Modeling inter-subject variability in fMRI activation location: A Bayesian hierarchical spatial model

Lei Xu[1,*], **Timothy D. Johnson**[2], **Thomas E. Nichols**[2,3,4], and **Derek E. Nee**[5]

[1] Department of Biostatistics, Vanderbilt University, Nashville, TN 37232

[2] Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109

[5] Department of Pyschology, University of Michigan, Ann Arbor, MI 48109

[3] GlaxoSmithKline Clinical Imaging Centre, Imperial College London, Hammersmith Hospital, W12 0NN, United Kingdom

[4] FMRIB Centre, Oxford University, Oxford, OX3 9DU, United Kingdom

Summary

The aim of this work is to develop a spatial model for multi-subject fMRI data. There has been extensive work on univariate modeling of each voxel for single and multi-subject data, some work on spatial modeling of single-subject data, and some recent work on spatial modeling of multi-subject data. However, there has been no work on spatial models that explicitly account for inter-subject variability in activation locations. In this work, we use the idea of activation centers and model the inter-subject variability in activation locations directly. Our model is specified in a Bayesian hierarchical frame work which allows us to draw inferences at all levels: the population level, the individual level and the voxel level. We use Gaussian mixtures for the probability that an individual has a particular activation. This helps answer an important question which is not addressed by any of the previous methods: What proportion of subjects had a significant activity in a given region. Our approach incorporates the unknown number of mixture components into the model as a parameter whose posterior distribution is estimated by reversible jump Markov Chain Monte Carlo. We demonstrate our method with a fMRI study of resolving proactive interference and show dramatically better precision of localization with our method relative to the standard mass-univariate method. Although we are motivated by fMRI data, this model could easily be modified to handle other types of imaging data.

Keywords

Bayesian hierarchical model; Functional brain mapping; Multi-subject fMRI data analysis; Reversible jump Markov chain Monte Carlo; Spatial mixture model

1 Introduction

Among methods for mapping brain function, functional magnetic resonance imaging (fMRI) is widely used. Conventionally a classical, mass-univariate approach is used where univariate time-series models are fit independently at each voxel. These models are used to

---

create images of parameter estimates and test statistics, which are then assessed for significance (Friston et al., 1995). Inference on mass-univariate models is made by assessing the statistic images, where the null hypothesis of no effect is tested at each voxel. Statistically significant voxels define regions of activation. Various methods on how to choose the threshold have been proposed (Genovese, Lazar, and Nichols, 2002; Nichols and Hayasaka, 2003). This computationally efficient method has several limitations. The approach does not model the spatial properties of the signal explicitly and can not incorporate any prior knowledge. It cannot account for mismatch in activation location and will only detect voxels with consistent change in activation. Multi-subject analyses are particularly problematic since, even after registration of the subjects' brain to a common atlas, there is residual variation in the anatomical landmarks; further, it has been shown that even if sulci and gyri are aligned there is variation in the functional landmarks (Morosan et al., 2001).

A lot of effort has been made on how to use the spatial information in the data to enhance signal detection. Some methods find a threshold that accounts for correlation in the null hypothesis statistic images (Worsley et al., 1996; Friston et al., 1993). The most general approach is to spatially smooth the data with a Gaussian kernel. However this tends to blur and change the shape of active regions. Alternatively, Descombes, Kruggel, and von Cramon (1998) proposed restoring the signal using a spatiotemporal Markov Random Field (MRF). The MRF is used to define prior knowledge of the signal. Polzehl and Spokoiny (2001) proposed a structural adaptive smoothing procedure, specifically the propagation-separation (PS) approach for time-series of images. Tabelow et al. (2006) provided a complete procedure for fMRI analysis using the PS approach and showed significant improvement on the information of the spatial extent and the shape of the activation region.

Many others have proposed Bayesian methods to integrate spatial modeling into the statistical analysis. Hartvig and Jensen (2000) used a spatial mixture model and achieve computational feasibility by formulating the model through the marginal distribution on a small grid of voxels. Hartvig (2002) proposed a regression based spatial model using the idea of "activation centers". He used a reversible jump algorithm to insert, delete and change an activation center given known variance parameters. A similar idea was later proposed by Lukic et al. (2007) using general Kernel based methods. Both methods implicitly assume that voxels are independent. Cosman, Fisher, and Wells (2004) used an Ising MRF as the prior for neural activity and Woolrich et al. (2005) developed a spatial mixture model using a discrete Markov random field (MRF) prior. Penny, Trujillo-Barreto, and Friston (2005) proposed a Gaussian MRF prior on the regression coefficient of a general linear model and approximated the posterior with the Variational Bayes method, while Penny, Flandin, and Trujillo-Barreto (2007) showed how model evidence can be approximated using a Bayesian framework and allows Analysis of Variance and Cluster of Interest analysis. Flandin and Penny (2007) replaced the Gaussian MRF prior with sparse spatial basis function prior. While these methods model spatial dependence in the signal, they are only for single subject analyses. Recently, some work on multi-subject analysis has been done. Beckman and Smith (2005) developed a time-subject-space three-way independent component analysis method. They treat space as a generic dimension and ignore spatial contiguity. Bowman et al. (2008) developed a Bayesian model for fMRI data that accounts for task-related connectivity. Briefly, they reduce the size of the problem by dividing the brain into a small number of regions and model the within and between regional correlation through subject and region specific random effects. While their model has the advantage of modeling long range correlations they use only the stage I BOLD response estimates at each voxel as the data. They also assume exchangeability for voxels within each region. In contrast, we use BOLD estimates and voxel locations and directly model activation locations.

We develop a Bayesian hierarchical model that improves on the standard methods. First, we use an explicit spatial model for activations without pre-smoothing. Second, ours is a multi-subject model that accounts for inter-subject heterogeneity in activation location about a population location. Third, we use Gaussian mixtures for the probability that individual activation locations belong to a population activation center and allows us to answer: "What proportion of subjects have a activity in a given region?" Our model can be used to make inference on activation patterns at all levels: the population level, the individual level and the voxel level. Specifically, we are interested in the population location of activations, population prevalence of activation at a location, and inter-subject variability in the location of activation. Throughout, we take the individual t-statistic images as the observed data, Data from three subjects from our motivating example can be found in the supplementary materials Web Appendix A. The paper is organized as follows. In Section 2, we introduce the spatial mixture model in a fully Bayesian framework. In Section 3, we describe the algorithm used for posterior inference and discuss how we summarize posterior inference in Section 4. In Section 5, we demonstrate our method using a real data set. We conclude the manuscript with a discussion and ideas for future work.

2 Bayesian Hierarchical Model

We begin with an overview of the model and notation (with all subscripts initially suppressed), after which we present the distribution of the data and priors in detail.

Our model is specified hierarchically, as illustrated in Figure 1. At the first level there are an unknown number of "population centers", μ, that follow, a priori, a homogenous spatial Poisson process defined over the brain. At the second level, an unknown number of "individual centers", ξ, are distributed as Gaussian mixtures whose component means are μ with covariance matrix Σ. At the third level, an unknown number of "individual component" means, η, are distributed as Gaussian mixtures whose component means are ξ with covariance matrix Φ. While the "individual centers" (level 2) are intended to model the location of a given subject's activation with a single point, heterogeneity in activation shape (i.e. non-Gaussian-blob-shapes) require the use of another level in the hierarchy to fit intra-subject activation shape. The "individual components" (level 3) are the locations of Gaussian mixtures used to fit an individual's activation about a given individual center. We assume that each subject is fit with an intra-subject fMRI signal model, producing scalar t-images of the fMRI blood oxygenation level dependent (BOLD) signal; we refer to these intra-subject summary measures as "the data". At the fourth level, we assume the data, **y**, for each subject are distributed as a Gaussian mixture with an unknown number of components whose means are θ with variances σ^2. The mixing weights for the datum at voxel v are proportional to the density at v of a multivariate normal distribution with mean η and covariance matrix R. There is one special component representing the constant background intensity for each subject. We now present the details of our model.

Level 4: data distribution

Let y_{jv} denote the t-statistic at voxel $v = 1, \ldots, V$, for subject $j = 1, \ldots, J$, where a single index v is used to represent the 3D voxel: $\mathbf{x}_v = (x_{1v}, x_{2v}, x_{3v})$. We assume y_{jv} has a mixture distribution:

$$f(y_{jv}|p_{jvl}, \theta_{jl}, \sigma_{jl}^2, c_j) = p_{jv0}\varphi(y_{jv};\theta_0, \sigma_0^2) + \sum_{l=1}^{c_j} p_{jvl}\varphi(y_{jv};\theta_{jl}, \sigma_{jl}^2)$$

(1)

where c_j is the number of mixture components for subject j, excluding the background component, $\varphi(a; b, c^2)$ is the density at a of a normal distribution with mean b and variance c^2, while θ_0 and σ_0^2 represent the background mean and variance. Conditional on the latent allocation variable, ω_{jv}, with $\Pr(\omega_{jv} = l) = p_{jvl}$, $l = 0, \ldots, c_j$, the likelihood is

$$f(y_{jv}|\theta_0, \sigma_0^2, \omega_{jv}=0)=\varphi(y_{jv};\theta_0, \sigma_0^2) \text{ for background,}$$
$$f(y_{jv}|\theta_{jl}, \sigma_{jl}^2, \omega_{jv}=l)=\varphi(y_{jv};\theta_{jl}, \sigma_{jl}^2) \text{ otherwise.}$$

(2)

Let $\varphi_2(\cdot; \mathbf{a}, \mathbf{B})$ denote the multivariate normal density with mean $\mathbf{a}$ and covariance matrix $\mathbf{B}$. The mixing weight $p_{jv0} \propto m$ and $p_{jvl} \propto \varphi_2(\mathbf{x}_v; \boldsymbol{\eta}_{jl}, \mathbf{R}_{jl})$ for $l > 0$ with $\sum_{l=0}^{c_j}p_{jvl}=1$. Here $\boldsymbol{\eta}_{jl}^{\mathrm{T}}=(\eta_{1jl}, \eta_{2jl}, \eta_{3jl})$ is the component l mean—spatial dependence is captured by these weights. Given m and $\mathbf{R}_{jl}$, the weights depend on the distance from the voxel to each component mean. Thus, neighbors tend to be more correlated than distant observations. (However, deviations from this assumption in functional neuroimaging data may exist; see Bowman et al. (2008). While our model does not account for such distant correlations, we do not anticipate they will impair the model fit or posterior estimation.) Note that an observation distant from all components centers, a priori, will belong to background with probability approximately 1. Furthermore, if a voxel and component mean are coincident and no other components nearby, then the a priori probability that this voxel belongs to the background is approximately $m/(|\mathbf{R}_{jl}^{-1}|^{1/2}/(2\pi)^{3/2}+m)$. A priori, we want $|E(\mathbf{R}_{jl}^{-1})|^{1/2}=(2\pi)^{3/2}$ and achieve this by taking $\mathbf{R}_{jl} \sim IW_{10}[(2\π/10) \mathbf{I}]$, where $IW_d(X)$ is the inverse Wishart distribution with d degrees of freedom and X is a symmetric positive definite matrix. Thus, a priori the probability that this voxel belongs to the background is approximately $m/(m + 1)$. Changing the value of m, therefore, changes the probability of belonging to the background. The number of mixing components, c_j, for each subject is not known and is estimated. A priori we assume c_j is Poisson with mean λ_c.

We assume the background mean, θ_0, follows a standard normal distribution, a priori, and reflects our belief that when no signal is present, the BOLD signal should have zero mean. However, by placing a distribution on θ_0, we allow for non-zero null contrasts—a quite common occurrence in fMRI data. For the intensity means of the active components, θ_{jl}, $l > 0$, we use a truncated normal distribution: $\theta_{jl} \sim \text{trunc}_{(0,\infty)}N(\lambda_\theta, \sigma_\theta^2)$. Hyperprior distributions are then placed on its mean and variance: $\lambda_\theta \sim N(3, 10^8)$ and $\sigma_\theta^2 \sim IG(.01, .01)$. We truncate the prior distribution of θ_{jl} at 0 to ensure that the voxels classified to activated regions have positive expectations: activated regions are defined as a positive change in the BOLD signal.

We place inverse gamma priors on σ_0^2 and σ_{jl}^2 to take advantage of conjugacy: $\sigma_0^2 \sim IG(.001, .001)$ and $\sigma_{jl}^2 \sim IG(2, \beta_\sigma)$. For each subject, most of the data will belong to the background and thus a vague prior can be placed on σ_0^2. For the variances of mixture components, σ_{jl}^2, the prior has finite first moment and infinite second moment. Furthermore, we place a vague hyperprior gamma distribution on the scale parameter, $\beta_\sigma \sim G(.001, .001)$, thereby reducing sensitivity of the posterior of σ_{jl}^2 on its prior. (We note here that hyperprior values that we feel result in vague prior distributions and hyperprior values that we feel are not application dependent will be stated immediately. Others, which we feel are are application dependent, will be discussed later in Section 5).

Level 3: individual component means

Individual component mean priors constitute level three. The prior of mean l for subject j, is a mixture of b_j multivariate normals.

$$\pi(\eta_{jl}|\phi_{jh}, \boldsymbol{\xi}_{jh}, \boldsymbol{\Phi}_{jh}, b_j) = \sum_{h=1}^{b_j} \phi_{jh}\varphi_2(\eta_{jl};\boldsymbol{\xi}_{jh}, \boldsymbol{\Phi}_{jh}).$$

(3)

We introduce a latent variable, s_{jl}, such that $\Pr(s_{jl} = h) = \phi_{jh}$, $h = 1, \ldots, b_j$ and, conditional on s_{jl}, each component mean is associated with a particular individual center: $[\boldsymbol{\eta}_{jl}|\boldsymbol{\xi}_{jh}, \boldsymbol{\Phi}_{jh}, s_{jl} = h] \sim N(\boldsymbol{\xi}_{jl}, \boldsymbol{\Phi}_{jl})$. Here, the ϕ_{jh} are mixing weights. A natural choice for the prior on ϕ_{jh} is a symmetric $b_j - 1$ dimensional Dirichlet distribution: $\phi_{jh}|b_j \sim D(1, 1, \ldots, 1)$. The number of individual centers, b_j, for each subject is not known and we assume, a priori, that b_j is Poisson with mean λ_b.

We give $\boldsymbol{\Phi}_{jh}$ an inverse Wishart with 10 degrees of freedom: $\boldsymbol{\Phi}_{jh} \sim IW_{10}(S_{\Phi}^{-1})$. A hyperprior distribution is then placed on S_{Φ}: $S_{\Phi} \sim IW_{10}(\boldsymbol{T}_{\Phi})$. These distributional forms are chosen to take advantage of conjugacy and to aid in reducing the dependence of the posterior of $\boldsymbol{\Phi}_{jh}$ on its prior. The degrees of freedom for both inverse Wishart distributions are set to 10, ensuring proper priors that are necessary for the reversible jump moves.

Level 2: individual centers

It typically takes several mixture components to adequately model an activation region. Thus, we introduce the second level in the hierarchy to model the centers of these activation regions. We assume, $\boldsymbol{\xi}_{jh}$, $h = 1, \ldots, b_j$ for subject j, is a mixture of c_p multivariate normals.

$$\pi(\boldsymbol{\xi}_{jh}|\psi_i, \mu_i, \sum_i, c_P) = \sum_{i=1}^{c_p} \psi_i\varphi_2(\boldsymbol{\xi}_{jh};\mu_i, \sum_i).$$

(4)

We introduce a latent variable, z_{jh}, such that $\Pr(z_{jh} = i) = \psi_i$, $i = 1, \ldots, c_p$. The prior on the mixing weights ψ_i is a symmetric $c_p - 1$ dimensional Dirichlet distribution. We give Σ_i an inverse Wishart distribution with 10 degrees of freedom: $\sum_i \sim IW_{10}(S_{\Sigma}^{-1})$. A hyperprior distribution is then placed on S_{Σ}: $S_{\Sigma} \sim IW_{10}(\boldsymbol{T}_{\Sigma})$. These distributional forms are chosen to take advantage of conjugacy and to aid in reducing the dependence of the posterior of $\boldsymbol{\Phi}_{jh}$ on its prior. The degrees of freedom for the inverse Wishart distributions are 10, ensuring proper priors distributions that are necessary for the reversible jump moves.

Thus far, we have only modeled the activation regions for each subject, independently, along with the individual centers. On average, it takes one individual center to model each activation region. However, it may take more, especially for oddly shaped activation regions. It is also possible that two or more spatially close activation regions may be considered as one individual center. Thus, at this level, there is some inherent local smoothing at the subject level.

Level 1: population centers

The final level of the hierarchy clusters the individual centers about population centers. The parameter $\mu_i^{\mathrm{T}} = (\mu_{1i}, \mu_{2i}, \mu_{3i})$ is the location of population activation center i, $i = 1, \ldots, c_p$. Let

A_j denote the volume of the brain of subject j. Set $A = \cup_{j=1}^{J} A_j$. (We note here that, although all subject's data have been mapped onto a common brain atlas, due to motion artifacts and field inhomogeneities, there are missing data. Typically, fMRI analyses are performed on the intersection of the A_j. By taking the union, we allow for the possibility that a population center is in a region where some subjects may have missing data. Note, however, that the spatial parameters at the higher levels of the hierarchy are confined to A_j, for subject j.) A priori, we assume that the population centers follow a homogenous spatial Poisson process with rate λ_p defined on A. Thus, $c_p \sim P(\lambda_p A)$ and $\pi(\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{c_p}/c_p) = A^{-c_p}$.

Although the prior distribution of the population centers is a homogenous point process, its posterior is not necessarily. It depends largely on the posterior distribution of the individual centers. From Equation (4) it is possible that two or more mixture components may be close enough in space that the mixture of these may result in a single mode in the posterior distribution. Thus, although we call c_p the number of population centers and $\boldsymbol{\mu}_i$ a population center location, the number of modes in the posterior of ξ_{jh} can different from the mode of c_p.

3 Posterior estimation

The full posterior distribution does not have an analytic solution. Thus, the posterior distribution is simulated via Markov chain Monte Carlo (MCMC) techniques. Reversible jump MCMC (RJMCMC) is used to estimate the number of individual components, individual centers and population centers. RJMCMC was developed by Green (1995) and can be viewed as a generalization of the Metropolis-Hastings (MH) algorithm (Hastings (1970)).

We propose to add a population center, an individual center or an individual component, each with probability 0.5 and propose to delete a population center, an individual center or individual component, each with probability 0.5 at every iteration of the algorithm. We over-sample the RJMCMC moves five times per iteration which results in better mixing of these parameters. When we propose a birth, the mixing component parameters are drawn from their prior distributions. New mixing weight ψ_* (φ_*) are drawn from Beta(1, c_p) (Beta(1, b_j)). We re-scale the old weights $\psi' = \psi(1 - \psi_*)$ such that all weights sum up to 1 (similarly for φ_*). The deletion move is the inverse of this construction.

Conditional on the number of population centers, the number of individual centers and the number of individual components, other parameters are updated via a Gibbs step or a random walk MH step. The variances in the proposal distribution for the MH steps are dynamically calibrated to obtain acceptance rates of approximately 35%. Following Fernandez and Green (2002) we use the marginal expression for the likelihood and for the priors of $\boldsymbol{\eta}_{jl}$ and ξ_{jh}, as specified in equations (1), (3) and (4) to obtain better mixing of the trans-dimensional moves. Details of the algorithm are given in Web Appendix D.

4 Summarizing posterior inference

Simultaneous visualization of the joint posterior distribution of all parameters is infeasible. Instead we view the distributions of certain univariate parameters and create images summarizing the posteriors of the various spatial parameters. In this section we review the approaches we use to understand our posterior and assess model fit.

We create a "Posterior Probability of Activation" image for each subject. This image is the voxel-wise (marginal) posterior probability of activation: $\Pr(\omega_{jv} > 0 | \mathbf{y})$.

We create an image of the "Individual Center Posterior": a voxelization (a discrete estimate of a continuous distribution) of the estimated individual center posterior mean: $\hat{E}(\tilde{\xi}|\mathbf{y})$, where $\tilde{\xi}$ is a center from a randomly chosen subject from the population. At each sweep we

evaluate $\sum_{i=1}^{c_p} \psi_i \varphi_2(\cdot;\mu_i, \sum_i)$ at each voxel v. Averaging this over sweeps creates an estimate of the posterior mean. The Individual Center Posterior shows the most likely location(s) of an activation center for a randomly chosen subject from the population. From it, we can see the inter-subject variability of individual centers about population centers. We take care not to over-interpret the relative intensities shown in this image. As can be seen in Equation (4), the mode heights in the Individual Center Posterior are affected by two quantities. First, all things equal, smaller diagonal elements in Σ_i indicates less inter-subject variability about a population center, resulting in a larger mode. Second, a mode's height will be relatively larger when more subjects have centers associated with a particular population center: the mixing weights will tend to be larger for population components that have more individual centers associated with them.

Similar to the "Individual Center Posterior" image, we create an image of the "Individual Center Scale". This is an image of the conditional (on population centers occurring in certain volumes) estimated posterior 95% credible ellipse (marginalized over z).

We also create an image to characterize the population centers. The "Population Center Location" image is a voxelization of the posterior rate function (counts per voxel) for $\boldsymbol{\mu}$. We estimate this by computing the 3D histogram of $\{\boldsymbol{\mu}_i\}$ for each iteration and then average this over iterations (see, e.g., Johnson, 2007, Equation (11)). The values represented in this image sum to the estimated posterior expectation of c_p. The modes seen in the Individual Center Posterior image roughly correspond to the most likely location of population centers. However one cannot ascertain the variability of population centers from the Individual Center Posterior image. These two images taken together show the most likely locations of the population centers and the most likely locations of an individual center from a randomly chosen subject from the population. They also show the relative variability of a population center and of the individual centers about each population center.

Quantitatively, we estimate the posterior probability that a population center exists in certain small volumes. To do so we adopt an idea from Johnson, 2007, Equation (14): find local maxima in the individual center posterior image and, for small volumes centered on these maxima, find the posterior probability that at least one μ_i exists in this volume. This probability approximates the probability that a population center exists in this volume.

The "Population Center Prevalence" image shows the fraction of subjects that are associated with particular population centers. Let $c_{ij} = \Sigma_h I(z_{jh} = i)$ denote the count of individual centers that subject j has associated with population center i. The count of subjects with

center i is then $c_i = \sum_{j=1}^{J} I(c_{ij} > 0)$, and the population prevalence is c_i/J. Conditional on a population center occurring in voxel v, this image is the posterior average of c_i/J after marginalizing over population centers.

For comparison with other methods we create two other images. The "Classical t-image" is the one-sample t-test on the BOLD effect magnitudes ($\mathbf{y}_v = \{y_{jv}\}_j$) at each voxel. We also create an image of $-\log_{10}$P-values for a one-sided, one sample t-test. The $-\log_{10}$ transformation makes for easier visualization, creating an image that is "bright" in voxels with evidence for an non-null effect magnitude.

5 An Application: Proactive Interference-Resolution

We demonstrate our model with data from a study investigating, among other things, proactive interference-resolution (Nee, Jonides, and Berman, 2007). Proactive interference is our difficulty to recall desired information due to interference from previously learned information. Our ability to resolve proactive interference from previously relevant, but no longer relevant information is a key factor in determining how much information we can retain in short-term memory. Thus, the neural mechanisms underlying proactive interference-resolution is central in our understanding of short-term memory.

An experimental paradigm commonly used to investigate proactive interference-resolution is the Recent Probes task. In this paradigm, subjects are given a small set of items (the target set) to remember over a short retention interval followed by a recognition probe. The recognition probe is either a subset (positive probe) of the target set or not (negative probe). The recognition probe can further be a member of a previous trial (recent probes) or not (non-recent probes). Subjects show slower reaction times and an increase in error rates when rejecting recent negative probes when compared to non-recent negative probes. This decrease in performance is taken as a marker of proactive interference. The left lateral prefrontal cortex (left LPFC) is a region in the brain that has been linked to the resolution of proactive interference.

5.1 Hyperprior parameter values

The joint prior distribution is factored hierarchically (see Web Appendix C). It is not possible to use fully non-informative priors in a mixture setting and have a proper posterior (Richardson and Green, 1997, Section 2.4). Nevertheless, we assess the sensitivity of the posterior. Results of these sensitivity analyses are given in Web Appendix B.

Level 4 hyperprior parameter values—Typically, fMRI experiments are designed such that only a small percentage of the brain (roughly 1%–5%) shows a significant increase in BOLD signal. Thus, a priori, we set $m = 19$. Given a voxel and individual component mean are coincident with no other components nearby, this choice of m results in the a priori probability of 0.95 that this voxel belongs to the background. All voxels that are distant from all components have an a priori probability near 1.0 of belonging to the background. This parameter controls the intensity level at which voxels are considered activated. We demonstrate this in our sensitivity analyses. With $m = 19$, the estimated posterior mean of λ_θ is 2.68. The mean of the number of mixing components is set to 25. The posterior is not sensitive to the value of λ_c as m has a much larger effect on the posterior distribution of c_j.

Level 3 hyperprior parameter values—We also expect that the number of individual activation centers will be small and so we set $\lambda_b = 5$.

Each individual activation region is typically fitted with several components via Equation (1). The means of these normal components are taken to be the individual activation region centers. The covariance matrices of these mixture components will control the size of the normal mixture components in Equation (3). For this application, we believe that the subjects' activation regions should be small. Thus, a priori, we desire $E(\Phi_{jh}) = I$. This translates to 0.2 cm standard deviation in the x, y, and z directions. It also results in an a priori 95% spherical confidence region with radius 0.557 cm (a volume of 0.724 cm^3, about the size of a Garbanzo bean). To achieve this a priori expected value for Φ_{jh}, we set $T_\Phi = (5/3) I$.

Level 2 hyperprior parameter values—Interest lies mainly in the left LPFC which is spatially well located. We expect most subjects will have individual centers in the left LPFC

that are reasonable close to each other. Hence, $T_\Sigma = 5/(3 \cdot 2.5^2) I$. This results in $E(\Sigma_i) = 2.5^2 I$ and an a priori 95% spherical confidence region with radius 1.392 cm (a volume of 11.31 cm^3, about the volume of a standard sized walnut).

Level 1 hyperprior parameter values—We expect that the number of population centers will be small. Thus we set $\lambda_p = 5/A$ (the a priori expected number of population centers in the section of the brain is 5).

5.2 Results

The data we use consists of unsmoothed fMRI images from eighteen subjects. Each image has voxel dimension $79 \times 95 \times 69$ and each voxel is 2mm on a side. The left LPFC is of particular interest and so we apply our model to a $79 \times 95 \times 7$ subset of the original data that contains the left LPFC. We run the algorithm for 12,000 iterations, discarding the first 2,000 iterations for burn-in. The acceptance rate for the population level birth/death RJMCMC is about 15%.

Web Figures 1–12 display all seven slices of the data from three randomly chosen subjects, along with the $-\log_{10}(p)$ images and the Posterior Probability of Activation images. In each of these figures, slices are arranged in columns. The top row displays the image intensity, or data. The second row shows the $-\log_{10}(p\text{-value})$ images. The third row in these figures show the posterior probability of activation of the corresponding slices. We can see from these figures that these three subjects have fairly high activation in the left LPFC. There is also evidence of high activation in various other regions of the brain. In particular, subject 6 shows some activation in the right LPFC in slices 43 and 44 (Web Figures 7 and 8).

The main focus of our work is at the population level. We conducted a classical group analysis on these data. We found no statistically significant activated voxels controlling for either a family-wise error or controlling the false discovery rate at 0.05. (Bonferroni corrected threshold: 7.8141; minimum FDR p-value: .2390). Although the aforementioned three subjects show a strong signal in the left LPFC, the signals vary in spatial location between subjects (c.f. Web Figures 1–12). This variation is great enough that the population based analyses did not result in statistical significance.

However, based on our model and algorithm there is strong evidence that there is a population center located in the left LPFC. Figures 2 through 4 show the middle three slices (40, 41 and 42) of the seven slice subset (slices 38–44) we analyze. Web Figures 13–19 show population level images for all seven slices in color, heat-map images. The posterior probability of a population center existing in the the $2.2 \times 2.2 \times 1.4$ cm neighborhood ($11 \times 11 \times 7$ voxels, volume = 6.776 cm^3) centered at voxel (17, 66, 4) is 0.9989 (see Panel A in Figures 2–3). The population prevalence in the left LPFC is about 0.55 (see Figures 2C, 3C and 4C). Furthermore, in the right LPFC the posterior probability of population center existing in the $2.2 \times 3.8 \times 1.4$ cm neighborhood ($11 \times 19 \times 7$ voxels, volume = 11.008 cm^3) centered at voxel (63, 64, 4) is 0.9033 suggesting that a homologous region in the right LPFC may also be involved in proactive interference-resolution. Such bilateral recruitment is often seen in tasks requiring cognitive control. However, we see that the inter-subject variability around the population right LPFC center is greater than that in the left LPFC (! see Panels B and D, Figures 2–4), consistent with the more common finding of left LPFC involvement in this task (Jonides and Nee, 2006). There is also evidence of a population center in the precuneus of the posterior parietal cortex. The posterior probability of a population center existing in the $3.0 \times 3.0 \times 1.4$ cm neighborhood ($15 \times 15 \times 7$ voxels, volume = 12.6 cm^3) centered at voxel (34, 26, 4) is 0.9366. This region is involved in the recollection of specific memory details such as contextual information (Wagner et al., 2005), which may be required to properly categorize and reject recent negative probes. Finally, we

note that there is also evidence of a population center in the anterior cingulate cortex (ACC) (centered at voxel (38, 66, 4)). The ACC is involved in a wide range of cognitive control tasks and is theorized to act as a conflict monitor, calling for increased control from the LPFC in situations of high conflict (Botvinick et al., 2001). The posterior probability of a population center existing in the $2.2 \times 2.2 \times 1.4$ cm neighborhood ($11 \times 11 \times !7$, volume = 6.776 cm^3) about this voxel is 0.9481.

Figures 2–4, in particular Panels A, B and D, demonstrate the ability of our algorithm to separate out inter-subject variability (individual center scale and individual center posterior images, Panels B, D) from the population center variability (population center location images, Panel A). This is most evident in Panel B, the individual center scale image. In the population center location image, there appears to be four major areas with evidence for a population center. We estimated the harmonic means of Σ_i conditional on a population center existing in the neighborhoods defined in the previous paragraph and evaluated the 95% posterior credible ellipse. From these ellipses and the posterior rate function, one can ascertain that the inter-subject variability is much larger than the variability in the population center locations. In slices 38, 39, 43 and 44, there is still evidence of inter-subject variability (c.f. Web Figures 13–19), but the population centers are confined mainly to the middle three slices, and for the most part, to slice 41. Figures 2–4(E,F) display the classical t-image and the $-\log_{10}(p)$ image. While these images similarly illustrate the spread of activation between subjects, they cannot quantify the spatial precision of the results as our method does, nor can they separate out the various sources of spatial variability.

5.3 Sensitivity analysis and simulation study

We ran both sensitivity analyses and a simulation study. The sensitivity analyses were carried out on the proactive interference-resolution data set. We assess the sensitivity to prior and hyperprior parameter values that may have an influence on the population level posterior distributions. We focus on the following parameters: λ_p and λ_b, the Poisson rate for the distribution of the number of population and individual centers; T_Σ, the inverse Wishart parameter for Σ; and m, the parameter controls the probability that a voxel belongs to the background. We found that changing λ_p, λ_b or m has little effect on the population center locations. Changing T_Σ which changes the expected value of Σ has a more pronounced effect. Increasing the variability causes a forth mode (between the left LPFC and ACC) to appear in the posterior population rate function. Decreasing the variability will increase the number of ! mixture components and make the modes other than the left LPFC less visible. Details of both the sensitivity analyses and simulation study can be found in Web Appendix B. (c_p) the mode of the estimated marginal posterior distribution of c_1 becomes 7 (29%); the mode of the estimated marginal posterior of c_p remains 8, but is more diffuse (53%). This influence of the Poisson mean on the posterior distribution of the number of components is expected, see, e.g., Green (1995).

Equal spread in the x and y directions. Σ_3 gives a population center with negatively correlated ($\rho = -0.5$) individual centers with equal spread in the x and y direction. The mean intensities of these three population level centers are 2, 2.5 and 3.

6 Future work and discussion

We have described a Bayesian mixture model for fMRI data analysis. The method considers the spatial structure of the signal, which is often ignored in frequentist approaches. Moreover this method extends the current Bayesian spatial modeling literature in two key ways. First, we consider multi-subject data, explicitly population centers and the dispersion of individual's response about those centers. Second, instead of assuming a normal shape model for activation magnitude, we assume a normal shape model for probability of

activation and assume homogeneous magnitude within component. We argue that this leads to a more flexible yet still interpretable parameterization.

The seven slice analysis takes approximately 8 hours of CPU time. The full 3D analysis takes approximately 21 hours of CPU time on a MAC 3.0 GHz Xserve Server. However, we do feel that there is more room for improving the efficiency.

The flexibility of our model does make for some interpretive limitations. While most users of fMRI conceive of activation "loci" (x,y,z locations), in our model the population centers are random variables $\boldsymbol{\mu}$ whose distribution can't easily be summarized. We visualize this distribution with the Population Center Location image, inspecting for modes and assessing the spread about modes.

A final limitation of our model is that we do not have explicit linking of population centers to individual centers. A fMRI practitioner would ideally like to point to population loci χ and ask "which subjects have evidence for that loci, and what is the pattern of activation in each subject corresponding to that loci." For any one MCMC sweep this connection is known through latent variables z_{jl} and ω_{jv}, but cannot be easily summarized over different sweeps as the number and location of population centers, as well as individual centers, may change.

In future work we would like to address these shortcomings, perhaps by introducing new latent variables that specifically correspond to investigators' notions of activation loci. For example, we could define a loci as "nearest center $\boldsymbol{\mu}_i$ to anatomical landmark X". Alternatively, local maxima on the classical t-statistic image could be used to label population centers; once population centers are identifiable, they can be used to track summary measures of interest. Our approach can be extended to model the negative BOLD response. This can be achieved by adding population and individual centers for negative BOLD responses.. Now each voxel will be assigned to the background, the positive center or the negative center. The mean intensity for negative BOLD responses will be restricted to negative values.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Attias, H. Advances in Neural Information Processing Systems. Cambridge, MA: MIT Press; 2000. A variational Bayesian framework for graphical models.

Beckman CF, Smith SM. Tensorial extensions of independent component analysis for multisubject fmri analysis. Neuroimage. 2005; 25:294–311. [PubMed: 15734364]

Botvinick MM, et al. Conflict monitoring and cognitive control. Psychological Review. 2001; 108:624–652. [PubMed: 11488380]

Bowman FD, et al. A bayesian hierarchical framework for spatial modeling of fmri data. Neuroimage. 2008; 39:146–156. [PubMed: 17936016]

Cosman ER, Fisher JW, Wells WM. Exact map activity detection in fmri using a glm with an ising spatial prior. Proc MICCAI'04. 2004; 2:703–710.

Descombes X, Kruggel F, von Cramon D. Spatio-temporal fmri analysis using markov random fields. IEEE transactions on medical imaging. 1998; 17:1028–1039. [PubMed: 10048860]

Fernandez C, Green P. Modelling spatially correlated data via mixtures:a bayesian approach. Journal of the Royal Statistical Society B. 2002; 64:805–826.

Flandin G, Penny WD. Bayesian fmri data analysis with sparse spatial basis function priors. NeuroImage. 2007; 34:1108–1125. [PubMed: 17157034]

Friston KJ, et al. Assessing the significance of focal activations using their spatial extent. Human Brain Mapping. 1993; 1:210–220.

Friston KJ, et al. Analysis of fmri time-series revisited. NeuroImage. 1995; 2:45–53. [PubMed: 9343589]

Genovese C, Lazar N, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. NeuroImage. 2002; 15:870–878. [PubMed: 11906227]

Green PJ. Reversible jump markov chain monte carlo computation and bayesian model determination. Biometrika. 1995; 82:711–732.

Hartvig NV. A stochastic geometry model for functional magnetic resonance images. Board of the Foundation of the Scandinavian Journal of Statistics. 2002; 29:333–353.

Hartvig NV, Jensen JL. Spatial mixture mixture modelling of fmri data. Human Brain Mapping. 2000; 11:233–248. [PubMed: 11144753]

Hastings WK. Monte carlo sampling methods using markov chains and their applications. Biometrika. 1970; 57:97–109.

Johnson TD. Analysis of pulsatile hormone concentration profiles with non-constant basal concentration: a Bayesian approach. Biometrics. 2007; 63:1207–1217. [PubMed: 18078483]

Jonides J, Nee DE. Brain mechanisms of porctive interference in working memory. Neuroscience. 2006; 139:181–193. [PubMed: 16337090]

Lukic AS, et al. Bayesian kernel methods for analysis of functional neuroimages. IEEE Transactions on Medical Imaging. 2007; 26:1613–1624. [PubMed: 18092732]

Miglioretti DL, McCulloch C, Zeger SL. Combining images across multiple subjects: a study of direct cortical electrical interference. Journal of the American Statistical Association. 2002; 97:125–135.

Morosan P, et al. Human primary auditory cortex: Cytoarchitectonic subdivisions and mapping into a spatial reference system. NeuroImage. 2001; 13:684–701. [PubMed: 11305897]

Nee DE, Jonides J, Berman MG. Neural mechanisms of proactive interference-resolution. NeuroImage. 2007; 38:740–751. [PubMed: 17904389]

Nichols T, Hayasaka. Controlling the familywise error rate in functional neuroimaging: A comparative review. Statistical Methods in Medical Research. 2003; 12:419–446. [PubMed: 14599004]

Penny WD, Flandin G, Trujillo-Barreto NJ. Bayesian comparison of spatially regularized general linear models. Human Brain Mapping. 2007; 28:275–293. [PubMed: 17133400]

Penny WD, Trujillo-Barreto NJ, Friston KJ. Bayesian fmri time series analysis with spatial priors. NeuroImage. 2005; 24:350–362. [PubMed: 15627578]

Polzehl J, Spokoiny V. Functional and dynamic magnetic resonance imaging using vector adaptive weights smoothing. Applied Statistics. 2001; 50:485–501.

Richardson S, Green PJ. On Bayseian analysis of mixtures with an unknown number of components (with discussion). J R Statist Soc B. 1997; 59:731–792.

Tabelow K, et al. Analyzing fmri experiments with structural adaptive smoothing procedures. NeuroImage. 2006; 33:55–62. [PubMed: 16891126]

Wagner AD, et al. Contributions to episodic memory retrieval. Trends in Cognitive Science. 2005; 9:445–453.

Woolrich MW, et al. Mixture models with adaptive spatial regularisation for segmentation with an application to fmri data. IEEE transactions on medical imaging. 2005; 24:1–11. [PubMed: 15638182]

Worsley K, et al. A unified statistical approach for determining significant signals in images of cerebral activation. Human Brain Mapping. 1996; 4:58–73. [PubMed: 20408186]
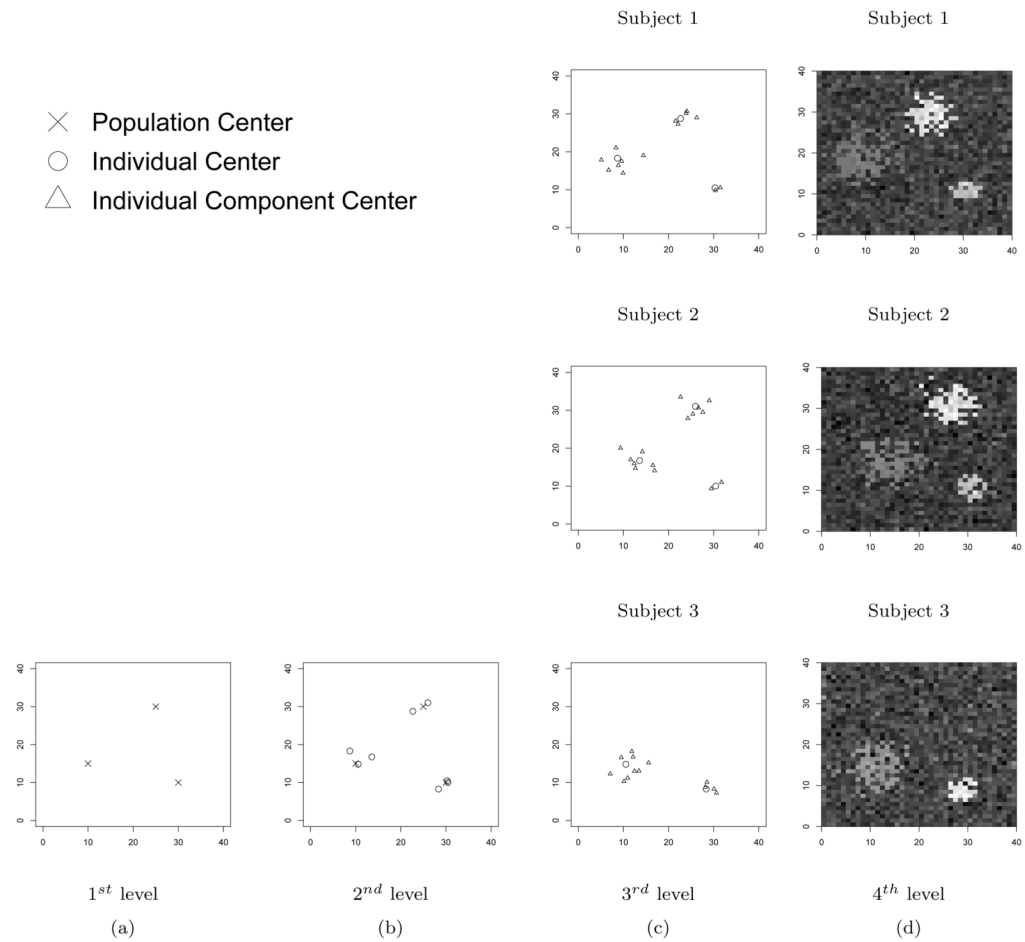
× Population Center
○ Individual Center
△ Individual Component Center

Figure 1.
Hierarchical structure of the spatial Bayesian model. (a) At the first level an unknown number of "population centers", $\boldsymbol{\mu}$, that follow a homogenous spatial Poisson process defined over the confines of the brain. (b) At the second level, an unknown number of "individual centers", $\boldsymbol{\xi}$, are distributed as Gaussian mixtures whose mixture component means are the population centers with covariance matrix $\boldsymbol{\Sigma}$. (c) At the third level, an unknown number of "individual component" means, $\boldsymbol{\eta}$, are distributed as Gaussian mixtures whose mixture component means are the individual centers with covariance matrix $\boldsymbol{\Phi}$. (d) At the final level, we assume the data, y, for each subject are distributed as a Gaussian mixture with an unknown number of mixing components whose means are θ with variances σ^2.
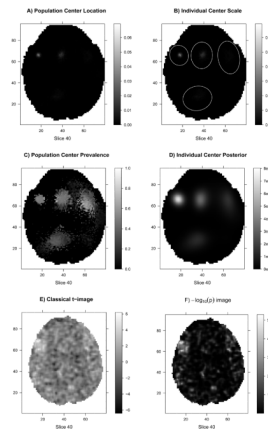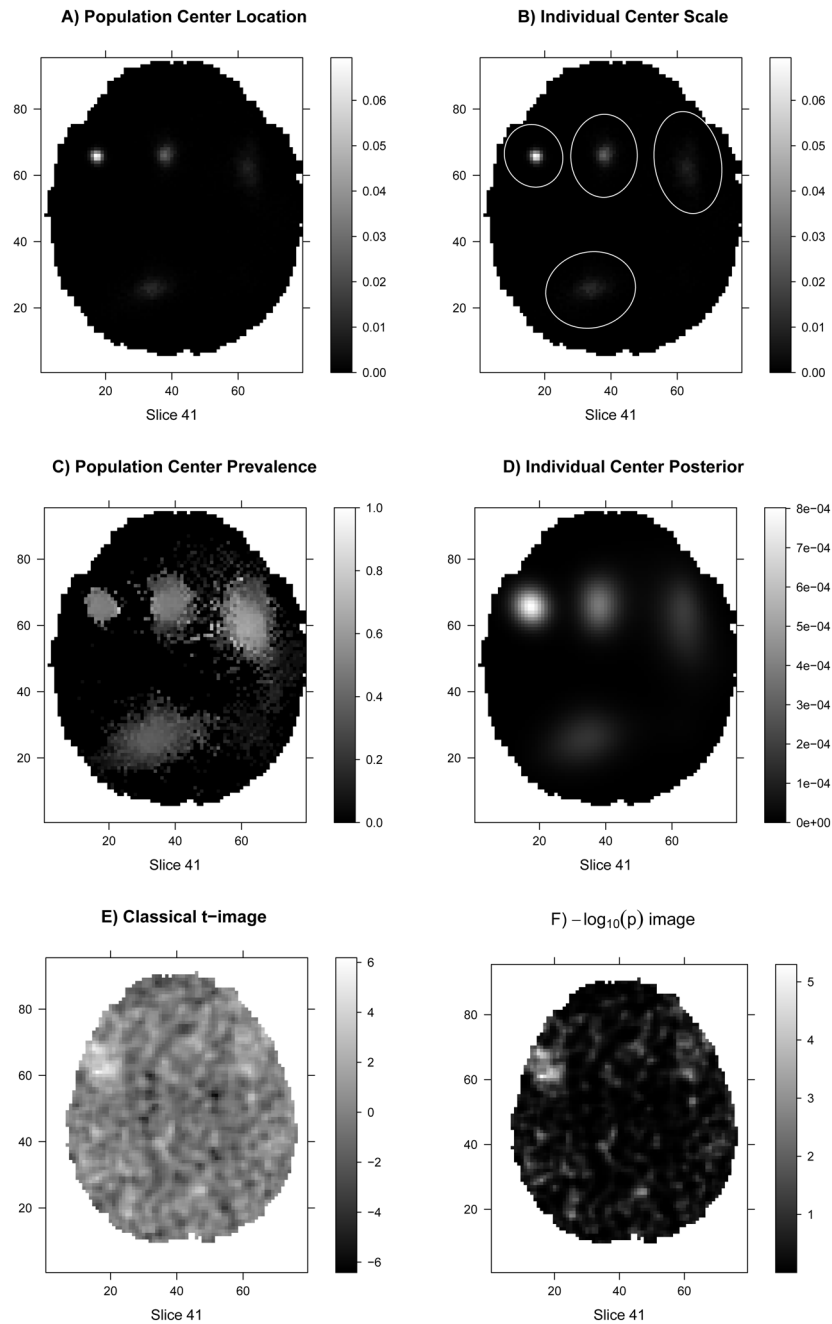
Figure 2.
Population level results. Slice 40. A) Population center locations. B) Individual center scale. C) Population center prevalence. D) Individual center posterior. E) Classical t-image. F) − log$_{10}$(p) image.

Figure 3.
Population level results. Slice 41. A) Population center locations. B) Individual center scale. C) Population center prevalence. D) Individual center posterior. E) Classical t-image. F) − $\log_{10}(p)$ image.
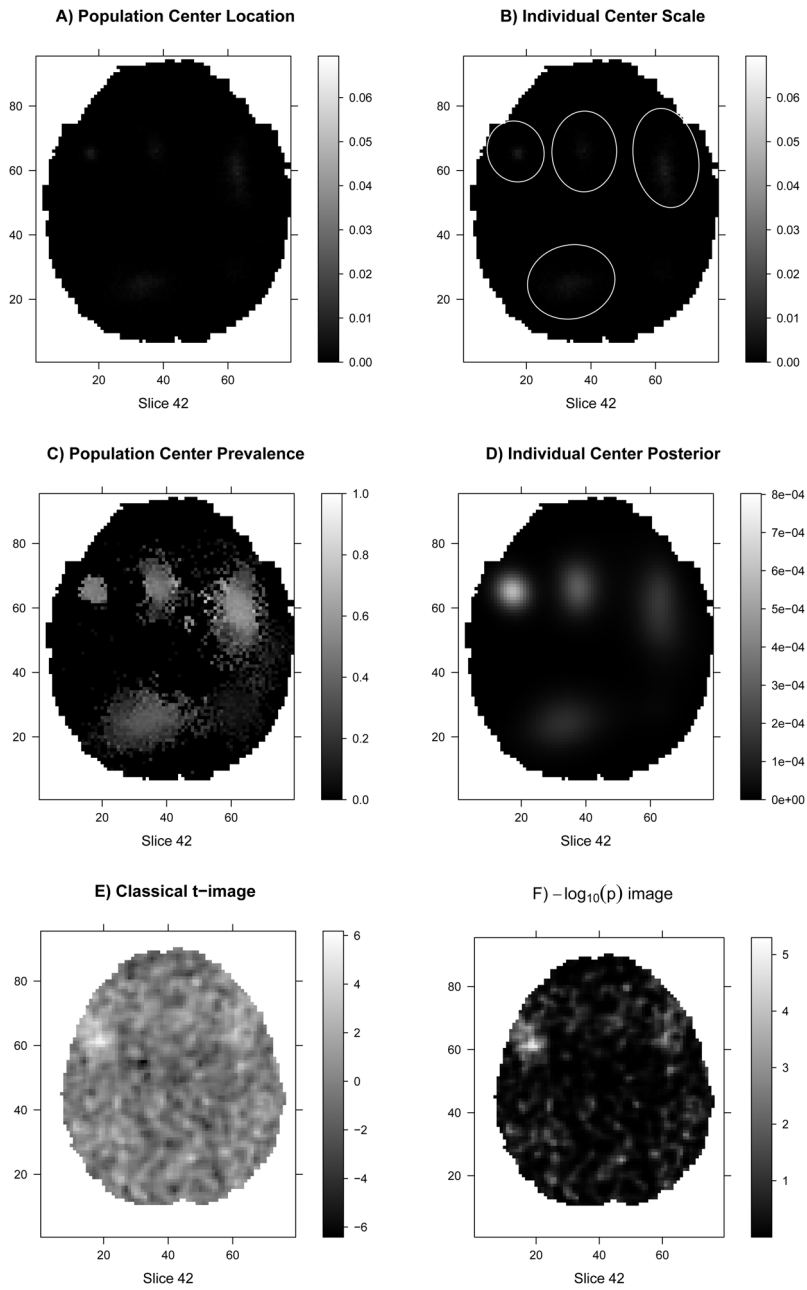
Figure 4.
Population level results. Slice 42. A) Population center locations. B) Individual center scale.
C) Population center prevalence. D) Individual center posterior. E) Classical t-image. F) −
$\log_{10}(p)$ image.