Neuroimage Meta Analysis Web Appendix

This Web Appendix contains supplementary material for *Meta Analysis of Functional Neuroimaging Data via Bayesian Spatial Point Processes*, by Jian Kang, Timothy D. Johnson, Thomas E. Nichols and Tor D. Wager

A Spatial Point Processes

The Cox process was introduced by Cox (1955) as the doubly stochastic Poisson process. For our purposes, suppose $Z(s)$ is a non-negative random field defined on the brain, $\mathcal{B} \subset \mathbb{R}^3$. If, conditional on $\lambda(s)$, $\mathbf{X}$ is a Poisson process with intensity $\lambda(s)$, denoted Poisson$(\mathcal{B}, \lambda)$, then (marginally) $\mathbf{X}$ is said to be a Cox process driven by λ. If $\mathbf{X}$ is a Cox process driven by λ and λ is restricted to be a constant function, we call $\mathbf{X}$ a homogeneous Cox process. Now, suppose $[\mathbf{X} \mid \lambda] \sim$ Poisson$(\mathcal{B}, \lambda)$. The density of this measure with respect to Lesbesgue measure on $\mathcal{B}$ does not exist. However, the density (or Radon-Nikodym derivative) does exist with respect to the probability measure induced by the unit rate Poisson process defined on $\mathcal{B}$ (Møller and Waagepetersen 2004) and is

$$\pi(\mathbf{x} \mid \lambda) = \exp\left[|\mathcal{B}| - \int_{\mathcal{B}} \lambda(s)ds\right] \prod_{x \in \mathbf{x}} \lambda(x)$$

where $|\mathcal{B}|$ denotes the volume of $\mathcal{B}$. If, further, we attach a mark, $m_x \in \mathcal{M}$ to each point $x \in \mathbf{X}$, then $(\mathbf{X}, \boldsymbol{M}) = \{(x, m_x) : x \in \mathbf{X}, m_x \in \mathcal{M}\}$ is a marked Cox process with intensity $\rho : \mathcal{B} \times \mathcal{M} \to \mathbb{R}^+ \cup \{0\}$. If the M_x are independent and identically distributed with density $\pi(m)$ and are also independent of the process $\mathbf{X}$, then $\rho(s, m) = \lambda(s)\pi(m)$. Now consider a point process $\mathbf{X}$. Conditional on $\mathbf{X} = \mathbf{x}$, suppose that associated with each $x \in \mathbf{x}$ is a process $\mathbf{Y}_x$ centered at x and that these processes are independent of one another. Then $\cup_{x \in \mathbf{x}} \mathbf{Y}_x$ is an independent cluster process. Scatter noise and outliers are often modeled by a homogeneous Poisson process with intensity ϵ independently of all $\mathbf{y}_x$ (van Lieshout and Baddeley 2002). Then the independent cluster process has modified intensity $\lambda(\cdot; \mathbf{x}) = \epsilon + \sum_{x \in \mathbf{x}} f_x(\cdot; x)$.

B Probability Equivalence Assumption

The *probability equivalence assumption*: the probability that a type 0 focus in study c clusters about a population center $z \in \mathbf{z}$ and the probability that a study center in study c clusters about the same population center are equal over all studies, for $c = 1, \ldots, C$. Let x, y be two points in the brain $\mathcal{B}$, then we have

$$\begin{aligned}
\Pr(x \in \mathbf{X}^0_{cz} \mid x \in \mathbf{X}^0_c) &= \Pr(y \in \mathbf{Y}_{cz} \mid y \in \mathbf{Y}_c) \\
\frac{\Pr(x \in \mathbf{X}^0_{cz})}{\Pr(x \in \mathbf{X}^0_c)} &= \frac{\Pr(y \in \mathbf{Y}_{cz})}{\Pr(y \in \mathbf{Y}_c)} \\
\frac{E[I(x \in \mathbf{X}^0_{cz})]}{E[I(x \in \mathbf{X}^0_c)]} &= \frac{E[I(y \in \mathbf{Y}_{cz})]}{E[I(y \in \mathbf{Y}_c)]} \qquad \text{(B.1)} \\
\frac{E[N_{\mathbf{X}^0_{cz}}(\mathcal{B})]}{E[N_{\mathbf{X}^0_c}(\mathcal{B})]} &= \frac{E[N_{\mathbf{Y}_{cz}}(\mathcal{B})]}{E[N_{\mathbf{Y}_c}(\mathcal{B})]}, \qquad \text{(B.2)}
\end{aligned}$$

where (B.2) follows from (B.1) by the following argument. Suppose there are two point processes $\mathbf{A}, \mathbf{B} \subset \mathcal{S}$ and $\mathbf{B} \neq \emptyset$, for any two points $\xi, \eta \in \mathcal{S}$, then $\Pr(\xi \in \mathbf{A}) = \Pr(\eta \in \mathbf{A})$ and $\Pr(\xi \in \mathbf{B}) = \Pr(\eta \in \mathbf{B}) > 0$, thus,

$$\frac{\Pr(\xi \in \mathbf{A})}{\Pr(\xi \in \mathbf{B})} = \frac{\Pr(\eta \in \mathbf{A})}{\Pr(\eta \in \mathbf{B})} \Rightarrow \frac{E[I(\xi \in \mathbf{A})]}{E[I(\xi \in \mathbf{B})]} = \frac{E[I(\eta \in \mathbf{A})]}{E[I(\eta \in \mathbf{B})]} \equiv R.$$

Therefore,

$$\frac{E[N_{\mathbf{A}}(\mathcal{S})]}{E[N_{\mathbf{B}}(\mathcal{S})]} = \frac{\sum_{\eta \in \mathcal{S}} E[I(\eta \in \mathbf{A})]}{\sum_{\eta \in \mathcal{S}} E[I(\eta \in \mathbf{B})]} = \frac{\sum_{\eta \in \mathcal{S}} R \cdot E[I(\eta \in \mathbf{B})]}{\sum_{\eta \in \mathcal{S}} E[I(\eta \in \mathbf{B})]} = R = \frac{E[I(\xi \in \mathbf{A})]}{E[I(\xi \in \mathbf{B})]}, \quad \text{for any } \xi \in \mathcal{S},$$

where $N_\bullet(\mathcal{S}) = \sum_{a \in \mathcal{S}} I(a \in \bullet)$.

According to the definitions of $\mathbf{X}^0_c$, $\mathbf{X}^0_{cz}$, $\mathbf{Y}_c$ and $\mathbf{Y}_{cz}$ and the intensity function (3) and (6) in Section 2.2, (B.2) implies, for $c = 1, \ldots, C$,

$$p_z \equiv \frac{\theta_{1c}\Phi(\mathcal{B}; z, \mathbf{\Sigma}_z)}{\epsilon_{1c}|\mathcal{B}| + \sum_{z \in \mathbf{z}} \theta_{1c}\Phi(\mathcal{B}; z, \mathbf{\Sigma}_z)} = \frac{\theta_{2c}\Phi(\mathcal{B}; z, \mathbf{\Sigma}_z)}{\epsilon_{2c}|\mathcal{B}| + \sum_{z \in \mathbf{z}} \theta_{2c}\Phi(\mathcal{B}; z, \mathbf{\Sigma}_z)}. \qquad \text{(B.3)}$$

By routine calculations, we have

$$\frac{\theta_{1c}}{\epsilon_{1c}} = \frac{\theta_{2c}}{\epsilon_{2c}} = \frac{1}{|\mathcal{B}|}\left(\frac{\Phi(\mathcal{B}; z, \mathbf{\Sigma}_z)}{p_z} - \sum_{z' \in \mathbf{z}} \Phi(\mathcal{B}; z', \mathbf{\Sigma}'_z)\right) \equiv q_z. \qquad \text{(B.4)}$$

From the definition of θ and ϵ in equation (8) in Section 2.2, we have

$$\frac{\theta}{\epsilon} = \frac{\sum_{c=1}^{C}(\theta_{1c}+\theta_{2c})}{\sum_{c=1}^{C}(\epsilon_{1c}+\epsilon_{2c})} = \frac{\sum_{c=1}^{C} q_z(\epsilon_{1c}+\epsilon_{2c})}{\sum_{c=1}^{C}(\epsilon_{1c}+\epsilon_{2c})} = q_z = \frac{\theta_{1c}}{\epsilon_{1c}} = \frac{\theta_{2c}}{\epsilon_{2c}}. \tag{B.5}$$

This further implies that we can define ρ_{1c} and ρ_{2c} as follows:

$$\rho_{1c} \equiv \frac{\epsilon_{1c}}{\epsilon} = \frac{\theta_{1c}}{\theta}, \qquad \rho_{2c} \equiv \frac{\epsilon_{2c}}{\epsilon} = \frac{\theta_{2c}}{\theta}, \tag{B.6}$$

such that

$$\sum_{c=1}^{C}(\rho_{1c}+\rho_{2c}) = \frac{\sum_{c=1}^{C}(\theta_{1c}+\theta_{2c})}{\theta} = \frac{\sum_{c=1}^{C}(\epsilon_{1c}+\epsilon_{2c})}{\epsilon} = 1. \tag{B.7}$$

Write $\boldsymbol{\rho} = (\rho_{11}, \cdots, \rho_{1C}, \rho_{21}, \cdots, \rho_{2C})$. The parameter set $\{\epsilon_{1c}, \theta_{1c}, \epsilon_{2c}, \theta_{2c}\}_{c=1}^{C}$ can be now be reparametrized by $\{\boldsymbol{\rho}, \theta, \epsilon\}$. Thus, the number of parameters is reduced by $2C-1$.

C Algorithm Details

In this section, we provide algorithm details and pseudo code. We first present the posterior distribution, and then discuss the details of the continuous time spatial birth-and-death processes for simulating from the posterior study activation center processes and the posterior population center process. We also provide details on updating all parameters in the hybrid MCMC algorithm. We then discuss some issues on Normal and Student-t probability computations, show how to estimate the posterior intensity function for the population center process and the 95% credible ellipsoids.

C.1 Posterior Distribution

The joint posterior of model parameters, including latent variables, is

$$\begin{aligned}
&\pi[\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_C, (\mathbf{y}, \boldsymbol{\Psi})_1, \ldots, (\mathbf{y}, \boldsymbol{\Psi})_C, (\mathbf{z}, \boldsymbol{\Sigma}), \boldsymbol{\epsilon}, \boldsymbol{\theta}, \boldsymbol{\eta}, \beta, \boldsymbol{T} \mid \mathbf{x}_1, \ldots \mathbf{x}_C] \\
&\quad \propto \prod_{c=1}^{C} \left\{ \pi\left[(\mathbf{x}, \boldsymbol{\delta})_c \mid \epsilon_{1c}, \theta_{1c}, \eta_c, (\mathbf{z}, \boldsymbol{\Sigma}), (\mathbf{y}, \boldsymbol{\Psi})_c\right] \pi(\epsilon_{1c})\pi(\eta_c)\pi(\theta_{1c}) \right\} \times \\
&\qquad \prod_{c=1}^{C} \left\{ \pi\left[(\mathbf{y}, \boldsymbol{\Psi})_c \mid \epsilon_{2c}, \theta_{2c}, (\mathbf{z}, \boldsymbol{\Sigma})\right] \pi(\epsilon_{2c})\pi(\theta_{2c}) \right\} \times \pi\left[(\mathbf{z}, \boldsymbol{\Sigma}) \mid \beta, \boldsymbol{T}\right] \pi(\beta)\pi(\boldsymbol{T}),
\end{aligned} \tag{C.1}$$

where $\boldsymbol{\epsilon} = (\epsilon_{11}, \ldots, \epsilon_{1C}, \epsilon_{21}, \ldots, \epsilon_{2C})$, $\boldsymbol{\theta} = (\theta_{11}, \ldots, \theta_{1C}, \theta_{21}, \ldots, \theta_{2C})$ and $\boldsymbol{\eta} = (\eta_1, \cdots, \eta_C)$. With the probability equivalence assumption, we have that $\epsilon_{dc} = \rho_{dc}\epsilon$ and $\theta_{dc} = \rho_{dc}\theta$ for $d = 1, 2$ and $c = 1, \ldots, C$, Thus, (C.1) is equivalent to

$$\begin{aligned}
&\pi[\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_C, (\mathbf{y}, \boldsymbol{\Psi})_1, \ldots, (\mathbf{y}, \boldsymbol{\Psi})_C, (\mathbf{z}, \boldsymbol{\Sigma}), \boldsymbol{\rho}, \boldsymbol{\eta}, \beta, \boldsymbol{T} \mid \mathbf{x}_1, \ldots \mathbf{x}_C] \\
&\quad \propto \prod_{c=1}^{C} \left\{ \pi\left[(\mathbf{x}, \boldsymbol{\delta})_c \mid \rho_{1c}, \eta_c, (\mathbf{z}, \boldsymbol{\Sigma}), (\mathbf{y}, \boldsymbol{\Psi})_c\right] \pi(\eta_c) \right\} \times \\
&\qquad \prod_{c=1}^{C} \left\{ \pi\left[(\mathbf{y}, \boldsymbol{\Psi})_c \mid \rho_{2c}, (\mathbf{z}, \boldsymbol{\Sigma})\right] \right\} \pi(\boldsymbol{\rho}) \times \pi\left[(\mathbf{z}, \boldsymbol{\Sigma}) \mid \beta, \boldsymbol{T}\right] \pi(\beta) \pi(\boldsymbol{T}). \qquad \text{(C.2)}
\end{aligned}$$

This is proportional to

$$\begin{aligned}
&\prod_{c=1}^{C} \left\{ \pi\left[\mathbf{x}_c^0 \mid \rho_{1c}, (\mathbf{z}, \boldsymbol{\Sigma})\right] \prod_{x \in \mathbf{x}_c^0} \pi(\delta_x = 0) \right\} \times \\
&\prod_{c=1}^{C} \left\{ \pi\left[\mathbf{x}_c^1 \mid \eta_c, (\mathbf{y}, \boldsymbol{\Psi})_c\right] \prod_{x \in \mathbf{x}_c^1} \pi(\delta_x = 1) \pi(\eta_c) \right\} \times \\
&\prod_{c=1}^{C} \left\{ \pi\left[(\mathbf{y}, \boldsymbol{\Psi})_c \mid \rho_{2c}, (\mathbf{z}, \boldsymbol{\Sigma})\right] \right\} \pi(\boldsymbol{\rho}) \times \pi\left[(\mathbf{z}, \boldsymbol{\Sigma}) \mid \beta, \boldsymbol{T}\right] \pi(\beta) \pi(\boldsymbol{T}). \qquad \text{(C.3)}
\end{aligned}$$

We propose an hybrid algorithm with two continuous time spatial birth-and-death processes embedded in a standard MCMC algorithm to sample from the posterior (C.3).

C.2 Imputation of Missing Type Indicators

The full conditional distribution of $\boldsymbol{\delta}_c$ is

$$\pi[\boldsymbol{\delta}_c \mid \cdot] \propto \prod_{d=0}^{1} \left\{ \pi[\mathbf{x}_c^d \mid \cdot] \prod_{x \in \mathbf{x}_c^d} \pi(\delta_x = d) \right\} \propto \prod_{d=0}^{1} \prod_{x \in \mathbf{x}_c^d} \lambda_{1c}^d(x; \cdot) \pi(\delta_x = d).$$

Furthermore, the full conditional probability mass function of δ_x for each $x \in \mathbf{x}_c$ is $\pi[\delta_x = d \mid \cdot] \propto \lambda_{1c}^d(x; \cdot)\pi(\delta_x = d)$ for $d = 0, 1$. We set $\pi(\delta_x = d) = 0.5$, thus,

$$\pi[\delta_x = d \mid \cdot] \;=\; \frac{\lambda_{1c}^d(x; \cdot)}{\lambda_{1c}(x; \cdot)}, \qquad \text{for } d = 0, 1. \qquad \text{(C.4)}$$

C.3 Spatial Birth-and-Death Processes

The spatial birth-and-death process (Preston 1975, Møller and Waagepetersen 2004), a continuous time Markov process whose transitions are either births or deaths, can be used to simulate spatial point processes.

C.3.1 General Procedure

Suppose we wish to construct a spatial birth-and-death process to simulate a latent point process $\mathbf{A}$ from its posterior $\pi(\mathbf{a} \mid \mathbf{u})$, where $\mathbf{u}$ is data. Preston (1975) showed that if the birth and death rates satisfy the detailed balance equation

$$\pi(\mathbf{a} \mid \mathbf{u}) b(\mathbf{a}, \xi) = \pi(\mathbf{a} \cup \{\xi\} \mid \mathbf{u}) d(\mathbf{a} \cup \{\xi\}, \{\xi\}), \tag{C.5}$$

then the chain is time reversible and that the spatial birth-and-death process has a unique equilibrium distribution $\pi(\mathbf{a} \mid \mathbf{u})$ to which it converges in distribution from any initial state. In (C.5) $b(\mathbf{a}, \xi)$ is the birth rate for adding a new point ξ to the current configuration, $\mathbf{a}$, of the point process $\mathbf{A}$, and $d(\mathbf{a}, \xi)$ denotes the death rate for removing a point ξ from $\mathbf{a}$. We adopt the birth rate suggested by van Lieshout and Baddeley (2002) using a mixture intensity, i.e. $b(\mathbf{a}, \xi) = \sum_{u \in \mathbf{u}} h(\xi; u)$, where $h(\xi; \cdot)$ is an intensity function. To satisfy the detailed balance equation, the death rate for removing ξ from $\mathbf{a}$ is $d(\mathbf{a}, \xi) = [\pi(\mathbf{a}/\{\xi\} \mid \mathbf{u})/\pi(\mathbf{a} \mid \mathbf{u})] \sum_{u \in \mathbf{u}} h(\xi; u)$. Let $B(\mathbf{a}) = \int b(\mathbf{a}, \xi) d\xi$ denote the total birth rate. Note that $B(\mathbf{a})$ does not depend on $\mathbf{a}$. Write $B = B(\mathbf{a})$. Let $D(\mathbf{a}) = \sum_{a \in \mathbf{a}} d(\mathbf{a}; a)$ denote the total death rate. Given current state $\mathbf{a}$, after an exponentially distributed time with rate $B + D(\mathbf{a})$, a birth is proposed with probability $B/\{B + D(\mathbf{a})\}$ by sampling ξ from the mixture density $\sum_{u \in \mathbf{u}} h(\xi; u)/B$. A death is proposed with probability $D(\mathbf{a})/\{B + D(\mathbf{a})\}$ and the point $a \in \mathbf{a}$ is removed with probability $d(\mathbf{a}; a)/D(\mathbf{a})$.

C.3.2 Details for Simulating the Posterior Study Activation Center Processes

From (C.2), for $c = 1, \ldots, C$, the full conditional posterior distribution of a marked study activation center process for study c is

$$\pi[(\mathbf{y}, \mathbf{\Psi})_c \mid \cdot] \propto \pi[\mathbf{x}_c^1 \mid \eta_c, (\mathbf{y}, \mathbf{\Psi})_c] \times \pi[(\mathbf{y}, \mathbf{\Psi})_c \mid \rho_{2c}, (\mathbf{z}, \mathbf{\Sigma})],$$

for which we construct a spatial birth-and-death process using the birth rate

$$b_c[(\mathbf{y}, \mathbf{\Psi})_c, (y, \mathbf{\Psi}_y)] = \sum_{x \in \mathbf{x}_c^1} \eta_c \phi_3(y; x, \mathbf{\Psi}_y) \pi(\mathbf{\Psi}_y), \tag{C.6}$$

where $\phi_3(\xi; \mu, \mathbf{\Sigma})$ denotes the 3D normal density at ξ with mean μ and covariance matrix $\mathbf{\Sigma}$. The death rate for removing $(y, \mathbf{\Psi}_y) \in (\mathbf{y}, \mathbf{\Psi})_c$ is then

$$\begin{aligned}
&d_c[(\mathbf{y}, \mathbf{\Psi})_c, (y, \mathbf{\Psi}_y)] \\
&\quad = \frac{\pi[\mathbf{x}_c^1 \mid \eta_c, (\mathbf{y}, \mathbf{\Psi})_c / \{(y, \mathbf{\Psi}_y)\}] \cdot \pi[(\mathbf{y}, \mathbf{\Psi})_c / \{(y, \mathbf{\Psi}_y)\} \mid \rho_{2c}, (\mathbf{z}, \mathbf{\Sigma})]}{\pi[\mathbf{x}_c^1 \mid \eta_c, (\mathbf{y}, \mathbf{\Psi})_c] \cdot \pi[(\mathbf{y}, \mathbf{\Psi})_c \mid \rho_{2c}, (\mathbf{z}, \mathbf{\Sigma})]} \sum_{x \in \mathbf{x}_c} \eta_c \phi_3(y; x, \mathbf{\Psi}_y) \pi(\mathbf{\Psi}_y) \\
&\quad = \frac{\exp\{\eta_c \Phi_3(\cdot; y, \mathbf{\Psi}_y)\}}{\prod_{x \in \mathbf{x}_c^1} \left[1 + \frac{\eta_c \phi_3(x; y, \mathbf{\Psi}_y)}{\lambda_{1c}^1(x; \cdot) - \eta_c \phi_3(x; y, \mathbf{\Psi}_y)}\right]} \cdot \frac{\eta_c \sum_{x \in \mathbf{x}_c} \phi_3(y; x, \mathbf{\Psi}_y)}{\lambda_{2c}(y; \cdot)},
\end{aligned} \tag{C.7}$$

where $\Phi_3(\cdot \mid \mu, \mathbf{\Sigma}) = \int_{\mathcal{B}} \phi_3(\xi \mid \mu, \mathbf{\Sigma}) d\xi$, is 3D normal probability over the brain $\mathcal{B}$ with mean μ and covariance matrix $\mathbf{\Sigma}$. Using (C.6), the total birth rate is

$$B_c = \int_{\mathcal{B}} \int_{\mathcal{M}} \sum_{x \in \mathbf{x}_c^1} \eta_c \phi_3(y; x, \mathbf{\Psi}_y) \pi(\mathbf{\Psi}_y) d\mathbf{\Psi}_y dy = \eta_c \sum_{x \in \mathbf{x}_c^1} \mathbf{T}_3(\cdot; x, \boldsymbol{S}/(d-2), d), \tag{C.8}$$

where $\mathcal{M}$ is the space of 3×3 symmetric positive definite matrices. $\mathbf{T}_3(\cdot; \mu, \boldsymbol{S}, d)$ is student-t probability over the brain $\mathcal{B}$ with mean μ, scale matrix $\boldsymbol{S}$ and d degrees of freedom. The total death rate is given by

$$D_c[(\mathbf{y}, \mathbf{\Psi})_c] = \sum_{(y, \mathbf{\Psi}_y) \in (\mathbf{y}, \mathbf{\Psi})_c} d_c[(\mathbf{y}, \mathbf{\Psi})_c, (y, \mathbf{\Psi}_y)]. \tag{C.9}$$

For a birth, we draw a new point $(y, \mathbf{\Psi}_y)$ from

$$\frac{b_c[(\mathbf{y}, \mathbf{\Psi})_c, (y, \mathbf{\Psi}_y)]}{B_c} = \sum_{x \in \mathbf{x}_c^1} \left\{ \frac{\phi_3(y; x, \mathbf{\Psi}_y) \pi(\mathbf{\Psi}_y)}{\mathbf{T}_3(\cdot; x, \boldsymbol{S}/(d-2), d)} \cdot \frac{\mathbf{T}_3(\cdot; x, \boldsymbol{S}/(d-2), d)}{\sum_{x' \in \mathbf{x}_c^1} \mathbf{T}_3(\cdot; x', \boldsymbol{S}/(d-2), d)} \right\}, \tag{C.10}$$

To draw from this mixture distribution, first draw

$$\mathbf{\Psi}_y \sim W^{-1}(\boldsymbol{S}, d), \tag{C.11}$$

then draw y from the following mixture distribution,

$$[y \mid \mathbf{\Psi}_y] \sim \sum_{x \in \mathbf{x}_c^1} v_x N_{\mathcal{B}}(x, \mathbf{\Psi}_y), \tag{C.12}$$

where $N_{\mathcal{B}}(\mu, \mathbf{\Sigma})$ denotes the 3D normal distribution truncated to $\mathcal{B}$ with mean μ and covariance matrix $\mathbf{\Sigma}$, and the weight $v_x = \mathbf{T}_3(\cdot; x, \boldsymbol{S}/(d-2), d)/B_c$.

The simulation time for this spatial birth-and-death process at each iteration of the hybrid algorithm is set to $1/B_c$, and the number of points in $\mathbf{y}_c$ is initially set to zero.

C.3.3 Details for Simulating the Posterior Population Center Process

Note that (B.6) implies that the intensity functions for $\mathbf{X}_0^c$, $\mathbf{Y}_c$ and $\mathbf{Y}$, say, λ_{1c}^0, λ_{2c} and λ have the following relationships: $\lambda_{1c}^0(y;\cdot) = \rho_{1c}\lambda(y;\cdot)$ and $\lambda_{2c}(y;\cdot) = \rho_{2c}\lambda(y;\cdot)$. Thus,

$$\begin{aligned}
&\prod_{c=1}^{C} \left\{ \pi[\mathbf{x}_c^0 \mid \rho_{1c}, (\mathbf{z}, \mathbf{\Sigma})] \cdot \pi[\mathbf{y}_c \mid \rho_{2c}, (\mathbf{z}, \mathbf{\Sigma})] \right\} \\
&\quad \propto \prod_{c=1}^{C} \left\{ \exp\left\{ -\int_{\mathcal{B}} \lambda_{1c}^0(x;\cdot) + \lambda_{2c}(x;\cdot) dx \right\} \prod_{x \in \mathbf{x}_c^0} \lambda_{1c}^0(x;\cdot) \prod_{y \in \mathbf{y}_c} \lambda_{2c}(y;\cdot) \right\} \\
&\quad \propto \pi[\mathbf{y} \mid (\mathbf{z}, \mathbf{\Sigma})] \cdot \prod_{c=1}^{C} \left\{ \rho_{1c}^{\mathbf{n}(\mathbf{x}_c^0)} \rho_{2c}^{\mathbf{n}(\mathbf{y}_c)} \right\}.
\end{aligned}$$

This implies that (C.3), the full joint posterior, is proportional to

$$\begin{aligned}
&\prod_{c=1}^{C} \left\{ \pi\left[\mathbf{x}_c^1 \mid \eta_c, (\mathbf{y}_c, \mathbf{\Psi}_c)\right] \pi(\eta_c) \rho_{1c}^{\mathbf{n}(\mathbf{x}_c^0)} \rho_{2c}^{\mathbf{n}(\mathbf{y}_c)} \right\} \prod_{y \in \mathbf{y}} \pi(\mathbf{\Psi}_y) \times \\
&\pi(\boldsymbol{\rho}) \pi\left[\mathbf{y} \mid (\mathbf{z}, \mathbf{\Sigma})\right] \pi\left[\mathbf{z} \mid \beta\right] \pi(\beta) \pi(\boldsymbol{T}) \prod_{z \in \mathbf{z}} \pi(\mathbf{\Sigma} \mid \boldsymbol{T}).
\end{aligned} \tag{C.13}$$

Thus, the full conditional posterior distribution of the population center process is

$$\pi[(\mathbf{z}, \mathbf{\Sigma}) \mid \cdot] \quad \propto \quad \pi[\mathbf{y} \mid (\mathbf{z}, \mathbf{\Sigma})] \pi[\mathbf{z} \mid \beta] \prod_{z \in \mathbf{z}} \pi(\mathbf{\Sigma}_z \mid \boldsymbol{T}),$$

for which we construct a spatial birth-and-death process using the birth rate,

$$b[(\mathbf{z}, \mathbf{\Sigma}), (z, \mathbf{\Sigma}_z)] = \beta\pi(\mathbf{\Sigma}_z \mid \boldsymbol{T}) \left[1 + \frac{\theta}{\epsilon} \sum_{y \in \mathbf{y}} \phi_3(y; z, \mathbf{\Sigma}_z)\right], \tag{C.14}$$

and the death rate, for removing a point $(z, \mathbf{\Sigma}_z) \in (\mathbf{z}, \mathbf{\Sigma})$, is

$$d[(\mathbf{z}, \mathbf{\Sigma}), (z, \mathbf{\Sigma}_z)] = \frac{\exp\{\theta\Phi_3(\cdot; z, \mathbf{\Sigma}_z)\}}{\prod_{y \in \mathbf{y}} \left[1 + \frac{\theta\phi_3(y;z,\mathbf{\Sigma}_z)}{\lambda(y;\cdot) - \theta\phi_3(y;z,\mathbf{\Sigma}_z)}\right]} \left[1 + \frac{\theta}{\epsilon} \sum_{y \in \mathbf{y}} \phi_3(y; z, \mathbf{\Sigma}_z)\right]. \tag{C.15}$$

The total birth rate is

$$B = \int_{\mathcal{B}} \int_{\mathcal{M}} b[(\mathbf{z}, \mathbf{\Sigma}), (z, \mathbf{\Sigma}_z)] d\mathbf{\Sigma}_z dz = \beta \left[|\mathcal{B}| + \frac{\theta}{\epsilon} \sum_{y \in \mathbf{y}} \mathbf{T}_3(\cdot; y, \boldsymbol{T}/(\nu - 2), \nu - 2)\right] \tag{C.16}$$

and the total death rate is

$$D(\mathbf{z}, \mathbf{\Sigma}) = \sum_{z \in \mathbf{z}} d[(\mathbf{z}, \mathbf{\Sigma}), (z, \mathbf{\Sigma}_z)] \tag{C.17}$$

For a birth, we draw a new point $(z, \mathbf{\Sigma}_z)$ from the mixture density

$$\frac{b[(\mathbf{z}, \mathbf{\Sigma}), (z, \mathbf{\Sigma}_z)]}{B} = \sum_{y \in \mathbf{y}} \left\{ \frac{\beta|\mathcal{B}|\pi(\mathbf{\Sigma}_z)}{B} + \frac{\phi_3(y; z, \mathbf{\Sigma}_z)\pi(\mathbf{\Sigma}_z)}{\mathbf{T}_3(\cdot; y, \frac{\boldsymbol{T}}{\nu-2}, \nu - 2)} \frac{\beta\mathbf{T}_3(\cdot; y, \frac{\boldsymbol{T}}{\nu-2}, \nu - 2)}{B} \right\}. \tag{C.18}$$

This implies that we first draw

$$\mathbf{\Sigma}_z \sim W^{-1}(\boldsymbol{S}, \nu), \tag{C.19}$$

then draw z from the following mixture distribution,

$$[z \mid \mathbf{\Sigma}_z] \sim w_\emptyset U_{\mathcal{B}} + \sum_{y \in \mathbf{y}} w_y N_{\mathcal{B}}(y, \mathbf{\Sigma}_z), \tag{C.20}$$

where $U_{\mathcal{B}}$ denotes the uniform distribution over the brain $\mathcal{B}$, and the weight $w_y = \beta|\mathcal{B}|/B$ when $y = \emptyset$ and $w_y = \beta B^{-1}\mathbf{T}_3(\cdot; y, \boldsymbol{T}/(\nu - 2), \nu - 2)$ if $y \in \mathbf{y}$.

The total simulation time for this spatial birth-and-death process at each iteration of the hybrid algorithm is set to $1/B$ and number of points in the population center process is set to zero as the initial state.

C.4 Standard MCMC Updates

Upon exiting the spatial birth-and-death processes, conditional on the number of points in $\mathbf{Y}_c$ and $\mathbf{Z}$, we use a standard MCMC algorithm to update the study activation center processes parameters and population center process parameters, i.e. $(\mathbf{y}, \boldsymbol{\Psi})_c$ and $(\mathbf{z}, \boldsymbol{\Sigma})$, as well as $\boldsymbol{\eta}$, $\boldsymbol{\rho}$, β and $\boldsymbol{T}$.

C.4.1 Update $(\mathbf{y}, \boldsymbol{\Psi})_c$ given $N_{\mathbf{Y}_c}(\mathcal{B}) = n_c$

Define a latent indicator variable $\alpha_x \in \mathbf{y}_c$ with prior $\pi(\alpha_x = y) = 1/n_c$ for each $x \in \mathbf{X}_c^1$ and $y \in \mathbf{y}_c$, such that $\mathbf{X}_{cy}^1 = \{x \in \mathbf{X}_c^1; \alpha_x = y\}$. Recall that $\mathbf{X}_{cy}^1$ is a process defined in Section 2.1 with intensity function $\eta_c \phi_3(x; y, \boldsymbol{\Psi}_y)$. Write $\boldsymbol{\alpha}_c = \{\alpha_x, x \in \mathbf{X}_c^1\}$. Note that $\pi[\mathbf{x}_c^1, \boldsymbol{\alpha}_c \mid \lambda_{1c}^1] \propto \prod_{y \in \mathbf{y}} \pi[\mathbf{x}_{cy} \mid \eta_c, y, \boldsymbol{\Psi}_y]$. Then the full conditional distribution of $(\mathbf{y}, \boldsymbol{\Psi})_c$ and $\boldsymbol{\alpha}_c$ is $\pi[(\mathbf{y}, \boldsymbol{\Psi})_c, \boldsymbol{\alpha}_c \mid \cdot] \propto \prod_{y \in \mathbf{y}_c} \left\{ \pi[\mathbf{x}_{cy} \mid \eta_c, \boldsymbol{\Psi}_y] \pi(\boldsymbol{\Psi}_y) \right\} \pi[\mathbf{y}_c \mid \rho_{2c}, (\mathbf{z}, \boldsymbol{\Sigma})]$. This further implies the full conditional of $y \in \mathbf{y}_c$ is

$$\pi[y \mid \cdot] \propto \exp\{-\eta_c \Phi_3(\cdot; y, \boldsymbol{\Psi}_y)\} \lambda_{2c}(y; \cdot)\, \phi_3 \left(y; \frac{\sum\limits_{x \in \mathbf{x}_c^1} x I(\alpha_x = y)}{\sum\limits_{x \in \mathbf{x}_c^1} I(\alpha_x = y)}, \frac{\boldsymbol{\Psi}_y}{\sum\limits_{x \in \mathbf{x}_c^1} I(\alpha_x = y)} \right).$$

Thus to update $y \in \mathbf{y}_c$, draw

$$y^* \sim N_{\mathcal{B}} \left[\frac{\sum_{x \in \mathbf{x}_c^1} x I(\alpha_x = y)}{\sum_{x \in \mathbf{x}_c^1} I(\alpha_x = y)}, \frac{\boldsymbol{\Psi}_y}{\sum_{x \in \mathbf{x}_c^1} I(\alpha_x = y)} \right], \tag{C.21}$$

and accept with probability

$$\min \left\{ 1, \exp\{\eta_c [\Phi_3(\cdot; y, \boldsymbol{\Psi}_y) - \Phi_3(\cdot; y^*, \boldsymbol{\Psi}_y)]\} \frac{\lambda_2(y^*; \cdot)}{\lambda_2(y; \cdot)} \right\}. \tag{C.22}$$

The full conditional of $\boldsymbol{\Psi}_y$ is

$$\pi[\boldsymbol{\Psi}_y \mid \cdot] \propto \exp\{-\eta_c \Phi_3(\cdot; y, \boldsymbol{\Psi}_y)\} \pi(\boldsymbol{\Psi}_y)\, \phi_3 \left(y; \frac{\sum\limits_{x \in \mathbf{x}_c^1} x I(\alpha_x = y)}{\sum\limits_{x \in \mathbf{x}_c^1} I(\alpha_x = y)}, \frac{\boldsymbol{\Psi}_y}{\sum\limits_{x \in \mathbf{x}_c^1} I(\alpha_x = y)} \right).$$

Thus to update $\boldsymbol{\Psi}_y$, draw

$$\boldsymbol{\Psi}_y^* \sim W^{-1} \left[\boldsymbol{S} + \sum_{x \in \mathbf{x}_1^c} (x - y)(x - y)^{\mathrm{T}} I(\alpha_x = y), d + \sum_{x \in \mathbf{x}_c^1} I(\alpha_x = y) \right], \tag{C.23}$$

and accept with probability

$$\min\left\{1, \exp\{-\eta_c[\Phi_3(\cdot; y, \mathbf{\Psi}_y^*) - \Phi_3(\cdot; y, \mathbf{\Psi}_y)]\}\right\}. \tag{C.24}$$

To update α_x, we have

$$\pi[\alpha_x = y \mid \cdot] \;\; = \;\; \frac{\phi_3(x; y, \mathbf{\Psi}_y)}{\sum_{y' \in \mathbf{y}_c} \phi_3(x; y', \mathbf{\Psi}_{y'})} \quad \forall\, y \in \mathbf{y}_c. \tag{C.25}$$

C.4.2 Update $(\mathbf{z}, \mathbf{\Sigma})$ given $N_{\mathbf{Z}}(\mathcal{B}) = m$

Define a latent indicator variable $\gamma_y \in \mathbf{z} \cup \{\emptyset\}$ with prior $\pi(\gamma_y = z) = 1/(m+1)$ for each $z \in \mathbf{z} \cup \{\emptyset\}$ and $y \in \mathbf{Y}$ such that $\mathbf{Y}_{cz} = \{y \in \mathbf{Y}; \gamma_y = z\}$. From Section 2.1, $\mathbf{Y}_{c\emptyset}$ is a homogeneous Poisson process with constant intensity ϵ and $\mathbf{Y}_{cz}$ for $z \in \mathbf{Z}$ is a point process with intensity $\theta\phi_3(y; z, \mathbf{\Sigma}_z)$. Write $\boldsymbol{\gamma} = \{\gamma_y, y \in \mathbf{Y}\}$. Note that $\pi[(\mathbf{y}, \boldsymbol{\gamma}) \mid \cdot] \propto \pi[\mathbf{y}_{c\emptyset}] \prod_{z \in \mathbf{z}} \pi[\mathbf{y}_{cz} \mid z, \mathbf{\Sigma}_z]$. Then the joint posterior distribution of $(\mathbf{z}, \mathbf{\Sigma})$ and $\boldsymbol{\gamma}$ given all other parameters is

$$\pi[(\mathbf{z}, \mathbf{\Sigma}), \boldsymbol{\gamma} \mid \cdot] \propto \prod_{z \in \mathbf{z}} \left\{\pi[\mathbf{y}_{cz} \mid z, \mathbf{\Sigma}_z] \pi[\mathbf{\Sigma}_z]\right\} \pi[\mathbf{z} \mid \beta].$$

Thus the full conditional of $z \in \mathbf{z}$ is

$$\pi[z \mid \cdot] \propto \exp\{-\theta\Phi_3(\cdot; z, \mathbf{\Sigma}_z)\} \phi_3\left(z; \frac{\sum_{y \in \mathbf{y}} y I(\gamma_y = z)}{\sum_{y \in \mathbf{y}} I(\gamma_y = z)}, \frac{\mathbf{\Sigma}_z}{\sum_{y \in \mathbf{y}} I(\gamma_y = z)}\right).$$

Thus to update $z \in \mathbf{z}$, draw

$$z^* \sim N_{\mathcal{B}}\left[\frac{\sum_{y \in \mathbf{y}} y I(\gamma_y = z)}{\sum_{y \in \mathbf{y}} I(\gamma_y = z)}, \frac{\mathbf{\Sigma}_z}{\sum_{y \in \mathbf{y}} I(\gamma_y = z)}\right], \tag{C.26}$$

and accept with probability

$$\min\{1, \exp\{\theta[\Phi_3(\cdot, z, \mathbf{\Sigma}_z) - \Phi_3(\cdot, z^*, \mathbf{\Sigma}_z)]\}\}. \tag{C.27}$$

The full conditional of $\mathbf{\Sigma}_z$ is

$$\pi[\mathbf{\Sigma}_z \mid \cdot] \propto \exp\{-\theta\Phi_3(\cdot; z, \mathbf{\Sigma}_z)\} \pi(\mathbf{\Sigma}_z \mid \boldsymbol{T})\, \phi_3\left(z; \frac{\sum\limits_{y \in \mathbf{y}} y I(\gamma_y = z)}{\sum\limits_{y \in \mathbf{y}} I(\gamma_y = z)}, \frac{\mathbf{\Sigma}_z}{\sum\limits_{y \in \mathbf{y}} I(\gamma_y = z)}\right).$$

Thus update $\boldsymbol{\Sigma}_z$ by first drawing

$$\boldsymbol{\Sigma}_z^* \sim W^{-1}\left[\boldsymbol{T} + \sum_{y\in\mathbf{y}}(y-z)(y-z)^{\mathrm{T}} I(\gamma_y = z), \nu + \sum_{y\in\mathbf{y}} I(\gamma_y = z)\right], \tag{C.28}$$

and accept with probability

$$\min\{1, \exp\{\theta[\Phi_3(\cdot, z, \boldsymbol{\Sigma}_z) - \Phi_3(\cdot, z, \boldsymbol{\Sigma}_z^*)]\}\}. \tag{C.29}$$

To update each γ_y for $y \in \mathbf{Y}$, we have

$$\pi[\gamma_y = \emptyset \mid \cdot] = \frac{\epsilon}{\epsilon + \theta\sum_{z'\in\mathbf{z}}\phi_3(y; z', \boldsymbol{\Sigma}_{z'})} \tag{C.30}$$

$$\pi[\gamma_y = z \mid \cdot] = \frac{\theta\phi_3(y; z, \boldsymbol{\Sigma}_z)}{\epsilon + \theta\sum_{z'\in\mathbf{z}}\phi_3(y; z', \boldsymbol{\Sigma}_{z'})} \quad \forall\, z \in \mathbf{z}, \tag{C.31}$$

C.4.3 Update $\boldsymbol{\eta}$, $\boldsymbol{\rho}$, β and $\boldsymbol{T}$

The full conditional for η_c for $c = 1, 2, \ldots, C$ is

$$\pi[\eta_c \mid \cdot] \propto \pi[\mathbf{x}_c^1 \mid \eta_c, (\mathbf{y}, \boldsymbol{\Psi})_c]\pi(\eta_c) \propto \exp\left\{-\eta_c\left[\sum_{y\in\mathbf{y}_c}\Phi_3(\cdot; y, \boldsymbol{\Psi}_y) + b_\eta\right]\right\}\eta_c^{\mathbf{n}(\mathbf{x}_c^1)+a_\eta}, \tag{C.32}$$

Thus to update η_c draw

$$\eta_c \sim G\left[\sum_{y\in\mathbf{y}_c}\Phi_3(\cdot; y, \boldsymbol{\Psi}_y) + b_\eta, \mathbf{n}(\mathbf{x}_c^1) + a_\eta\right], \tag{C.33}$$

According to (C.13), the full conditional for $\boldsymbol{\rho}$ is

$$\pi[\boldsymbol{\rho} \mid \cdot] \propto \prod_{c=1}^{C}\left\{\rho_{1c}^{\mathbf{n}(\mathbf{x}_c^0)}\rho_{2c}^{\mathbf{n}(\mathbf{y}_c)}\right\}\pi(\boldsymbol{\rho}) \propto \prod_{c=1}^{C}\left\{\rho_{1c}^{\mathbf{n}(\mathbf{x}_c^0)+\alpha_{1c}}\rho_{2c}^{\mathbf{n}(\mathbf{y}_c)+\alpha_{2c}}\right\}, \tag{C.34}$$

Thus update $\boldsymbol{\rho}$ by drawing

$$\boldsymbol{\rho} \sim D[\mathbf{n}(\mathbf{x}_1^0) + \alpha_{11}, \ldots, \mathbf{n}(\mathbf{x}_C^0) + \alpha_{1C}, \mathbf{n}(\mathbf{y}_1) + \alpha_{21}, \ldots, \mathbf{n}(\mathbf{y}_C) + \alpha_{2C}]. \tag{C.35}$$

The full conditional for β is $\pi[\beta \mid \cdot] \propto \exp\{-\beta(|\mathcal{B}| + b_\beta)\}\,\beta^{\mathbf{n}(\mathbf{z})+a_\beta}$. Thus draw

$$\beta \sim G[\mathbf{n}(\mathbf{z}) + \alpha_\beta, |\mathcal{B}| + b_\beta]. \tag{C.36}$$

The full conditional of $\boldsymbol{T}$ is $\pi[\boldsymbol{T} \mid \cdot] \propto \pi(\boldsymbol{T})\prod_{z\in\mathbf{z}}\pi(\boldsymbol{\Sigma}_z \mid \boldsymbol{T})$. Therefore, draw

$$\boldsymbol{T} \sim W^{-1}\left[(\boldsymbol{T}_0^{-1} + \sum_{z\in\mathbf{z}}\boldsymbol{\Sigma}_z^{-1}), \nu_0 + \nu\mathbf{n}(\mathbf{z})\right]. \tag{C.37}$$

C.5 Normal and T Probability Computation

The total birth rate and death rate in the spatial birth-and-death process, as well as many standard MCMC updates involve the evaluation of 3D normal and student-t probabilities over the brain. It is difficult to directly evaluate these probabilities over arbitrary regions such as the brain. Thus, we resort to Monte Carlo simulation of these probabilities. Since more than half of the normal and t probabilities in our motivating example are close to 1 we consider the following approximation. If the 99% credible ellipsoid of the target distribution lies completely within the brain we set the probability to 0.995, otherwise, we estimate it via Monte Carlo simulation. We set the Monte Carlo sample size $n = 500$ using the estimated standard error, $\sqrt{p(1-p)/n}$, where p is the true probability. This results in a maximum absolute value of the Monte Carlo error (twice the standard error) of 0.05 when $p = 0.5$.

C.6 Posterior Intensity Estimation

Let $(\mathbf{z}^{(k)}, \boldsymbol{\Sigma}^{(k)})$ be the posterior draw of the population center process at the kth iteration after burn-in, for $k = 1, \ldots, K$. To obtain its intensity function, we combine all of the draws of locations, i.e. $\cup_{k=1}^{K} \mathbf{z}^{(k)}$, which are then smoothed with a mixture of Dirichlet process priors model (a non-parametric Bayesian model used for density smoothing/estimation, Escobar and West (1995)). We rescale the density by multiplying the posterior mean number of population centers, i.e. $\sum_{k=1}^{K} \mathbf{n}(\mathbf{z}^{(k)})/K$ to estimate the intensity.

Also, using $(\mathbf{z}^{(k)}, \boldsymbol{\Sigma}^{(k)})$, we estimate the activation center process intensity function by $K^{-1} \sum_{k=1}^{K} \lambda(y; \mathbf{z}^{(k)}, \boldsymbol{\Sigma}^{(k)})$, where $\lambda(y; \cdot)$ is defined by equation (9) in Section 2.2.

Let $\boldsymbol{\rho}^{(k)}$ be the posterior draw at the kth iteration after burn-in. To obtain a posterior predictive intensity $\tilde{\lambda}$ for a new study, given $\boldsymbol{\rho}^{(k)}$ and $(\mathbf{z}^{(k)}, \boldsymbol{\Sigma}^{(k)})$, we randomly pick $c^{(k)} \in \{1, \ldots, C\}$, then simulate a Poisson point process $\mathbf{y}_0^{(kl)}$ with intensity $\rho_{1c^{(k)}}^{(k)} \lambda(y; \mathbf{z}^{(k)}, \boldsymbol{\Sigma}^{(k)})$, for $l = 1, 2, \ldots, L$. For the kth iteration, $\tilde{\lambda}^{(k)}$ with a voxel v is estimated by

$$\tilde{\lambda}^{(k)}(y; \cdot) = \frac{\sum_{l=1}^{L} \sum_{y_0 \in \mathbf{y}^{(kl)}} I[y_0 \in v]}{L|v|}, \text{ for } y \in v,$$

Then $\tilde{\lambda}(y;)$ is estimated by $\sum_{k=1}^{K} \tilde{\lambda}^{(k)}(y;\cdot)/K$.

C.7 Credible Ellipsoid Computation

Conditional on the event that there is exactly one population center, $z \in \mathbf{Z}$, in the region of interest, the posterior distribution of z can be approximated by a normal distribution $N(\mu_z, \mathbf{\Lambda}_z)$, where $(\mu_z, \mathbf{\Lambda}_z)$ can be simulated via a mixture of Dirichlet process priors model based on the posterior sample of $\mathbf{Z}$. The 95% credible ellipsoid for the population centers is $\mathrm{CR}_z^p = \{x : (x-\mu_z)^{\mathrm{T}} \mathbf{\Lambda}_z^{-1} (x-\mu_z) \leq \chi^2_{0.95,3}\}$, where $\chi^2_{0.95,3}$ is the 0.95 quantile of χ^2 distribution with 3 degrees of freedom.

Conditional on the same event, the posterior distribution of an activation center $y \in (\mathbf{Y}_z = \cup_{c=1}^{C} \mathbf{Y}_{cz})$ that is associated with the population center z located in the region of interest, is a normal distribution with mean z and covariance matrix $\mathbf{\Sigma}_z$, which can be estimated using the posterior draws of the population center process. Thus, the 95% credible ellipsoid for the activation centers is $\mathrm{CR}_z^a = \{x : (x-z)^{\mathrm{T}} \mathbf{\Sigma}_z^{-1} (x-z) \leq \chi^2_{0.95,3}\}$.

C.8 Pseudo Code

Starting with an initial state of all the parameters, repeat the following steps for a pre-specified total number of iterations:

1. Update $\boldsymbol{\delta}$ according to (C.4).

2. For $c = 1, \ldots, C$, run the spatial birth-and-death process for study center processes.

   2.1. Compute B_c according to (C.8); set $\tau_c = 1/B_c$; set $t = 0$; set $(\mathbf{y}, \mathbf{\Psi})_c = (\{\emptyset\}, \{\emptyset\})$;

   2.2. Compute $d_c[\cdot, (y, \mathbf{\Psi}_y)]$ for all $(y, \mathbf{\Psi}_y) \in (\mathbf{y}, \mathbf{\Psi})_c$ and D_c according to (C.7) and (C.9).

   2.3. Draw $r \sim U[0,1]$. If $r < B_c/(B_c + D_c)$, then draw $(y, \mathbf{\Psi}_y)$ according to (C.11) and (C.12), and set $(\mathbf{y}, \mathbf{\Psi})_c = (\mathbf{y} \cup \{y\}, \mathbf{\Psi} \cup \{\mathbf{\Psi}_y\})_c$, else select $(y, \mathbf{\Psi}_y)$ from $(\mathbf{y}, \mathbf{\Psi})_c$ with probability $d_c[\cdot, (y, \mathbf{\Psi}_y)]/D_c$ and set $(\mathbf{y}, \mathbf{\Psi})_c = (\mathbf{y} \backslash \{y\}, \mathbf{\Psi} \backslash \{\mathbf{\Psi}_y\})_c$.

2.4. Draw a sojourn time s from an exponential distribution with rate $B_c + D_c$; set $t = t + s$, if $t < \tau_c$, then go to 2.2.

3. Update each $(y, \mathbf{\Psi}_y) \in (\mathbf{y}, \mathbf{\Psi})_c$, for $c = 1, \ldots, C$ based on (C.21) and (C.23).

4. Update η_c, for $c = 1, \ldots, C$ according to (C.33).

5. Update $\boldsymbol{\rho}$ from (C.35).

6. Run the spatial birth-and-death process for the posterior population center process:

   6.1. Compute B based on (C.16); set $\tau_c = 1/B$; set $t = 0$; set $(\mathbf{z}, \mathbf{\Sigma}) = (\{\emptyset\}, \{\emptyset\})$.

   6.2. Compute $d[\cdot, (z, \mathbf{\Sigma}_z)]$ for all $(z, \mathbf{\Sigma}_z) \in (\mathbf{z}, \mathbf{\Sigma})$ and D based on (C.15) and (C.17).

   6.3. Draw $r \sim U[0, 1]$. If $r < B/(B + D)$, then draw $(z, \mathbf{\Sigma}_z)$ according to (C.19) and (C.20), and set $(\mathbf{z}, \mathbf{\Sigma}) = (\mathbf{z} \cup \{z\}, \mathbf{\Sigma} \cup \{\mathbf{\Sigma}_z\})$, else select $(z, \mathbf{\Sigma}_z)$ from $(\mathbf{z}, \mathbf{\Sigma})$ with probability $d[\cdot, (z, \mathbf{\Sigma}_z)]/D$ and set $(\mathbf{z}, \mathbf{\Sigma}) = (\mathbf{z} \backslash \{z\}, \mathbf{\Sigma} \backslash \{\mathbf{\Sigma}_z\})$.

   6.4. Draw a sojourn time s from an exponential distribution with rate $B + D$; set $t = t + s$; if $t < \tau$, then go to 6.2.

7. Update each $(z, \mathbf{\Sigma}_z) \in (\mathbf{z}, \mathbf{\Sigma})$ based on (C.26) and (C.28).

8. Update β from (C.36).

9. Update $\boldsymbol{T}$ from (C.37).

D Simulation Studies

In this section, we conduct two simulation studies. In the first, we investigate the sensitivity of the prior parameter settings in our model for the emotion data. In the second, we study sensitivity of the model specification by simulating three data sets based on different models.

D.1 Sensitivity to Priors

Our primary interest is how the posterior inference (including the number, location and variability) of the population centers varies with different informative prior specifications. We keep the non-informative priors that $\boldsymbol{\rho} \sim D(0.5, \ldots, 0.5)$ and $\eta_c \sim G(0.001, 0.001)$ for any $c = 1, 2, \cdots, C$ and let $\beta|\mathcal{B}| \sim G(a_\beta, 0.001)$, which implies that $\text{var}(\beta|\mathcal{B}|)/E(\beta|\mathcal{B}|) = 1000$, thus the prior for β is relatively vague. We consider nine scenarios where scenario 1 has the same prior set up as in Section 3 in the manuscript, other scenarios vary the settings of $E[N_{\mathbf{Z}}(\mathcal{B})]$, $E[N_{\mathbf{A}}(\mathcal{B})]/E[N_{\mathbf{Y}}(\mathcal{B})]$, $E[N_{\mathbf{U}_c}(\mathcal{B})]$, $E[\boldsymbol{\Sigma}_z]$ and $E[\boldsymbol{\Psi}_y]$ (see Table 1 for a summary of the different scenarios), where $\mathbf{U}_c = \mathbf{Y}_c \cup \mathbf{X}_c^0$, is the activation centers in study c, $\mathbf{Y} = \cup_{c=1}^{C} \mathbf{U}_c$, represents the activation centers over all studies, and $\mathbf{A} = \cup_{c=1}^{C} \cup_{z \in \mathbf{z}} (\mathbf{X}_{cz} \cup \mathbf{Y}_{cz})$, denotes the activation centers that cluster at the population level.

Table 1: Sensitivity Analysis Priors.

Scenario	[a] $E[N_{\mathbf{Z}}(\mathcal{B})]$	[b] $E[N_{\mathbf{U}_c}(\mathcal{B}))]$	[c] $\frac{E[N_{\mathbf{A}}(\mathcal{B})]}{E[N_{\mathbf{Y}}(\mathcal{B})]}$	$E[\boldsymbol{\Sigma}_z]$	$E[\boldsymbol{\Psi}_y]$	$\boldsymbol{T}_0$	[d] θ	[d] $\epsilon\|\mathcal{B}\|$
1	30	5.0	0.80	$4\boldsymbol{I}$	$\boldsymbol{I}$	$0.8\boldsymbol{I}$	58.27	437.00
2	25	5.0	0.80	$4\boldsymbol{I}$	$\boldsymbol{I}$	$0.8\boldsymbol{I}$	69.92	437.00
3	35	5.0	0.80	$4\boldsymbol{I}$	$\boldsymbol{I}$	$0.8\boldsymbol{I}$	49.94	437.00
4	30	5.0	0.75	$4\boldsymbol{I}$	$\boldsymbol{I}$	$0.8\boldsymbol{I}$	54.62	546.25
5	30	5.0	0.85	$4\boldsymbol{I}$	$\boldsymbol{I}$	$0.8\boldsymbol{I}$	61.91	327.75
6	30	5.0	0.80	$8\boldsymbol{I}$	$\boldsymbol{I}$	$1.6\boldsymbol{I}$	58.27	437.00
7	30	5.5	0.80	$4\boldsymbol{I}$	$\boldsymbol{I}$	$0.8\boldsymbol{I}$	64.09	480.70
8	30	4.5	0.80	$4\boldsymbol{I}$	$\boldsymbol{I}$	$0.8\boldsymbol{I}$	52.44	393.30
9	30	5.0	0.80	$4\boldsymbol{I}$	$2\boldsymbol{I}$	$0.8\ \boldsymbol{I}$	58.27	437.00

[a] A priori expected number of population centers.
[b] A priori expected number of activation centers per study.
[c] A priori proportion of act. centers that cluster about a population center.
[d] Derived values from footnotes *a*, *b* and *c*.

We simulate the posterior distribution with 20,000 iterations after a burn-in of 2,000 iterations. Table 2 shows descriptive statistics on the posterior distribution of the number of population centers after burn-in. As anticipated, scenarios 2 and 7 result in a decrease in the posterior mode of the number of the population centers. This is because scenario 2 has

a smaller prior mean number of population centers while scenario 7 has a larger prior mean number of activation centers that clusters about population centers. However, the change is not dramatic. The posterior mode of the number of population centers for all nine scenarios is around 42, with a range from 36 to 49.

Table 2: The summary statistics on the posterior number of population centers

Scenario	Min	Max	Mean	s.d.	Mode
1	38	47	42.60	1.4	42
2	32	40	35.90	1.2	36
3	44	55	49.10	1.5	49
4	37	48	42.40	1.4	42
5	38	47	42.50	1.4	42
6	37	49	42.60	1.5	43
7	34	42	37.60	1.3	37
8	43	55	48.30	1.6	48
9	36	45	40.10	1.3	40

In Figure 1 we compare sensitivity of the activation center posterior intensity, and in Figure 2 we compare sensitivity of the population center posterior intensity. Figure 3 compares the marginal credible ellipses for both population centers and activation centers. From these figures, we see that the posterior expected intensities and marginal credible ellipses are qualitatively quite similar. Table 3 summarizes the estimated location (x, y, z), the Euclidean distance in locations between scenario 1 and other scenarios, the volume of the 95% credible ellipsoid and the volume of the 95% credible interval of population centers as well as activation centers conditional on the population centers located in the amygdalae. The estimated number of population centers and the estimated volume of the credible ellipsoids is somewhat sensitive to the choice of prior on β, θ and ϵ, but the posterior intensity and the estimated location of the population centers are stable to various prior specifications. The maximum distance in credible ellipsoid locations for population center between scenario 1 and other scenarios is around 1mm (0.5 voxel).

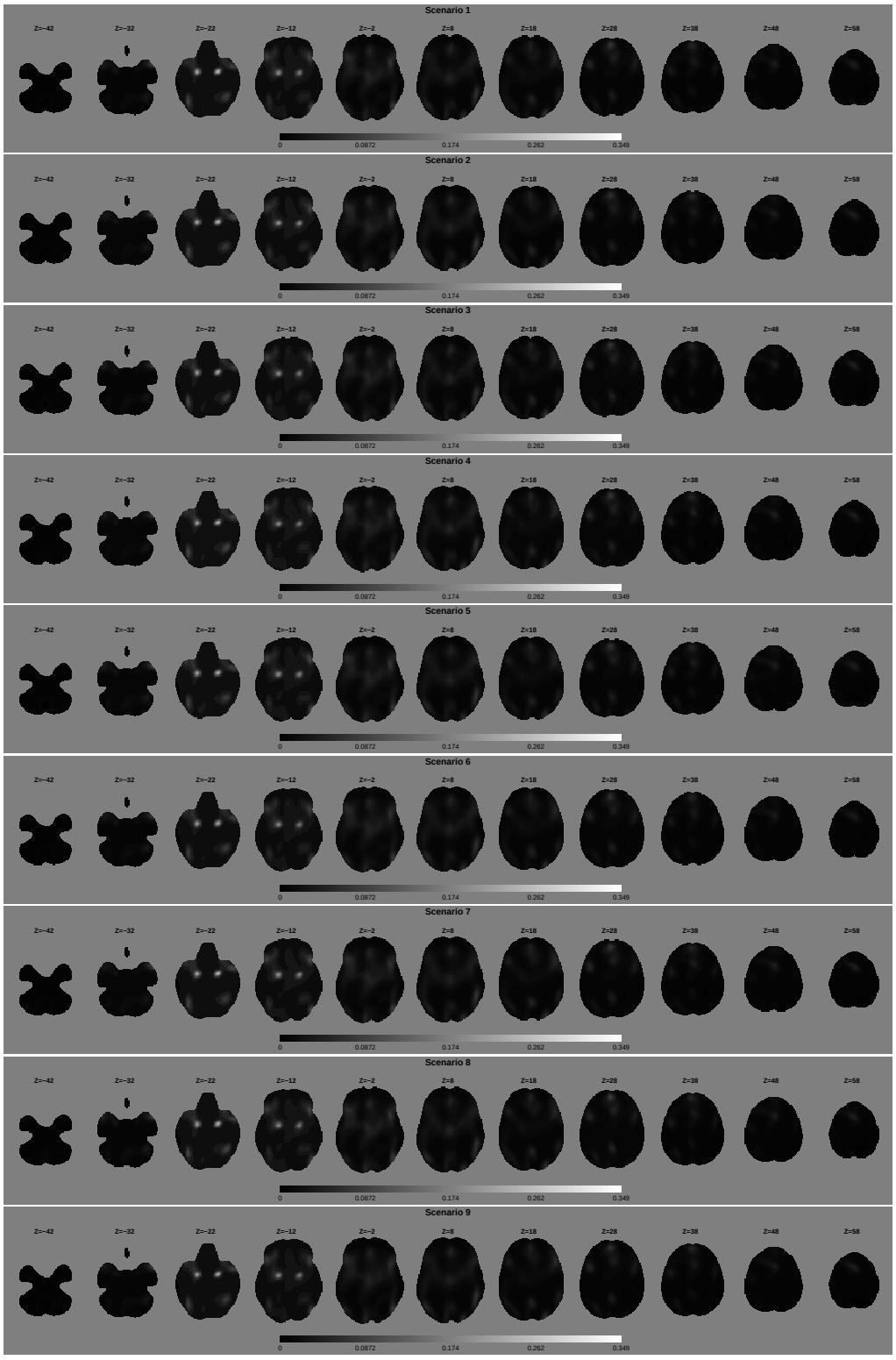

Figure 1: Sensitivity analysis results: comparisons of the posterior activation center intensity on 11 slices of the brain from $Z = -42$mm to $Z = 58$mm.

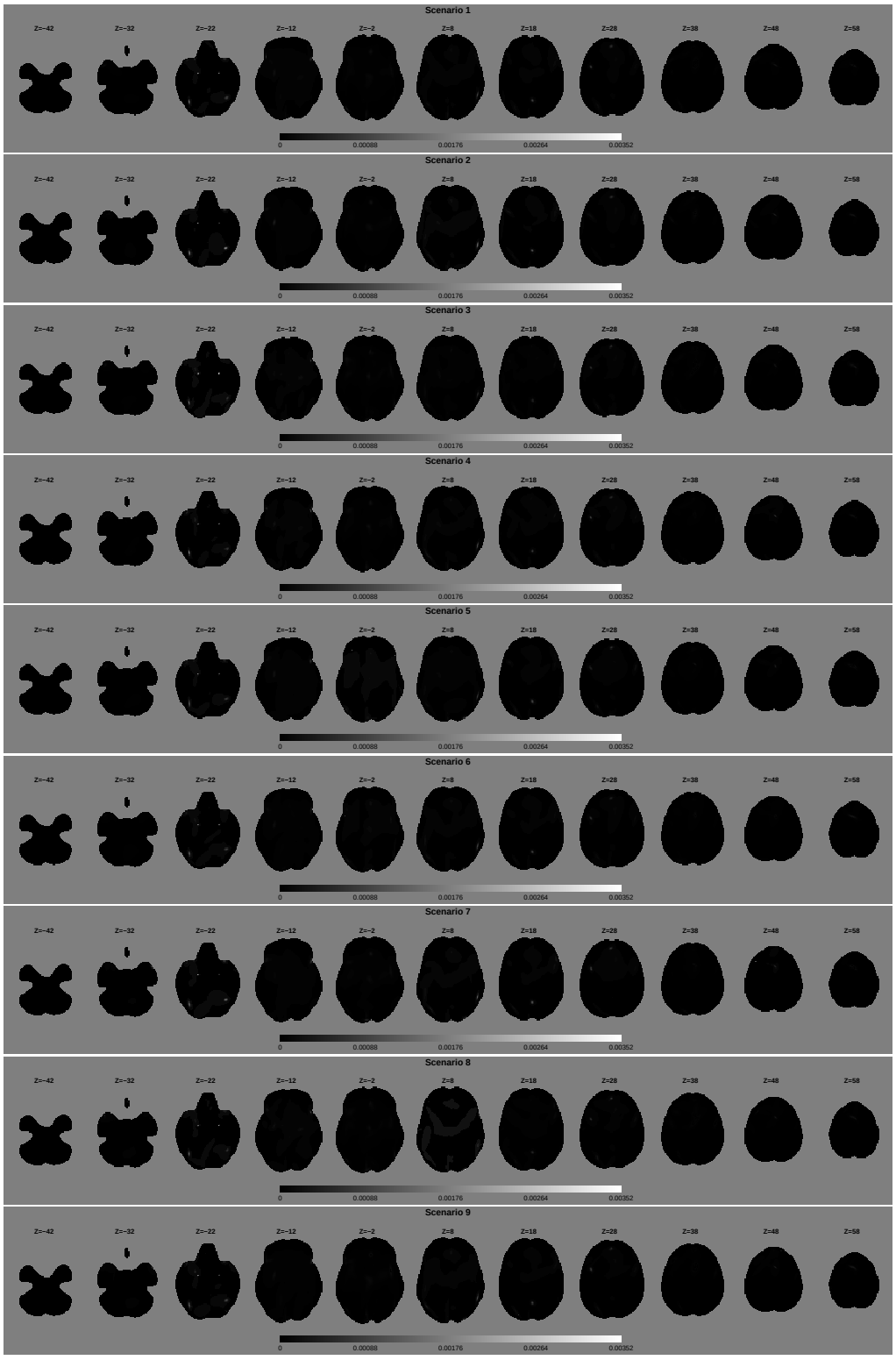

Figure 2: Sensitivity analysis results: comparisons of the population center intensity on 11 slices of the brain from $Z = -42$mm to $Z = 58$mm.

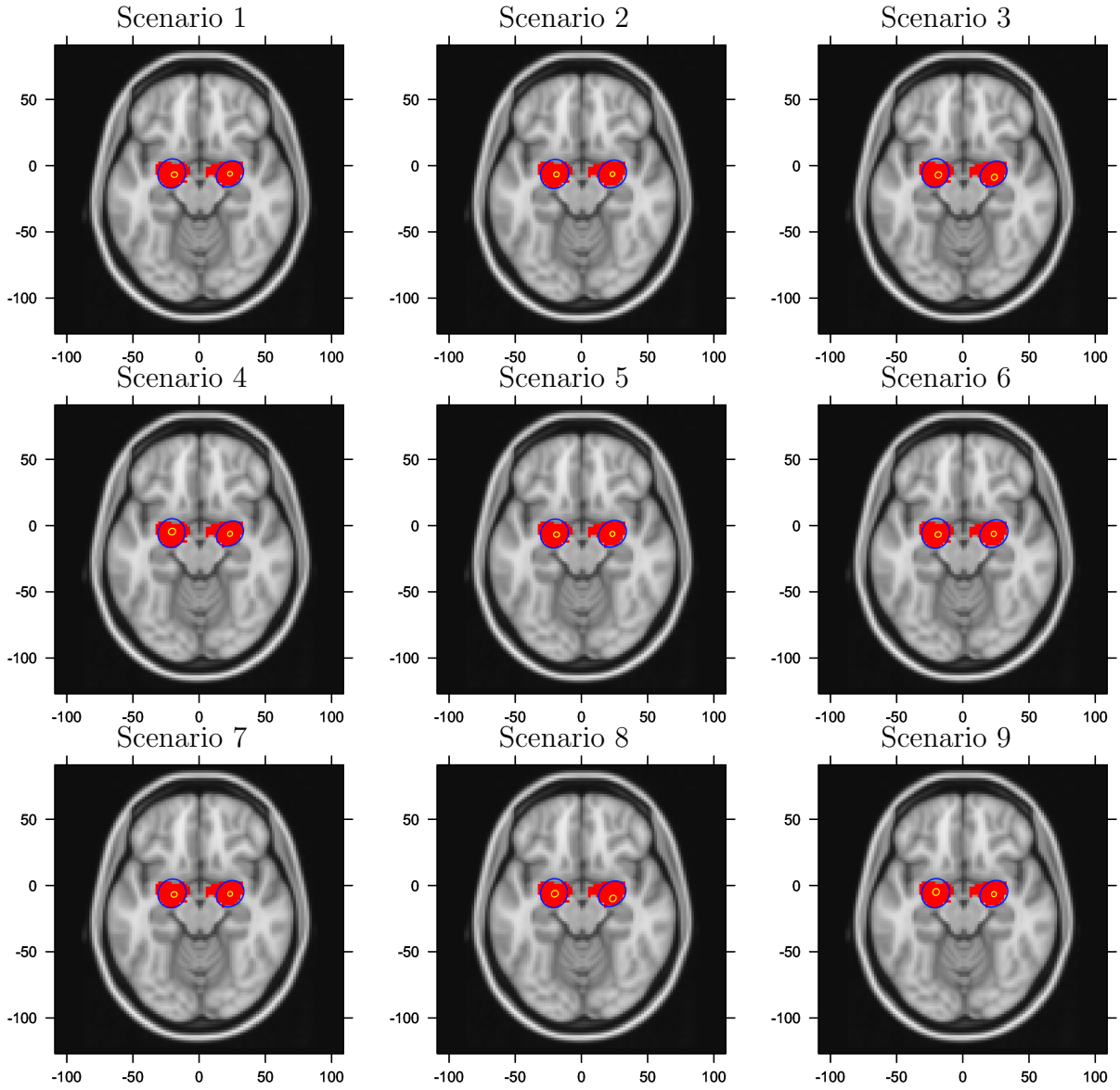

Figure 3: Sensitivity analysis results: comparisons of the 95% marginal credible ellipses of both population centers (blue circles) and activation centers (yellow circles).

Table 3: The estimated location (x, y, z), the Euclidean distance in locations between scenario 1 and other scenarios (Dist.), the volume of 95% credible ellipsoid and the volume 95% credible interval of population center as well as activation centers conditional on the population centers located in the amygdalae.

Scenario	Amygdala Region	Location (x, y, z)	Dist.	Volume	95% CI
1	L. Pop. Ctr.	(-18.8, -6.9,-19.8)	0.00	86.80	[80.1, 94.1]
	L. Act. Ctr.	(-20.8, -6.0,-19.0)	0.00	12885.80	[6093.0, 47684.1]
	R. Pop. Ctr.	(23.2, -6.2,-20.3)	0.00	35.70	[33.1, 38.6]
	R. Act. Ctr.	(23.2, -6.3,-19.9)	0.00	8299.00	[4623.9, 13586.7]
2	L. Pop. Ctr.	(-19.0, -6.6,-19.8)	0.30	60.00	[55.4, 64.8]
	L. Act. Ctr.	(-20.3, -6.4,-19.1)	0.60	12582.80	[6736.1, 41708.8]
	R. Pop. Ctr.	(23.4, -6.5,-20.0)	0.40	35.10	[32.5, 37.9]
	R. Act. Ctr.	(23.3, -6.4,-19.8)	0.10	9118.20	[5332.8, 13965.5]
3	L. Pop. Ctr.	(-18.5, -7.0,-19.8)	0.30	123.90	[110.7, 142.5]
	L. Act. Ctr.	(-20.8, -5.6,-19.0)	0.40	11782.70	[5488.1, 35089.0]
	R. Pop. Ctr.	(23.7, -6.8,-20.3)	0.80	71.60	[65.0, 78.6]
	R. Act. Ctr.	(23.1, -6.3,-19.7)	0.20	8217.70	[4272.6, 14578.1]
4	L. Pop. Ctr.	(-18.5, -6.6,-20.3)	0.60	160.80	[91.4, 193.0]
	L. Act. Ctr.	(-21.0, -5.8,-19.1)	0.30	12420.30	[5881.2, 44945.2]
	R. Pop. Ctr.	(23.2, -6.3,-20.2)	0.20	41.50	[37.2, 63.4]
	R. Act. Ctr.	(23.2, -6.2,-19.9)	0.10	8229.20	[4449.7, 14538.9]
5	L. Pop. Ctr.	(-18.9, -6.8,-19.8)	0.10	75.40	[69.4, 82.0]
	L. Act. Ctr.	(-20.5, -6.1,-19.2)	0.40	12651.90	[6244.3, 43772.6]
	R. Pop. Ctr.	(23.3, -6.4,-20.0)	0.40	42.80	[37.7, 103.2]
	R. Act. Ctr.	(23.3, -6.3,-19.7)	0.20	8929.50	[4907.4, 14888.7]
6	L. Pop. Ctr.	(-18.7, -6.8,-19.7)	0.20	88.20	[81.1, 95.9]
	L. Act. Ctr.	(-20.6, -6.0,-19.1)	0.20	13065.90	[6560.5, 41233.8]
	R. Pop. Ctr.	(23.3, -6.3,-20.2)	0.20	44.40	[41.1, 48.0]
	R. Act. Ctr.	(23.3, -6.2,-19.8)	0.20	9638.40	[5371.7, 16713.5]
7	L. Pop. Ctr.	(-19.0, -6.8,-19.8)	0.20	83.10	[76.4, 90.2]
	L. Act. Ctr.	(-20.6, -6.0,-19.1)	0.20	12590.50	[6323.7, 41233.9]
	R. Pop. Ctr.	(23.3, -6.4,-20.1)	0.30	33.80	[31.4, 36.5]
	R. Act. Ctr.	(23.3, -6.3,-19.9)	0.10	8043.00	[4966.5, 12771.3]
8	L. Pop. Ctr.	(-18.2, -6.3,-19.5)	0.90	147.90	[119.0, 252.6]
	L. Act. Ctr.	(-20.8, -5.8,-19.0)	0.10	12408.30	[5566.6, 42291.3]
	R. Pop. Ctr.	(23.6, -6.8,-19.3)	1.20	86.60	[78.1, 95.0]
	R. Act. Ctr.	(23.1, -6.3,-19.7)	0.20	8537.70	[4473.1, 13830.1]
9	L. Pop. Ctr.	(-18.3, -7.0,-20.3)	0.60	133.50	[98.8, 151.5]
	L. Act. Ctr.	(-20.7, -5.9,-19.2)	0.20	13087.00	[6236.0, 44717.0]
	R. Pop. Ctr.	(23.4, -6.6,-19.8)	0.70	42.60	[36.6, 65.4]
	R. Act. Ctr.	(23.3, -6.3,-19.8)	0.20	8358.00	[4868.9, 13905.9]

D.2 Sensitivity to Model Specification

Simulation A: We simulate data, or foci, $\mathbf{x}$, the centers of activation regions, $\mathbf{y}$, and the population centers from our model with the following parameters. The mean number of population centers is $\beta|\mathcal{B}| = 5$ and the number of studies is $C = 15$. For each study, we set the mean number of study centers associated with the population centers as $(1-\sum_{c=1}^{C}\rho_{1c})\theta/C = 1/15$, where $\rho_{1c} = 8/125$ and $\theta = 50$. The mean number of foci, per population center, that do not cluster with a study center is $\sum_{c=1}^{C}\rho_{1c}\theta/C = 4/3$. The mean number of the multiple foci per activation region is $\eta_c = 5$, for all $c = 1, \ldots, C$. The mean number of activation centers that do not cluster is $\epsilon|\mathcal{B}| = 30$. The covariance matrix that describes the variability of activation centers about population centers is set to $\mathbf{\Sigma}_z = 9\boldsymbol{I}$ for all $z \in \mathbf{z}$. The covariance matrix for multiply reported foci that cluster about study centers is $\mathbf{\Psi}_y = \boldsymbol{I}$ for all $y \in \mathbf{y}$. Based on the above settings, the mean number of observed foci is 315. The simulated data are shown in Figure 4. The hyper prior values are $E[\beta|\mathcal{B}|] = 2$, $\theta = 100$, $\epsilon|\mathcal{B}| = 50$, $E(\eta_c) = 10$, $E(\rho_{1c}) = 0.03$ for all $c = 1, \cdots, C$, $E[\mathbf{\Sigma}_z] = 4\boldsymbol{I}$, for all $z \in \mathbf{z}$ and $E(\mathbf{\Psi}_y) = \boldsymbol{I}$ for all $y \in \mathbf{y}$.

Results: The posterior mean number of population centers is 5. The estimated posterior marginal intensity function of the activation centers is shown in Figure 4 from which we can identify the 5 clusters. Also, we can see that the data and the intensity are well matched. The estimated posterior marginal intensity function of the population centers is also shown in Figure 4. Clearly, the intensity is highly concentrated around the 5 true population locations.

We conclude that if the data are generated from our model, then our method provides very accurate results even when then priors are biased from the truth. Next, we investigate how the proposed method is robust to model mis-specification.

Simulation B: We set population centers $\mathbf{z} = \{z_1, z_2\}$, where $z_1 = (22, -6, -18)$ and $z_2 = (-20, -6, -18)$, i.e. the centers of the amygdalae. For each population center $z_i \in \mathbf{z}, i = 1, 2$, we draw 50 foci and 5 study centers from a uniform distribution over spheres

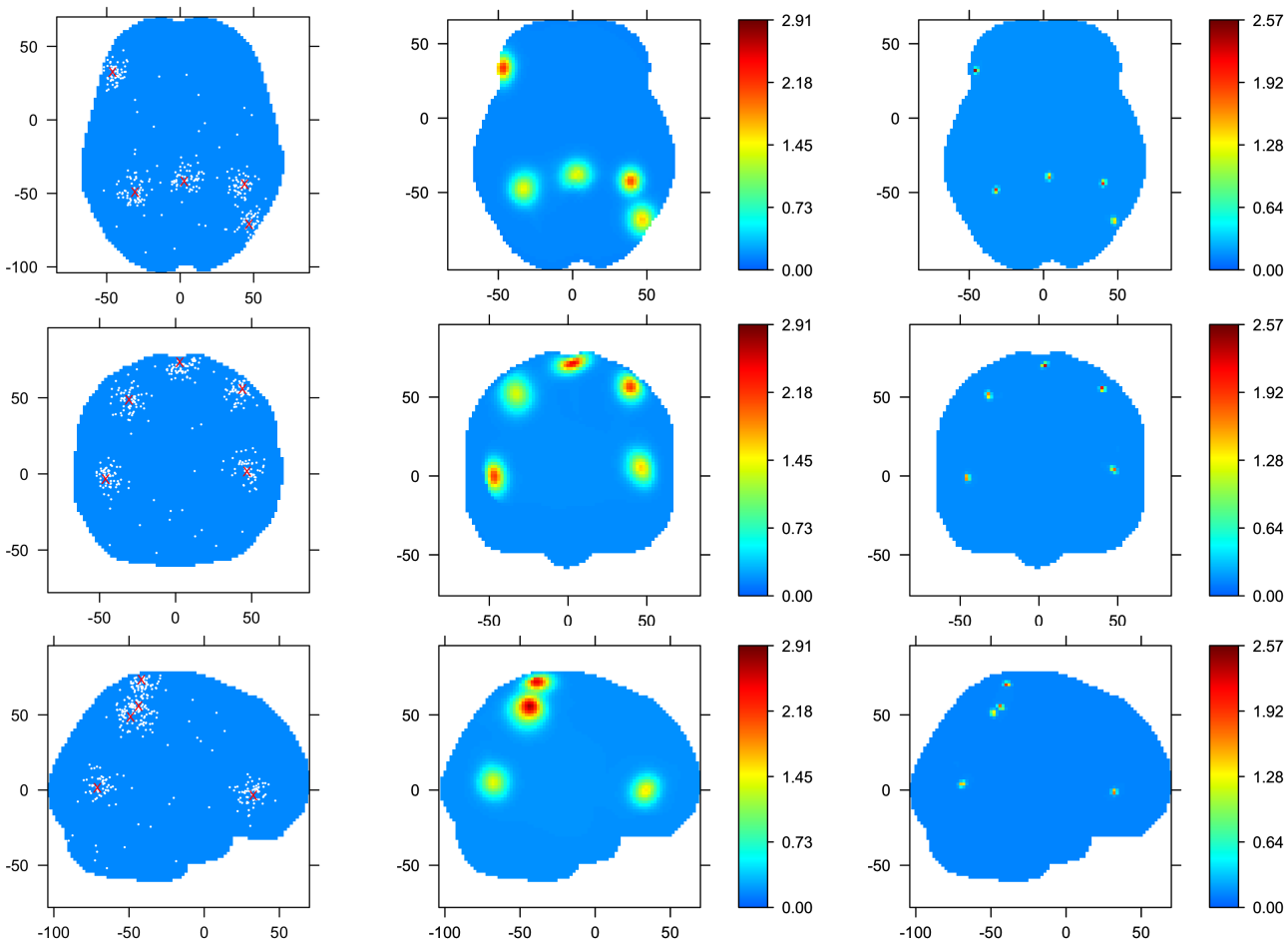

Figure 4: Simulation A results: the images in the first column show the simulated foci (white points) and the true locations of population centers (red X), projected unto the XY, XZ and YZ plane for the top, middle and bottom, respectively. The images in the middle column show the posterior intensity of the activation centers, integrated over the Z, Y and X directions for the top, middle and bottom, respectively. The images in the last column show the posterior intensity of the population centers, each integrated over one dimension as in the middle column.

with centers z_i and radius $R_{\mathbf{\Sigma}} = 5$. For each activation center, $y_j \in \mathbf{y}, j = 1, 2, \cdots, 10$, we draw 10 foci from a uniform distribution over spheres with centers y_j, and radius $R_\Lambda = 1$. The uniform distributions are in direct contrast with the Gaussian distributions assumed in our model. We set the number of studies to $C = 5$ and randomly assign each study two activation centers and 40 foci, for a total of 200 foci. Hyper prior values are $E(\beta|\mathcal{B}|) = 2$, $\epsilon|\mathcal{B}| = 60$, $\theta = 100$, $E(\mathbf{\Sigma}_z) = 4\boldsymbol{I}$ for all $z \in \mathbf{z}$, $E(\mathbf{\Psi}_y) = \boldsymbol{I}$ for all $y \in \mathbf{y}$, $E(\eta_c) = 10$ and $E(\rho_{1c}) = 0.033$ for all $c = 1, 2, \ldots, C$.

Results are shown in Figure 5. The estimated posterior marginal intensity of activation centers clearly show there are two activation regions and match the truth well. Also, in Figure 5, the estimated posterior marginal intensity of the population centers is highly concentrated on the two points. The above results imply that the proposed model is robust to this model mis-specification.

Simulation C: For this simulation, we do not set population centers and study centers. Rather, we directly simulate foci, $\mathbf{x} = \{x_1, \cdots, x_{350}\}$. For $i = 1, 2, \cdots, 300$, we simulate them from the following function:

$$x_i = x_0 + r \begin{pmatrix} \sin(\psi_i)\cos(\varphi_i) \\ \sin(\psi_i)\sin(\varphi_i) \\ \cos(\psi_i) \end{pmatrix} \tag{D.1}$$

where $x_0 = (46, 55, 46)$, $\psi_i \sim \mathrm{U}(-0.5\pi, 0.5\pi)$, $\varphi_i \sim \mathrm{U}(0.25\pi, 0.5\pi)$ and $r \sim \mathrm{U}(20, 25)$. For $i = 301, \cdots, 350$, the x_i are drawn uniformly over the brain and are considered as noise. We then randomly assign the 350 foci to 20 studies. The hyper-prior values are $E(\beta|\mathcal{B}|) = 5$, $\epsilon|\mathcal{B}| = 60$, $\theta = 50$, $E(\mathbf{\Sigma}_z) = 100\boldsymbol{I}$ for all $z \in \mathbf{z}$, $E(\mathbf{\Psi}_y) = \boldsymbol{I}$ for all $y \in \mathbf{Y}$, $E(\eta_c) = 10$ and $E(\rho_{1c}) = 0.033$ for all $c = 1, 2, \ldots, C$.

Figure 6 compares the estimated posterior marginal intensity of activation centers with the data. Figure 6 also compares the posterior intensity of the "population centers"—contrasted against the data as there are no true population centers. Our model picks about 7 population centers. Obviously, in this case, the "population centers" are driven by the data—our model clusters the data about these "population centers" although the data generating

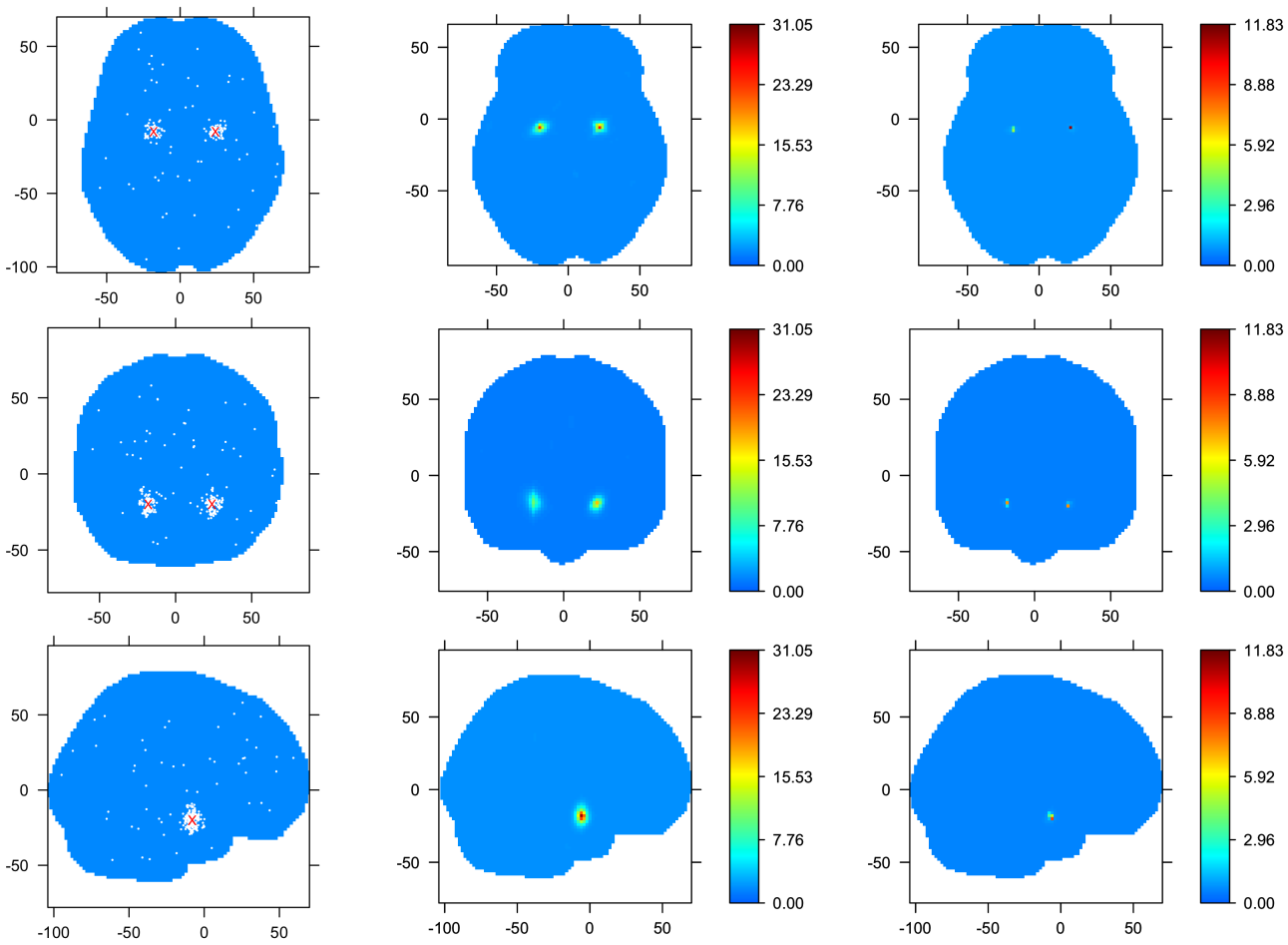

Figure 5: Simulation B results: The first column shows the simulated foci and the true locations of population centers. The middle column show the posterior intensity of the activation centers, and the last column shows the posterior intensity of the population centers. See Figure 4 for display conventions.

mechanism assumes no such centers. The posterior intensity of the activation centers is still well estimated, however, care must be taken in the interpretation of the population level parameters. They exist solely to fit the data.

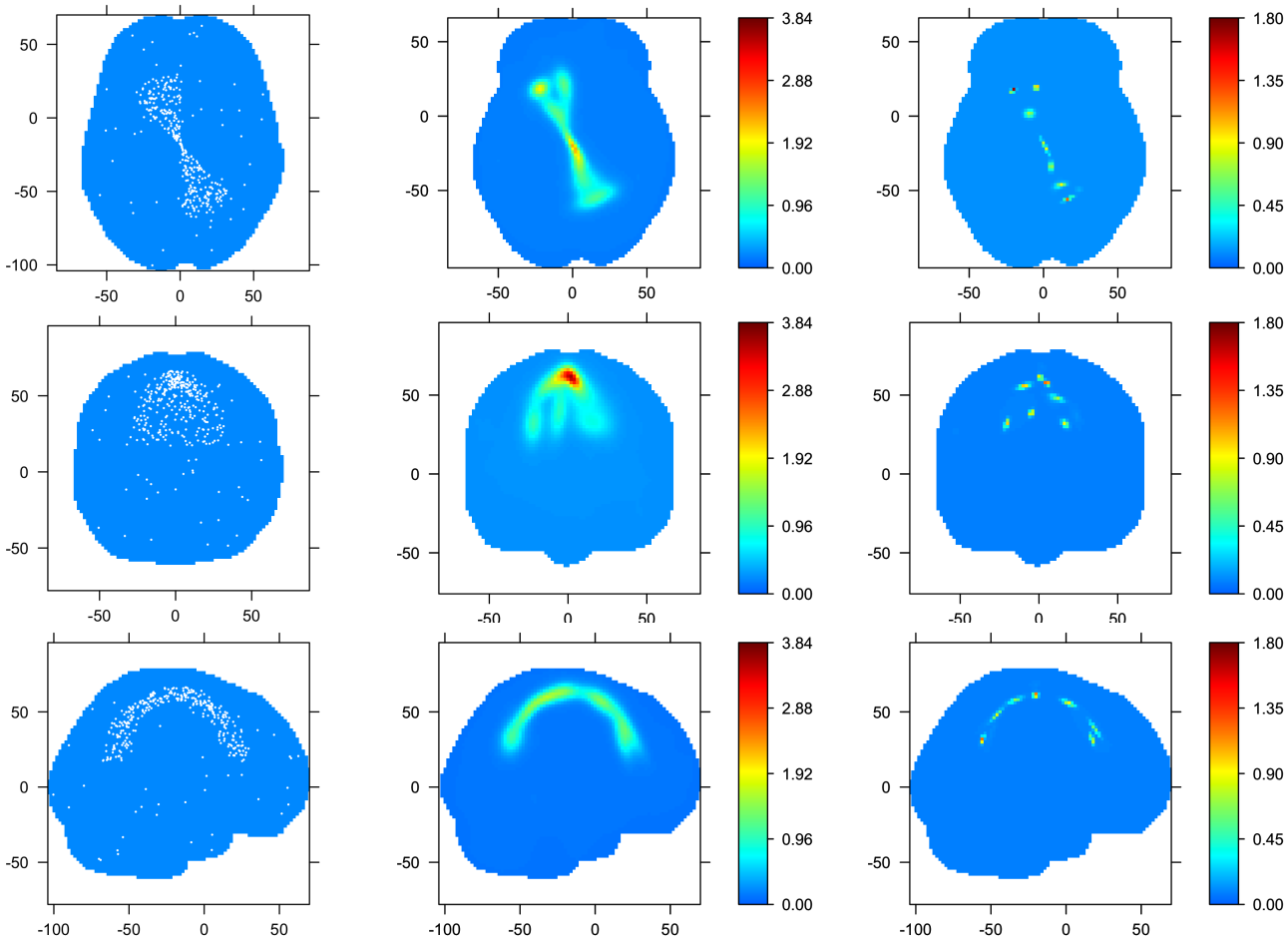

Figure 6: Simulation C results. The first column shows the simulated foci. The middle column shows the posterior intensity of activation centers, and the last columns shows the posterior intensity of the "population centers". See Figure 4 for display conventions.

E Whole Brain Results

We emphasize here that the model fits intensity functions to the entire brain and not only on ROIs. ROIs, however, are a useful way to summarize the posterior intensity functions. Table 4 summarizes the posterior mean number of population centers occurring in various ROI as well as the posterior probability that at least one population center occurs in said ROI. Note that most regions are bilateral, specifically the amygdala consists of a pair of symmetric regions. Most ROIs have strong evidence for one or more population centers. ROI's are from the CIC Atlas (Tziortzi et al. 2010), a revised and hierarchical version of the Harvard-Oxford atlas (Desikan et al. 2006).

Table 4: The posterior expected number of population centers in different ROIs and the posterior probability that at least one population centers is located in the ROI.

Region ($\mathcal{R}$)		$\mathrm{E}[N_{\mathbf{Z}}(\mathcal{R}) \mid \mathbf{X}]$	$\Pr[N_{\mathbf{Z}}(\mathcal{R}) \geq 1 \mid \mathbf{X}]$
Occipital Lobe	Occipital Pole	4.02	>0.999
	Calcarine Cortex	2.00	>0.999
	Cuneus	0.00	0.000
	Lingual Gyrus	1.00	>0.999
	Occipital Fusiform Gyrus	1.00	>0.999
Insular & Temporal Lobe	Insular Cortex	1.00	>0.999
	Anterior Temporal Pole	2.71	>0.999
	Superior Temporal Gyrus	0.41	0.412
	Middle Temporal Gyrus	1.00	>0.999
	Inferior Temporal Gyrus	0.00	0.000
	Parahippocampal Gyrus	0.00	0.000
	Temporal Fusiform Gyrus	1.00	>0.999
	Amygdala	2.05	>0.999
	Hippocampus	0.00	0.000
Frontal Lobe	Precentral Gyrus	0.00	0.000
	DorsoLateral Frontal Cortex	5.00	>0.999
	Medial Frontal Cortex	2.16	>0.999
	Frontal Operculum	0.98	0.982
	Orbitofrontal	4.00	>0.999
	Supplementary Motor Area	0.00	0.000
Cingulate Cortex	Posterior Cingulate Gyrus	0.00	0.000
	Anterior Cingulate	2.00	>0.999
Parietal Lobe	Postcentral Gyrus	0.00	0.000

	Parietal Lobule	0.00	0.000
	Supramarginal Gyrus	0.00	0.000
	Angular Gyrus	0.00	0.000
	Precuneous Cortex	0.00	0.000
	Parietal Operculum Cortex	0.00	0.000
Basal Ganglia	Globus Pallidus	0.00	0.000
	Striatum	1.00	>0.999
Thalamus & Brainstem	Thalamus	0.91	0.913
	Midbrain	0.00	0.000
	Pons	0.00	0.000
	Medulla	0.00	0.000
Cerebellum	Ventrolateral Cerebellum	0.00	0.000
	Medial Cerebellum	0.00	0.000
	Dorsal Cerebellum	2.00	>0.999
Other	Cerebral White Matter	6.00	>0.999
	Lateral Ventricle	0.00	0.000

We also show the posterior standard deviation of activation center intensity function in Figure 7. Each of 11 slices is a 2mm thick slice, and intensity functions shown in each slice are integrated over the 2mm slice. The axial slices are separated by 10 mm.

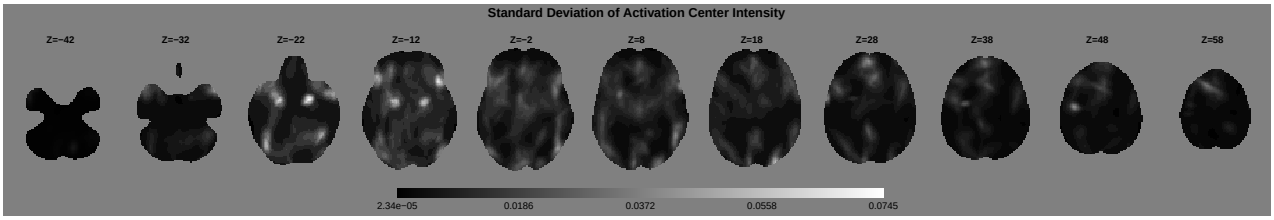

Figure 7: The posterior standard deviation of the activation center intensity on 11 axial slices of the brain (from $Z = -42$mm to $Z = 58$mm).

References

Cox, D. R. (1955), "Some statistical models related with series of events," *Journal of the Royal Statistical Society Series B*, 17, 129–164.

Desikan, R., Sgonne, F., Fischl, B., Quinn, B., Dickerson, B., Blacker, D., Buckner, R., Dale, A., Maguire, R., Hyman, B., Albert, M., and Killiany, R. (2006), "An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest," *Neuroimage*, 31, 968–980.

Escobar, M. D. and West, M. (1995), "Bayesian density estimation and inference using mixtures," *Journal of the American Statistical Association*, 90, 577–588.

Møller, J. and Waagepetersen, R. P. (2004), *Statistical Inference and Simulation for Spatial Point Processes*, Chapman and Hall/CRC.

Preston, C. J. (1975), "Spatial birth-and-death processes," *Bulletin of the International Statistical Institute*, 46, 371–391.

Tziortzi, A. C., Searle, G. E., Tzimopoulou, S., Salinas, C., Beaver, J. D., Jenkinson, M., Laruelle, M., Rabiner, E. A., and Gunn, R. N. (2010), "Imaging dopamine receptors in humans with [(11)C]-(+)-PHNO: Dissection of D3 signal and anatomy." *NeuroImage*.

van Lieshout, M. N. M. and Baddeley, A. J. (2002), "Extrapolating and interpolating spatial patterns," in *Spatial Cluster Modelling*, eds. Lawson, A. B. and Denison, D. G. T., Chapman & Hall/CRC, chap. 4, pp. 61–86.