

[Read](#)[View source](#)[View history](#)

Search



# Statistical parametric mapping (SPM)

 Recommend this on Google

Guillaume Flandin and Karl J. Friston (2008), Scholarpedia,  
3(4):6232.

doi:10.4249/scholarpedia.6232

revision #91821 [[link to/cite this article](#)]

- **Dr. Guillaume Flandin**, Wellcome Trust Centre for Neuroimaging, London, UK
- **Karl J. Friston**, Wellcome Department of Imaging Neuroscience, London, UK

## Post-publication activity

Curator: [Guillaume Flandin](#)



**Statistical parametric mapping** is the application of Random Field Theory to make inferences about the topological features of statistical processes that are continuous functions of space or time. It is usually used to identify regionally specific effects (*e.g.*, [brain](#) activations) in [neuroimaging](#) data to characterize functional anatomy and disease-related changes.

Contents [[hide](#)]

## 1 Statistical parametric mapping

### 1.1 The general linear model

### 1.2 Testing for contrasts

## 2 Topological inference and the theory of random fields

### 2.1 Anatomically closed hypotheses

Typesetting math: 100% [open hypotheses and levels of inference](#)

[3 Recommended reading](#)

[4 References](#)

[5 External links](#)

[6 See also](#)

## Statistical parametric mapping

---

Brain mapping studies are usually analyzed with some form of statistical parametric mapping. This entails the construction of continuous statistical processes to test hypotheses about regionally specific effects (Friston et al. 1991). Statistical Parametric Maps (SPM) are images or fields with values that are, under the null hypothesis, distributed according to a known probability density function, usually the Student's  $t$  or  $F$ -distributions. These are known colloquially as  $t$ - or  $F$ -maps. The success of statistical parametric mapping is due largely to the simplicity of the idea. Namely, one analyses each and every voxel (*i.e.*, image volume element) using any standard (univariate) statistical test, usually based on a General Linear Model (GLM) of the data. The resulting statistics are assembled into an image - the SPM. SPMs are interpreted as continuous statistical processes by referring to the probabilistic behaviour of random fields (Adler 1981, Worsley et al. 1992, Friston et al. 1994, Worsley et al. 1996). Random fields model both the univariate probabilistic characteristics of an SPM and any non-stationary spatial **covariance** structure. 'Unlikely' topological features of the SPM, like peaks or clusters, are interpreted as regionally specific effects, attributable to the experimental manipulation. In short, the GLM is used to explain continuous (image) data in exactly the same way as in conventional analyses of discrete data. Random Field Theory (RFT) is used to resolve the multiple-comparison problem when making inferences over the volume analysed. RFT provides a method for adjusting  $p$ -values for the search volume and plays the same role for SPMs as the Bonferroni correction for discrete statistical tests.

### The general linear model

Statistical analysis  
of imaging data

Statistical parametric map (SPM)

corresponds to inverting generative models of data. Inferences are then pursued using statistics that assess the significance of interesting effects. A brief review of the literature may give the impression that there are numerous ways to analyze neuroimaging time-series (e.g., from Positron emission tomography (PET), functional magnetic resonance imaging (fMRI) and

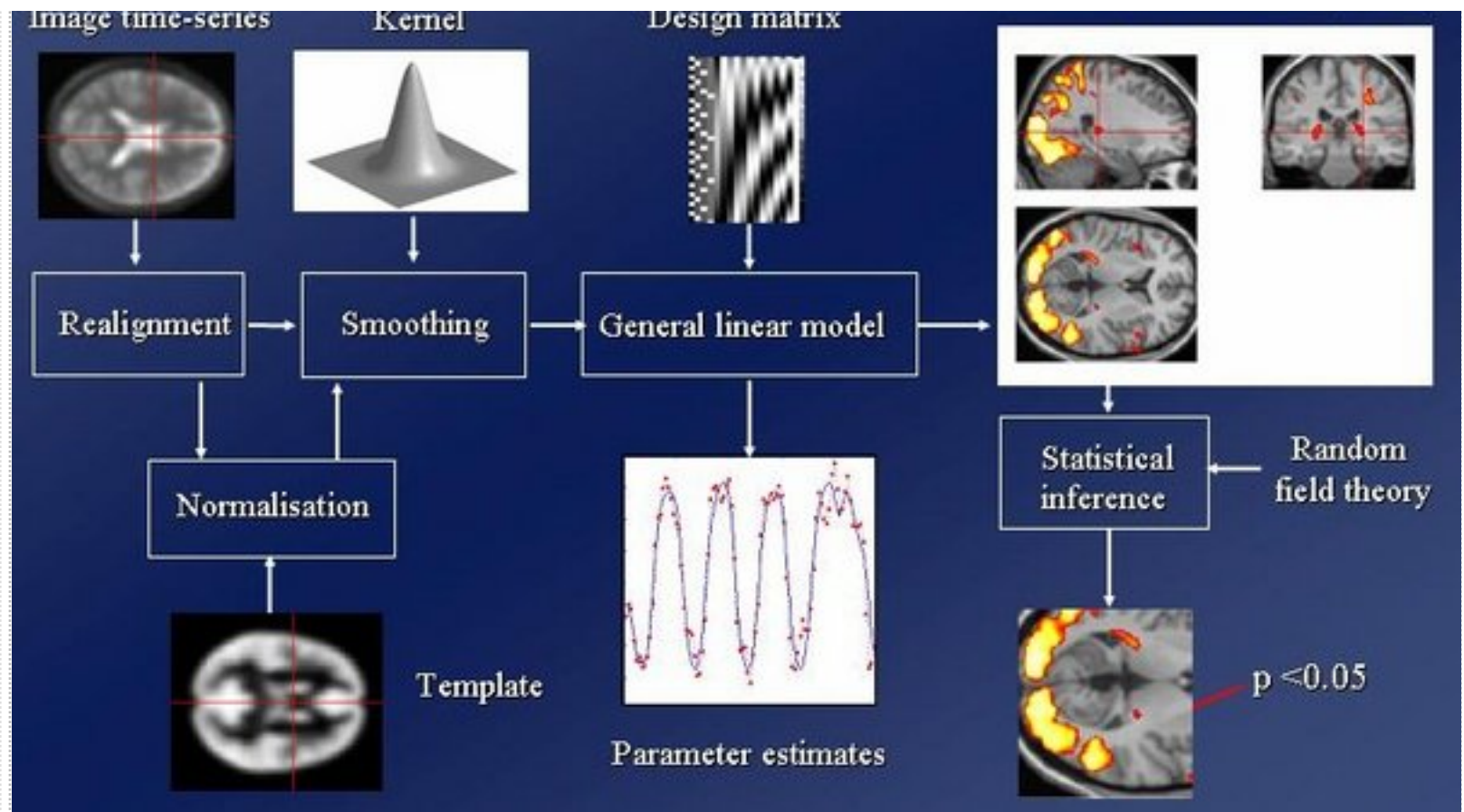


Figure 1: This schematic depicts the transformations that start with an imaging data sequence and end with a statistical parametric map (SPM). An SPM can be regarded as an 'X-ray' of the significance of regional effects. Voxel-based analyses require the data to be in the same anatomical space: this is effected by realigning the data. After realignment, the images are subject to non-linear warping so that they match a spatial model or template that already conforms to a standard anatomical space. After smoothing, the general linear model is employed to estimate the parameters of a temporal model (encoded by a design matrix) and derive the appropriate univariate test statistic at every voxel. The test statistics (usually  $t$  or  $F$ -statistics) constitute the SPM. The final stage is to make statistical inferences on the basis of the SPM and Random Field Theory and characterize the responses observed using the fitted responses or parameter estimates.

electroencephalography (EEG)). This is not the case; with very few exceptions, every analysis is a variant of the general

linear model. This includes: simple  $t$ -tests on scans assigned to one condition or another, correlation coefficients between observed responses and boxcar stimulus functions in fMRI, inferences made using multiple linear regression and evoked responses estimated using linear time-invariant models. Mathematically, they are all formally identical and can be implemented with the same equations and [algorithms](#). The only thing that distinguishes among them is the design matrix encoding the temporal model or experimental design.

The general linear model is an equation  $Y = X\beta + \epsilon$  that expresses the observed response variable in terms of a linear combination of explanatory variables  $X$  plus a well behaved error term (see Figure 1 and Friston et al. 1995). The general linear model is variously known as 'analysis of covariance' or 'multiple regression analysis' and subsumes simpler variants, like the ' $t$ -test' for a difference in means, to more elaborate linear convolution models such as finite impulse response (FIR) models. The matrix that contains the explanatory variables (e.g. designed effects or confounds) is called the *design matrix*. Each column of the design matrix corresponds to an effect one has built into the experiment or that may confound the results. These are referred to as explanatory variables, covariates or regressors. The example in Figure 1 relates to an fMRI study of visual stimulation, under four conditions. The effects on the response variable are modelled in terms of functions of the presence of these conditions (*i.e.* boxcars smoothed with a hemodynamic response function) and constitute the first four columns of the design matrix. There then follows a series of terms that are designed to remove or model low-frequency variations in signal due to artefacts such as aliased biorhythms and other drift terms. The final column is whole brain activity. The relative contribution of each of these columns is assessed using standard maximum likelihood and inferences about these contributions are made using  $t$  or  $F$ -statistics, depending upon whether one is looking at a particular linear combination (*e.g.*, a subtraction), or all of them together.

## Testing for contrasts

The GLM can be used to implement a vast range of statistical analyses. The issue is therefore not the mathematics but the formulation of a design matrix appropriate to the study design and inferences sought. The design matrix can contain both covariates and indicator variables. Each column has an associated unknown or free parameter. Some of these parameters will be of interest (*e.g.* the effect of a particular sensorimotor or cognitive condition or the regression coefficient of hemodynamic responses on reaction time). The remaining parameters will pertain to confounding effects (*e.g.* the effect of

being a particular subject or the regression slope of voxel activity on global activity). Inferences about the parameter estimates are made using their estimated variance. This allows one to test the null hypothesis that all the estimates are zero using the  $F$ -statistic or that some contrast or linear mixture (*e.g.* a subtraction) of the estimates is zero using an  $SPM\{t\}$ . An example of a contrast weight vector would be  $[-1 \ 1 \ 0 \ \dots]$  to compare the difference in responses evoked by two conditions, modelled by the first two regressors in the design matrix. Sometimes several contrasts of parameter estimates are jointly interesting. For example, when using polynomial (Büchel et al. 1996) or basis function expansions of some experimental factor. In these instances, the  $SPM\{F\}$  is used and is specified with a matrix of contrast weights that can be thought of as a collection of ‘ $t$ -contrasts’. An  $F$ -contrast may look like:

$$\begin{bmatrix} -1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \end{bmatrix}$$

This would test for the significance of the first or second parameter estimates. The fact that the first weight is negative has no effect on the test because the  $F$ -statistic is based on sums of squares.

In most analyses the design matrix contains indicator variables or parametric variables encoding the experimental manipulations. These are formally identical to classical analysis of covariance (*i.e.* ANCOVA) models. An important instance of the GLM, from the perspective of time-series data (*e.g.*, functional magnetic resonance imaging or fMRI), is the linear time-invariant model or convolution model.

## Topological inference and the theory of random fields

---

Classical inference

---

using SPMs can be of two sorts, depending on whether one knows where to look in advance. With an anatomically constrained hypothesis, about effects in a particular brain region, the uncorrected  $p$ -value associated with the height or extent of that region in the SPM can be used to test the hypothesis. With an anatomically open hypothesis (*i.e.* a null hypothesis that there is no effect anywhere in a specified

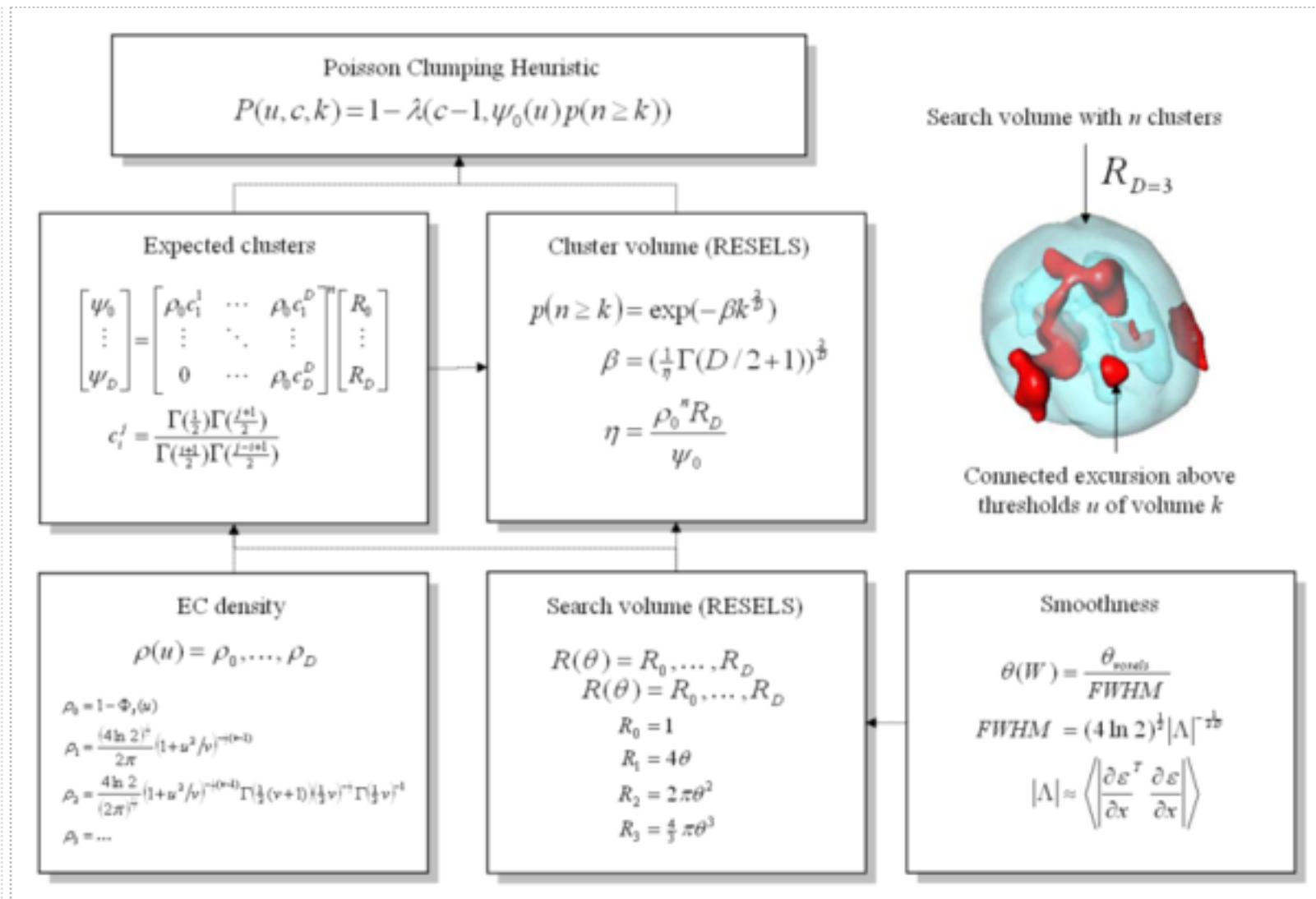


Figure 2: Schematic illustrating the use of Random Field Theory in making inferences about SPMs. If one knew precisely where to look, then inference can be based on the value of the statistic at the specified location in the SPM. However, generally, one does not have a precise anatomical *prior*, and an adjustment for multiple dependent comparisons has to be made to the  $p$ -values. These corrections use distributional approximations from RFT. This schematic deals with a general case of  $n$  SPM $\{t\}$  whose voxels all survive a common threshold  $u$  (*i.e.* a conjunction of  $n$  component SPMs). The central probability, upon which all peak, cluster or set-level inferences are made, is the probability  $P(u, c, k)$  of getting  $c$  or more clusters with  $k$  or more *RESELS* (resolution elements) above this threshold. By assuming that clusters behave like a

volume) a correction for multiple dependent comparisons is necessary. The theory of random fields provides a way of adjusting the  $p$ -value that takes into account the fact that

multidimensional Poisson point-process (*i.e.*, the Poisson clumping heuristic),  $P(u, c, k)$  is determined simply: the distribution of  $c$  is Poisson with an expectation that corresponds to the product of the expected number of clusters, of any size, and the probability that any cluster will be bigger than  $k$  RESELS. The latter probability depends on the expected number of RESELS per cluster  $\eta$ . This is simply the expected supra-threshold volume, divided by the expected number of clusters. The expected number of clusters  $\psi_0$  is estimated with the Euler characteristic (EC) (effectively the number of blobs minus the number of holes). This depends on the EC density for the statistic in question (with degrees of freedom  $\nu$ ) and the RESEL counts. The EC density is the expected EC per unit of  $D$ -dimensional volume of the SPM where the volume of the search is given by the RESEL counts. RESEL counts are a volume measure that has been normalized by the smoothness of the SPMs component error fields ( $\epsilon$ ), expressed in terms of the full width at half maximum ( $FWHM$ ). In this example equations for a sphere of radius  $\epsilon$  are given.  $\Psi$  denotes the cumulative density function for the statistic in question.

neighbouring voxels are not independent, by virtue of continuity in the original data. Provided the data are smooth the RFT adjustment is less severe (*i.e.* is more sensitive) than a Bonferroni correction for the number of voxels. As noted above RFT deals with the multiple comparisons problem in the context of continuous, statistical fields, in a way that is analogous to the Bonferroni procedure for families of discrete statistical tests. There are many ways to appreciate the difference between RFT and Bonferroni corrections. Perhaps the most intuitive is to consider the fundamental difference between an SPM and a collection of discrete  $t$ -values. When declaring a peak or cluster of the SPM to be significant, we refer collectively to all the voxels associated with that feature. The false positive rate is expressed in terms of peaks or clusters, under the null hypothesis of no activation. This is not the expected false positive rate of voxels. If the SPM is smooth, one false positive peak may be associated with hundreds of voxels. Bonferroni correction controls the expected number of false positive *voxels*, whereas RFT controls the expected number of false positive *peaks*. Because the number of peaks is always less than the number of voxels, RFT can use a lower threshold, rendering it much more sensitive. In fact, the number of false positive voxels is somewhat irrelevant because it is a function of smoothness. The RFT correction discounts voxel size by expressing the search volume in terms of smoothness or resolution elements (*RESELS*), see Figure 2. This intuitive perspective is expressed formally in terms of differential topology using the *Euler characteristic* (Worsley et al. 1992). At high thresholds the Euler characteristic corresponds to the number peaks above threshold.

There are only two assumptions underlying the use of the RFT:

- The error fields (but not necessarily the data) are a reasonable lattice approximation to an underlying random field with a multivariate Gaussian distribution,
- These fields are continuous, with an analytic autocorrelation function.

In practice, for neuroimaging data, the inference is appropriate if 1) the threshold chosen to define the blobs is high enough such that the expected Euler characteristic is close to the number of blobs, which for cluster size tests would be around a Z score of three 2) the lattice approximation is reasonable, which implies a smoothness about three times the voxel size on each space axis, 3) the errors of the specified statistical model are normally distributed, which implies that the model is not misspecified.

A common misconception is that the autocorrelation function has to be Gaussian. It does not. The only way RFT might not be valid is if at least one of the above assumptions does not hold.

## **Anatomically closed hypotheses**

When making inferences about regional effects (*e.g.* activations) in SPMs, one often has some idea about where the activation should be. In this instance a correction for the entire search volume is inappropriate. However, a problem remains in the sense that one would like to consider activations that are 'near' the predicted location, even if they are not exactly coincident. There are two approaches one can adopt: pre-specify a small search volume and make the appropriate RFT correction (Worsley et al. 1996) or use the uncorrected  $p$ -value based on spatial extent of the nearest cluster (Friston 1997). This probability is based on getting the observed number of voxels, or more, in a given cluster (conditional on that cluster existing). Both these procedures are based on distributional approximations from RFT.

## **Anatomically open hypotheses and levels of inference**

To make inferences about regionally specific effects the SPM is thresholded, using some height and spatial extent thresholds that are specified by the user. Corrected  $p$ -values can then be derived that pertain to various topological



features of the excursion set (*i.e.* subset of the SPM above threshold):

- Set-level inference: the number of activated regions (*i.e.*, number of connected subsets above some height and volume threshold),
- Cluster-level inference: the number of activated voxels (*i.e.*, volume) comprising a particular connected subset (*i.e.*, cluster),
- Peak-level inference: the height of maxima within that cluster.

These  $p$ -values are corrected for the multiple dependent comparisons and are based on the probability of obtaining  $c$ , or more, clusters with  $k$ , or more, voxels, above a threshold  $u$  in an SPM of known or estimated smoothness. This probability has a reasonably simple form (see Figure 2 for details).

Set-level refers to the inference that the number of clusters comprising an observed activation profile is highly unlikely to have occurred by chance and is a statement about the activation profile, as characterized by its constituent regions.

Cluster-level inferences are a special case of set-level inferences, that obtain when the number of clusters  $c = 1$ . Similarly peak-level inferences are special cases of cluster-level inferences that result when the cluster can be small (*i.e.*  $k = 0$ ). One usually observes that set-level inferences are more powerful than cluster-level inferences and that cluster-level inferences are generally more powerful than peak-level inferences. The price paid for this increased sensitivity is reduced localizing power. Peak-level tests permit individual maxima to be identified as significant features, whereas cluster and set-level inferences only allow clusters or sets of clusters to be identified. Typically, people use peak-level inferences and a spatial extent threshold of zero. This reflects the fact that characterizations of functional anatomy are generally more useful when specified with a high degree of anatomical precision.

## Recommended reading

---

- K.J. Friston, J.T. Ashburner, S.J. Kiebel, T.E. Nichols and W.D. Penny (2006). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier, London.

### Internal references

- Valentino Braitenberg (2007) [Brain](#). *Scholarpedia*, 2(11):2918.
- Paul L. Nunez and Ramesh Srinivasan (2007) [Electroencephalogram](#). *Scholarpedia*, 2(2):1348.
- William D. Penny and Karl J. Friston (2007) [Functional imaging](#). *Scholarpedia*, 2(5):1478.
- Seiji Ogawa and Yul-Wan Sung (2007) [Functional magnetic resonance imaging](#). *Scholarpedia*, 2(10):3105.

## References

---

- R.J. Adler (1981). *The geometry of random fields*. Wiley New York.
- C. Büchel, R.S.J. Wise, C.J. Mummary, J.-B. Poline and K.J. Friston (1996). Nonlinear regression in parametric activation studies. *NeuroImage*, 4:60-66.
- K.J. Friston, C.D. Frith, P.F. Liddle and R.S.J. Frackowiak (1991). Comparing functional (PET) images: the assessment of significant change. *Journal of Cerebral Blood Flow and Metabolism*, 11:690-699.
- K.J. Friston, K.J. Worsley, R.S.J. Frackowiak, J.C. Mazziotta and A.C. Evans (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*, 1:214-220.
- K.J. Friston, A.P. Holmes, K.J. Worsley, J.-B. Poline, C.D. Frith and R.S.J. Frackowiak (1995). Statistical Parametric Maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2:189-210.
- K.J. Friston (1997). Testing for anatomical specified regional effects. *Human Brain Mapping*, 5:133-136.
- K.J. Worsley, A.C. Evans, S. Marrett and P. Neelin (1992). A three-dimensional statistical analysis for rCBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism*, 12:900-918.
- K.J. Worsley, S. Marrett, P. Neelin, A.C. Vandal, K.J. Friston and A.C. Evans (1996). A unified statistical approach or determining significant signals in images of cerebral activation. *Human Brain Mapping*, 4:58-73.

## External links

---

- <http://www.fil.ion.ucl.ac.uk/spm/>

# See also

[Brain, Neuroimaging, Functional magnetic resonance imaging](#)

Sponsored by: [Eugene M. Izhikevich](#), Editor-in-Chief of Scholarpedia, the peer-reviewed open-access encyclopedia

Reviewed by: [Anonymous](#)

Accepted on: [2008-03-24 12:11:24 GMT](#)

Category: [Multiple Curators](#)



[Main page](#)   [Propose a new article](#)   [About](#)   [Random article](#)   [Help](#)   [F.A.Q.'s](#)   [Blog](#)

Focal areas

[Astrophysics](#)  
[Celestial mechanics](#)  
[Computational neuroscience](#)  
[Computational intelligence](#)  
[Dynamical systems](#)  
[Physics](#)

[Touch](#)

"*Statistical parametric mapping (SPM)*" by [Guillaume Flandin and Karl J. Friston](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License](#). Permissions beyond the scope of this license are described in the [Terms of Use](#)

Activity



[Privacy policy](#)   [About Scholarpedia](#)   [Disclaimers](#)

