CHAPTER

1

A short history of SPM

K. Friston

INTRODUCTION

For a young person entering imaging neuroscience it must seem that the field is very large and complicated, with numerous approaches to experimental design and analysis. This impression is probably compounded by the abundance of TLAs (three-letter-acronyms) and obscure terminology. In fact, most of the principles behind design and analysis are quite simple and had to be established in a relatively short period of time at the inception of brain mapping. This chapter presents an anecdotal perspective on this period. It serves to explain why some ideas, like *t*-maps or, more technically, statistical parametric maps, were introduced and why other issues, like global normalization, were crucial, even if they are not so important nowadays.

The history of human brain mapping is probably shorter than many people might think. Activation studies depend on imaging changes in brain state within the same scanning session. This was made possible using short-half-life radiotracers and positron emission tomography (PET). These techniques became available in the eighties (e.g. Herscovitch *et al*., 1983) and the first activation maps appeared soon after (e.g. Lauter *et al*., 1985; Fox *et al*., 1986). Up until this time, regional differences among brain scans had been characterized using hand-drawn regions of interest (ROI), reducing hundreds of thousands of voxels to a handful of ROI measurements, with a somewhat imprecise anatomical validity. The idea of making voxel-specific statistical inferences, through the use of statistical parametric maps, emerged in response to the clear need to make inferences about brain responses without knowing where those responses were going to be expressed. The first *t*-map was used to establish functional specialization for colour processing in 1989 (Lueck *et al*., 1989). The underlying methodology was described in a paper entitled: 'The relationship between global and local changes in PET scans' (Friston *et al*., 1990). This may seem an odd title to introduce statistical parametric mapping (SPM) but it belies a key motivation behind the approach.

Statistical maps versus regions of interest

Until that time, images were usually analysed with analysis of variance (ANOVA) using ROI averages. This approach had become established in the analysis of autoradiographic data in basic neuroscience and metabolic scans in human subjects. Critically, each region was treated as a level of a factor. This meant that the regional specificity of a particular treatment was encoded in the region by treatment interaction. In other words, a main effect of treatment *per se* was not sufficient to infer a regionally specific response. This is because some treatments induced a global effect that was expressed in all the ROIs. Global effects were, therefore, one of the first major conceptual issues in the development of SPM. The approach taken was to treat global activity as a confound in a separate analysis of covariance (ANCOVA) *at each voxel*, thereby endowing inference with a regional specificity that could not be explained by global changes. The resulting SPMs were like X-rays of region-specific changes and, like X-rays, are still reported in maximum-intensity projection format (known colloquially as glass-brains). The issue of regional versus global changes and the validity of global estimators were debated for several years, with many publications in the specialist literature. Interestingly, it is a theme that enjoyed a reprise with the advent of functional magnetic resonance imaging (fMRI) (e.g. Aguirre *et al*., 1998) and still attracts some research interest today.

Adopting a voxel-wise ANCOVA model paved the way for a divergence between the mass-univariate approach used by SPM (i.e. a statistic for each voxel) and multivariate models used previously. A subtle but

Statistical Parametric Mapping, by Karl Friston et al.
ISBN–13: 978-0-12-372560-8 ISBN–10: 0-12-372560-7

important motivation for mass-univariate approaches was the fact that a measured haemodynamic response in one part of the brain may differ from the response in another, *even if the underlying neuronal activation was exactly the same*. This meant that the convention of using region-by-condition interactions as a test for regionally specific effects was not tenable. In other words, even if one showed that two regions activated differently in terms of measured haemodynamics, this did not mean there was a regionally specific difference at the neuronal or computational level. This issue seems to have escaped the electroencephalography (EEG) community, who still use ANOVA with region as a factor, despite the fact that the link between neuronal responses and channel measurements is even more indeterminate than for metabolic imaging. However, the move to voxel-wise, whole-brain analysis entailed two special problems: the problem of registering images from different subjects so that they could be compared on a voxel-by-voxel basis and the multiple-comparisons problem that ensued.

Spatial normalization

The pioneering work of the St Louis group had already established the notion of a common anatomical or stereotactic space (Fox *et al*., 1988) in which to place subtraction or difference maps, using skull X-rays as a reference. The issue was how to get images into that space efficiently. Initially, we tried identifying landmarks in the functional data themselves to drive the registration (Friston *et al*., 1989). This approach was dropped almost immediately because it relied on landmark identification and was not a hundred per cent reproducible. Within a year, a more reliable, if less accurate, solution was devised that matched images to a template without the need for landmarks (Friston *et al*., 1991a). The techniques for spatial normalization using template- or model-based approaches have developed consistently since that time and current treatments regard normalization as the inversion of generative models for anatomical variation that involve warping templates to produce subject-specific images (e.g. Ashburner and Friston, 2005).

Topological inference

Clearly, performing a statistical test at each voxel engendered an enormous false positive rate when using unadjusted thresholds to declare activations significant. The problem was further compounded by the fact that the data were not spatially independent and a simple Bonferroni correction was inappropriate (PET and SPECT (single photon emission computerized tomography) data are inherently very smooth and fMRI had not been invented at this stage). This was the second major theme that occupied people trying to characterize functional neuroimaging data. What was needed was a way of predicting the probabilistic behaviour of SPMs, under the null hypothesis of no activation, which accounted for the smoothness or spatial correlations among voxels. From practical experience, it was obvious that controlling the false positive rate of voxels was not the answer. One could increase the number of positive voxels by simply making the voxels smaller but without changing the topology of the SPM. It became evident that conventional control procedures developed for controlling family-wise error (e.g. the Bonferroni correction) had no role in making inferences on continuous images. What was needed was a new framework in which one could control the false positive rate of the regional effects themselves, noting a regional effect is a topological feature, not a voxel.

The search for a framework for topological inference in neuroimaging started in the theory of stochastic processes and level-crossings (Friston *et al*., 1991b). It quickly transpired that the resulting heuristics were the same as established results from the theory of random fields. Random fields are stochastic processes that conform very nicely to realizations of brain scans under normal situations. Within months, the technology to correct *p*-values was defined within random field theory (Worsley *et al*., 1992). Although the basic principles of topological inference were established at this time, there were further exciting mathematical developments with extensions to different sorts of SPMs and the ability to adjust the *p*-values for small bounded volumes of interest (see Worsley *et al*., 1996). Robert Adler, one of the world's contemporary experts in random field theory, who had abandoned it years before, was understandably very pleased and is currently writing a book with a protégé of Keith Worsley (Adler and Taylor, in preparation).

Statistical parametric mapping

The name 'statistical parametric mapping' was chosen carefully for a number of reasons. First, it acknowledged the TLA of 'significance probability mapping', developed for EEG. Significance probability mapping involved creating interpolated pseudo-maps of *p*-values to disclose the spatiotemporal organization of evoked electrical responses (Duffy *et al*., 1981). The second reason was more colloquial. In PET, many images are derived from the raw data reflecting a number of different physiological parameters (e.g. oxygen metabolism, oxygen extraction fraction, regional cerebral blood flow etc.). These were referred to as parametric maps. All parametric maps are non-linear functions of the original data. The

distinctive thing about *statistical* parametric maps is that they have a known distribution under the null hypothesis. This is because they are predicated on a statistical model of the data (as opposed to a physiological parametric model).

One important controversy, about the statistical models employed, was whether the random fluctuations or error variance was the same from brain region to brain region. We maintained that it was not (on common sense grounds that the frontal operculum and ventricles were not going to show the same fluctuations in blood flow) and adhered to voxel-specific estimates of error. For PET, the Montreal group considered that the differences in variability could be discounted. This allowed them to pool their error variance estimator over voxels to give very sensitive SPMs (under the assumption of stationary error variance). Because the error variance was assumed to be the same everywhere, the resulting *t*-maps were simply scaled subtraction or difference maps (see Fox *et al*., 1988). This issue has not dogged fMRI, where it is generally accepted that error variance is voxel-specific.

The third motivation for the 'statistical parametric mapping' was that it reminded people they were using parametric statistics that assume the errors are additive and Gaussian. This is in contradistinction to non-parametric approaches that are generally less sensitive, more computationally intensive, but do not make any assumptions about the distribution of error terms. Although there are some important applications of non-parametric approaches, they are generally a specialist application in the imaging community. This is largely because brain imaging data conform almost exactly to parametric assumptions by the nature of image reconstruction, post-processing and experimental design.

THE PET YEARS

In the first few years of the nineties, many landmark papers were published using PET and the agenda for a functional neuroimaging programme was established. SPM proved to be the most popular way of characterizing brain activation data. It was encoded in Matlab and used extensively by the MRC Cyclotron Unit at the Hammersmith Hospital in the UK and was then distributed to collaborators and other interested units around the world. The first people outside the Hammersmith group to use SPM were researchers at NIH (National Institutes of Health, UDA) (e.g. Grady *et al*., 1994). Within a couple of years, SPM had become the community standard for analysing PET activation studies and the assumptions behind SPM were largely taken for granted. By this stage, SPM was synonymous with the general linear model and random field theory. Although originally framed in terms of ANCOVA, it was quickly realized that any general linear model could be used to produce an SPM. This spawned a simple taxonomy of experimental designs and their associated statistical models. These were summarized in terms of subtraction or categorical designs, parametric designs and factorial designs (Friston *et al*., 1995a). The adoption of factorial designs was one of the most important advances at this point. The first factorial designs focused on adaptation during motor learning and studies looking at the interaction between a psychological and pharmacological challenge in psychopharmacological studies (e.g. Friston *et al*., 1992). The ability to look at the effect of changes in the level of one factor on activations induced by another led to a rethink of cognitive subtraction and pure insertion and the appreciation of context-sensitive activations in the brain. The latitude afforded by factorial designs is reflected in the fact that most studies are now multifactorial in nature.

THE fMRI YEARS

In 1992, at the annual meeting of the Society of Cerebral Blood Flow and Metabolism in Miami, Florida, Jack Belliveau presented, in the first presentation of the opening session, provisional results using photic stimulation with fMRI. This was quite a shock to the imaging community that was just starting to relax: most of the problems had been resolved, community standards had been established and the way forward seemed clear. It was immediately apparent that this new technology was going to reshape brain mapping radically, the community was going to enlarge and established researchers were going to have to re-skill. The benefits of fMRI were clear, in terms of the ability to take many hundreds of scans within one scanning session and to repeat these sessions indefinitely in the same subject. Some people say that the main advances in a field, following a technological breakthrough, are made within the first few years. Imaging neuroscience must be fairly unique in the biological sciences, in that exactly five years after the inception of PET activation studies, fMRI arrived. The advent of fMRI brought with it a new wave of innovation and enthusiasm.

From the point of view of SPM, there were two problems, one easy and one hard. The first problem was how to model evoked haemodynamic responses in fMRI timeseries. This was an easy problem to resolve because SPM could use any general linear model, including convolution models of the way haemodynamic responses were caused (Friston *et al*., 1994). Stimulus functions encoding the occurrence of a particular event or experimental

state (e.g. boxcar-functions) were simply convolved with a haemodynamic response function (HRF) to form regressors in a general linear model (*cf* multiple linear regression).

Serial correlations

The second problem that SPM had to contend with was the fact that successive scans in fMRI time-series were not independent. In PET, each observation was statistically independent of its precedent but, in fMRI coloured time-series, noise rendered this assumption invalid. The existence of temporal correlations originally met with some scepticism, but is now established as an important aspect of fMRI time-series. The SPM community tried a series of heuristic solutions until it arrived at the solution presented in Worsley and Friston (1995). This procedure, also known as 'pre-colouring', replaced the unknown endogenous autocorrelation by imposing a known autocorrelation structure. Inference was based on the Satterthwaite conjecture and is formally identical to the non-specificity correction developed by Geisser and Greenhouse in conventional parametric statistics. An alternative approach was 'pre-whitening' which tried to estimate a filter matrix from the data to de-correlate the errors (Bullmore *et al*., 2001). The issue of serial correlations, and more generally non-sphericity, is still important and attracts much research interest, particularly in the context of maximum likelihood techniques and empirical Bayes (Friston *et al*., 2002).

New problems and old problems

The fMRI community adopted many of the developments from the early days of PET. Among these were the use of the standard anatomical space provided by the atlas of Talairach and Tournoux (1988) and conceptual issues relating to experimental design and interpretation. Many debates that had dogged early PET research were resolved rapidly in fMRI; for example, 'What constitutes a baseline?' This question, which had preoccupied the whole community at the start of PET, appeared to be a non-issue in fMRI with the use of well-controlled experimental paradigms. Other issues, such as global normalization were briefly revisited, given the different nature of global effects in fMRI (multiplicative) relative to PET (additive). However, one issue remained largely ignored by the fMRI community. This was the issue of adjusting *p*-values for the multiplicity of tests performed. While people using SPM quite happily adjusted their *p*-values using random field theory, others seemed unaware of the need to control false positive rates. The literature now entertained reports based on uncorrected *p*-values, an issue which still confounds editorial decisions today. It is interesting to contrast this, historically, with the appearance of the first PET studies.

When people first started reporting PET experiments there was an enormous concern about the rigor and validity of the inferences that were being made. Much of this concern came from outside the imaging community who, understandably, wanted to be convinced that the 'blobs' that they saw in papers (usually *Nature* or *Science*) reflected true activations as opposed to noise. The culture at that time was hostile to capricious reporting and there was a clear message from the broader scientific community that the issue of false positives had to be resolved. This was a primary motivation for developing the machinery to adjust *p*-values to protect against family-wise false positives. In a sense, SPM was a reaction to the clear mandate set by the larger community, to develop a valid and rigorous framework for activation studies. In short, SPM was developed in a culture of scepticism about brain mapping that was most easily articulated by critiquing its validity. This meant that the emphasis was on specificity and reproducibility, as opposed to sensitivity and flexibility. Current standards for reporting brain mapping studies are much more forgiving than they were at its beginning, which may explain why recent developments have focused on sensitivity (e.g. Genovese *et al*., 2002).

The convolution model

In the mid-nineties, there was lots of fMRI research; some of it was novel, some recapitulating earlier findings with PET. From a methodological point of view, notable advances included the development of event-related paradigms that furnished an escape from the constraints imposed by block designs and the use of retinotopic mapping to establish the organization of cortical areas in human visual cortex. This inspired a whole sub-field of cortical surface mapping that is an important endeavour in early sensory neuroimaging. For SPM there were three challenges that needed to be addressed:

Temporal basis functions

The first involved a refinement of the models of evoked responses. The convolution model had become a cornerstone for fMRI with SPM. The only remaining issue was the form of the convolution kernel or haemodynamic response function that should be adopted and whether the form changed from region to region. This was resolved simply by convolving the stimulus function with not one response function but several [basis

functions]. This meant that one could model condition, voxel and subject-specific haemodynamic responses using established approaches. Temporal basis functions (Friston *et al*., 1995b) were important because they allowed one to define a family of HRFs that could change their form from voxel to voxel. Temporal basis functions found an important application in the analysis of event-related fMRI. The general acceptance of the convolution model was consolidated by the influential paper of Boynton a year later (Boynton *et al*., 1996). However, at this time, people were starting to notice some non-linearities in fMRI responses (Vazquez and Noll, 1998) that were formulated, in the context of SPM, as a Volterra series expansion of the stimulus function (Friston *et al*., 1998). This was simple because the Volterra series can be formulated as another linear model (compare with a Taylor expansion). These Volterra characterizations would later be used to link empirical data and balloon models of haemodynamic responses.

Efficiency and experimental design

The second issue that concerned the developers of SPM arose from the growing number and design of event-related fMRI studies. This was the efficiency with which responses could be detected and estimated. Using an analytical formulation, it was simple to show that the boxcar paradigms were much more efficient that event-related paradigms, but event-related paradigms could be made efficient by randomizing the occurrence of particular events such that they 'bunched' together to increase experimental variance. This was an interesting time in the development of data analysis techniques because it enforced a signal processing perspective on the general linear models employed.

Hierarchical models

The third area motivating the development of SPM was especially important in fMRI and reflects the fact that many scans can be obtained in many individuals. Unlike in PET, the within-subject scan-to-scan variability can be very different from the between-subject variability. This difference in variability has meant that inferences about responses in a single subject (using within-subject variability) are distinct from inferences about the population from which that subject was drawn (using between-subject variability). More formally, this distinction is between fixed- and random-effects analyses. This distinction speaks to hierarchical observation models for fMRI data. Because SPM only had the machinery to do single-level (fixed-effects) analyses, a device was required to implement random-effects analyses. This turned out to be relatively easy and intuitive: subject-specific effects were estimated in a first-level analysis and the contrasts of parameter estimates (e.g. activations) were then re-entered into a second-level SPM analysis (Holmes and Friston, 1998). This recursive use of a single-level statistical model is fortuitously equivalent to multilevel hierarchical analyses (compare with the summary statistic approach in conventional statistics).

Bayesian developments

Understanding hierarchical models of fMRI data was important for another reason: these models support empirical Bayesian treatments. Empirical Bayes was one important component of a paradigm shift in SPM from classical inference to a Bayesian perspective. From the late nineties, Bayesian inversion of anatomical models had been a central part of spatial normalization. However, despite early attempts (Holmes and Ford, 1993), the appropriate priors for functional data remained elusive. Hierarchical models provided the answer, in the form of empirical priors that could be evaluated from the data themselves. This evaluation depends on the conditional dependence implicit in hierarchical models and brought previous maximum likelihood schemes into the more general Bayesian framework. In short, the classical schemes SPM had been using were all special cases of hierarchical Bayes (in the same way that the original ANCOVA models for PET were special cases of the general linear models for fMRI). In some instances, this connection was very revealing, for example, the equivalence between classical covariance component estimation using restricted maximum likelihood (i.e. ReML) and the inversion of two-level models with expectation maximization (EM) meant we could use the same techniques used to estimate serial correlations to estimate empirical priors on activations (Friston *et al*., 2002).

The shift to a Bayesian perspective had a number of motivations. The most principled was an appreciation that estimation and inference corresponded to Bayesian inversion of generative models of imaging data. This placed an emphasis on generative or forward models for fMRI that underpinned work on biophysical modelling of haemodynamic responses and, indeed, the framework entailed by dynamic causal modelling (e.g. Friston *et al*., 2003; Penny *et al*., 2004). This reformulation led to more informed spatiotemporal models for fMRI (e.g. Penny *et al*., 2005) that effectively estimate the optimum smoothing by embedding spatial dependencies in a hierarchical model. It is probably no coincidence that these developments coincided with the arrival of the Gatsby Computational Neuroscience Unit next to the Wellcome Department of Imaging Neuroscience. The Gatsby housed several experts in Bayesian inversion and

machine learning and the Wellcome was home to many of the SPM co-authors.

The second motivation for Bayesian treatments of imaging data was to bring the analysis of EEG and fMRI data into the same forum. Source reconstruction in EEG and MEG (magnetoencephalography) is an ill-posed problem that depends explicitly on regularization or priors on the solution. The notion of forward models in EEG-MEG, and their Bayesian inversion had been well established for decades and SPM needed to place fMRI on the same footing.

THE MEG-EEG YEARS

At the turn of the century people had started applying SPM to source reconstructed EEG data (e.g. Bosch-Bayard *et al*., 2001). Although SPM is not used widely for the analysis of EEG-MEG data, over the past five years most of the development work in SPM has focused on this modality. The motivation was to accommodate different modalities (e.g. fMRI-EEG) within the same analytic and anatomical framework. This reflected the growing appreciation that fMRI and EEG could offer complementary constraints on the inversion of generative models. At a deeper level, the focus had shifted from generative models of a particular modality (e.g. convolution models for fMRI) and towards models of neuronal dynamics that could explain any modality. The inversion of these models corresponds to true multimodal fusion and is the aim of recent and current developments within SPM.

In concrete terms, this period saw the application of random field theory to SPMs of evoked and induced responses, highlighting the fact that SPMs can be applied to non-anatomical spaces, such as space-peristimulus-time or time-frequency (e.g. Kilner *et al*., 2005). It has seen the application of hierarchical Bayes to the source reconstruction problem, rendering previous heuristics, like L-curve analysis, redundant (e.g. Phillips *et al*., 2002) and it has seen the extension of dynamic causal modelling to cover evoked responses in EEG-MEG (David *et al*., 2006).

This section is necessarily short because the history of SPM stops here. Despite this, a lot of the material in this book is devoted to biophysical models of neuronal responses that can, in principle, explain any modality. Much of SPM is about the inversion of these models. In what follows, we try to explain the meaning of the more important TLAs entailed by SPM (Table 1-1).

TABLE 1-1 Some common TLAs

TLA Three letter acronym
SPM Statistical parametric map(ping)
GLM General linear model
RFT Random field theory
VBM Voxel-based morphometry
FWE Family-wise error
FDR False discovery rate
IID Independent and identically distributed
MRI Magnetic resonance imaging
PET Positron emission tomography
EEG Electroencephalography
MEG Magnetoencephalography
HRF Haemodynamic response function
IRF Impulse response function
FIR Finite impulse response
ERP Event-related potential
ERF Event-related field
MMN Mis-match negativity
PPI Psychophysiological interaction
DCM Dynamic causal model
SEM Structural equation model
SSM State-space model
MAR Multivariate autoregression
LTI Linear time invariant
PEB Parametric empirical Bayes
DEM Dynamic expectation maximization
GEM Generalized expectation maximization
BEM Boundary-element method
FEM Finite-element method

REFERENCES

Adler RJ, Taylor JE Random fields and geometry. In preparation - To be published by Birkhauser.

Aguirre GK, Zarahn E, D'Esposito M (1998) The inferential impact of global signal covariates in functional neuroimaging analyses. *NeuroImage* **8**: 302–06

Ashburner J, Friston KJ (2005) Unified segmentation. *NeuroImage* **26**: 839–51

Bosch-Bayard J, Valdes-Sosa P, Virues-Alba T *et al*. (2001) 3D statistical parametric mapping of EEG source spectra by means of variable resolution electromagnetic tomography (VARETA). *Clin Electroencephalogr* **32**: 47–61

Boynton GM, Engel SA, Glover GH *et al*. (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* **16**: 4207–21

Bullmore ET, Long C, Suckling J *et al*. (2001) Colored noise and computational inference in neurophysiological (fMRI) time series analysis: resampling methods in time and wavelet domains. *Hum Brain Mapp* **12**: 61–78

David O, Kiebel SJ, Harrison LM *et al*. (2006). Dynamic causal modeling of evoked responses in EEG and MEG. *NeuroImage*, **30**(4): 1255–7

Duffy FH, Bartels PH, Burchfiel JL (1981) Significance probability mapping: an aid in the topographic analysis of brain electrical activity. *Electroencephalogr Clin Neurophysiol* **51**: 455–62

Fox PT, Mintun MA, Raichle ME *et al*. (1986) Mapping human visual cortex with positron emission tomography. *Nature* **323**: 806–9

Fox PT, Mintun MA, Reiman EM *et al*. (1988) Enhanced detection of focal brain responses using intersubject averaging and change-distribution analysis of subtracted PET images. *J Cereb Blood Flow Metab* **8**: 642–53

Friston KJ, Passingham RE, Nutt JG *et al*. (1989) Localisation in PET images: direct fitting of the intercommissural (AC-PC) line. *J Cereb Blood Flow Metab* **9**: 690–705

Friston KJ, Frith CD, Liddle PF *et al*. (1990) The relationship between global and local changes in PET scans. *J Cereb Blood Flow Metab* **10**: 458–66

Friston KJ, Frith CD, Liddle PF *et al*. (1991a) Plastic transformation of PET images. *J Comput Assist Tomogr* **15**: 634–39

Friston KJ, Frith CD, Liddle PF *et al*. (1991b) Comparing functional (PET) images: the assessment of significant change. *J Cereb Blood Flow Metab* **11**: 690–09

Friston KJ, Frith C, Passingham RE *et al*. (1992) Motor practice and neurophysiological adaptation in the cerebellum: a positron tomography study. *Proc R Soc Lond Series B* **248**: 223–28

Friston KJ, Jezzard PJ, Turner R (1994) Analysis of functional MRI time-series. *Hum Brain Mapp* **1**: 153–71

Friston KJ, Holmes AP, Worsley KJ *et al*. (1995a) Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* **2**: 189–210

Friston KJ, Frith CD, Turner R *et al*. (1995b) Characterizing evoked hemodynamics with fMRI. *NeuroImage* **2**: 157–65

Friston KJ, Josephs O, Rees G *et al*. (1998) Nonlinear event-related responses in fMRI. *Magnet Reson Med* **39**: 41–52

Friston KJ, Glaser DE, Henson RN *et al*. (2002) Classical and Bayesian inference in neuroimaging: applications. *NeuroImage* **16**: 484–512

Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. *NeuroImage* **19**: 1273–302

Genovese CR, Lazar NA, Nichols T (2002) Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage* **15**: 870–78

Grady CL, Maisog JM, Horwitz B *et al*. (1994) Age-related changes in cortical blood flow activation during visual processing of faces and location. *J Neurosci* **14**:1450–62

Herscovitch P, Markham J, Raichle ME (1983) Brain blood flow measured with intravenous H2(15)O. I. Theory and error analysis. *J Nucl Med* **24**: 782–89

Holmes A, Ford I (1993) A Bayesian approach to significance testing for statistic images from PET. In *Quantification of brain function, tracer kinetics and image analysis in brain PET*, Uemura K, Lassen NA, Jones T *et al*. (eds). *Excerpta Medica, Int. Cong.* Series no. **1030**: 521–34

Holmes AP, Friston KJ (1998) Generalisability, random effects and population inference. *NeuroImage* **7**: S754

Kilner JM, Kiebel SJ, Friston KJ (2005) Applications of random field theory to electrophysiology. *Neurosci Lett* **374**: 174–78

Lauter JL, Herscovitch P, Formby C *et al*. (1985) Tonotopic organization in human auditory cortex revealed by positron emission tomography. *Hear Res* **20**: 199–205

Lueck CJ, Zeki S, Friston KJ *et al*. (1989) The colour centre in the cerebral cortex of man. *Nature* **340**: 386–89

Penny WD, Stephan KE, Mechelli A *et al*. (2004) Comparing dynamic causal models. *NeuroImage* **22**: 1157–72

Penny WD, Trujillo-Barreto NJ, Friston KJ (2005) Bayesian fMRI time series analysis with spatial priors. *NeuroImage* **24**: 350–62

Phillips C, Rugg MD, Friston KJ (2002) Systematic regularization of linear inverse solutions of the EEG source localization problem. *NeuroImage* **17**: 287–301

Talairach P, Tournoux J (1988) *A stereotactic coplanar atlas of the human brain*. Thieme, Stuttgart

Vazquez AL, Noll DC (1998) Nonlinear aspects of the BOLD response in functional MRI. *NeuroImage* **7**: 108–18

Worsley KJ, Evans AC, Marrett S *et al*. (1992) A three-dimensional statistical analysis for CBF activation studies in human brain. *J Cereb Blood Flow Metab* **12**: 900–18

Worsley KJ, Friston KJ (1995) Analysis of fMRI time-series revisited – again. *NeuroImage* **2**: 173–81

Worsley KJ, Marrett S, Neelin P *et al*. (1996) A unified statistical approach of determining significant signals in images of cerebral activation. *Hum Brain Mapp* **4**: 58–73