same point spread function as the errors of the model. Hence, the smoothness parameter of the spread

function is estimated from the residuals of the fit, taken for this purpose as proxies of the errors. The formalization adopted here makes explicit the random field assumption in terms of a restriction on $\mathbf{G}$ in a multivariate model; the result of equation (5) may be taken for granted when $\mathbf{G}$ is a stationary random field, but experimenters may not consider this result equally obvious when $\mathbf{G}$ expresses another spatial covariance pattern.

2.2 Extension to other designs

For the rest perfusion data that motivated the present study, the covariance $\mathbf{S}$ at the second level may be decomposed into two distinct sources: acquisition-to-acquisition (i.e. from successive CBF estimates), and subject-to-subject, each presenting its own spatial pattern. Both contribute to the covariance of the parametric map, but with different weights:

$$\mathrm{Var}(\mathbf{z}_{\mathrm{SPM}}) = \mathbf{W}_A^{\frac{1}{2}}\mathbf{Q}\mathbf{W}_A^{\frac{1}{2}} + \mathbf{W}^{\frac{1}{2}}\mathbf{G}\mathbf{W}^{\frac{1}{2}}, \qquad \mathbf{W}_A + \mathbf{W} = \mathbf{I}_v. \qquad (6)$$

Here, $\mathbf{Q}$ and $\mathbf{G}$ are the subject-to-subject and acquisition-to-acquisition correlation matrices, and $\mathbf{W}_A$, $\mathbf{W}$ are diagonal matrices of influence weights regulating the relative contribution of $\mathbf{Q}$ and $\mathbf{G}$. Note that, since $\mathbf{Q}$ and $\mathbf{G}$ differ, the spatial patterns of $\mathbf{z}_{\mathrm{SPM}}$ merge two distinct pattern sources with different intensities according to the respective weights, varying across voxels. A formal analysis (Appendix C) reveals these weights to depend on the number of acquisitions, with a rate that depends on the relative size of acquisition-to-acquisition variance across the volume. Increasing the number of acquisitions reduces the relative influence of the acquisition-to-acquisition component $\mathbf{G}$, but the influence of the subject-to-subject variance $\mathbf{Q}$ increases, and is present irrespective of the size of the sample defined by the number of participants. Hence, while recruiting more subjects cannot eliminate large-scale nonstationarity due to subject-to-subject variation, increasing the number of acquisitions will gradually degrease the relative influence of acquisition patterns in the merge. As shown in Appendix C, the influence of acquisition-to-acquisition correlation $\mathbf{G}$ in rest perfusion images acquired for 8 min or more is likely to be felt only in circumscribed white matter areas or at the base of the brain.

One approach to understanding this result, and applying the same reasoning to more general mixed models to determine the sources of the patterns of covariance, is to view the t or z statistics that have been considered so far for the parametric maps as specializations of F tests, and make use of the insight given by ANOVA tables provided by standard texts of experimental design. The voxelwise sample standard error $\sqrt{\hat{\sigma}_k^2 \mathbf{c}' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{c}}$ that was used to scale the voxelwise contrast estimate $\mathbf{C}'\hat{\boldsymbol{\beta}}$ and obtain a parametric map in the previous section corresponds, in the F test, to the 'background' mean sum of squares (Snedecor and Cochran 1967), i.e. the one used in the denominator of the relevant F ratio, of an ANOVA table. The expectation of these mean sums of squares is given by a weighted sum of the relevant variance terms, and this is the origin of the weights of equation (6), in which two sources of variance are merged (Appendix C, equations C.1-C.5; in fact, the ANOVA result is the starting point to derive these equations). The value of this approach is that it provides a simple way to determine the origin of possible sources of the spatial patterns of covariance and their relative influence, at least under additivity of the random components and for balanced designs, by looking up the F ratio corresponding in the table to the contrast brought to the second level. If one's aim is only to identify the structural sources of variance in the map, depending on the test and which factors are fixed and which are random, it is not necessary to produce tables replicating the covariates and the contrasts of interest exactly, or derive the corresponding algebraic expression.

As an example, consider a rest perfusion imaging study of individual differences in which subjects were divided into groups. For convenience, the ANOVA table for this design is reproduced in Table 1. Here, one can see that the F ratio to test an effect of groups, MS_{GROUPS}/MS_{SUBJ}, has at the denominator the mean squares between subjects MS_{SUBJ}, whose expectation (given by the E(MS) column in the table) is a weighted sum of subject-to-subject and acquisition-to-acquisition variance, as in equation (6). Contrast this result with testing the

activation of a task administered as a number of "treatments" within subjects. Table 2 is the ANOVA table of a repeated measures experiment with different levels of a within-subjects factor of experimental treatments. Here, tests on the within-subjects factor are conducted on the ratio $MS_{TREAT}/MS_{SUBJ \times TREAT}$. The expectation of the denominator mean squares $MS_{SUBJ \times TREAT}$ contains, besides acquisition-to-acquisition variance, the random interaction experimental treatment × subjects. Further application of this approach reveals that in the activation study tests of individual differences in task activations are done on $MS_{SUBJ \times TREAT}$, but tests of individual differences in average perfusion levels are done on MS_{SUBJ} (likely to be similar, even if not identical, to the MS_{SUBJ} of a baseline perfusion study; see Winer et al. 1991, pp. 326-333). Note that acquisition-to-acquisition variance is present in all these mean squares terms, arising from the precision of estimation at the first level; however, its relative influence may be limited due to the large number of acquisitions of MRI studies. The main contribution to these mean squares terms is therefore given by the other variance term, which depends on the test at hand. Because the spatial covariance of different random components does not, in general, generate the same spatial pattern, the confounding in the map may be different. These patterns may also differ from those of single-subject activation studies, or activation studies in which subjects are treated as a fixed nuisance effect, as in these cases the test the effect of interest is carried out on MS_w, which is constituted exclusively by acquisition-to-acquisition variance. Analogous reasoning may be useful to identify the source of spatial patterns arising in more complex designs.

TABLES 1, 2

3. Material and methods

3.1 Data acquisition

Perfusion images were acquired with the CASL technique described in Wang et al. (2005) at 3T for 8 min. while participants were resting with eyes closed in the scanner (resulting in 60 perfusion images). The inversion pulse was given for 2 sec, followed by a delay of 1 sec

before the image was acquired. Details of the acquisition protocols may be found in Viviani et al. (2009). The smoothing kernel size was FWHM = 8 mm.

3.2 Spatial distribution of overthreshold voxels in rest perfusion CASL data

Principal component analysis on the correlation matrix was carried out on the images averaged in each subject (for an estimate of the components of covariance at second level), or on the centered images in each subject (for an estimate of the components of covariance arising from first level). A singular value decomposition was carried out on the empirical volume × volume correlation matrix, and the components of the voxel × voxel matrix were subsequently recovered (see the appendix of Viviani et al. 2005). Computations were carried out in MATLAB R2006b (The Mathworks, Natick, MA) using the BLAS and LAPACK routines provided with that software, running on a 64-bit Athlon processor (Advanced Micro Devices, Sunnyvale, CA) under Windows XP (Microsoft, Redmond, WA).

To compute rates of overthreshold voxels, averaged images in each subject based on $n = 60$ acquisitions were resampled at random over subjects from a pool of 295 subjects to obtain balanced groups of sample size $a = 24$, on which a two-sample t test was computed. The t maps were thresholded at the predefined significance level $p = 0.001$, uncorrected, and clusters larger than the predefined size of 200 voxels were retained. These values were chosen as representative of those encountered in practice. The resampling was repeated 8000 times, during which a count was kept in each voxel of the number of times it was included in one of these clusters, and the proportion of these events was used to generate the images in the Results section. In the simulations with ranked data, ranking was carried out in each trial in each voxel separately (in a sample of 48 subjects, data in each voxels were converted to numbers from 1 to 48, preserving their order). In simulations with correction for inhomogeneity of smoothness, cluster size was defined as the sum of RESELs in the overthreshold voxels, using equation (7) of Hayasaka et al. (2004). The RESEL estimate was computed with the function *spm_est_smoothness*, version 2.7, in the SPM2 package (Kiebel et

al. 1999). For reasons of computational feasibility, the RESEL estimate was computed once on the whole sample of 295 average baseline images; in real applications, this estimate may be more imprecise than in the present study. To maintain comparability, cluster extent thresholds were divided by the average voxel RESEL value. To assess the effect of voxel size in the distribution of clusters, the same procedure was repeated with cluster sizes 0, 50, and 100 voxels. The images of the figures were generated with the freely available software MRIcroN (www.sph.sc.edu/comd/rorden/mricron/).

Simulations with artificial data were carried out on the same number of volumes, in the same size and smoothness. Random fields were obtained by smoothing white noise with the same function *spm_smooth* used on perfusion data, as detailed in Viviani et al. (2007). Because there were no clusters of size 200 after 8000 trials, the cluster size was reduced to 50 and the cluster-defining threshold was raised to $p = 0.005$, uncorrected.

3.3 Logistic regression of overthreshold count

To model the inhomogeneous binomial process resulting from thresholding parametric maps, the counts of overthreshold voxels were fitted to the first four component images in a logistic regression (with an overdispersion term to account for spatial dependency) using the function *glm* of the freely available software R, version 2.9.2 (obtained from http://www.r-project.org, The R Foundation for Statistical Computing, Vienna, Austria). Each component image provides a predictor in each voxel for the probability of observing overthreshold signal. Because common measures of explained variance, such as R^2, have no equivalent counterpart in nonlinear models, we made use of methods introduced by Gelman and Pardoe (2007) to summarize the relative predictive effect of component images (see also Gelman and Hill 2007, pp. 101-104, and pp. 466-473 for rationale and examples). Briefly, let f be the inverse logit function, $f(x) = 1/[1+\exp(-x)]$. The inverse logit of the linear predictor of a logistic regression $f(X\hat{\beta})$ computes the fitted probability of obtaining a success at each trial for given values of the predictors X and fitted coefficients $\hat{\beta}$. Let $X^{(j)}$ be the same set of

predictors, except for the predictor variable *j*, which is set to zero. Then the difference

$f(X\hat{\beta}) - f(X^{(j)}\hat{\beta})$ gives the net contribution of the predictor variable *j* to the fitted

probability. This difference was computed for each voxel *i* using the values of the predictor

variables that obtain in each voxel (i.e. the components) to draw the box plots of the figures.

Significance values are not reported here because they may be misleading (the predictors were

obtained from the same dataset where the resampling took place). Computational feasibility

precluded the use of 2nd order resampling methods.

3.4 Spatial distribution of cluster peaks

In the study where counts of clusters maxima (peaks) were held, the clusters were selected

with the same cluster-defining threshold of $p = 0.001$, but the number of resampling trials was

brought to 80 000 to obtain images with sufficient contrast. To visualize local

inhomogeneities of smoothness, a RESEL map was computed with the same function

spm_est_smoothness used to compute the correction for cluster size.

4. Results

4.1 Spatial distribution of overthreshold voxels in rest perfusion CASL data

In the treatment of spatial covariance of parametric maps presented in the Theory section, it

was shown that this covariance approximates the spatial correlation of the original data, and

that it does not vanish with increasing sample sizes. For this reason, a principal component

analysis of the data collected under the null is informative of the patterns that may emerge

when thresholding the parametric map. The practical consequences of this result depend on

the prominence of nonstationary components in the data, in particular large-scale covariance

patterns. Here, we investigate empirically the impact of spatial covariance on perfusion data

using a well-characterized sample of baseline perfusion images (Viviani et al. 2009) collected

for 8 min. during rest with the continuous arterial spin labelling technique (Wang et al. 2005).

First, we carried out a component analysis of the correlation matrices of these images

(in Viviani et al. 2009 the covariance was analysed, but the results presented here indicate that

it is the correlation that is relevant for the parametric maps). The first component is positive across the whole volume, and affects grey more than white matter (Figure 2A). The first component loads weakly in the posterior part of the brain, which is covered by the second component (Figure 2B). The third and fourth components, not shown here for brevity, capture left-right asymmetries in perfusion and further variation between grey and white matter (for more details and interpretation, see Viviani et al. 2009).

FIGURE 2

Second, we computed statistical parametric maps repeatedly for t tests comparing two randomly selected groups of perfusion images, using thresholds commonly employed to display images or define clusters. A count was kept of the number of times each voxel reached the threshold. Figure 3 shows that the count was not homogenous across the volume, as one would expect if the stationary model were true. It is apparent that the first two components of the correlation matrix of the data played an important role in determining the probability that a voxel is included in a cluster. For voxels included in a covariance pattern, the probability of cluster detection increased about fivefold. A very similar distribution of clusters was obtained in the opposite direction, i.e. for clusters of reduced signal, as one would expect from a random generating mechanism under a symmetric distribution of data (not shown here for brevity). Because the first component is strongly associated with global perfusion values (Viviani et al. 2009), t tests that adjust for global perfusion by including it as a covariate (as those of Figure 1) were much less affected by the first component; as a result, the influence of the second component is much more apparent (Figure 3B; Spearman correlation $r = 0.38$, $p < 0.001$, permutation test).

To verify the extent to which the notion of local isotropy represented the spatial covariance characterizing these data, we carried out the same study using the RESEL correction for cluster size (Hayasaka et al. 2004), displayed in Figure 3C. The resulting overthreshold rate map is not very different from the one without correction. This suggests

that increased smoothness in gray matter is not the main reason for the increased cluster count. For comparison, Figure 3D shows the spatial distribution of clusters in artificial data generated according to the Gaussian random field distribution.

FIGURE 3

4.2 Logistic regression of overthreshold count

To quantify the relative influence of the covariance components, we carried out a logistic regression of the voxel counts. The voxel values of the first four principal component images were used as predictors of the overthreshold count rates in each voxel, and the logistic regression was used to obtain the net contribution of each component to the fitted count rate (Figure 4; see Methods for computational details and references). The first plot on the left, 'Positive clusters', contains the same data as in Figure 3A. One can see that the first component fits a large part of the observed count rates, shown in the first blue box from the left (the average fitted net count rate due to the first component alone is $0.57^0/_{00}$, while the average observed count rate in the data is $0.64^0/_{00}$). The fitted count rates due to the remaining components are progressively smaller (average fitted net count rates of 0.20, 0.08, and $0.05^0/_{00}$). Note however that the net fitted rates are considerable for subsets of voxels, marked with circles as outliers in the plots. This is due to the fact that these remaining components load on more specific groups of voxels than component 1. After RESEL correction, (second plot from left), the fitted rates decrease slightly (with average fitted net count rates of 0.51 and $0.13^0/_{00}$ for the first two components, relative to an average observed rate of $0.61^0/_{00}$), but remain far from those obtained in artificial data where the stationary Gaussian distribution holds (far right), which range around the values of $0.01\text{-}0.03^0/_{00}$.

FIGURES 4, 5, 6

A possible alternative explanation of the observed inhomogenous count rates of Figure 3 would be based on differences in the distribution of the data across the volume. For example, regions with higher kurtosis are likely to reach the threshold more often (Viviani et

al. 2007). To exclude this eventuality, the computation of the counts was repeated after ranking the data in each voxel separately (Figure 4, 'Ranked data'). The plot shows that ranking failed to remove the net rate due to the first two components in the logistic regression (average fitted net count rates of 0.32 and $0.11^0/_{00}$ for the first two components, relative to an average observed rate of $0.35^0/_{00}$).

In Figure 3 it is apparent that, since regions bound together by the covariance are large in these data, the detected clusters are larger than those of artificial data. This suggests that filtering clusters by size may influence the contribution of the covariance components to the count rates. This intuition is confirmed by the results shown in Figure 5. When the cluster filter was 400 voxels, the average fitted net count rate of the first component was $0.50^0/_{00}$, relative to an average observed rate of $0.53^0/_{00}$. When there was no cluster filter, the average fitted net count rate of the first component was only $0.48^0/_{00}$ relative to an observed rate of $1.02^0/_{00}$. This is proportionally less than when clusters are filtered, but still much more than in artificial stationary data.

4.3 Spatial distribution of cluster peaks

The influence on cluster detection rates of large-scale nonstationarity in these data raises the issue of how cluster peaks are affected. We investigated whether the incidence of cluster peaks follows that of whole clusters, or whether it is more similar to the incidence of maxima over the whole volume, which depends on local smoothness only (Taylor and Worsley 2007). Differences in local smoothness for rest perfusion data are shown in Figure 6A, as quantified by the RESEL density map. In Figure 6B, a count was taken in each voxel of when the voxel was a cluster peak in each resample. The count rate shows cluster peaks to follow the RESEL map closely, unlike the count rate of overthreshold voxels of the previous figures.

5. Discussion

Even when not applying random field theory in computing significance values, experimenters often adopt the random field model when thinking about their data. Under the assumption of

stationarity accompanying the random field model, the spatial distribution of errors is like that of smoothed white noise. The resulting expectation is that false positive activations are likely to be randomly and uniformly distributed throughout the brain. If the stationarity assumption does not hold, however, the spatial distribution of the thresholded signal may follow the main patterns of spatial covariation, instead of resembling a random set of blobs scattered across the volume of interest. When many voxels may be expected to survive thresholding by chance, the spatial pattern they form, which follows their covariance, may confound the interpretation of the thresholded map by superimposing on any existing true effects.

Applying cluster-defining thresholds that may be encountered in practice to our rest perfusion data, we found that the probability of activation under the null was several times larger for voxels bound together by spatial covariance. The preponderance of large-scale nonstationarity in these data was demonstrated by the minor impact of the RESEL correction on this finding (Hayasaka et al. 2004). The importance of spatial covariance should be seen in relation to the fact that its influence of cannot be made arbitrarily small by increasing sample size, as one would usually expect from statistics obtained from consistent estimators. The problem has its origin in how parametric maps are obtained, i.e. by scaling coefficient estimates by voxelwise variances only instead of scaling by the whole covariance (which is not estimable in data of this dimensionality) as one would in a fully multivariate approach.

The stationary random field model also justifies the notion that setting a size threshold to reported clusters improves the chances of the surviving clusters to be true (Forman et al. 1995). Unfortunately, in the presence of the large-scale covariance of our data, this practice had the opposite effect: the association between detected clusters and covariance patterns increased. Hence, in the presence of large-scale nonstationarity, arguments commonly adduced in support of the validity of uncorrected significance levels, such as filtering cluster size, rest on shaky ground. The findings presented here also highlight the danger of the use in

observational studies of uncorrected significance levels, which are often adopted even for data whose distributional properties are not well documented (Beschoner et al. 2005).

Another consequence of nonstationarity is the unreliability of the intuition that if enough data are collected, or enough studies compared, patterns emerging in parametric maps must reflect true effects, even if thresholded at uncorrected levels. Replications of results, referring to parametric maps thresholded in this way, may be contaminated by replications of the covariance pattern of the data. The spatial covariance of the parametric map is retained by increasing sample size, like its smoothness; however, practitioners might unwittingly apply different expectations when considering the effect of sample size on smoothness and on other aspects of spatial covariance.

The importance of these spatial confounds on the interpretation of results will depend in general on the design of the study. In a mixed effects model, several random components may constitute the variance against which an effect of interest is tested; all of these may be at the origin of distinct spatial patterns, which will be present, with different weights, in the covariance of the parametric map. A careful consideration of the origin of these variance components may be required to correctly interpret the visual appearance of the thresholded map and the seriousness of the associated confounding pattern.

Some of the covariance patterns visualized in the overthreshold maps of this study are interpretable in terms of the shape of the volumes outside the brain parenchyma, the properties of the vascular tree, or of the arterial spin labelling technique with which data were acquired. The dissociation between the anterior and posterior part of the brain, shown in Figure 2, follows the boundaries of the vascular territory of the cerebral artery (Viviani et al. 2009), and other areas identified by the acquisition-to-acquisition components, especially at the base of the brain or in the cortex (Viviani et al. 2010b), may arise from labelled spins that have not yet reached the capillary bed, or from interactions between the timing of the labelling pulse and the rhythm of blood flow (transit time effects, Alsop and Detre 1996;

Gonzalez-At et al. 2000; Hernandez-Garcia et al. 2005; Lee et al. 2007; Mildner et al. 2005).

Fortunately, at least under the customary additivity assumption on random effects, increasing

the number of acquisitions may reduce the influence of this source of variance and the

associated patterns on the parametric map. In a rest perfusion study, the remaining source of

confounds are the patterns arising from subject-to-subject variance.

In an activation study, which evaluates the effect of an experimental paradigm, the

patterns of covariance would mainly be those of the interaction experimental

treatment × subjects. If the influence of a confounded session-by-session variance term is

negligible, as is the case for BOLD-EPI data (Smith et al. 2005; Wei et al. 2004), the random

effect of the experimental factor × subjects interaction arises from an experimental

manipulation, thus somewhat limiting the possible influences on the effect to covariates that

are related to the task. This conclusion is supported by current investigations of subject

variability in task-dependent cortical activation, which show it to be consistent with the

involvement of specific networks, notwithstanding the variation in the precise localization of

the task effect (Kherif et al. 2008; Miller and Van Horn 2007; Noppeney et al. 2006; Seghier

et al. 2008).

The relative specificity of the interaction component may be contrasted with the subject-

to-subject variance against which individual differences are tested in studies of signal mean

levels. The subject-by-subject variance of these studies, which is purely observational, may be

associated with a large number of possible uncontrolled covariates affecting the measurement,

i.e. all those that for whatever reason change the individual signal across subjects. Because we

cannot eliminate spatial covariance simply by acquiring larger samples, it might be difficult to

determine whether the observed spatial patterns describe a true effect of the variable of

interest, or emerge from an incidental association with such uncontrolled covariates. These

difficulties would also apply to observational studies with an analogous design acquired with

other neuroimaging techniques, if large-scale spatial nonstationarity is present in the data.

The results reported here empirically confirm the remarkable dissociation of clusters and cluster peaks in their sensitivity to large-scale covariance. Even for nonstationary data far from Gaussian as those examined here (CBF estimates are a ratio of random variables), cluster peaks followed the distribution of maxima described by random field theory, which only depends on the local smoothness of the image, not on large-scale stationarity (Taylor and Worsley 2007). Furthermore, since tests based on the distribution of maxima give strong control of the error rate, the frequency with which voxels reach the threshold due only to random variation (which follows the spatial joint distribution and gives rise to the problems described here) is bound by the significance level of the test.

The different susceptibility of maxima and clusters to spatial covariance has a potentially wide-ranging implication, underscoring the importance of the result by Taylor and Worsley (2007). In the presence of a true effect of the predictor variable, the voxelwise test statistic must increase with sample size due to the consistency of common estimators, and sooner or later a rejection of the null at voxel level must be obtained using the maximal statistic. Permutation of this statistic (Bullmore et al. 2001; Holmes et al. 1996; Nichols and Holmes 2001) also delivers correct voxel-level significance values irrespective of the voxelwise and the joint distribution of the data. Therefore, if the question of interest may be answered by the detection of a punctiform effect, the technical means to overcome the spatial covariance problem by increasing sample size are widely available, and consist in the adoption of maxima-based statistics or voxel-level corrections. Interesting in this respect is the recent proposal of false-discovery rate approaches based on cluster maxima but controlling for error rates less than strongly (Chumbley et al. 2009). In contrast, inference on regional effects appears not to escape its intrinsically multivariate nature, showing the difficulty of addressing the problem of regional inference by refining correction thresholds with weak control of the error rate (Genovese et al. 2002; Hayasaka et al. 2004; Heller et al.

2006). The challenge these approaches face is that of accounting for complex joint distributions in the data, without modelling them in full (Efron 2008, 2010).

Acknowledgments

Appendices

Appendix A. Multivariate formulation of the 'massively univariate' statistical parametric map model.

In this section the multivariate derivation of the voxel-by-voxel model is presented. This is done in part for completeness, and in part to clarify in detail to what extent the 'massive univariate' and the multivariate approaches coincide. We consider here the general case in which the acquisitions may also be temporally correlated, as according to an $n \times n$ autocorrelation matrix $\mathbf{P}$, which is assumed to be separable from the spatial covariance $\mathbf{S}$. Using the same notation and stacking scheme as in the main text, the model is

$$\mathbf{y} = (\mathbf{I}_v \otimes \mathbf{X})\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad \mathrm{Var}(\boldsymbol{\varepsilon}) = \mathbf{S} \otimes \mathbf{P} = \mathbf{K}^{\frac{1}{2}}\mathbf{G}\mathbf{K}^{\frac{1}{2}} \otimes \mathbf{P},$$

and the generalized least squares estimate $\hat{\boldsymbol{\beta}}$,

$$\hat{\boldsymbol{\beta}} = \left((\mathbf{I}_v \otimes \mathbf{X})'(\mathbf{G} \otimes \mathbf{P})^{-1}(\mathbf{I}_v \otimes \mathbf{X})\right)^{-1}(\mathbf{I}_v \otimes \mathbf{X})'(\mathbf{G} \otimes \mathbf{P})^{-1}\mathbf{y}$$

$$= (\mathbf{G}^{-1} \otimes \mathbf{X}'\mathbf{P}^{-1}\mathbf{X})^{-1}(\mathbf{G}^{-1} \otimes \mathbf{X}'\mathbf{P}^{-1})\mathbf{y}$$

$$= \left(\mathbf{I}_v \otimes (\mathbf{X}'\mathbf{P}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{P}^{-1}\right)\mathbf{y},$$

where in the main text, $\mathbf{P} = \mathbf{I}_n$.

The remarkable fact about this estimate, which is best linear unbiased, is that it makes no use of the spatial covariance, when it is separable from $\mathbf{P}$. However, the sampling variance of $\hat{\boldsymbol{\beta}}$ does depend on the spatial covariance $\mathbf{S}$:

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}) = \left((\mathbf{I}_v \otimes \mathbf{X})'(\mathbf{S} \otimes \mathbf{P})^{-1}(\mathbf{I}_v \otimes \mathbf{X})\right)^{-1}$$

$$= \left(\mathbf{S}^{-1} \otimes \mathbf{X}'\mathbf{P}^{-1}\mathbf{X}\right)^{-1}$$

$$= \mathbf{S} \otimes (\mathbf{X}'\mathbf{P}^{-1}\mathbf{X})^{-1},$$

where the formula used in the main text is retrieved again by replacing $\mathbf{P}$ with $\mathbf{I}_n$. This model goes back at least to S. N. Roy, who started with the standard multivariate model $\mathbf{Y} = \mathbf{XB} + \mathbf{E}$, and introduced the stacking scheme to demonstrate the independence of the estimates from $\mathbf{S}$ (Roy et al. 1971, pp. 27-28).

Let $\mathbf{C} = (\mathbf{I}_v \otimes \mathbf{c})$ be a linear operator to test a hypothesis separately in each voxel in the contrast image $\mathbf{C}'\hat{\boldsymbol{\beta}} = \mathbf{0}$. The spatial covariance of the contrast image is

$$\mathrm{Var}(\mathbf{C}'\hat{\boldsymbol{\beta}}) = \mathbf{C}'\mathrm{Var}(\hat{\boldsymbol{\beta}})\mathbf{C}$$

$$= (\mathbf{I}_v \otimes \mathbf{c})'[\mathbf{S} \otimes (\mathbf{X}'\mathbf{P}^{-1}\mathbf{X})^{-1}](\mathbf{I}_v \otimes \mathbf{c})$$

$$= \mathbf{S} \otimes \mathbf{c}'(\mathbf{X}'\mathbf{P}^{-1}\mathbf{X})^{-1}\mathbf{c}.$$

Appendix B. Covariance of the statistical map in the 'massively univariate' model.
Let the model be that of eq. (1) of the main text, and $\mathbf{c}$ the vector of contrasts to test the hypothesis $\mathbf{C}'\hat{\boldsymbol{\beta}} = (\mathbf{I}_v \otimes \mathbf{c})'\hat{\boldsymbol{\beta}} = \mathbf{0}$ in each voxel. In each voxel, the variance of the contrast image $\mathbf{C}'\hat{\boldsymbol{\beta}}$ is $\sigma_k^2 \mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}$ (i.e., the diagonal elements of $\mathbf{S} \otimes \mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}$). For known voxelwise variances σ_k^2 collected on the diagonal of the matrix $\mathbf{K}$, the parametric map is therefore given by

$$\mathbf{z}_{\mathrm{SPM}} = \left[\mathbf{K} \otimes \mathbf{c}'\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{c}\right]^{-\frac{1}{2}} \mathbf{C}'\hat{\boldsymbol{\beta}}.$$

This expression scales $\mathbf{C'\hat{\beta}}$ by the square root of $\sigma_k^2 \mathbf{c'(X'X)^{-1}c}$ in each voxel separately.

Hence, making use of the result $\mathrm{Var}(\mathbf{C'\hat{\beta}}) = \left(\mathbf{S} \otimes \mathbf{c'(X'X)^{-1}c}\right)$ of the previous section,

$$\mathrm{Var}(\mathbf{z}_{\mathrm{SPM}}) = \left[\mathbf{K} \otimes \mathbf{c'}\left(\mathbf{X'X}\right)^{-1}\mathbf{c}\right]^{-\frac{1}{2}} \mathrm{Var}(\mathbf{C'\hat{\beta}}) \left[\mathbf{K} \otimes \mathbf{c'}\left(\mathbf{X'X}\right)^{-1}\mathbf{c}\right]^{-\frac{1}{2}}$$

$$= \left(\mathbf{K}^{-\frac{1}{2}} \otimes \frac{1}{\sqrt{\mathbf{c'}\left(\mathbf{X'X}\right)^{-1}\mathbf{c}}}\right) \left(\mathbf{S} \otimes \mathbf{c'(X'X)^{-1}c}\right) \left(\mathbf{K}^{-\frac{1}{2}} \otimes \frac{1}{\sqrt{\mathbf{c'}\left(\mathbf{X'X}\right)^{-1}\mathbf{c}}}\right)$$

$$= \mathbf{K}^{-\frac{1}{2}}\mathbf{S}\mathbf{K}^{-\frac{1}{2}} \otimes 1$$

$$= \mathbf{G}.$$

Appendix C. Extension to two nested random components.

In the situation that motivated this study, the first level estimate is given by the average of the voxelwise signal y_{ij} over the acquisitions j ($j = 1, 2, \dots n$) in each subject i ($i = 1, 2, \dots a$). Hence, the sampling structure of the data includes two random components, assumed to be independent: a subject-to-subject variance σ_A^2, representing the variation of a random variable A (the mean signal in each subject), and an acquisition-to-acquisition variance σ^2, this latter representing variation in the errors (to simplify notation, we omit the subscript indexing voxels when the context allows it). Using standard ANOVA results, the model and the sampling variance of this average signal $\bar{y}_{i\cdot}$, when brought to the second level, will be in each voxel

$$\bar{y}_{i\cdot} = \mu + X_i\beta + \eta_i, \ \mathrm{Var}(\eta) = \sigma_A^2 + \frac{1}{n}\sigma^2, \tag{C.1}$$

where μ is the intercept, X_i the second-level predictor with its associated voxelwise effect β, and η_i are the errors at the second level. The model equation of the second level is structurally the same as in the previous section, except that the errors variance $\mathrm{Var}(\eta)$ is now composed by the subject-to-subject variance σ_A^2, and its contamination by acquisition-to-acquisition variance $n^{-1}\sigma^2$, reflecting the imprecision with which the mean signal is estimated

in each subject. One may therefore apply the previous analysis, now referring to second-level

estimation, after replacing the covariance of the errors.

Let $\mathbf{K}_A$ and $\mathbf{K}$ be the $v \times v$ matrices collecting at the diagonal the values of σ_A^2 and σ^2

at each voxel, and $\mathbf{Q}$ and $\mathbf{G}$ the spatial correlation $v \times v$ matrices relative to subject-to-subject

and acquisition-to-acquisition variation, respectively. Using the same data-stacking scheme as

previously and an analogous notation, the model at the second level for the data in all voxels

is now

$$\overline{\mathbf{y}} = (\mathbf{I}_v \otimes \mathbf{X})\boldsymbol{\beta} + \boldsymbol{\eta},$$

$$\mathrm{Var}(\boldsymbol{\eta}) = \left(\mathbf{K}_A^{\frac{1}{2}}\mathbf{Q}\mathbf{K}_A^{\frac{1}{2}} + \frac{1}{n}\mathbf{K}^{\frac{1}{2}}\mathbf{G}\mathbf{K}^{\frac{1}{2}} \right) \otimes \mathbf{I}_a.$$

$$(C.2)$$

Here, the expression for the covariance of the second-level errors $\boldsymbol{\eta}$ is given by the same

weighted sum of the variance components of the voxel-by-voxel model (C.1), assuming

additivity of the spatial correlations $\mathbf{Q}$ and $\mathbf{G}$ (the coefficients of the spatial covariances are

those of the random effects in the expected mean squares of the ANOVA model, scaled by n).

After replacing the spatial covariance $\mathbf{S}$ of equation (4) with that of $\boldsymbol{\eta}$, one obtains

$$\mathrm{Var}(\mathbf{C}'\hat{\boldsymbol{\beta}}) = \left(\mathbf{K}_A^{\frac{1}{2}}\mathbf{Q}\mathbf{K}_A^{\frac{1}{2}} + \frac{1}{n}\mathbf{K}^{\frac{1}{2}}\mathbf{G}\mathbf{K}^{\frac{1}{2}} \right) \otimes \mathbf{c}'\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{c}.$$

$$(C.3)$$

The developoment parallels that of Appendix B after replacing variances and covariances of

$\mathrm{Var}(\boldsymbol{\varepsilon})$ with those of $\mathrm{Var}(\boldsymbol{\eta})$. For known $\mathbf{K}_A$ and $\mathbf{K}$, the parametric map is given by

$$\mathbf{z}_{\mathrm{SPM}} = \left[\left(\mathbf{K}_A + \frac{1}{n}\mathbf{K} \right) \otimes \mathbf{c}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{c} \right]^{-\frac{1}{2}} \mathbf{C}'\hat{\boldsymbol{\beta}}.$$

Hence, making use of equation (C.3) for $\mathrm{Var}(\mathbf{C}'\hat{\boldsymbol{\beta}})$,

$$\mathrm{Var}(\mathbf{z}_{\mathrm{SPM}}) = \left[\left(\mathbf{K}_A + \frac{1}{n}\mathbf{K} \right) \otimes \mathbf{c}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{c} \right]^{-\frac{1}{2}} \mathrm{Var}\left(\mathbf{C}'\hat{\boldsymbol{\beta}}\right) \left[\left(\mathbf{K}_A + \frac{1}{n}\mathbf{K} \right) \otimes \mathbf{c}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{c} \right]^{-\frac{1}{2}}$$

$$= \left(\mathbf{K}_A + \frac{1}{n}\mathbf{K} \right)^{-\frac{1}{2}} \left(\mathbf{K}_A^{\frac{1}{2}}\mathbf{Q}\mathbf{K}_A^{\frac{1}{2}} + \frac{1}{n}\mathbf{K}^{\frac{1}{2}}\mathbf{G}\mathbf{K}^{\frac{1}{2}} \right) \left(\mathbf{K}_A + \frac{1}{n}\mathbf{K} \right)^{-\frac{1}{2}} \otimes 1$$

$$= \left(\mathbf{K}_A + \frac{1}{n}\mathbf{K} \right)^{-\frac{1}{2}} \mathbf{K}_A^{\frac{1}{2}}\mathbf{Q}\mathbf{K}_A^{\frac{1}{2}} \left(\mathbf{K}_A + \frac{1}{n}\mathbf{K} \right)^{-\frac{1}{2}} +$$

$$+ \frac{1}{n}\left(\mathbf{K}_A + \frac{1}{n}\mathbf{K} \right)^{-\frac{1}{2}} \mathbf{K}^{\frac{1}{2}}\mathbf{G}\mathbf{K}^{\frac{1}{2}} \left(\mathbf{K}_A + \frac{1}{n}\mathbf{K} \right)^{-\frac{1}{2}}$$

$$= \left(\mathbf{I}_V + n^{-1}\frac{\mathbf{K}}{\mathbf{K}_A} \right)^{-\frac{1}{2}} \mathbf{Q} \left(\mathbf{I}_V + n^{-1}\frac{\mathbf{K}}{\mathbf{K}_A} \right)^{-\frac{1}{2}} + \left(\mathbf{I}_V + n\frac{\mathbf{K}_A}{\mathbf{K}} \right)^{-\frac{1}{2}} \mathbf{G} \left(\mathbf{I}_V + n\frac{\mathbf{K}_A}{\mathbf{K}} \right)^{-\frac{1}{2}}.$$

Using this result, the spatial covariance of the resulting parametric map $\mathbf{z}_{\mathrm{SPM}}$ may be written

as a weighted sum of the spatial correlations,

$$\mathrm{Var}(\mathbf{z}_{\mathrm{SPM}}) = \mathbf{W}_A^{\frac{1}{2}}\mathbf{Q}\mathbf{W}_A^{\frac{1}{2}} + \mathbf{W}^{\frac{1}{2}}\mathbf{G}\mathbf{W}^{\frac{1}{2}}, \tag{C.4}$$

where the weighting diagonal matrices $\mathbf{W}_A$ and $\mathbf{W}$ are given by

$$\mathbf{W}_A = \left(\mathbf{I}_v + n^{-1}\frac{\mathbf{K}}{\mathbf{K}_A} \right)^{-1}, \quad \mathbf{W} = \left(\mathbf{I}_v + n\frac{\mathbf{K}_A}{\mathbf{K}} \right)^{-1}, \tag{C.5}$$

$\mathbf{K}/\mathbf{K}_A$ is the diagonal matrix collecting the variance ratio at each voxel $\sigma^2/\sigma_A^2 = \gamma$, and

$$\mathbf{W}_A + \mathbf{W} = \mathbf{I}_v.$$

These weights reflect the relative importance of subject-to-subject and acquisition-to-acquisition variance in the influence of the respective correlation patterns. The relative contribution of the acquisition-to-acquisition correlation $\mathbf{G}$ as determined by its weight $\mathbf{W}$, however, is discounted by the number of acquisitions at the first level n. From equation (C.5), one can see that with increasing n, $\mathbf{W}_A$ approaches $\mathbf{I}_v$, while $\mathbf{W}$ vanishes. Conversely, for $n = 1$, $\mathbf{W}_A$ is the matrix collecting on its diagonal the intraclass correlation at each voxel $\rho = \sigma_A^2/(\sigma_A^2 + \sigma^2)$, while $\mathbf{W}$ has on the diagonal $1 - \rho$. Note that, as before, the number of subjects in the sample does not affect the spatial covariance of the parametric map in the limiting case in which the voxelwise variances are known.

To show that $\mathbf{W}_A + \mathbf{W} = \mathbf{I}_v$, note that $\mathbf{W}_A$ and $\mathbf{W}$ are diagonal matrices in which each diagonal element is a function of the voxelwise variances. Hence, each of these diagonal elements in the sum $\mathbf{W}_A + \mathbf{W}$ is

$$\left(1 + \frac{\sigma^2}{n\sigma_A^2}\right)^{-1} + \left(1 + \frac{n\sigma_A^2}{\sigma^2}\right)^{-1}$$

$$= \left(\frac{n\sigma_A^2 + \sigma^2}{n\sigma_A^2}\right)^{-1} + \left(\frac{\sigma^2 + n\sigma_A^2}{\sigma^2}\right)^{-1}$$

$$= \frac{n\sigma_A^2}{n\sigma_A^2 + \sigma^2} + \frac{\sigma^2}{n\sigma_A^2 + \sigma^2} = 1.$$

Therefore, the fractional contribution of subject-to-subject and acquisition-to-acquisition variation to the covariance of the parametric map is given by $\mathbf{W}_A$ and $\mathbf{W}$ themselves:

$$\frac{\mathbf{W}_A}{\mathbf{W}_A + \mathbf{W}} = \mathbf{W}_A, \quad \frac{\mathbf{W}}{\mathbf{W}_A + \mathbf{W}} = \mathbf{W}.$$

In practical applications, the sum of the weights may be larger than unity since the voxelwise variances are estimated from the data.

In Figure 7A, the theoretical fractional contributions of acquisition-to-acquisition (in green) and subject-to-subject (in blue) covariance were plotted for the values of variance ratio $\gamma = 10$ and $\gamma = 100$, using equation (C.5). These values are representative of variance ratios in the gray and white matter, respectively (Figure 7B, 'Variance ratio'), for rest perfusion data of 8 min (60 acquisitions). One can see that in the gray matter the number of acquisitions necessary to damp acquisition-to-acquisition variance is reached quickly. In the white matter, by contrast, there are circumscribed areas where this random effect remains influential. However, with the possible exception of the anterior temporal lobe and the orbitofrontal cortex below $z = -20$, these areas are too small to form large clusters. Figure 7C, 'Component #1', shows that the map of the first spatial component of acquisition-to-acquisition variance has little contrast where the fractional ratio is high. The only possible exception is, again, the

base of the brain, and the tissue surrounding the vertices of the ventricle horns. We conclude that these circumscribed regions are vulnerable to displaying spatial patterns arising from acquisition-to-acquisition variation in rest perfusion imaging studies.

FIGURE 7

Tables

Table 1.

Source of variation	Degrees of freedom	MS	E(MS)
Between groups	$a-1$	$\mathrm{MS_{GROUPS}=SS_{GROUPS}}\,/\,(a-1)$	$\sigma^2 + n\sigma^2_{\mathrm{SUBJ}} + bn\kappa^2_{\mathrm{GROUPS}}$
Between subjects	$a(b-1)$	$\mathrm{MS_{SUBJ}=SS_{SUBJ}}\,/\,a(b-1)$	$\sigma^2 + n\sigma^2_{\mathrm{SUBJ}}$
Within	$ab(n-1)$	$\mathrm{MS}_w=\mathrm{SS}_w\,/\,[ab(n-1)]$	σ^2

Explanation of symbols. SS: sums of squares; MS, E(MS): mean squares, expected mean squares; a: number of groups; b: number of subjects per group; n: number of acquisitions per subject; $\kappa^2_{\mathrm{GROUPS}}$ variation due to group effect (fixed); σ^2_{SUBJ}: subject-to-subjects variance (random); σ^2: residual error variance (acquisition-to-acquisition variance).

Table 2.

Source of variation	Degrees of freedom	MS	E(MS)
Between subjects	$a-1$	$\mathrm{MS_{SUBJ}}=\mathrm{SS_{SUBJ}}/(a-1)$	$\sigma^2 + bn\sigma^2_{\mathrm{SUBJ}}$
Between treatments within subjects	$b-1$	$\mathrm{MS_{TREAT}}=\mathrm{SS_{TREAT}}/(b-1)$	$\sigma^2 + n\sigma^2_{\mathrm{SUBJ\times TREAT}} + an\kappa^2_{\mathrm{TREAT}}$
Interaction	$(a-1)(b-1)$	$\mathrm{MS_{SUBJ\times TREAT}}=$ $=\mathrm{SS_{SUBJ\times TREAT}}/[(a-1)(b-1)]$	$\sigma^2 + n\sigma^2_{\mathrm{SUBJ\times TREAT}}$
Within	$ab(n-1)$	$\mathrm{MS}_w=\mathrm{SS}_w/[ab(n-1)]$	σ^2

Explanation of symbols. SS: sums of squares; MS, E(MS): mean squares, expected mean squares; a: number of subjects; b: number of treatments, or levels, in the experimental factor (task); n: number of acquisitions per treatment in each subject; σ^2_{SUBJ}: subject-to-subject variance (random); $\kappa^2_{\mathrm{TREAT}}$: variation due to treatments within-subjects (fixed); $\sigma^2_{\mathrm{SUBJ\times TREAT}}$: subjects × treatment interaction (random); σ^2: residual error variance (acquisition-to-acquisition variance).

Figure legends

Figure 1. Thresholded t maps from models including global perfusion as covariate. A: Regression of rest perfusion values on working memory capacity scores, measured by the OSPAN test (Engle et al. 1999) in a sample of N=146. B: t maps in perfusion differences according to the serotonin transporter promoter genotype; the reduction in perfusion shown here affects carriers of the short 'risk' allele (sample size N=183). C: reduction in perfusion in individual with higher perfusion scores at the ERQ scale (Gross and John 2003), in a sample of N=29. D: the second principal component of rest perfusion data shows dissociation in perfusion values between the posterior and the anterolateral part of the brain, roughly corresponding to the areas where the t maps show a large cluster. All slices were taken at $z = -10$, Montreal Neurological Institute coordinates.

Figure 2. Principal component analysis of rest perfusion data. First (A) and second (B) correlation components, computed from the whole sample of 295 average rest perfusion images. The top and bottom parts of the brain were removed to avoid differences in coverage to affect the estimate of the main components. The scale and sign of these images are arbitrary. The position of the slice is given in Montreal Neurological Institute coordinates.

Figure 3. Overthreshold rates. A, B: count rates (in percent) of voxels included in clusters selected by an extent of 200 voxels and an uncorrected significance value $p = 0.001$ from 8000 two-sample t tests on randomly selected groups of 12 subjects each. In B, the count was repeated after including the global perfusion signal as a covariate (as in the analyses of Figure 1). C: the same statistic, after correcting for local anisotropy. D: count rates in artificial images randomly generated according to the Gaussian random field distribution. To obtain approximately comparable count rates, the artificial images were thresholded at $p = 0.005$, cluster size 50 voxels.

Figure 4. Observed and fitted overthreshold rates. Box plots of observed rates (marked with 'Obs' on the x axis) and net rates of individual components (#1 to #4 on the x axis) contributing to the fitted probabilities of reaching the cluster-defining threshold. The data contributing to the plots are those of all voxels in a volume; circles are outlier voxels where the cluster-defining threshold is reached more often. The first box on the left in each plot (in blue) contains the observed count rates across the volume, in per thousand. Because the uncorrected threshold was $p = 0.001$, unity on the x axis corresponds to the expected uncorrected overthreshold rate of $1^0/_{00}$, if clusters had not been filtered by size. All remaining boxes in the plot (in black) summarize the net count rate contributed by each of the four components in the logistic fit, i.e. the extent to which each component can predict the occurrence of overthreshold voxels. The horizontal line at the centre of the boxes is the median net fitted rate due to the component over the whole volume, and the spread reflects its variation from voxel to voxel. From left to right, the plots refer to clusters for increased signal (also shown in Figure 3A), to ranked data, to clusters computed with RESEL correction, and to artificial data following the Gaussian random field distribution (also shown in Figure 3D). The artificial data were obtained at the lower cluster-defining threshold $p = 0.005$ and cluster size 50 voxels.

Figure 5. Influence of cluster extent threshold on observed and fitted overthreshold rates. Plot of observed and net fitted count rates, computed for parametric maps in which clusters were filtered by progressively larger extent size (from left to right, starting with no cluster size filter). One can see that increasing the cluster size filter leads to progressively larger influence of the spatial components in count rates.

Figure 6. Overthreshold rates of cluster peaks. A: RESEL density map, showing decreased local smoothness in white matter. B: count rates (per ten thousand) of cluster peaks computed from 80 000 t test. Unlike cluster data, the spatial distribution of these peaks shows an association with the RESEL density map.

Figure 7. Influence of acquisition-to-acquisition covariance on overthreshold rates. A: plots of the fractional contributions of acquisition-to-acquisition (in green) and subject-to-subject (in blue) to the spatial covariance of a parametric are plotted for values of variance ratios representative of gray and white matter. One can see that the contribution of acquisition-to-acquisition variation decreases with increasing numbers of acquisitions. Representative maps of these variance ratios for rest perfusion data acquired for 8 min (60 acquisitions) are shown in part B. In part C, bottom, the first component of acquisition-to-acquisition variance, showing the most important patterns that may be expected from this random effect.

References

Abler, B., Hofer, C., Viviani, R., 2008. Habitual emotion regulation strategies and baseline brain perfusion. NeuroReport 19, 21-24.

Alsop, D.C., Detre, J.A., 1996. Reduced transit-time sensitivity in noninvasive magnetic resonance imaging of human cerebral blood flow. J. Cereb. Blood Flow Metab. 16, 1236-1249.

Alsop, D.C., Detre, J.A., 1998. Multi section cerebral blood flow MR imaging with continuous arterial spin labeling. Radiology 208, 410-416.

Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry - The methods. NeuroImage 11, 805-821.

Beschoner, P., Richter, S., Lo, H., Sim, E.J., Baron, K., Osterfeld, N., Horn, A.B., Viviani, R., 2008. Baseline brain perfusion and working memory capacity: A neuroimaging study. NeuroReport 19, 1803-1807.

Beschoner, P., Riecker, A., Viviani, R., 2005. A review of magnetic resonance tomography findings in schizophrenia: Magnetization transfer ratio and diffusion-weighted imaging. Nervenheilkunde 3, 190-197.

Bullmore, E., Suckling, J., Brammer, M., 2001. In praise of tedious permutation. In: Moore, M. (ed). Spatial Statistics: Methodological Aspects and Applications. Springer, New York.

Chumbley, J.R., Worsley, K., Flandin, G., Friston, K.J., 2009. Topological FDR for neuroimaging. NeuroImage 49, 3057-3064.

Detre, J.A., Leigh, J.S., Williams, D.S., Koretsky, A.P., 1992. Perfusion imaging. Magn. Reson. Med. 3, 454-462.

Efron, B., 2008. Microarrays, empirical Bayes and the two-groups model. Statist. Sci. 23, 1-22.

Efron, B., 2010. The future of indirect evidence. Statist. Sci. (in press). Available from http://www-stat.stanford.edu/~ckirby/brad/papers/.

Engle, R.W., Tuholski, S.W., Laughlin, J.E., Conway, A.R.A., 1999. Working memory, short-term memory and general fluid intelligence: A latent variable approach. J. Exp. Psychol., Gen. 128, 309-331.

Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C., 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. Magn. Reson. Med. 33, 636-647.

Friston, K.J., Worsley, K.J., Frackowiak, R.S.J., Mazziotta, J.C., Evans, A.C., 1994. Assessing the significance of focal activations using their spatial extent. Hum. Br. Mapping 1, 214-220.

Gelman, A., Hill, J., 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, Cambridge (UK).

Gelman, A., Pardoe, I., 2007. Average predictive comparisons for models with nonlinearity, interactions, and variance components. Sociolog. Methodol. 37, 23-51.

Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. NeuroImage 15, 870-878.

Gonzalez-At, J.B., Alsop, D.C., Detre, J.A., 2000. Cerebral perfusion and arterial transit time changes during task activation determined with continuous arterial spin labeling. Magn. Reson. Med. 43, 739-746.

Gross, J.J., John, O.P., 2003. Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. J. Pers. Soc. Psychol. 85, 348-362.

Hayasaka, S., Phan, K.L., Liberzon, I., Worsley, K.J., Nichols, T.E., 2004. Nonstationary cluster-size inference with random field and permutation methods. NeuroImage 22, 676-687.

Heller, R., Stanley, D., Yekuteli, D., Rubin, N., Benjamini, Y., 2006. Cluster-based analysis of FMRI data. NeuroImage 33, 599-608.

Hernandez-Garcia, L., Lee, G.R., Vazquez, A.L., Yip, C.Y., Noll, D.C., 2005. Quantification of perfusion fMRI using a numerical model of arterial spin labeling that accounts for dynamic transit time effects. Magn. Reson. Med. 54, 955-964.

Holmes, A.P., Blair, R.C., Watson, J.D.G., Ford, I., 1996. Nonparametric analysis of statistic images from functional mapping experiments. J. Cereb. Blood Flow Metab. 16, 7-22.

Kherif, F., Josse, G., Seghier, M.L., Price, C.J., 2008. The main sources of intersubject variability in neuronal activation for reading aloud. J. Cogn. Neurosci. 21, 654-668.

Kiebel, S.J., Poline, J.B., Friston, K.J., Holmes, K.J., 1999. Robust smoothness estimation in statistical parametric maps using standardized residuals from the general linear model. NeuroImage 10, 756-766.

Kirchheiner, J., Seeringer, A., Godoy, A.L., Maier, C., Beschoner, P., Sim, E.J., Viviani, R., 2010. *CYP2D6* in the brain: Genotype effects on resting brain perfusion. Mol. Psychiatry doi:10.1038/mp.2010.42 (epub ahead of print).

Lee, G.R., Hernandez-Garcia, L., Noll, D.C., 2007. Functional imaging with Turbo-CASL: Transit time and multislice imaging considerations. Magn. Reson. Med. 57, 661-669.

Mardia, K.V., Kent, J.T., Bibby, J.M., 1979. Multivariate Analysis. Academic Press, London.

Mildner, T., Möller, H.E., Driesel, W., Norris, D.G., Trampel, R., 2005. Continuous arterial spin labeling at the human common carotid artery. NMR Biomed. 18, 19-23.

Miller, M.B., Van Horn, J.D., 2007. Individual variability in brain activations associated with episodic retrieval: A role for largs-scale databases. Int. J. Psychophysiol. 63, 205-213.

Nichols, T.E., Holmes, A.P., 2001. Nonparametric permutation tests for functional neuroimaging: A primer with examples. Hum. Br. Mapping 15, 1-25.

Noppeney, U., Penny, W.D., Price, C.J., Flandin, G., Friston, K.J., 2006. Identification of degenerate neuronal systems based on intersubject variability. NeuroImage 30, 885-890.

Penny, W., Friston, K., 2003. Mixtures of general linear models for functional neuroimaging. IEEE Trans. Med. Imag. 22, 504-514.

Poldrack, R.A., Fletcher, P.C., Henson, R.N., Worsley, K.J., Brett, M., Nichols, T.E., 2008. Guidelines for reporting an fMRI study. NeuroImage 40, 409-414.

Poline, J.-B., Mazoyer, B.M., 1993. Analysis of individual positron emission tomography activation maps by detection of high signal-to-noise pixel clusters. J. Cereb. Blood Flow Metab. 13, 425-437.

Roy, S.N., Gnanadesikan, R., Srivastava, J.N., Fowlkes, E.B., Lee, E.T., 1971. Analysis and Design of Certain Quantitative Multiresponse Experiments. Pergamon Press, Oxford.

Seghier, M.L., Lee, H.L., Schofield, T., Ellis, C.L., Price, C.J., 2008. Inter-subject variability in the use of two different neuronal networks for reading aloud familiar words. NeuroImage 42, 1226-1236.

Smith, S.M., Beckmann, C.F., Ramnani, N., Wollrich, M.W., Bannister, P.R., Jenkinson, M., Matthews, P.M., McGonigle, D.J., 2005. Variability in fMRI: A reexamination of inter-session differences. Hum. Br. Mapping 24, 248-257.

Snedecor, G.W., Cochran, W.G., 1967. Statistical Methods (6th ed.). The Iowa State University Press, Ames (Iowa).

Taylor, J.E., Worsley, K.J., 2007. Detecting sparse signals in random fields, with an application to brain mapping. J. Am. Statist. Assoc. 102, 913-928.

Viviani, R., Beschoner, P., Ehrhard, K., Schmitz, B., Thöne, J., 2007. Non-normality and transformations of random fields, with an application to voxel-based morphometry. NeuroImage 35, 121-130.

Viviani, R., Beschoner, P., Jaeckle, T., Hipp, P., Kassubek, J., Schmitz, B., 2007. The bootstrap and cross-validation in neuroimaging applications: Estimation of the distribution of extrema of random fields for single volume tests, with an application to ADC maps. Hum. Br. Mapping 28, 1075-1088.

Viviani, R., Beschoner, P., Lo, H., Richter, S., Haffer, S., Osterfeld, N., Thöne, J., Sim, E.J., 2010b. Components of acquisition-to-acquisition variance in continuous arterial spin labelling (CASL) imaging. BMC Neurosci. 11, 30.

Viviani, R., Grön, G., Spitzer, M., 2005. Functional principal component analysis of fMRI data. Hum. Br. Mapping 24, 109-129.

Viviani, R., Sim, E.J., Lo, H., Beschoner, P., Osterfeld, N., Maier, C., Seeringer, A., Godoy, A., Rosa, A., Comas, D., Kirchheiner, J., 2010a. Baseline brain perfusion and the serotonin transporter promoter polymorphism. Biol. Psychiatry 67, 317-322.

Viviani, R., Sim, E.J., Lo, H., Richter, S., Haffer, S., Osterfeld, N., Thöne, J., Beschoner, P., 2009. Components of variance in brain perfusion and the design of studies of individual differences: The baseline study. NeuroImage 46, 12-22.

Wager, T.D., Lindquist, M., Kaplan, L., 2007. Meta-analysis of functional neuroimaging data: Current and future directions. Soc. Cogn. Aff. Neurosci. 2, 150-158.

Wang, J., Zhang, Y., Wolf, R.L., Roc, A.C., Alsop, D.C., Detre, J.A., 2005. Amplitude-modulated continuous arterial spin-labeling 3.0-T perfusion MR imaging with a single coil: Feasibility study. Radiology 235, 218-228.

Wei, X., Yoo, S.S., Dickey, C.C., Zou, K.H., Guttmann, C.R.G., Panych, L.P., 2004. Functional MRI of auditory verbal working memory: Long-term reproducibility analysis. NeuroImage 21, 1000-1008.

Williams, D.S., Detre, J.A., Leigh, J.S., Koretsky, A.P., 1992. Magnetic resonance imaging of perfusion using spin-inversion of arterial water. Proc. Natl Acad. Sci. USA 89, 212-216.

Winer, B.J., Brown, D.R., Michels, K.M., 1991. Statistical Principles in Experimental Design (3rd ed.). McGraw Hill, New York.

Worsley, K.J., Andermann, M., Koulis, T., MacDonald, D., Evans, A.C., 1999. Detecting changes in non-isotropic images. Hum. Br. Mapping 8, 98-101.

Worsley, K.J., Marrett, S., Neelin, P., Evans, A.C., 1992. A three-dimensional statistical analysis for CBF activation studies in human brain. J. Cereb. Blood Flow Metab. 12, 900-918.
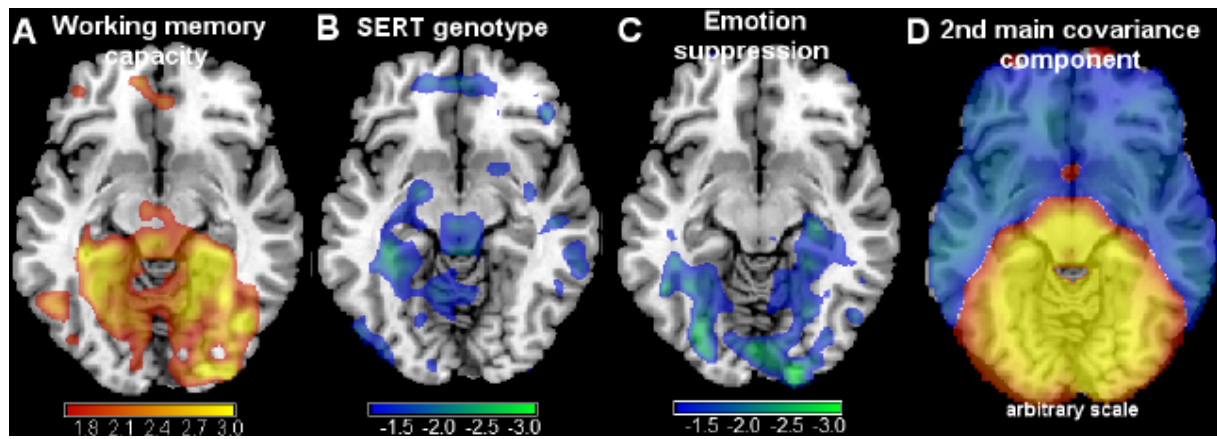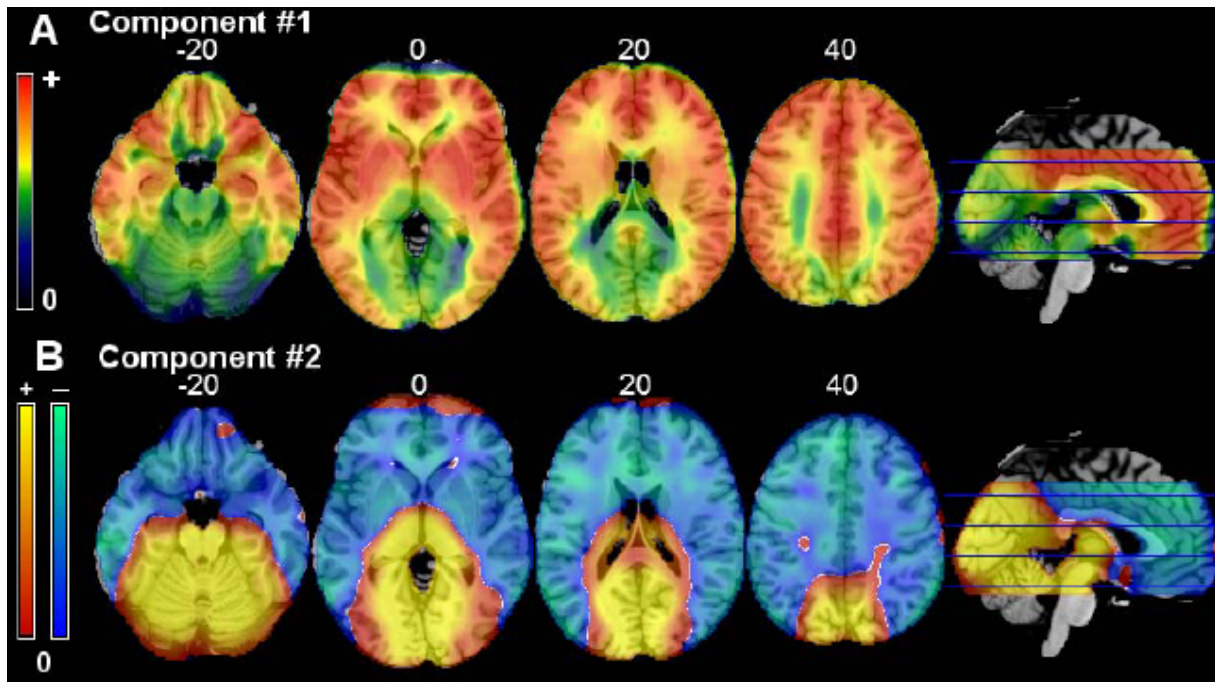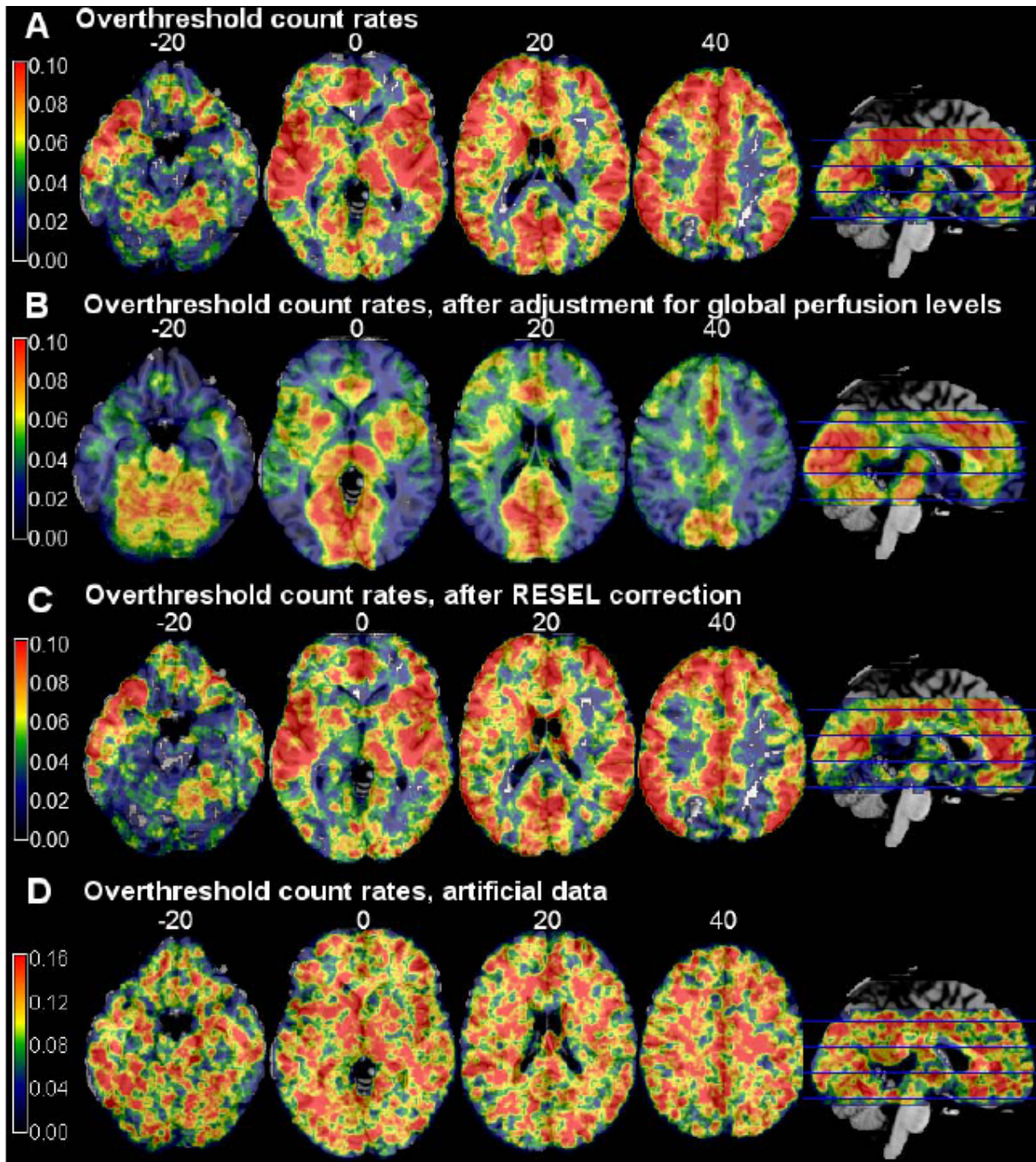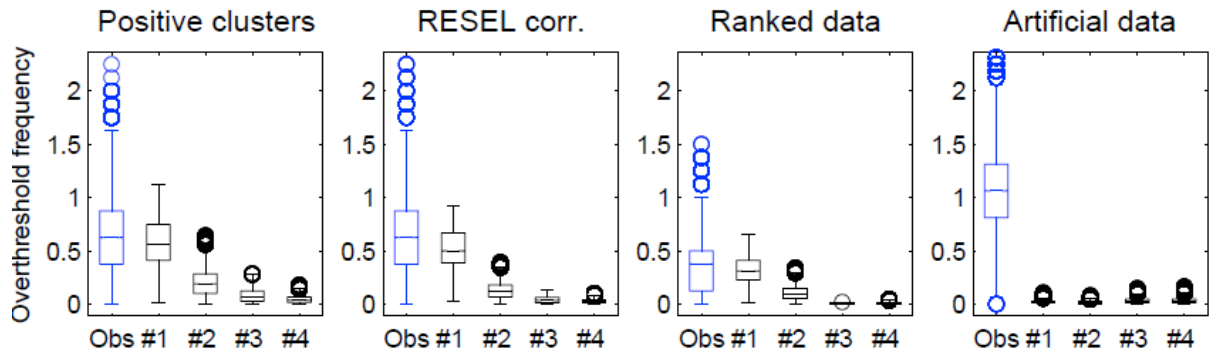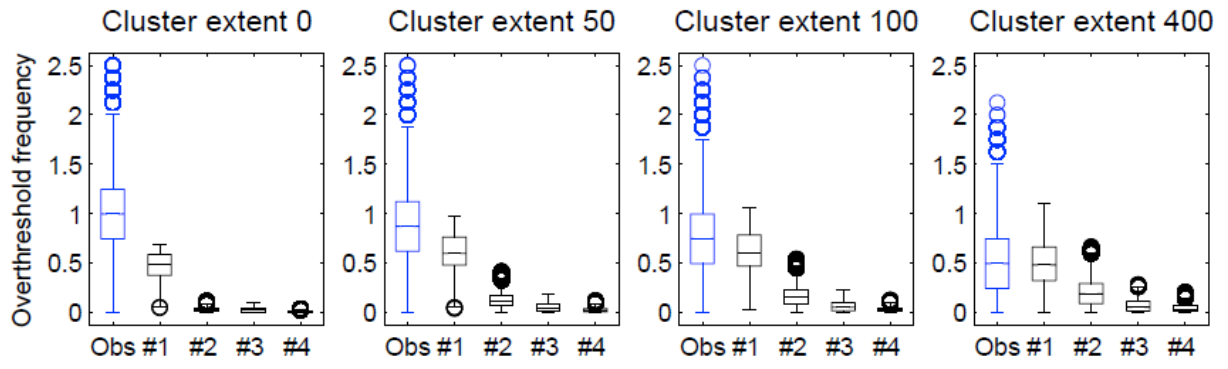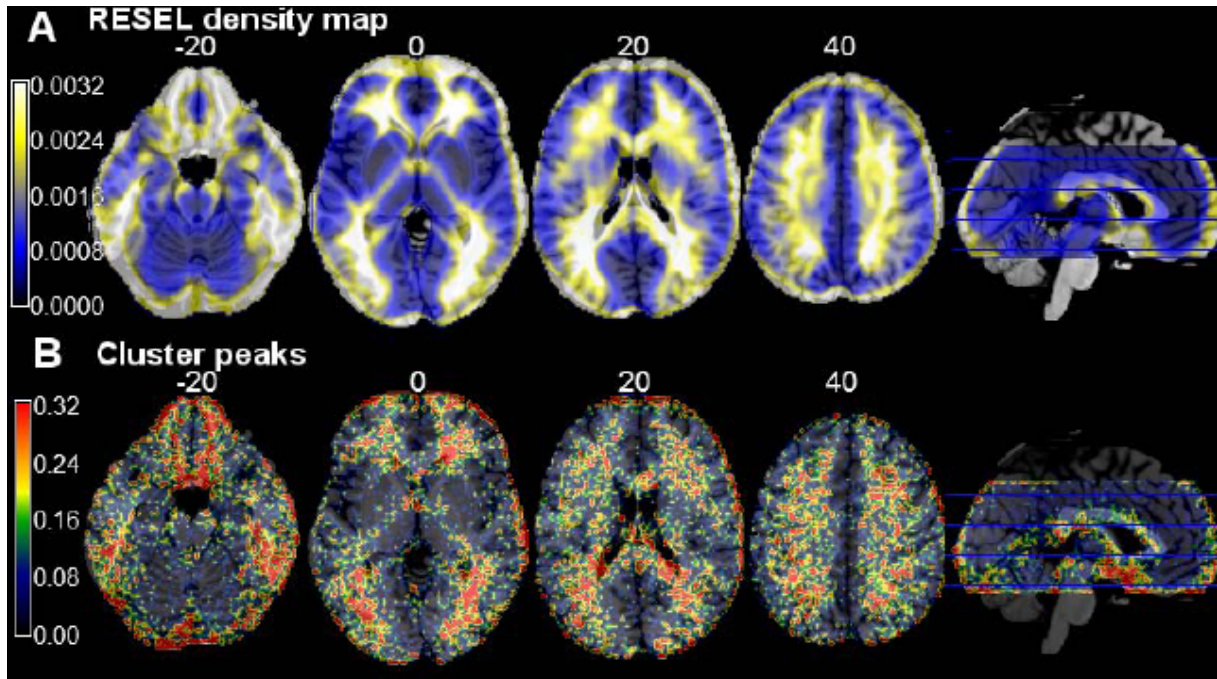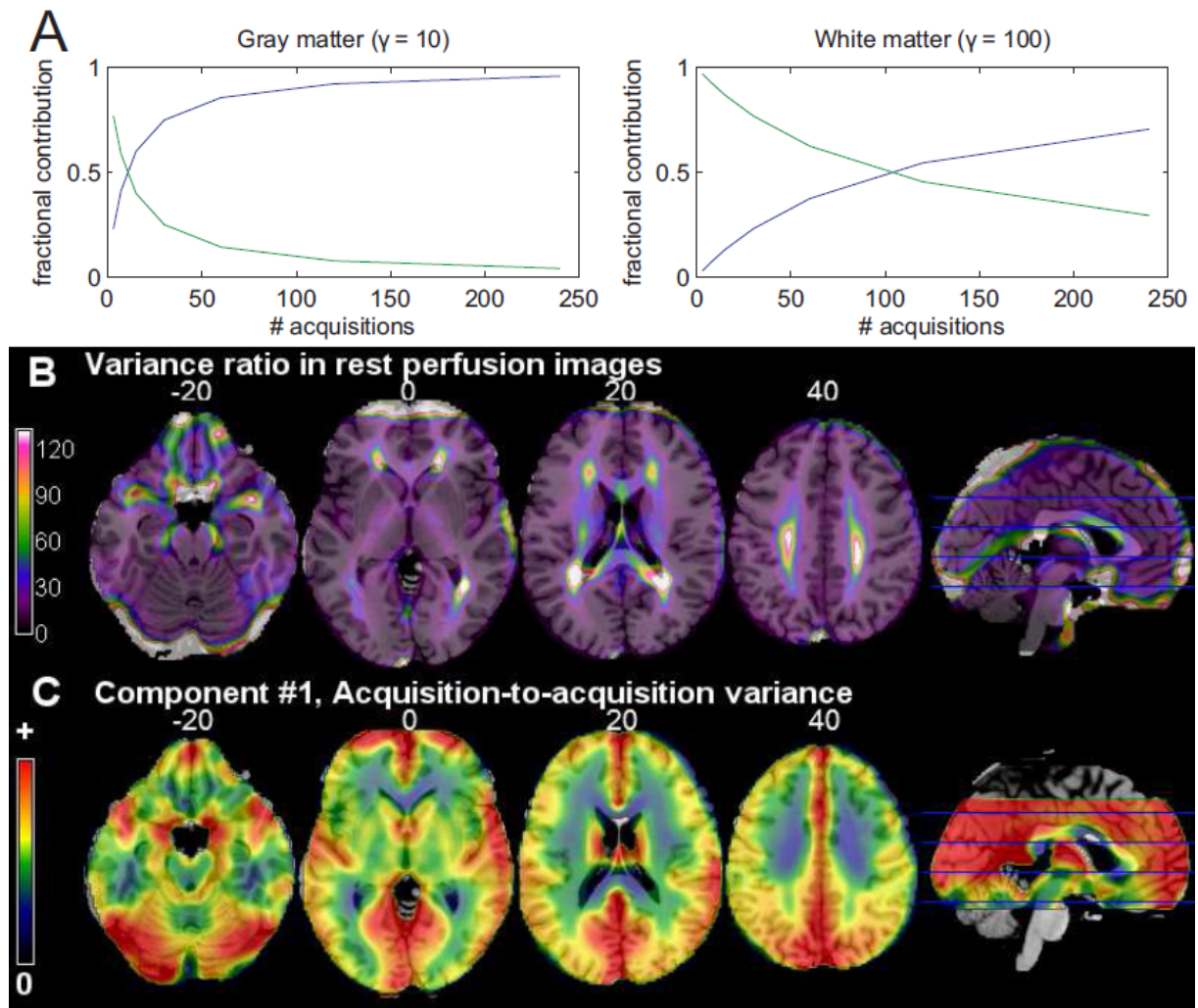
Figures

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7