

# Bayesian models for functional magnetic resonance imaging data analysis

Linlin Zhang,<sup>1</sup> Michele Guindani<sup>2</sup> and Marina Vannucci<sup>1\*</sup>

Functional magnetic resonance imaging (fMRI), a noninvasive neuroimaging method that provides an indirect measure of neuronal activity by detecting blood flow changes, has experienced an explosive growth in the past years. Statistical methods play a crucial role in understanding and analyzing fMRI data. Bayesian approaches, in particular, have shown great promise in applications. A remarkable feature of fully Bayesian approaches is that they allow a flexible modeling of spatial and temporal correlations in the data. This article provides a review of the most relevant models developed in recent years. We divide methods according to the objective of the analysis. We start from spatiotemporal models for fMRI data that detect task-related activation patterns. We then address the very important problem of estimating brain connectivity. We also touch upon methods that focus on making predictions of an individual's brain activity or a clinical or behavioral response. We conclude with a discussion of recent integrative models that aim at combining fMRI data with other imaging modalities, such as electroencephalography/magnetoencephalography (EEG/MEG) and diffusion tensor imaging (DTI) data, measured on the same subjects. We also briefly discuss the emerging field of imaging genetics. © 2014 Wiley Periodicals, Inc.

#### How to cite this article:

*WIREs Comput Stat* 2015, 7:21–41. doi: 10.1002/wics.1339

**Keywords:** Bayesian statistics; brain connectivity; classification and prediction; fMRI; spatiotemporal activation models

## INTRODUCTION

Functional magnetic resonance imaging (fMRI) is a noninvasive neuroimaging method that measures blood-oxygen level-dependent (BOLD) signals in the brain *in vivo*. Neural activity is associated with localized changes in metabolism. As a brain area becomes active, e.g., in response to a task, there is an increase in local oxygen consumption and, consequently, more oxygen-rich blood flows to the active brain area.

Thus, activated brain areas show a relative increase in oxyhemoglobin and a relative decrease in deoxyhemoglobin, as the increased supply of oxygen outpaces the increased demand for it. BOLD signals measure metabolic activity in the brain as the difference between the oxyhemoglobin and deoxyhemoglobin levels arising from changes in local blood flow.

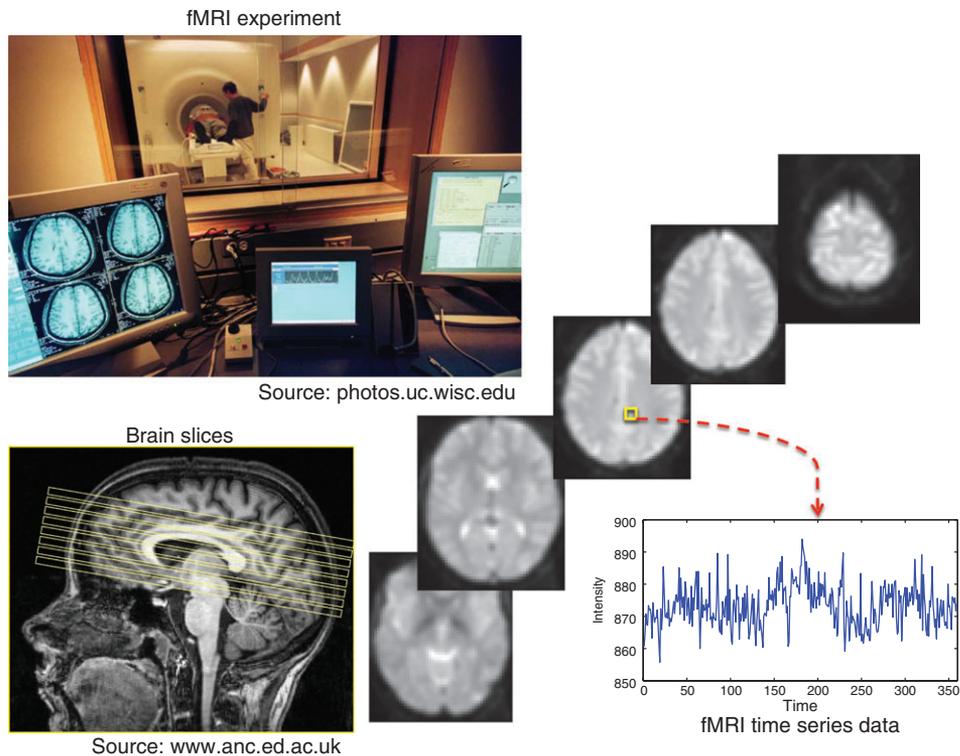
Figure 1 illustrates a typical fMRI experiment. Distributed three-dimensional (3D) maps of localized brain activity are measured over time while the subject lies in the MRI scanner. Scans are typically acquired every 2–3 seconds, with each scan arranged in a 3D array of volume elements (or 'voxels'). Time series of BOLD responses are produced at every voxel, as the temporal evolution of brain activity at that location. An fMRI experiment on a single subject can yield hundreds of scans in a single session. fMRI data are

\*Correspondence to: marina@rice.edu

<sup>1</sup>Department of Statistics, Rice University, Houston, TX, USA

<sup>2</sup>Department of Biostatistics, UT M.D. Anderson Cancer Center, Houston, TX, USA

Conflict of interest: The authors have declared no conflicts of interest for this article.



**FIGURE 1** | Typical fMRI experiment: 3D maps are acquired over time while the subject lies in the scanner, producing time series of fMRI BOLD responses measured at each brain voxel. Selected 2D arrays, corresponding to axial slices across the third dimension, are shown.

therefore massive collection of hundreds of thousands of time series, arising from spatially distinct locations.

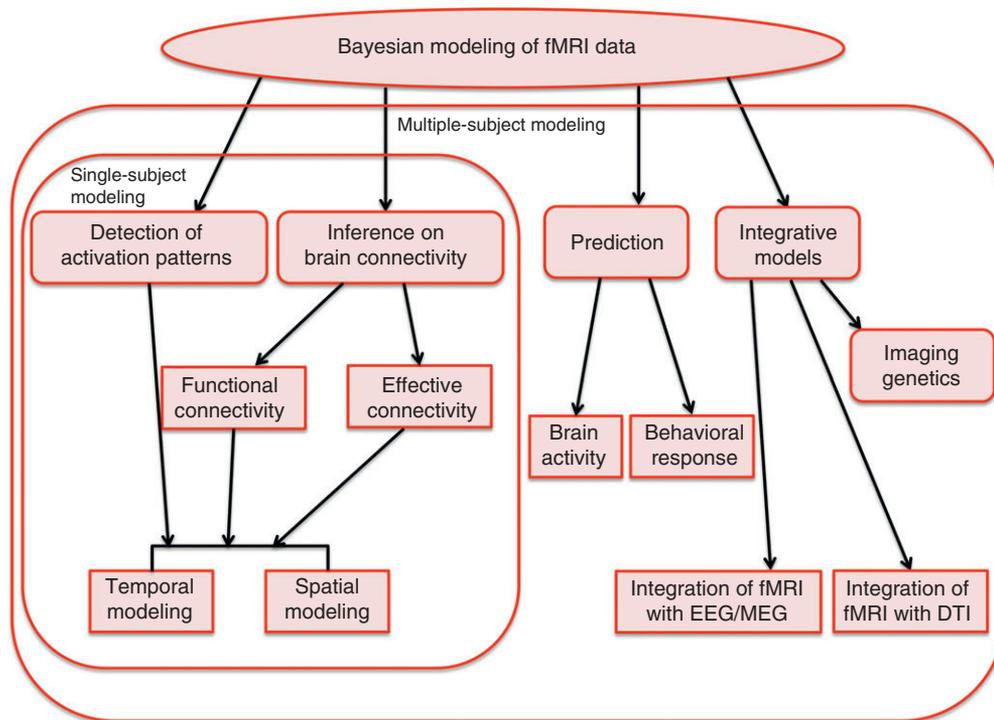
Statistical methods play a crucial role in the analysis of fMRI data,<sup>1–4</sup> due to the complex spatial and temporal correlation structure of the data, as well as their large dimensionality. Early approaches to the analysis of such data would calculate voxel-wise t-test or ANOVA statistics and/or fit a linear model at each voxel after a series of preprocessing steps, including scanner drift and motion corrections, adjustment for cardiac and respiratory-related noise, normalization and spatial smoothing, see Huettel et al.,<sup>5</sup> Chapter 10. However, spatial correlation is expected in a voxel-level analysis of fMRI data because the response at a particular voxel is likely to be similar to the responses of neighboring voxels. Also, correlation among voxels does not necessarily decay with distance. All this makes single-voxel approaches not appropriate, as the test statistics across voxels are not independent. In addition, serious multiplicity issues arise, due to the large dimensionality of the data.

This article provides a review of the most relevant Bayesian modeling approaches to fMRI data analysis that have been developed in recent years. Bayesian approaches have a great potential in applications as they allow a flexible modeling of spatial and temporal correlations in the data.<sup>6</sup> We divide methods

according to the objective of the analysis. We start from spatiotemporal models that estimate task-related activation patterns. In a typical task-related fMRI experiment, the whole brain is scanned at multiple times while a subject performs a series of tasks. The objective of the analysis is then to detect those brain regions that get activated by the external stimulus. We discuss general linear and nonlinear models, as well as mixture models, for both single- and multiple-subject studies.

Another important task in fMRI studies, which has received increased interest in recent years, is to infer brain connectivity. In general terms, connectivity looks at how brain regions interact with each other and how information is transmitted between them, with the aim of uncovering the actual mechanisms of how our brain functions. We discuss Bayesian approaches for both functional (undirected) and effective (directed) connectivity, as first defined by Friston.<sup>7</sup> Functional connectivity studies seek to identify multiple brain areas that exhibit similar temporal profiles, either task-related or at rest, while effective connectivity seeks to estimate the directed influence of one brain region on another.

In the last part of our review, we touch upon methods that focus on making predictions of an individual's brain activity or a clinical or behavioral



**FIGURE 2** | Outline of the Bayesian methods for fMRI data reviewed in this article. Methods are divided according to the objective of the analysis.

response and then conclude the article with a discussion of recent integrative models that aim at combining fMRI data with other imaging modalities, such as electroencephalography/magnetoencephalography (EEG/MEG) and diffusion tensor imaging (DTI) data, measured on the same subjects. We also briefly discuss the emerging field of imaging genetics. Figure 2 shows an outline of our review.

## DETECTION OF ACTIVATED BRAIN REGIONS

Detecting brain regions activated by an external stimulus or condition is probably the most common objective in fMRI studies. Neuronal activation in response to a stimulus occurs in milliseconds and cannot be observed directly. However, neuronal activation is followed by the metabolic process which increases blood flow and volume in the activated areas, and can therefore be measured by fMRI.

### Single-Subject Modeling

In a typical task-related fMRI experiment, the whole brain is scanned at multiple times while a subject performs a series of tasks, and a time series of BOLD response is acquired for each voxel of the brain.

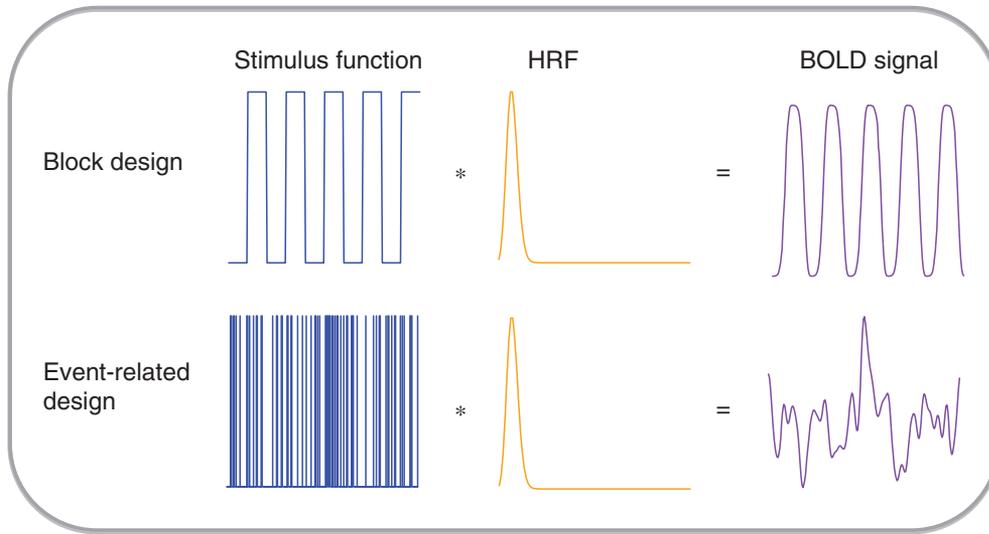
The most common statistical model of a time series of BOLD responses relies on the Gaussian

linear model, known as general linear model (GLM) in the fMRI literature, as first proposed by Friston et al.<sup>8</sup> This models the observed fMRI signal as the underlying BOLD response plus a noise component. Let  $Y_v$  be the  $T \times 1$  response vector of time-series data for voxel  $v$ , for  $v = 1, \dots, V$ , with  $T$  the number of time points and  $V$  the number of voxels, and let  $X_v$  be the  $T \times p$  design matrix, with  $p$  being the number of experimental tasks or input stimuli. We write the voxel-wise GLM as

$$Y_v = X_v \beta_v + \epsilon_v, \quad (1)$$

where  $\beta_v = (\beta_{v,1}, \dots, \beta_{v,p})^T$  is a  $p \times 1$  vector of regression coefficients and  $\epsilon_v$  is a  $T \times 1$  error vector. The error component in (1) captures random noise and various nuisance components due to the hardware as well as subject-related physiological noise.

Let's consider the underlying BOLD response component  $X_v \beta_v$  in model (1). This captures the relationship between the vascular response and the stimulus as follows. When measuring the change in the metabolism of BOLD contrast due to an outside stimulus, the MR signal gets delayed hemodynamically.<sup>9</sup> Such a hemodynamic response is typically referred to as the hemodynamic response function (HRF). A widely used model to account for the lapse of time between the stimulus onset and the vascular response looks at the BOLD signal as the convolution of the



**FIGURE 3** | Typical modeling of the BOLD signal at a given voxel, for both block and event-related designs. The BOLD signal is modeled as the convolution of the experimental stimulus and the HRF.

stimulus pattern with the HRF. This implies that in model (1) each column (task or input stimulus) of  $\mathbf{X}_v$  is modeled as

$$\int_0^t x(s) h_v(t-s) ds, \quad (2)$$

where  $x(s)$  represents the external time-dependent stimulus function for that particular task, which is known and corresponds to the experimental paradigm (e.g., a vector defined with elements set to 1 when the stimulus is ‘on’ and 0 when it is ‘off’). Figure 3 depicts such BOLD signal modeling for two commonly used experimental designs, block, and event-related.

Several models for the HRF  $h_v(t)$  have been proposed. Early proposals included Poisson<sup>8</sup> functions of the type

$$h(t) = \frac{\exp(-d) d^{t-1}}{(t-1)!}, \quad (3)$$

Gaussian<sup>10</sup> functions and gamma functions<sup>11</sup> of the type

$$h(t) = \theta_2 (\theta_2 t)^{\theta_1-1} \exp(-\theta_2 t) / \Gamma(\theta_1). \quad (4)$$

More common choices include the Canonical HRF, defined as the difference of two gamma functions<sup>12</sup>

$$h(t) = (t/d_1)^{a_1} \exp(-(t-d_1)/b_1) - c (t/d_2)^{a_2} \times \exp(-(t-d_2)/b_2), \quad (5)$$

and the inverse logit function,<sup>13</sup> which is generated as a superposition of three separate inverse logit

functions, that is,

$$h(t) = \alpha_1 L((t-T_1)/D_1) + \alpha_2 L((t-T_2)/D_2) + \alpha_3 L((t-T_3)/D_3), \quad (6)$$

with  $L(x) = 1/(1 + e^{-x})$ .

Several Bayesian approaches to models of type (1) have been investigated in the literature and successfully applied to fMRI data.<sup>14–29</sup>

These employ hierarchical models that make explicit assumptions on the model parameters, allowing inference via posterior densities. A remarkable feature of such approaches is their flexibility in modeling temporal and spatial correlation features of the data.

### Temporal Modeling

Clearly, some of the temporal correlation of fMRI data is captured via the modeling of the HRF, as described above. As temporal characteristics of the HRF vary across brain voxels, and across subjects, some authors have employed Bayesian models of type (1) where the parameters of the HRFs are voxel-dependent,<sup>24,27,29</sup> while Xia et al.<sup>28</sup> have proposed to model the HRF at each voxel nonparametrically.

A significant amount of work has been done in capturing temporal correlation in fMRI data via the choice of the noise structure. One approach, often used in the classical literature on fMRI data, is to prewhiten the data by obtaining an initial estimate of the autocorrelation structure, based on the data, and then removing this correlation by applying

a transformation to the data.<sup>8</sup> Other approaches, particularly in the Bayesian literature, look at directly modeling the error term. The most common choice is to impose an autoregressive structure of order  $q$  (AR( $q$ )) on  $\epsilon_v$  in (1),

$$\epsilon_{v,t} = \sum_{j=1}^q w_{v,j} \epsilon_{v,t-j} + z_{v,t}, \quad (7)$$

with  $\mathbf{w}_v = (w_{v,1}, \dots, w_{v,q})^T$  a  $q \times 1$  vector of AR coefficients and  $z_{v,t}$  a white noise, assuming prior distributions on the AR coefficients.<sup>21–23,27</sup> An important source of variability in the signal results from scanner drift, which induces slow changes in voxel intensity over time (low-frequency noise), and physiological noise, due to patient motion, respiration, and heart-beat causing fluctuations in signal across both space and time. To account for this, a level shift or deterministic trend, modeled, e.g., as a  $p$ th order polynomial function, can be included in model (1).<sup>3,30</sup> Alternatively, wavelet transforms can be used to filter noise.<sup>31</sup> In the Bayesian literature, Jeong et al.<sup>19</sup> and Zhang et al.<sup>29</sup> considered a general error structure

$$\epsilon_v \sim N(0, \Sigma_v), \quad (8)$$

with  $\Sigma_v(m, n) = [\gamma(|m - n|)]$  and  $\gamma(b)$  the autocovariance function of the process generating the data, and modeled the correlated noise as being from a  $1/f$  long memory process.<sup>32</sup> The authors applied discrete wavelet transforms (DWT) to model (1), transforming the data into the following model in the wavelet domain

$$\mathbf{Y}_v^* = \mathbf{X}_v^* \beta_v + \epsilon_v^*, \quad (9)$$

with  $\mathbf{Y}_v^* = \mathbf{WY}_v$ ,  $\mathbf{X}_v^* = \mathbf{WX}_v$ ,  $\epsilon_v^* = \mathbf{W}\epsilon_v$ , and  $\mathbf{W}$  an orthogonal  $T \times T$  matrix representing the wavelet transform, and performing inference on the model parameters based on the transformed data. Wavelet transforms have the advantage of ‘whitening’ the data, i.e., reducing the dense covariance matrix structure of the long memory to i.i.d. errors  $\epsilon_v^*$ .<sup>33–35</sup>

### Spatial Modeling

Spatial correlation is expected in a voxel-level analysis of fMRI data because the response at a particular voxel is likely to be similar to the responses of neighboring voxels. In Bayesian modeling, spatial dependence between brain voxels is captured by imposing spatial priors on the model parameters. Many approaches use Gaussian Markov random field

(GMRF) priors on the  $j$ th regression coefficient vector  $\beta_{(j)} = (\beta_{1,j}, \dots, \beta_{V,j})^T$ <sup>16,24</sup> of the type

$$p(\beta_{(j)} | \lambda) \propto \exp\left(-\frac{1}{2} \lambda \beta_{(j)}^T \mathbf{Q} \beta_{(j)}\right), \quad (10)$$

with precision matrix  $\mathbf{Q}$  having elements

$$\mathbf{Q}_{v,k} = \begin{cases} n_v, & v = k \\ -1, & v \sim k \\ 0, & \text{otherwise,} \end{cases} \quad (11)$$

with  $n_v$  the number of neighbors of voxel  $v$ , and with  $v \sim k$  denoting that voxels  $v$  and  $k$  are neighbors. Prior (10) with precision matrix  $\mathbf{Q}$  as in (11) is equivalent to

$$p(\beta_{(j)} | \lambda) \propto \exp\left\{-\frac{1}{2} \lambda \sum_{v \sim k} (\beta_{v,j} - \beta_{k,j})^2\right\}, \quad (12)$$

from which we have

$$\beta_{v,j} | \beta_{-v,j}, \lambda \sim N\left(\frac{1}{n_v} \sum_{k \sim v} \beta_{k,j}, \frac{1}{n_v \lambda}\right), \quad (13)$$

where  $\beta_{-v,j} = \{\beta_{l,j}; l \neq v\}$ . Similarly, Penny et al.<sup>23</sup> considered a spatial prior on the regression coefficient vector  $\beta_{(j)}$  of the type

$$\beta_{(j)} \sim N\left(0, \alpha_j^{-1} (\mathbf{S}^T \mathbf{S})^{-1}\right) \\ \alpha_j \sim Ga(a, b), \quad (14)$$

with  $\mathbf{S}$  a  $V \times V$  spatial kernel matrix equal to the Laplacian operator  $L$ , i.e., if  $\mathbf{w}_{(j)} = L\beta_{(j)}$  then  $w_{v,j}$  is equal to the sum of the differences between  $\beta_{v,j}$  and its neighbors, for  $v = 1, \dots, V$ , and  $\alpha_j$  a spatial precision parameter. Each element of  $\mathbf{w}_{(j)}$  follows a zero-mean Gaussian distribution with precision  $\alpha_j$ . Other spatial prior constructions on the regression coefficients that have been investigated in the literature include the diffusion-based spatial priors of Harrison et al.<sup>18</sup> and the conditional autoregressive (CAR) priors of Harrison and Green,<sup>17</sup> while Flandin and Penny<sup>14</sup> used sparse spatial basis function (SSBF) priors on wavelet-based regression coefficients.

Alternative modeling approaches to those described above look at the task of selecting activated voxels as a variable selection problem, that is the identification of the nonzero  $\beta_{v,j}$  in model (1). A common class of priors adopted in the Bayesian literature on variable selection specifies mixture distributions with a spike at zero (commonly called *spike-and-slab*

priors) on the regression coefficients.<sup>36–38</sup> For model (1) we can write

$$\beta_{v,j} \sim \gamma_{v,j} N(0, \tau_{v,j}) + (1 - \gamma_{v,j}) I(\beta_{v,j} = 0), \quad (15)$$

with  $\gamma_v = (\gamma_{v,1}, \dots, \gamma_{v,p})$  binary indicators representing the activation status, i.e.,  $\beta_{v,j} = 0$  if  $\gamma_{v,j} = 0$ , for inactive voxels, and  $\beta_{v,j} \neq 0$  if  $\gamma_{v,j} = 1$ , for active voxels, for  $j = 1, \dots, p$ , and with  $I(A)$  the indicator function equal to 1 if  $A$  is true and 0 otherwise.<sup>20,21,25,29</sup> Kalus et al.<sup>20</sup> considered the case  $p = 1$  and specify a spatial probit model for the prior probabilities of activation, that is  $P(\gamma_v = 1) = \Phi(\alpha_v)$ , with  $\Phi$  the standard normal cdf and  $\alpha = (\alpha_1, \dots, \alpha_V)^T$  following a GMRF prior. Specifically, the authors considered a first order intrinsic GMRF (IGMRF) prior

$$p(\alpha | \xi^2) \propto (\xi^2)^{-(V-1)/2} \exp\left(-\frac{1}{2\xi^2} \alpha^T Q \alpha\right), \quad (16)$$

with  $Q$  as in (11) and  $\xi^2$  a variance parameter determining the degree of smoothness, or a CAR prior of the type

$$\alpha | \tau^2, \xi^2 \sim N(0, \xi^2 P^{-1}), \quad (17)$$

with  $P = I + \tau^2 Q$  and  $\tau^2 > 0$  modeled by a positively truncated Normal prior. Alternatively, Zhang et al.<sup>29</sup> specified a Markov Random Field (MRF) prior on  $\gamma_v$ , parameterizing its conditional probability as

$$p(\gamma_v | d, e, \gamma_k, k \in N_v) \propto \exp\left(\gamma_v \left(d + e \sum_{k \in N_v} \gamma_k\right)\right), \quad (18)$$

with  $N_v$  the set of neighboring voxels of voxel  $v$ ,  $d \in (-\infty, \infty)$  a sparsity parameter controlling the expected prior number of activated voxels and  $e > 0$  the smoothing parameter which affects the probability of identifying a voxel as active according to the activation of its neighbors. Smith and Fahrmeir<sup>25</sup> and Lee et al.<sup>21</sup> considered spatial MRF priors for the case  $p > 1$ , incorporating anatomical prior information as well as spatial interaction between voxels. They wrote the prior on  $\gamma = \{\gamma_{v,j}, v = 1, \dots, V, j = 1, \dots, p\}$  as

$$\pi(\gamma) = \prod_{j=1}^p \pi(\gamma_{(j)}), \text{ where } \gamma_{(j)} = (\gamma_{1,j}, \dots, \gamma_{V,j})^T \text{ and}$$

$$p(\gamma_{(j)}) \propto \exp\left\{\sum_{v=1}^V \alpha_{v,j}(\gamma_{v,j}) + \sum_{v \sim k} \theta_{v,k,j} \omega_{v,k} I(\gamma_{v,j} = \gamma_{k,j})\right\}, \quad (19)$$

with  $\sum_{v=1}^V \alpha_{v,j}(\gamma_{v,j})$ , called the ‘external field’, capturing anatomical prior information, typically as a linear

combination of the parameters  $\alpha_{v,j}(\gamma_{v,j}) = \alpha_{v,j} \gamma_{v,j}$ , with scalars  $\alpha_{v,j}$  fixed a priori, and with the second term in the exponential function being the interaction effect of neighboring voxels  $v$  and  $k$ , with pre-specified weights  $\omega_{v,k}$  (typically, it is assumed  $\theta_{v,k,j} = \theta_j$ ). Smith et al.<sup>26</sup> and Xia et al.<sup>28</sup> also considered a spatial prior of type (19).

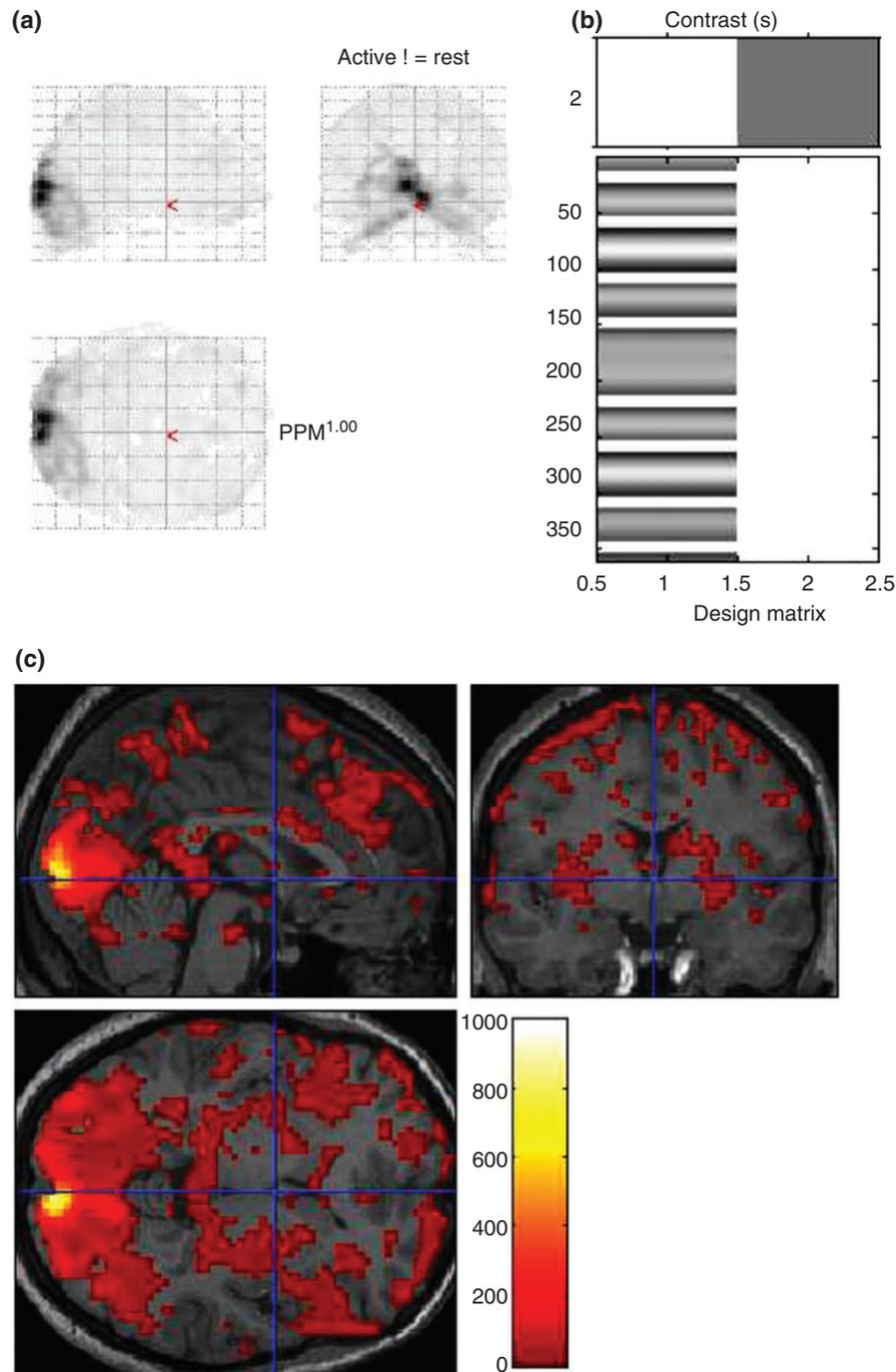
As a point of summary, among the contributions we have described above, Lee et al.<sup>21</sup> and Penny et al.<sup>23</sup> model both temporal and spatial correlations but assume pre-specified HRFs, while Zhang et al.<sup>29</sup> also include the estimation of the HRF. Woolrich et al.<sup>27</sup> impose a nonseparable space-time vector autoregressive (VAR) structure on the error term of the model and incorporate the estimation of the HRF. Flandin and Penny,<sup>14</sup> Gössl et al.,<sup>16</sup> Harrison and Green,<sup>17</sup> Kalus et al.<sup>20</sup> Quirós et al.,<sup>24</sup> Smith and Fahrmeir,<sup>25</sup> Smith et al.<sup>26</sup> make use of spatial priors on the model parameters but assume independent error terms. Gössl et al.<sup>16</sup> and Quirós et al.<sup>24</sup> also incorporate the estimation of the HRF.

### MCMC and Scalability

Many of the Bayesian approaches described above achieve posterior inference via numerical integration methods, such as Markov Chain Monte Carlo (MCMC) sampling algorithms.<sup>16,19–21,24–29</sup> These models are often fit to single 2D slices, as the large dimensionality of the data makes it impossible to model the entire 3D map of the data at once. Nevertheless, MCMC requires a large amount of computer time, even when inference is limited to single slices. This has motivated many authors to investigate alternative techniques for Bayesian inference that do not rely on numerical integration. A common approach is to employ Variational Bayes (VB) methods. Penny et al.<sup>22</sup> first proposed the use of a VB method for inference in a GLM of type (1) with AR errors, and several other authors have adopted VB methods since then.<sup>14,17,23,39</sup> The basic idea of VB methods is to approximate the true posterior density with an analytically tractable form by using a factorization of the posterior distribution over the model parameters.

### Posterior Probability Maps

For posterior inference, the major goal is to produce a spatial mapping of the activated brain regions. This can be achieved via inference on the regression parameters  $\beta_v$ 's in model (1). In particular, activations can be detected by constructing posterior probability maps (PPMs) based on the estimated regression parameters, e.g., as done by Friston et al.<sup>40</sup> and



**FIGURE 4** | (a) An Example of PPMs generated with the software SPM8 (<http://www.fil.ion.ucl.ac.uk/spm>). (b) Design matrix. (c) Overlay of  $\chi^2$  statistic values showing regions where activity is different between active and rest conditions.

Friston and Penny.<sup>15</sup> These authors proposed to detect activations by mapping the estimates of the model parameters at each voxel of single slices of imaging data and then thresholding the corresponding conditional posterior probabilities at a specified confidence level. Specifically, at each voxel the conditional posterior probability that a particular effect,

specified by a contrast weight vector  $\mathbf{w}$ , exceeds some threshold  $\kappa$  is calculated as

$$p = 1 - \Phi \left( \frac{\kappa - \mathbf{w}^T \mathbf{M}_{\beta_v | Y}}{\sqrt{\mathbf{w}^T \mathbf{C}_{\beta_v | Y} \mathbf{w}}} \right), \quad (20)$$

with  $M_{\beta_v|Y}$  and  $C_{\beta_v|Y}$  the posterior mean and covariance of the parameter  $\beta_v$ , respectively. PPMs can be displayed as images, see Figure 4 for example.

When spike-and-slab priors of type (15) are employed, PPMs can be obtained directly by thresholding the activation probabilities  $p(\gamma_{v,j} = 1|Y)$ . In particular, an individual voxel can be classified as active if  $p(\gamma_{v,j} = 1|Y) > Q$ , and as inactive otherwise, with  $Q$  a threshold to be chosen. There are several ways to decide the value of the threshold. Smith et al.,<sup>26</sup> Smith and Fahrmeir<sup>25</sup> and Lee et al.<sup>41</sup> suggested  $Q = 0.8722$ , following Raftery<sup>42</sup> who considered the statistic  $-2 \log(1 - p(\gamma_{v,j} = 1|Y))/p(\gamma_{v,j} = 1|Y)$  approximately distributed  $\chi^2(1)$  and solved the threshold at a critical value of 3.841, for a  $P$  value of 0.05. Kalus et al.<sup>20</sup> suggested choosing the threshold based on the Bayesian false discovery rate (FDR).<sup>43</sup> The decision on the optimal threshold can also be formulated in a compound decision theoretic framework, by minimizing a loss function defined as a linear combination of false positive and false negative counts.<sup>44</sup> Sun et al.<sup>45</sup> have recently shown that procedures thresholding posterior probability maps allow to control also the frequentist FDR in large-scale spatial multiple testing.

### Nonlinear and Mixed Models

Many investigators have acknowledged the presence of nonlinearities in BOLD responses, particularly for event-related designs,<sup>46,47</sup> both across brain regions and stimuli. Among the Bayesian contributions, Genovese<sup>48</sup> presented a nonlinear Bayesian hierarchical model that included the estimation of the drift function and the HRF, assuming independent error terms and without taking into account spatial correlation. Also, Yue et al.<sup>49</sup> proposed a Bayesian adaptive spatial smoothing approach to capture nonstationary spatial correlation, with a Gaussian smoothing kernel varying across space and time. Their model can be written as

$$y_{jk} = f(u_j, u_k) + \varepsilon_{jk}, \tag{21}$$

for  $j = 1, \dots, n_1$  and  $k = 1, \dots, n_2$ , with  $y_{jk}$  the fMRI response data observed at location  $[u_j, u_k]$ , at a given time point, and with  $f$  an unknown function representing the smoothed image. The authors applied their model to the raw fMRI data at each time point, independently, imposing a spatially adaptive IGMRF prior on the function  $f$  and assuming independent errors.

Another class of models that has been quite successful for the analysis of fMRI data, particularly in the Bayesian literature, is mixture models. Here the idea is to characterize the spatial distribution of the data via a (possibly infinite) mixture of distributions,

each capturing a distinct cluster of activations. These models are often applied to processed data, either ‘contrast’ maps, obtained by estimating the  $\beta$  coefficients of a GLM fitted to the fMRI time-series data, or simple z-statistic images. Woolrich et al.<sup>50</sup> first considered finite mixture models with adaptive spatial regularization priors on the model parameters. Their model was implemented in the software FSL. Other authors have considered infinite mixture model with Dirichlet process (DP) priors that enable learning on the number of components from the data.<sup>51,52</sup>

A formal definition of a DP, a stochastic process commonly used in Bayesian nonparametric inference, was first given by Ferguson.<sup>53</sup> Here it suffices to consider a DP as a prior on a class of probability distribution. Let  $G$  denote such random probability measure on the distribution space, with

$$G \sim \text{DP}(\eta, G_0), \tag{22}$$

indicating that the model depends on two parameters, the base measure  $G_0$  and the total mass parameter  $\eta$ . The base measure  $G_0$  is the prior mean of  $G$ , i.e.,  $E(G) = G_0$ . Typically, the unknown  $G$  is centered around a known parametric model, while the total mass parameter  $\eta$  determines the variation of the random measure around the prior mean, with smaller values of  $\eta$  implying higher uncertainty. Any realization  $G$  from a DP defines a discrete distribution almost surely. Let  $\varphi_i | G \sim G, i = 1, \dots, n$ , be an i.i.d. sample from a distribution  $G$ , then  $G$  can be almost surely written as a mixture of point masses,

$$G = \sum_{b=1}^{\infty} w_b \delta_{\varphi_b^*}, \tag{23}$$

with  $w_b = V_b \prod_{j=1}^{b-1} (1 - V_j), V_j \sim \text{Beta}(1, \eta), j = 1, \dots, b$ ,

and atoms  $\varphi_b^* \sim G_0, b = 1, \dots, \infty$ . The discreteness of the DP is best appreciated by looking at the predictive distribution of  $\varphi_i$ , conditional on all the other values  $\varphi_{-i} = \{\varphi_j : j \neq i\}$ ,

$$\varphi_i | \varphi_{-i} \sim \frac{\eta G_0 + \sum_{j \neq i} \delta_{\varphi_j}}{n - 1 + \eta}, \tag{24}$$

which gives positive probability to ties, therefore implying that the  $\varphi_i$ 's can form clusters. The parameter  $\eta$  acts as a weight, i.e., the larger  $\eta$ , the higher the probability that  $\varphi_i$  is sampled from the base measure  $G_0$  and thus the larger the number of clusters.

For fMRI data analysis, Kim et al.<sup>52</sup> used mixture models with DP priors to model processed data of the type  $y_\nu, \nu = 1, \dots, V$ , with  $y_\nu$  the estimate of the  $\beta$  coefficient of a GLM fitted to fMRI time-series data at voxel  $\nu$  at position  $\mathbf{x}_\nu = (x_{\nu 1}, x_{\nu 2})$ ,

$$p(y_\nu | \mathbf{x}_\nu, \theta) = \sum_{c \in C} p(y_\nu | c, \mathbf{x}_\nu, \theta) p(c | \mathbf{x}_\nu, \theta), \quad (25)$$

with  $C$  a set of component labels,  $p(y_\nu | c, \mathbf{x}_\nu, \theta)$  a Gaussian-shaped surface model for each of the mixture components, and with a DP prior imposed on the component label  $c_\nu, \nu = 1, \dots, V$ . Also, Johnson et al.<sup>51</sup> considered an infinite mixture model applied to z-scores  $y_\nu, \nu = 1, \dots, V$ . The model is conditional upon a latent activation state  $c_\nu \in \{-1, 0, 1\}, \nu = 1, \dots, V$ , with labels  $-1, 0$ , and  $1$  denoting three classes/states as deactivated, null, and activated, respectively. With a nonparametric hidden Markov random field model (Potts model) imposed on  $c$ , the model is given by

$$\begin{aligned} y_\nu | c_\nu = m, \varphi_\nu &\sim F_m(\varphi_\nu) \\ \varphi_\nu | c_\nu = m, G_m &\sim G_m \\ G_m | G_{m0}, \eta_m &\sim \text{DP}(\eta_m, G_{m0}), \end{aligned} \quad (26)$$

with  $F_m$  a Gaussian distribution with parameters  $\varphi_\nu = \{\mu_\nu, \sigma_\nu^2\}$  and  $G_{m0}$  a Gaussian distribution with mean  $\mu_{m0}$  and  $\sigma_{m0}^2$ .

### Multiple-Subject Modeling

Spatiotemporal models of type (1) have been also extended to multiple subjects.<sup>54,55</sup> The model becomes

$$\mathbf{Y}_{iv} = \mathbf{X}_{iv} \beta_{iv} + \mathbf{E}_{iv}, \quad (27)$$

with  $\beta_{iv} = (\beta_{iv1}, \dots, \beta_{ivp})$  and  $\beta_{ivk}$  the effect corresponding to condition  $k$  on voxel  $\nu$  for subject  $i$ .

The computational challenge of fitting spatiotemporal models of type (27) to voxel-wise data on multiple subjects has motivated researchers to adopt approaches where voxels are grouped into regions of interest (ROIs) and ‘summary statistics’ are calculated for each ROI, so that inference at the group level is based on the summary statistics from the lower level. This approach has also been dominant in the frequentist literature, under the name of group analysis, where GLM-based estimates of the regression parameters obtained at the voxel level are treated as summary statistics at the group level.<sup>56</sup> One Bayesian approach to group analysis was proposed by Su et al.,<sup>57</sup> who used a hierarchical model with shrinkage estimation

of residual variance by combining information across voxels.

Other Bayesian approaches have used two-stage modeling. Bowman et al.<sup>54</sup> combined whole-brain voxel-by-voxel modeling and ROI analysis within a Bayesian hierarchical framework aiming at the detection of task-related activated brain regions. At the first stage a voxel-wise GLM is fitted for each subject, assuming serially correlated errors and a pre-specified HRF. At the second stage, the authors considered an anatomical parcellation of the brain consisting of  $G$  regions and defined a contrast fMRI BOLD response vector associated with stimulus  $j$  as  $\beta_{igj} = (\beta_{ig(1)j}, \dots, \beta_{ig(V_g)j})^T$ , with  $V_g$  the number of voxels in region  $g = 1, \dots, G$ . They then fit a spatial hierarchical model of the type

$$\begin{aligned} \beta_{igj} | \mu_{gj}, \alpha_{igj}, \sigma_{gj}^2 &\sim \text{Normal}(\mu_{gj} + 1\alpha_{igj}, \sigma_{gj}^2 \mathbf{I}) \\ \mu_{gj} | \lambda_{gj}^2 &\sim \text{Normal}(1\mu_{0gj}, \lambda_{gj}^2 \mathbf{I}) \\ \sigma_{gj}^{-2} &\sim \text{Gamma}(a_0, b_0) \\ \alpha_{ij} | \Gamma_j &\sim \text{Normal}(0, \Gamma_j) \\ \lambda_{gj}^{-2} &\sim \text{Gamma}(c_0, d_0) \\ \Gamma_j^{-1} &\sim \text{Wishart}((h_0 \mathbf{H}_{0j})^{-1}, h_0), \end{aligned} \quad (28)$$

with  $\mu_{gj} = (\mu_{g(1)j}, \dots, \mu_{g(V_g)j})^T$ , and  $\alpha_{ij} = (\alpha_{i1j}, \dots, \alpha_{iGj})^T$ .

A different two-stage modeling approach was put forward by Sanyal and Ferreira,<sup>55</sup> who first fit a GLM of type (27), assuming independent errors and regressor  $X_{iv}$  as convolution of the stimulus function with an empirically derived subject-specific HRF. The authors first estimated the regression coefficients by an empirical Bayes methodology and then transformed the estimated standardized coefficients via DWT to obtain a model in the wavelet space, where they imposed spike-and-slab priors on the wavelet coefficients, extending the wavelet basis prior of Flandin and Penny<sup>14</sup> from one subject to multiple subjects.

Bayesian mixture models have also been successfully applied in multi-subject fMRI studies, to capture clusters of activation. Xu et al.<sup>58</sup> developed a spatial model for multiple-subject fMRI data to capture the inter-subject variability in activation locations. The authors considered scalar t-images, obtained by fitting an intra-subject fMRI model to each subject, and modeled those via a Gaussian mixture with an unknown number of components. Bayesian nonparametric methods for infinite mixtures were employed by

Thirion et al.<sup>59</sup> to model the spatial positions of brain regions. For subject  $s$  and ROI  $j$ , their model can be written as

$$\begin{aligned}
 t_j^s | z_j^s, (\mu_k)_{k=1}^\infty, \sigma^2 &\sim N(\mu_{z_j^s}, \sigma^2 \mathbf{I}), \\
 (\mu_k)_{k=1}^\infty | G &\sim G, \\
 z_j | \pi &\sim \pi, \\
 \pi_b &= \pi_b' \prod_{l=1}^{b-1} (1 - \pi_l'), \quad \pi_b' \sim \text{Beta}(1, \eta), \quad (29)
 \end{aligned}$$

with  $t_j^s$  the spatial coordinates of the center of the areas related to  $a_j^s$  which is the corresponding ROI for subject  $s$ ,  $j=1, \dots, I(s)$ , and with  $z_j^s$  denoting the cluster to which  $a_j^s$  is associated. The model infers spatial ROIs activations at group level while computing inter-subject correspondence via Bayesian network models. Jbabdi et al.<sup>60</sup> imposed a hierarchical DP mixture model on voxel-wise connectivity scores.

### BRAIN CONNECTIVITY

While constructing maps of brain regions activated by specific stimuli is certainly of major interest in imaging studies, another important task, which has received increased interest in recent years, is to infer brain connectivity. In general terms, connectivity looks at how brain regions interact with each other and how information is transmitted between them, with the aim of uncovering the actual mechanisms of how our brain functions. General interest in connectivity studies may be in comparing connectivity properties among subgroups of subjects and between different scanning sessions. Such studies often look at resting-state data, as opposed to task data. Also, of major interest is to understand the role that connectivity patterns, and their disruption, play in mental health disorders and brain diseases.

As defined in the fMRI literature,<sup>7</sup> two types of connectivity can be inferred based on fMRI data: Functional connectivity is defined as the undirected association, or temporal correlation, between BOLD signals from spatially remote brain regions, while effective connectivity is the directed influence of one brain region on other regions. Friston<sup>61</sup> provided a nice review on functional and effective connectivity.

### Functional Connectivity

Functional connectivity aims at identifying parts of the brain showing similar temporal characteristics and, as such, can be quantified using statistical measures of dependence among remote neurophysiological events.

In the classical literature, simple approaches to capture functional connectivity are based on temporal correlations between ROIs, or between a ‘seed’ region and other voxels throughout the brain.<sup>62</sup> Alternative approaches include clustering methods, to partition the brain into regions that exhibit similar temporal characteristics, and multivariate methods for dimension reduction, such as principal components analysis (PCA)<sup>63</sup> and independent components analysis (ICA),<sup>64,65</sup> which determine spatial patterns that account for most of the variability in the time-series data. Approaches that allow to estimate partial correlations between predefined ROIs have also been proposed, e.g., by using the graphical Lasso (GLasso), which estimates a sparse precision matrix.<sup>66,67</sup>

Initial efforts in the development of Bayesian methods to assess functional connectivity were put forward by Patel et al.,<sup>68,69</sup> who dichotomized the time-series data based on a threshold to indicate presence or absence of elevated activity at a given time point and then modeled the relationship between pairs of distinct brain regions by comparing expected joint and marginal probabilities of elevated neural activity. In their two-stage modeling approach, Bowman et al.<sup>54</sup> employed a measure of the strength of task-related intra-regional (or short-range) connectivity based on model (28) defined as

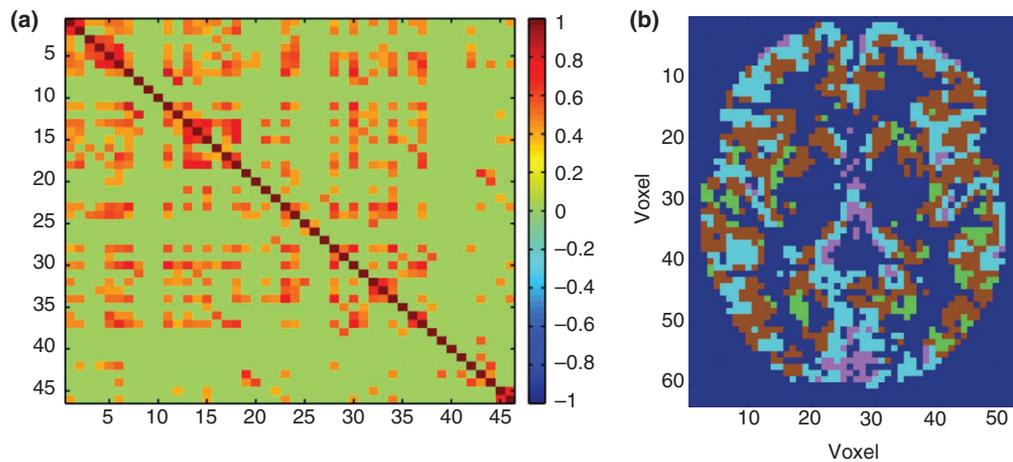
$$\rho_{gj}^{(w)} = \frac{\gamma_{gg}^{(j)}}{\gamma_{gg}^{(j)} + \sigma_{gj}^2}, \quad (30)$$

where  $\rho_{gj}^{(w)}$  reflects the similarity in the neural activity between voxels within a given brain region  $g$  and  $\gamma_{gg}^{(j)}$  is  $g$ th diagonal element of  $\Gamma_j$ . The authors also defined an inter-regional (or long-range) connectivity between regions  $g_1$  and  $g_2$  (see Figure 5(a)) as

$$\rho_{g_1 g_2 j}^{(b)} = \frac{\gamma_{g_1 g_2}^{(j)}}{\sqrt{(\gamma_{g_1 g_1}^{(j)} + \sigma_{g_1 j}^2)(\gamma_{g_2 g_2}^{(j)} + \sigma_{g_2 j}^2)}}. \quad (31)$$

In their estimation approach to spatiotemporal Bayesian models of type (1), Zhang et al.<sup>29</sup> allow clustering of spatially remote voxels that exhibit fMRI time series with similar characteristics (see Figure 5(b)), by imposing DP prior on the parameters of long memory error term (8). The induced clustering can be viewed as an aspect of functional connectivity, as it naturally captures statistical dependencies among remote neurophysiological events.

In addition, there has been recent interest in dynamic functional connectivity models that investigate temporal dynamic interactions among brain



**FIGURE 5** | Functional connectivity. (a) Matrix of posterior estimates of inter-regional correlations (Bowman et al.<sup>54</sup>); (b) Posterior clustering map of spatially remote voxels (Zhang et al.<sup>29</sup>).

regions. Zhang et al.<sup>70</sup> proposed a dynamic Bayesian variable partition model that simultaneously infers global functional interaction patterns within brain networks and their temporal transition boundaries.

### Effective Connectivity

Nongenerative modeling measures of connectivity based on statistical dependence, such as temporal correlation, can be affected by spurious results, as they can change between conditions or groups simply due to changes in, e.g., the signal-to-noise ratio (SNR) in the data.<sup>61</sup> Effective connectivity employs biologically plausible generative models of a typically small network of connected brain regions, assessing the statistical significance of the individual directed connections and effectively modeling the SNR. Being activity-dependent, effective connectivity is dynamic, that is, time-varying, in nature. Effective connectivity refers to causal dependence, as opposed to simple association. Commonly used approaches therefore include many of the methods typically employed to represent causal analysis. The most successful have been Structural Equation Modeling (SEM),<sup>71,72</sup> Dynamic Causal Modeling (DCM),<sup>73</sup> VAR models,<sup>74</sup> Granger causality (GC)<sup>41</sup> and Bayesian networks (BNs).<sup>75</sup> It should be pointed out, however, that even though such methods allow inference on directed connections between brain regions, none of them is able to measure physiological causality, as the direct influence of one region on another, which is ultimately where the scientific interest lies.

SEM was proposed for use in econometrics, and then applied to functional brain imaging data by McIntosh and Gonzalez-Lima.<sup>72</sup> Most of the existing methods use frequentist approaches, with the exception of Scheines et al.<sup>76</sup>

GC was first defined by Granger<sup>77</sup> for temporally structured data, such as time-series data. In its most general terms, GC does not rely on the prior specification of a structural model, but it is rather based on the idea that causes always precede effects. Therefore, past signal values from one brain region can be used to predict current values in another region. However, methods that directly model GC cannot be applied to fMRI data, due mostly to the mismatch between the sampling interval of the data and the much faster timings of the neurodynamics events,<sup>78,79</sup> resulting in GC mainly estimating causal interactions in the observed BOLD signals, rather than in the underlying neuronal responses. GC also does not properly take into account the experimentally induced modulatory effects while estimating causal interactions.<sup>80</sup>

A certain type of GC can be expressed in a state-space form via VAR models, as these models nicely account for time-varying parameters.<sup>81</sup> Let  $y_g(t)$  be the measured BOLD response for the  $g$ th region at time  $t$ , with  $g = 1, \dots, G$  and  $t = 1, \dots, T$ , and let  $x_g(t)$  be the expected BOLD response. Then a model for the observed fMRI signal can be specified as

$$y_g(t) = \alpha_g + x_g(t) \beta_g(t) + \varepsilon_g(t), \quad (32)$$

where  $\alpha_g$  and  $\beta_g(t)$  are the baseline and time-varying activation coefficients for the  $g$ th region at the  $t$ th time point. Assume that  $x_g(\cdot) = x(\cdot)$ ,  $\beta_g(t)$  can be modeled in terms of the noise-free BOLD response in the other regions at the previous time point  $t - 1$  as

$$\beta_g(t) = x(t-1) \left( \sum_{l=1}^G \gamma_{gl}(t) \beta_l(t-1) \right) + \omega_g(t), \quad (33)$$

where  $\omega_g(t)$  are independent zero-mean normal distributed errors and the effective connectivity parameter  $\gamma_{gl}(t)$  is the influence of the  $l$ th region on the  $g$ th region at time  $t$ . Equations (32) and (33) specify a type of VAR model. The objective is to make inference on the parameters  $\gamma_{gl}(t)$ , that capture effective connectivity. In the Bayesian literature, Bhattacharya et al.<sup>82</sup> proposed a symmetric random walk model for  $\gamma_{gl}(t)$ ,

$$\gamma_{gl}(t) = \gamma_{gl}(t - 1) + \delta_{gl}(t), \quad (34)$$

with  $\delta_{gl}$  independent zero-mean normal distributed errors, while Bhattacharya and Maitra<sup>83</sup> used non-parametric DP priors on  $\Gamma_{gl} = (\gamma_{gl}(1), \dots, \gamma_{gl}(T))^T$ ,

$$\begin{aligned} \Gamma_{gl} &\sim G^{(T)}, \\ G^{(T)} &\sim DP\left(\tau, G_0^{(T)}\right), \end{aligned} \quad (35)$$

with  $g, l = 1, \dots, G$ , and with  $G^{(T)}$  a T-variate distribution with mean  $G_0^{(T)}$  being the T-variate distribution implied by a standard time series such as AR(1). Yu et al.<sup>84</sup> used a slightly different model, by imposing a VAR-type structure on the noise term in (32), defining effective connectivity in terms of the corresponding VAR coefficients, and then using spike-and-slab priors on the VAR parameters to select the effective connectivities. Their model also leads to a measure of conditional and overall functional connectivity between ROIs based on precision matrix of the temporally uncorrelated noise component. As inference on VAR models is computationally challenging, due to the large size of the model space, applications of such models are typically done by considering a relatively small number of pre-selected regions, on single subjects. Recently, Gorrostiti et al.<sup>85</sup> have developed a Bayesian hierarchical VAR model for effective connectivity in multiple subjects, accounting for the variability in the connectivity structure within and between subjects.

Other much more sophisticated classes of state-space models have recently been developed to model effective connectivity. For example, Ryali et al.<sup>86</sup> proposed a class of multivariate dynamical models that used VAR state-space models incorporating both intrinsic and modulatory causal interactions. Intrinsic interactions reflect causal influences independent of external stimuli and task conditions, while modulatory interactions reflect context dependent influences. Causal interactions are modeled at the level of latent signals, rather than at the level of the observed BOLD-fMRI signals.

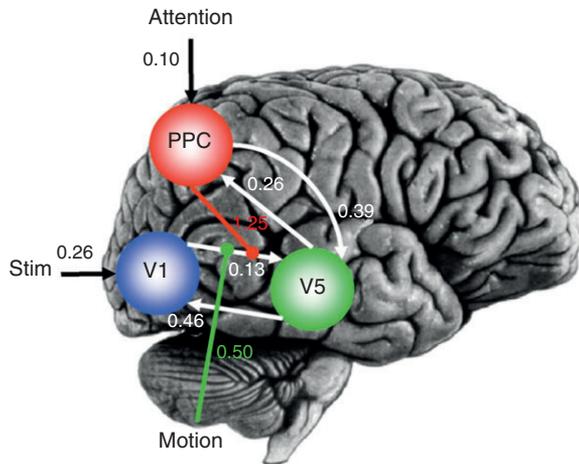
Directed acyclic graphs, or BNs, represent putative causal links between a set of variables in terms

of conditional independence. Dynamic Bayesian networks (DBNs), which takes into account the dynamic nature of the process, make use of VAR-type time-series modeling to represent GC. These approaches have been recently used to reveal effective connectivity among brain regions.<sup>87–91</sup> Kim et al.<sup>87</sup> used a discrete dynamic Bayesian network (dDBN) to discriminate brain regions between schizophrenic patients and healthy controls, measuring effective connectivity by the relative likelihood of correlations between brain regions in one group versus another. Rajapakse and Zhou<sup>91</sup> employed DBNs and Li et al.<sup>88</sup> compared multiple-subject approaches that either pool the data or learn a separate BN for each subject or place the same BN structure on each of the subjects. Li et al.<sup>89</sup> and Li et al.<sup>90</sup> inferred effective connectivity among multiple resting-state networks (RSNs) in the brain by using a group independent component analysis (ICA) first, to identify the RSNs, and then applied a BN learning approach to infer the conditional dependencies among RSNs.

Another popular approach to effective connectivity is DCM, originally proposed by Friston et al.<sup>73</sup> Stephan et al.<sup>92</sup> offered a nice review of the approach and its applications to fMRI data. DCM models interactions among brain regions directly at the neuronal level using fMRI time series at the hemodynamic level. These models heavily rely on complex biological assumptions, such as how the neuronal states enter a region-specific hemodynamic model to produce the BOLD responses. These assumptions have not yet been adequately verified.<sup>93</sup> The basic idea under DCM is to treat the brain as a nonlinear dynamic system with multiple inputs and outputs. Effective connectivity is parameterized in terms of the coupling among unobserved neuronal activity in different regions. The coupling parameters are estimated via perturbing the system, adapting to the fMRI experimental inputs (e.g., stimulus function), and measuring the response. The general mathematical framework of DCM is based on differential equations. More specifically, DCM for fMRI models state changes in a system (or network) of  $n$  interactive brain regions, with each region being represented by a state variable, via a bilinear differential equation of the type

$$\frac{dz(t)}{dt} = F(x, u, \theta) = \left( \mathbf{A} + \sum_{j=1}^m u_j \mathbf{B}^j \right) z + \mathbf{C}u, \quad (36)$$

where  $\theta = (\mathbf{A}, \mathbf{B}, \mathbf{C})$  are the neuronal parameters defining connectivity, or coupling, and interactions between brain regions. In particular, the matrix  $\mathbf{A}$  is a fixed connectivity matrix representing the intrinsic connectivity among the brain regions in the absence of inputs,  $\mathbf{B}^j$



**FIGURE 6 |** Effective connectivity. Maximum a posteriori estimates of parameters measuring effective connectivity in an fMRI study on attention to motion (Stephan et al.<sup>98</sup>).

represents changes in connectivity induced by the  $j$ th input and the matrix  $C$  reflects the strength of extrinsic influence of the inputs on neuronal activity. Based on state equation (36), the parameters  $\theta$  can be written as

$$A = \frac{\partial F}{\partial z} \Big|_{u=0}, \quad B^j = \frac{\partial^2 F}{\partial z \partial u_j}, \quad C = \frac{\partial F}{\partial u} \Big|_{z=0}. \quad (37)$$

DCM models are quite complex in structure and inference is usually infeasible with more than a few regions. Bayesian approaches have been particularly helpful for model parameter estimation.<sup>94–98</sup> Typically, Normal priors are placed on the model parameters and an optimization scheme is used to estimate parameters that maximize the posterior probability. The posterior density is then used to make inferences about the significance of the connections between various brain regions. Stephan et al.<sup>98</sup> proposed a nonlinear extension of DCM, augmenting the state equation (36) with additional nonlinear terms representing how connection strengths change due to the activity of other brain regions (see Figure 6). Daunizeau et al.<sup>94</sup> and Li et al.<sup>97</sup> used stochastic dynamic models that accommodate for random fluctuations in neuronal states, allowing application to resting-state fMRI data. Friston et al.<sup>96</sup> and Friston and Penny<sup>95</sup> used DCM for network discovery, to detect the dependence graph structure which best fits the observed fMRI data.

## CLASSIFICATION AND PREDICTION

Another important task in studies based on fMRI data is the ability to do classification or prediction. Some studies look at predicting individuals’ brain activity. For example, the prediction of post-treatment brain

activity may be of interest to clinicians as a guide to individualized treatment selection. Other studies consider the prediction of a clinical or a behavioral response.

For prediction of brain activity, Guo et al.<sup>99</sup> developed a two-stage hierarchical Bayesian model using patient’s pre-treatment scans of fMRI, in combination with other relevant patient characteristics, to predict the brain activity of the patient following a specific treatment. At the first stage, the authors considered a GLM for the dependent variable  $Y_i(v) = (Y'_{i1}(v), Y'_{i2}(v))'$  with  $Y_{i1}(v)$  and  $Y_{i2}(v)$  the  $(T_1 \times 1)$  pre- and  $(T_2 \times 1)$  post-treatment serial BOLD responses for subject  $i$ , measured at voxel  $v$ , which is given by

$$\begin{bmatrix} Y_{i1}(v) \\ Y_{i2}(v) \end{bmatrix} = \begin{bmatrix} X_{iv1}^{(1)} & 0 \\ 0 & X_{iv2}^{(1)} \end{bmatrix} \begin{bmatrix} B_{i1}(v) \\ B_{i2}(v) \end{bmatrix} + \begin{bmatrix} Z_{iv1}^{(1)} & 0 \\ 0 & Z_{iv2}^{(1)} \end{bmatrix} \times \begin{bmatrix} \alpha_{i1}(v) \\ \alpha_{i2}(v) \end{bmatrix} + \begin{bmatrix} \epsilon_{i1}^{(1)}(v) \\ \epsilon_{i2}^{(1)}(v) \end{bmatrix}, \quad (38)$$

where  $X_{ij}^{(1)}, j = 1, 2$  are design matrices convolved with a HRF and  $Z_{ij}^{(1)}, j = 1, 2$  are high-pass filtering matrices. At the second stage, the subject-specific effects  $B_{ij}(v), j = 1, 2$  are modeled via a linear model with design matrix containing covariates including treatment assignment and other relevant patient characteristics, to capture the association between pre- and post-treatment neuroimaging measurements. Prediction is done via the conditional distribution of the post-treatment effects  $B_{i2}(v)$  given the pre-treatment effects  $B_{i1}(v)$ . Derado et al.<sup>100</sup> proposed an extension of the model of Bowman et al.<sup>54</sup> that takes into account the spatial correlations between neighboring brain regions and intra-regional voxels in addition to capturing the temporal correlations between scans. Even though they applied the model in a study using positron emission tomography (PET) data, their proposed method is applicable to fMRI studies as well.

Linear regression models have been used for the prediction of clinical or behavioral outcomes based on fMRI data. Michel et al.<sup>101</sup> developed a multiclass sparse Bayesian regression model, based on a clustering of the voxels, for the prediction of cognitive states. van Gerven et al.<sup>102</sup> employed a Bayesian logistic regression model with a multivariate Laplace prior on the regression coefficients to predict experimental conditions (i.e., handwritten digits), based on BOLD response data. Morcom and Friston<sup>103</sup> used a multivariate Bayesian model to infer which activity patterns predict memory formation.

Furthermore, scalar-on-image regression models of the type

$$Y = X\eta + \epsilon, \tag{39}$$

with  $Y$  a  $n \times 1$  response variable (for instance, measurements of subjects' emotion) and  $X$  a  $n \times p$  matrix of imaging-related predictors at  $p$  voxels, have recently attracted significant interest. Goldsmith et al.<sup>104</sup> considered spike-and-slab priors on  $\eta$ , with an Ising prior on the selection indicator. Li et al.<sup>105</sup> proposed priors of the type

$$\eta_j | (\gamma_j, G) \sim (1 - \gamma_j) \delta_0 + \gamma_j G, \quad G \sim DP(\alpha, G_0), \tag{40}$$

to select voxels that are predictive of the subjects' response while simultaneously achieving clustering of similar regression coefficients.

### INTEGRATIVE IMAGING

An important recent trend in the literature on fMRI data is the use of multi-modal techniques, that is, combining measurements originated from multiple imaging methods, to overcome the limitations when only one modality is used and to aid estimation and prediction.<sup>106</sup> Also, nowadays more and more studies look at collecting both imaging and genetics data on the same subjects, making it possible to develop statistical models that aim at linking neural activity across multiple individuals to their genetic information.<sup>107,108</sup> As patterns of brain connectivity in fMRI scans are known to be related to the subjects' genome, the ability to model the link between the imaging and genetic components could indeed lead to improved diagnostics and therapeutic interventions.

#### Multi-Modal Techniques

A number of Bayesian methods have been proposed for the integration of fMRI data with other neuroimaging techniques that provide information on the brain function or structure. EEG<sup>109</sup> is a functional neuroimaging technique that achieves a direct recording of the brain's electrical activity via multiple electrodes placed on the scalp that measure voltage fluctuations from ionic current flows within the brain's neurons. MEG<sup>110</sup> maps brain activity by measuring the magnetic fields resulting from electrical current in the brain via magnetic field sensors. Unlike fMRI, EEG and MEG have a very high temporal resolution (in the order of milliseconds) but poor spatial resolution. In addition, DTI<sup>111</sup> is a noninvasive magnetic resonance imaging technique that measures the diffusion of water in biological tissues to produce neural

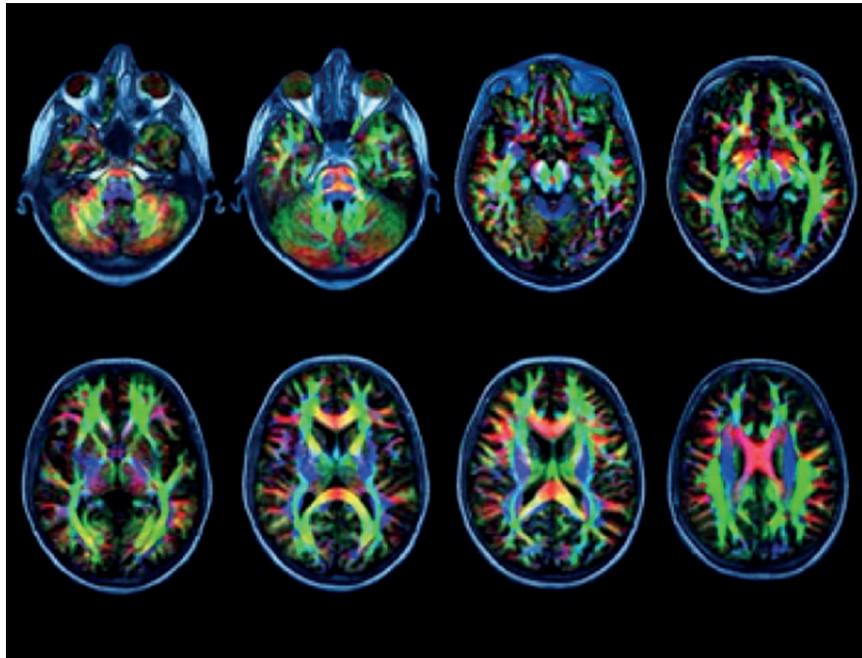
tract images of the brain *in vivo*, therefore capturing structural information. Clearly, the ability to combine neuroimaging techniques could enable researchers to obtain a more comprehensive understanding of the brain.

The complementary characteristics of temporal and spatial resolutions of EEG/MEG and fMRI techniques makes their integration highly desirable. Jorge et al.<sup>112</sup> presented a review on integrative methods. Information from EEG/MEG data has been incorporated in some of the modeling approaches for fMRI data that we have previously described. For example, Kalus et al.<sup>113</sup> extended the approach in Kalus et al.<sup>20</sup> by using EEG-informed spatial priors in their Bayesian variable selection approach to detect brain activation. Specifically, they relate the prior activation probabilities to a latent predictor stage  $\zeta = (\zeta_1, \dots, \zeta_V)^T$  via a probit link  $p(\gamma_v = 1) = \Phi(\zeta_v)$ , with  $\Phi$  the standard normal cdf and  $\zeta_v$  consisting of an intercept term and an EEG effect, that is

$$\zeta_v = \zeta_{0,v} + \zeta_{EEG,v} = \begin{cases} \zeta_{0,v}, & \text{if predictor } 0 \\ \zeta_{0,v} + \zeta_G J_v, & \text{if predictor } glob, \\ \zeta_{0,v} + \zeta_v J_v, & \text{if predictor } flex \end{cases} \tag{41}$$

where  $J_v, v = 1, \dots, V$  is the continuous spatial EEG information and where 0, *glob* and *flex* indicate three types of predictors: predictor 0 contains a spatially-varying intercept  $\zeta_0 = (\zeta_{0,1}, \dots, \zeta_{0,V})^T$ , and corresponds to an fMRI activation detection scheme without incorporating EEG information; predictor *glob* contains a global EEG effect  $\zeta_G$  in addition to the intercept; predictor *flex* contains a spatially-varying EEG effect  $\zeta = (\zeta_1, \dots, \zeta_V)^T$ . With an IGMRF, the priors of  $\zeta_0$  and  $\zeta$  are of the form in (16). Other approaches focus on incorporating information from fMRI into models for EEG/MEG data. A common approach to EEG/MEG data looks at source localization as an ill-posed inverse problem and uses other imaging techniques to determine prior information that impose constraints on source activity and locations.<sup>114–119</sup> Automatic relevance determination (ARD) priors on the variance of the source current at each source location were used by Sato et al.,<sup>119</sup> fMRI-based activity maps as priors on the source location by Phillips et al.,<sup>118</sup> Mattout et al.<sup>117</sup> and Henson et al.,<sup>115</sup> exponential distribution as fMRI-based prior on the source locations by Jun et al.<sup>116</sup> Babajani-Feremi et al.<sup>120</sup> explored variational Bayesian expectation maximization methods for the estimation of the model parameters.

A truly joint model of EEG/MEG and event-related fMRI data was first proposed by



**FIGURE 7** | Example of diffusion tensor imaging (DTI) data.

Daunizeau et al.<sup>121</sup> These authors consider a linear system of equations of the type

$$\begin{aligned} \mathbf{M} &= \mathbf{G}\mathbf{J} + \mathbf{E} \\ \mathbf{Y} &= \mathbf{B}\mathbf{h} + \mathbf{F}, \end{aligned} \quad (42)$$

where  $\mathbf{M}$  represents the  $p \times t_1$  matrix of EEG data, with  $p$  the number of sensors and  $t_1$  the number of time points,  $\mathbf{G}$  is a  $p \times n$  matrix associated with the position and orientation of the dipoles,  $\mathbf{J}$  is a  $n \times t_1$  matrix of the unknown time courses of the dipoles,  $\mathbf{Y}$  is the  $t_2 \times n$  matrix of voxel-wise fMRI data, with  $t_2$  the number of time points and  $n$  the number of voxels,  $\mathbf{h}$  is the  $k \times n$  matrix of unknown HRFs at each voxel and  $\mathbf{B}$  is the  $t_2 \times k$  design matrix. The authors considered a finite parcellation of the cortical surface into  $q$  anatomically and functionally homogeneous regions, associating a time course to each region  $P_i$  ( $i = 1, \dots, q$ ). They then defined a bioelectric event-related response for source  $J$  as

$$\mathbf{J} = \text{Diag}(\mathbf{w}^{\text{EEG}}) \mathbf{C}\mathbf{X} + \mathbf{R}, \quad (43)$$

with  $\mathbf{X}$  being an unknown  $q \times t_1$  matrix of the  $q$  time courses,  $\mathbf{C}$  the known  $n \times q$  matrix describing the cortex parcelling ( $C_{ji} = 1$  if  $j \in P_i$ , and  $C_{ji} = 0$  otherwise),  $\mathbf{w}^{\text{EEG}}$  a  $n \times 1$  unknown vector describing the spatial profile of each active cortical source, and  $\mathbf{R}$  a residual bioelectric activity. Similarly, they specify a hemodynamic event-related response of the type

$$\mathbf{h} = \mathbf{Z}\mathbf{C}^T \text{Diag}(\mathbf{w}^{\text{fMRI}}) + \mathbf{L}, \quad (44)$$

with  $\mathbf{Z}$  an unknown  $k \times q$  matrix of the HRF temporal shape of the  $q$  regions,  $\mathbf{w}^{\text{fMRI}}$  an unknown  $n \times 1$  vector associated with the spatial profile of the hemodynamic activity sources and  $\mathbf{L}$  a residual term. The author assume  $\mathbf{w}^{\text{EEG}} = \mathbf{w}^{\text{fMRI}} = \mathbf{w}$  and specify priors for the cortical currents and the HRF, as well as spatial priors based on a spatial Laplacian on  $\mathbf{w}$ . See also Ou et al.<sup>122</sup> and Luessi et al.<sup>123</sup> for extensions to more general and flexible prior models.

An interesting avenue for future research is the development of methods for the integration of fMRI and DTI data.<sup>124</sup> DTI is an MRI technique that provides information regarding the structure of white matter in the brain (see Figure 7). Axons, neuron fibers that serve as lines of transmission in the nervous system, form bundles of textured fibers in the white matter. This extensive system of white-matter bundles directly links some brain structures. DTI noninvasively maps these white-matter fiber tracts in the brain by measuring the diffusion of water molecules, therefore providing a measure of the so-called anatomical (or structural) connectivity, which refers to how different brain regions are physically connected. Many of the existing approaches to modeling functional connectivity by supplementing fMRI data with information from DTI data use frequentist models,<sup>125–127</sup> while Iyer et al.<sup>128</sup> employ BNs informed by DTI data.

### Imaging Genetics

It is generally known that human brain mapping and connectivity can be affected by the individual's genetic

characteristics. Studies that allow to investigate how particular subsets of polymorphisms can affect functional brain activity are of paramount importance. In addition, such studies could facilitate the identification of the genetic determinants of complex brain-related disorders such as autism, dementia, and schizophrenia. Imaging genetics refers, in particular, to situation where structural and functional neuroimaging techniques are applied to study subjects carrying genetic risk variants that relate to a psychiatric disorder.

Recently proposed approaches for integrative analyses of fMRI and genetic data use frequentist methods, such as classical regression models or PCA-ICA dimension reduction techniques.<sup>129–131</sup> In the Bayesian literature, Stingo et al.<sup>132</sup> proposed a hierarchical mixture model based on ROI summary measures of BOLD signal intensities measured on schizophrenic patients and healthy subjects. Their model incorporates spatial MRF priors for the selection of features (e.g., ROIs) that discriminate schizophrenic from healthy controls and mixture components that depend on selected covariates (e.g., single nucleotide polymorphisms—SNPs) measured on the individual subjects. Posterior inference results into the simultaneous selection of a set of discriminatory ROIs and the relevant SNPs, together with the reconstruction of the correlation structure of the selected regions. Salazar et al.<sup>133</sup> proposed a joint model of ordered/categorical questionnaire data, fMRI and SNP data. Their approach aims at predicting questionnaire answers based on fMRI and SNP data and brain responses to external stimuli based on SNP data and answers to questionnaires.

Imaging genetic studies have great potential for the discovery of biomarkers implicated in brain functions and are expected to be the topic of much research in the future. One interesting area of research, e.g., is the study of family data, in particular twin studies. With family data, the genetic similarity (on average) between the subjects within a family allows to infer the heritability of a phenotype. Even though this type of experimental design has been used primarily on structural data, some contributions also exist in fMRI studies, see e.g., Glahn et al.<sup>134</sup> and van den Berg et al.<sup>135</sup> for an early Bayesian model.

## CONCLUSIONS

There has been a continued interest in the use of fMRI data, and this has motivated a very rapid development of statistical techniques for the analysis of such data. In this review article we have focused on Bayesian methods. Unlike many of the classical inferential techniques, Bayesian models allow for flexibility, mainly via spatial and adaptive priors that can readily

incorporate external information, and can be easily fitted via full MCMC or approximate computational techniques.

In this review article, we have divided methods according to the objective of the analysis. First, we have described spatiotemporal hierarchical models for the estimation of the task-related activation patterns. We have then addressed methods for functional and effective brain connectivity. We have also touched upon methods for prediction of a psychological condition or a treatment response and, finally, have presented a discussion of models that aim at combining multi-modal imaging techniques, particularly fMRI data with EEG/MEG and DTI data. We have also briefly discussed the emerging field of imaging genetics. The latter topics, on integrative models, are quite recent and certainly represent interesting avenues for further developments of Bayesian methodologies. Indeed, many more important contributions are to be expected by Bayesian statisticians working in the area of fMRI data analysis.

Throughout the article we have highlighted that fMRI experiments produce massive amount of spatially and temporally correlated data, posing challenges to statistical analysis, for both classical and Bayesian procedures. Even though posterior inference can be often carried out via standard MCMC methods, the dimensionality of the data may limit the practical use of Bayesian methodologies. We have briefly discussed alternative posterior approximation methods, such as VB, which drastically reduce the computation times. Further research is needed in the development of those estimation schemes, especially in order to accurately assess the trade-off between the gain in computation time and the precision of the resulting estimates. For example, while VB methods may lead to good approximations of the posterior means, it is generally well-understood that they may underestimate the posterior variance and also poorly estimate the correlation structure of the data.<sup>136,137</sup>

Many software tools have been developed for fMRI data analysis. The most popular are FMRIB Software Library (FSL) <http://fsl.fmrib.ox.ac.uk/fsl>, Statistical Parametric Mapping (SPM) <http://www.fil.ion.ucl.ac.uk/spm>, Analysis of Functional NeuroImages (AFNI) <http://afni.nimh.nih.gov> and Brain Voyager <http://www.brainvoyager.com>. Among those, both FSL and SPM include Bayesian methods. For example, the models of Friston et al.<sup>40</sup> Penny et al.<sup>22,23</sup> are implemented in SPM and the method of Woolrich et al.<sup>138</sup> in FSL, allowing the option of either a full MCMC or a VB algorithm for posterior inference. In SPM, e.g., spatiotemporal models of type (1) can be fitted with priors as Unweighted Graph Laplacian

(UGL), Weighted Graph Laplacian (WGL), GMRF, etc., and with AR or independent error terms. A list of choices for the HRF are also available.

In spite of our best efforts, the coverage of Bayesian approaches for fMRI data analysis we have presented in this article is, of course, not comprehensive. For example, we have not discussed meta-analysis, which is an important area in fMRI. Since multi-subject studies often have limited sample sizes, it

is important to adopt strategies to determine whether task-related changes in brain activity, or networks of activated brain regions, are consistent across studies. Some of the recent methodological developments use Bayesian hierarchical models with spatial point processes<sup>139,140</sup> and Bayesian nonparametric regression models.<sup>141</sup> Studies with larger sample sizes are also beginning to emerge, as the result of collaborative efforts among various teams of researchers.

## ACKNOWLEDGMENTS

We wish to thank the Editor and the three referees for their constructive comments. We also thank Brian Caffo, Vince Calhoun, Ciprian Crainiceanu, Ivor Cribben, Marco Ferreira, Nicole Lazar, Hernando Ombao and Hongtu Zhu for comments on an earlier version of this article.

## REFERENCES

- Bowman FD. Brain imaging analysis. *Annu Rev Stat Appl* 2014, 1:61–85.
- Lazar NA. *The Statistical Analysis of Functional MRI Data*. New York, Statistics for Biology and Health: Springer; 2008.
- Lindquist M. The statistical analysis of fMRI data. *Stat Sci* 2008, 230:439–464.
- Poldrack RA, Mumford JA, Nichols TE. *Handbook of fMRI Data Analysis*. New York: Cambridge University Press; 2011.
- Huettel SA, Song AW, McCarthy G. *Functional Magnetic Resonance Imaging*, vol. 1. Sunderland, MA: Sinauer Associates; 2004.
- Woolrich M. Bayesian inference in fMRI. *Neuroimage* 2012, 62:801–810.
- Friston KJ. Functional and effective connectivity in neuroimaging: a synthesis. *Hum Brain Mapp* 1994, 2:56–78.
- Friston KJ, Jezzard P, Turner R. Analysis of functional MRI time-series. *Hum Brain Mapp* 1994, 10:153–171.
- Buxton R, Frank L. A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation. *J Cereb Blood Flow Metab* 1997, 170:64–72.
- Friston KJ, Holmes AP, Poline JB, Grasby PJ, Williams SCR, Frackowiak RSJ, Turner R. Analysis of fMRI time-series revisited. *Neuroimage* 1995, 20:45–53.
- Lange N, Zeger SL. Non-linear Fourier time series analysis for human brain mapping by functional magnetic resonance imaging. *J R Stat Soc Ser C Appl Stat* 1997, 460:1–29.
- Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R. Event-related fMRI: characterizing differential responses. *Neuroimage* 1998, 70:30–40.
- Lindquist M, Wager T. Validity and power in hemodynamic response modeling: a comparison study and a new approach. *Hum Brain Mapp* 2007, 280:764–784.
- Flandin G, Penny W. Bayesian fMRI data analysis with sparse spatial basis function priors. *Neuroimage* 2007, 340:1108–1125.
- Friston KJ, Penny W. Posterior probability maps and SPMs. *Neuroimage* 2003, 190:1240–1249.
- Gössl C, Auer D, Fahrmeir L. Bayesian spatio-temporal inference in functional magnetic resonance imaging. *Biometrics* 2001, 570:554–562.
- Harrison L, Green G. A Bayesian spatiotemporal model for very large data sets. *Neuroimage* 2010, 500:1126–1141.
- Harrison L, Penny W, Daunizeau J, Friston KJ. Diffusion-based spatial priors for functional magnetic resonance images. *Neuroimage* 2008, 410:408–423.
- Jeong J, Vannucci M, Ko K. A wavelet-based Bayesian approach to regression models with long memory errors and its application to fMRI data. *Biometrics* 2013, 69:184–196.
- Kalus S, Sämman P, Fahrmeir L. Classification of brain activation via spatial Bayesian variable selection in fMRI regression. *Adv Data Anal Classification* 2014, 8:63–83.
- Lee K, Jones GL, Caffo BS, Bassett SS. Spatial Bayesian variable selection models on functional magnetic resonance imaging time-series data. *Bayesian Anal* 2014, 90:699–732.
- Penny W, Kiebel S, Friston KJ. Variational Bayesian inference for fMRI time series. *Neuroimage* 2003, 190:727–741.

23. Penny W, Trujillo-Barreto N, Friston KJ. Bayesian fMRI time series analysis with spatial priors. *Neuroimage* 2005, 240:350–362.
24. Quirós A, Diez R, Gamerman D. Bayesian spatiotemporal model of fMRI data. *Neuroimage* 2010, 490:442–456.
25. Smith M, Fahrmeier L. Spatial Bayesian variable selection with application to functional magnetic resonance imaging. *J Am Stat Assoc* 2007, 1020:417–431.
26. Smith M, Pütz B, Auer D, Fahrmeier L. Assessing brain activity through spatial Bayesian variable selection. *Neuroimage* 2003, 200:802–815.
27. Woolrich M, Jenkinson M, Brady J, Smith S. Fully Bayesian spatio-temporal modeling of fMRI data. *IEEE Trans Med Imaging* 2004, 230:213–231.
28. Xia J, Liang F, Wang Y. fMRI analysis through Bayesian variable selection with a spatial prior. In: *Proceedings of the 6th IEEE International Symposium on Biomedical Imaging*, 2009:714–717.
29. Zhang L, Guindani M, Versace F, Vannucci M. A spatio-temporal nonparametric Bayesian variable selection model of fMRI data for clustering correlated time courses. *Neuroimage* 2014, 95:162–175.
30. Worsley KJ, Liao CH, Aston J, Petre V, Duncan GH, Morales F, Evans AC. A general statistical analysis for fMRI data. *Neuroimage* 2002, 150:1–15.
31. Weaver JB, Xu Y, Healy DM, Cromwell LD. Filtering noise from images with wavelet transforms. *Magn Reson Med* 1991, 21:288–295.
32. Smith A, Lewis B, Ruttimann U, Ye F, Sinwell T, Yang Y, Duyn J, Frank J. Investigation of low frequency drift in fMRI signal. *Neuroimage* 1999, 90:526–533.
33. Bullmore E, Fadili J, Maxim V, Şendur L, Whitcher B, Suckling J, Brammer M, Breakspear M. Wavelets and functional magnetic resonance imaging of the human brain. *Neuroimage* 2004, 23:234–249.
34. Fadili MJ, Bullmore ET. Wavelet-generalised least squares: a new BLU estimator of linear regression models with  $1/f$  errors. *Neuroimage* 2002, 15:217–232.
35. Meyer FG. Wavelet-based estimation of a semiparametric generalized linear model of fMRI time-series. *IEEE Trans Med Imaging* 2003, 220:315–322.
36. Brown PJ, Vannucci M, Fearn T. Multivariate Bayesian variable selection and prediction. *J R Stat Soc Ser B Stat Methodol* 1998, 600:627–641.
37. George EI, McCulloch RE. Approaches for Bayesian variable selection. *Stat Sinica* 1997, 70:339–373.
38. Sha N, Vannucci M, Tadesse MG, Brown PJ, Dragoni I, Davies N, Roberts TC, Contestabile A, Salmon M, Buckley C, et al. Bayesian variable selection in multinomial probit models to identify molecular signatures of disease stage. *Biometrics* 2004, 600:812–819.
39. Woolrich M, Behrens T, Smith S. Constrained linear basis sets for HRF modelling using variational Bayes. *Neuroimage* 2004, 210:1748–1761.
40. Friston K, Penny W, Phillips C, Kiebel S, Hinton G, Ashburner J. Classical and Bayesian inference in neuroimaging: theory. *Neuroimage* 2002, 16:465–483.
41. Goebel R, Roebroeck A, Kim D, Formisano E. Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. *Magn Reson Imaging* 2003, 210:1251–1261.
42. Raftery AE. *Hypothesis Testing and Model Selection*. Springer; 1996, 163–187.
43. Newton MA, Noueiry A, Sarkar D, Ahlquist P. Detecting differential gene expression with a semiparametric hierarchical mixture method. *Biostatistics* 2004, 50:155–176.
44. Müller P, Parmigiani G, Rice K. FDR and Bayesian multiple comparisons rules. In: Bernardo JM, Bayarri MJ, Berger JO, Dawid AP, Heckerman D, Smith AFM, West M, eds. *Bayesian Statistics 8*. Oxford: Oxford University Press; 2007.
45. Sun W, Reich BJ, Tony Cai T, Guindani M, Schwartzman A. False discovery control in large-scale spatial multiple testing. *J R Stat Soc Ser B Stat Methodol*. In press.
46. Friston KJ, Mechelli A, Turner E, Price CJ. Nonlinear responses in fMRI: the Balloon model, Volterra kernels and other hemodynamics. *Neuroimage* 2000, 9:416–429.
47. Wager TD, Vazquez A, Hernandez L, Noll DC. Accounting for nonlinear bold effects in fmri: parameter estimates and a model for prediction in rapid event-related studies. *Neuroimage* 2005, 25:206–218.
48. Genovese CR. A Bayesian time-course model for functional magnetic resonance imaging data. *J Am Stat Assoc* 2000, 950:691–703.
49. Yue Y, Loh J, Lindquist MA. Adaptive spatial smoothing of fMRI images. *Stat Interf* 2010, 3:3–13.
50. Woolrich M, Behrens T, Beckmann C, Smith S. Mixture models with adaptive spatial regularization for segmentation with an application to fMRI data. *IEEE Trans Med Imaging* 2005, 240:1–11.
51. Johnson T, Liu Z, Bartsch A, Nichols T. A Bayesian non-parametric Potts model with application to pre-surgical fMRI data. *Stat Methods Med Res* 2013, 220:364–381.
52. Kim S, Smyth P, Stern H. A nonparametric Bayesian approach to detecting spatial activation patterns in fMRI data. In: *Proceedings of the 9th International Conference on Medical Image Computing and Computer Assisted Intervention*, 2006:217–224.
53. Ferguson TS. A Bayesian analysis of some nonparametric problems. *Ann Stat* 1973, 10:209–230.

54. Bowman F, Caffo B, Bassett S, Kilts C. A Bayesian hierarchical framework for spatial modeling of fMRI data. *Neuroimage* 2008, 390:146–156.
55. Sanyal N, Ferreira M. Bayesian hierarchical multi-subject multiscale analysis of functional MRI data. *Neuroimage* 2012, 630:1519–1531.
56. Holmes AP, Friston KJ. Generalisability, random effects & population inference. *Neuroimage* 1998, 7:S754.
57. Su S, Caffo B, Garrett-Mayer E, Bassett S. Modified test statistics by inter-voxel variance shrinkage with an application to fMRI. *Biostatistics* 2009, 100:219–227.
58. Xu L, Johnson TD, Nichols TE, Nee DE. Modeling inter-subject variability in fMRI activation location: A Bayesian hierarchical spatial model. *Biometrics* 2009, 650:1041–1051.
59. Thirion B, Tucholka A, Keller M, Pinel P, Roche A, Mangin J, Poline J. High level group analysis of fMRI data based on Dirichlet process mixture models. *Inf Process Med Imaging* 2007, 20:482–494.
60. Jbabdi S, Woolrich M, Behrens T. Multiple-subjects connectivity-based parcellation using hierarchical Dirichlet process mixture models. *Neuroimage* 2009, 440:373–384.
61. Friston KJ. Functional and effective connectivity: a review. *Brain Connectivity* 2011, 10:13–36.
62. Zalesky A, Fornito A, Bullmore E. On the use of correlation as a measure of network connectivity. *Neuroimage* 2012, 60:2096–2106.
63. Andersen AH, Gash DM, Avison MJ. Principal component analysis of the dynamic response measured by fMRI: a generalized linear systems framework. *Magn Reson Imaging* 1999, 170:795–815.
64. Calhoun VD, Adali T, Pearlson GD, Pekar JJ. A method for making group inferences from functional MRI data using independent component analysis. *Hum Brain Mapp* 2001, 140:140–151.
65. McKeown M, Makeig S, Brown G, Jung T, Kindermann S, Bell A, Sejnowski T. Analysis of fMRI data by blind separation into independent spatial components. *Hum Brain Mapp* 1998, 6:160–188.
66. Cribben I, Haraldsdottir R, Atlas LY, Wager TD, Lindquist MA. Dynamic connectivity regression: determining state-related changes in brain connectivity. *Neuroimage* 2012, 61:907–920.
67. Varoquaux G, Gramfort A, Poline JB, Thirion B, Zemel R, Shawe-Taylor J. Brain covariance selection: better individual functional connectivity models using population prior. In: Zemel R, Shawe-Taylor J, eds. *Advances in Neural Information Processing Systems*. Vancouver, Canada: Curran Associates, Inc.; 2010: 2334–2342.
68. Patel R, Bowman F, Rilling J. A Bayesian approach to determining connectivity of the human brain. *Hum Brain Mapp* 2006, 270:462–470.
69. Patel R, Bowman F, Rilling J. Determining hierarchical functional networks from auditory stimuli fMRI. *Hum Brain Mapp* 2006, 270:462–470.
70. Zhang J, Li X, Li C, Lian Z, Huang X, Zhong G, Zhu D, Li K, Jin C, Hu X. Inferring functional interaction and transition patterns via dynamic Bayesian variable partition models. *Hum Brain Mapp* 2014, 35:3314–3331.
71. Büchel C, Friston KJ. Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cereb Cortex* 1997, 70:768–778.
72. McIntosh A, Gonzalez-Lima F. Structural equation modeling and its application to network analysis in functional brain imaging. *Hum Brain Mapp* 1994, 20:2–22.
73. Friston KJ, Harrison L, Penny W. Dynamic causal modelling. *Neuroimage* 2003, 190:1273–1302.
74. Harrison L, Penny W, Friston KJ. Multivariate autoregressive modeling of fMRI time series. *Neuroimage* 2003, 190:1477–1491.
75. Zheng X, Rajapakse JC. Learning functional structure from fMR images. *Neuroimage* 2006, 310:1601–1613.
76. Scheines R, Hoijtink H, Boomsma A. Bayesian estimation and testing of structural equation models. *Psychometrika* 1999, 64:37–52.
77. Granger C. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 1969, 36:424–438.
78. Smith SM, Bandettini PA, Miller KL, Behrens TEJ, Friston KJ, David O, Liue T, Woolrich M, Nichols TE. The danger of systematic bias in group-level FMRI-lag-based causality estimation. *Neuroimage* 2012, 59:1228–1229.
79. Smith SM, Miller KL, Salimi-Khorshidi G, Webster M, Beckmann C, Nichols T, Ramsey J, Woolrich M. Network modeling methods for FMRI. *Neuroimage* 2011, 54:875–891.
80. Friston K. Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biol* 2009, 7:e33.
81. Havlicek M, Jan J, Brazdil M, Calhoun VD. Dynamic Granger causality based on Kalman filter for evaluation of functional network connectivity in fMRI data. *Neuroimage* 2010, 53:65–77.
82. Bhattacharya S, Ho M, Purkayastha S. A Bayesian approach to modeling dynamic effective connectivity with fMRI data. *Neuroimage* 2006, 30:794–812.
83. Bhattacharya S, Maitra R. A nonstationary non-parametric Bayesian approach to dynamically modeling effective connectivity in functional magnetic resonance imaging experiments. *Ann Appl Stat* 2011, 50:1183–1206.

84. Yu Z, Ombao H, Prado R, Burke E. A Bayesian model for activation and connectivity in task-related fMRI data. Submitted for publication.
85. Gorrostiti C, Fiecas M, Ombao H, Burke E, Cramer S. Hierarchical vector auto-regressive models and their applications to multi-subject effective connectivity. *Front Comput Neurosci* 2013, 7.
86. Ryali S, Supekar K, Chen T, Menon V. Multivariate dynamical systems models for estimating causal interactions in fMRI. *Neuroimage* 2011, 54:807–823.
87. Kim D, Burge J, Lane T, Pearlson G, Kiehl K, Calhoun VD. Hybrid ICA-Bayesian network approach reveals distinct effective connectivity differences in schizophrenia. *Neuroimage* 2008, 420:1560–1568.
88. Li J, Wang Z, Palmer S, McKeown M. Dynamic Bayesian network modeling of fMRI: a comparison of group-analysis methods. *Neuroimage* 2008, 410:398–407.
89. Li R, Chen K, Fleisher A, Reiman E, Yao L, Wu X. Large-scale directional connections among multi resting-state neural networks in human brain: a functional MRI and Bayesian network modeling study. *Neuroimage* 2011, 560:1035–1042.
90. Li R, Wu X, Chen K, Fleisher A, Reiman E, Yao L. Alterations of directional connectivity among resting-state networks in Alzheimer disease. *Am J Neuroradiol* 2012, 340:340–345.
91. Rajapakse J, Zhou J. Learning effective brain connectivity with dynamic Bayesian networks. *Neuroimage* 2007, 370:749–760.
92. Stephan K, Harrison L, Kiebel S, David O, Penny W, Friston KJ. Dynamic causal models of neural system dynamics: current state and future extensions. *J Biosci* 2007, 320:129–144.
93. Valdes-Sosa PA, Roebroeck A, Daunizeau J, Friston KJ. Effective connectivity: influence, causality and biophysical modeling. *Neuroimage* 2011, 58:339–361.
94. Daunizeau J, Friston KJ, Kiebel S. Variational Bayesian identification and prediction of stochastic nonlinear dynamic causal models. *Physica D* 2009, 2380:2089–2118.
95. Friston KJ, Penny W. Post hoc Bayesian model selection. *Neuroimage* 2009, 560:2011.
96. Friston KJ, Li B, Daunizeau J, Stephan K. Network discovery with DCM. *Neuroimage* 2011, 560:1202–1221.
97. Li B, Daunizeau J, Stephan K, Penny W, Hu D, Friston KJ. Generalised filtering and stochastic DCM for fMRI. *Neuroimage* 2011, 580:442–457.
98. Stephan K, Kasper L, Harrison L, Daunizeau J, den Ouden H, Breakspear M, Friston KJ. Nonlinear dynamic causal models for fMRI. *Neuroimage* 2008, 420:649–662.
99. Guo Y, Bowman FD, Kilts C. Predicting the brain response to treatment using a Bayesian hierarchical model with application to a study of schizophrenia. *Hum Brain Mapp* 2008, 290:1092–1109.
100. Derado G, Bowman FD, Zhang L. Predicting brain activity using a Bayesian spatial model. *Stat Methods Med Res* 2013, 220:382–397.
101. Michel V, Eger E, Keribin C, Thirion B. Multiclass sparse Bayesian regression for fMRI-based prediction. *J Biomed Imaging* 2011, 2011.
102. van Gerven MAJ, Cseke B, de Lange FP, Heskes T. Efficient Bayesian multivariate fMRI analysis using a sparsifying spatio-temporal prior. *Neuroimage* 2010, 500:150–161.
103. Morcom A, Friston KJ. Decoding episodic memory in ageing: a Bayesian analysis of activity patterns predicting memory. *Neuroimage* 2012, 590:1772–1782.
104. Goldsmith J, Huang L, Crainiceanu CM. Smooth scalar-on-image regression via spatial Bayesian variable selection. *J Comput Graph Stat* 2014, 230:46–64.
105. Li F, Zhang T, Wang Q, Gonzalez MZ, Maresh EL, Coan J. Spatial Bayesian variable selection and grouping in high-dimensional scalar-on-image regressions. Submitted for publication.
106. Calhoun VD, Lemieux L. Neuroimage: special issue on multimodal data fusion. *Neuroimage*. In press.
107. Liu J, Calhoun VD. A review of multivariate analyses in imaging genetics. *Front Neuroinform* 2014, 8:Article 29.
108. Thompson PM, Ge T, Glahn DC, Jahanshad N, Nichols TE. Genetics of the connectome. *Neuroimage* 2013, 80:475–488.
109. Niedermeyer E, da Silva F. *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Philadelphia: Lippincott Williams & Wilkins; 2005.
110. Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev Mod Phys* 1993, 650:413–497.
111. Le Bihan D, Mangin J, Poupon C, Clark C, Pappata S, Molko N, Chabriat H. Diffusion tensor imaging: concepts and applications. *J Magn Reson Imaging* 2001, 130:534–546.
112. Jorge J, van der Zwaag W, Figueiredo P. EEG-fMRI integration for the study of human brain function. *Neuroimage*. In press.
113. Kalus S, Sämann P, Czisch M, Fahrmeir L. fMRI activation detection with EEG priors. Technical Report 146, University of Munich, 2013.
114. Daunizeau J, Grova C, Mattout J, Marrelec G, Clonda D, Goulard B, Pélégriani-Issac M, Lina J, Benali H. Assessing the relevance of fMRI-based prior in the EEG inverse problem: a Bayesian model comparison approach. *IEEE Trans Signal Process* 2005, 530:3461–3472.

115. Henson R, Flandin G, Friston KJ, Mattout J. A parametric empirical Bayesian framework for fMRI-constrained MEG/EEG source reconstruction. *Hum Brain Mapp* 2010, 310:1512–1531.
116. Jun S, George J, Kim W, Paré-Blagoev J, Plis S, Ranken D, Schmidt D. Bayesian brain source imaging based on combined MEG/EEG and fMRI using MCMC. *Neuroimage* 2008, 400:1581–1594.
117. Mattout J, Phillips C, Penny W, Rugg M, Friston KJ. MEG source localization under multiple constraints: an extended Bayesian framework. *Neuroimage* 2006, 300:753–767.
118. Phillips C, Mattout J, Rugg M, Maquet P, Friston KJ. An empirical Bayesian solution to the source reconstruction problem in EEG. *Neuroimage* 2005, 240:997–1011.
119. Sato M, Yoshioka T, Kajihara S, Toyama K, Goda N, Doya K, Kawato M. Hierarchical Bayesian estimation for MEG inverse problem. *Neuroimage* 2004, 230:806–826.
120. Babajani-Feremi A, Bowyer S, Moran J, Elisevich K, Soltanian-Zadeh H. Variational Bayesian framework for estimating parameters of integrated E/MEG and fMRI model. In: *Proceedings of SPIE 7262, Medical Imaging: Biomedical Applications in Molecular, Structural, and Functional Imaging*, Orlando, FL, 2009;72621T-1–72621T-11.
121. Daunizeau J, Grova C, Marrelec G, Mattout J, Jbabdi S, Pélégrini-Issac M, Lina J, Benali H. Symmetrical event-related EEG/fMRI information fusion in a variational Bayesian framework. *Neuroimage* 2007, 360:69–87.
122. Ou W, Nummenmaa A, Ahveninen J, Belliveau J, Hämäläinen M, Golland P. Multimodal functional imaging using fMRI-informed regional EEG/MEG source estimation. *Neuroimage* 2010, 520:97–108.
123. Luessi M, Babacan S, Molina R, Booth J, Katsaggelos A. Bayesian symmetrical EEG/fMRI fusion with spatially adaptive priors. *Neuroimage* 2011, 550:113–132.
124. Zhu D, Zhang T, Jiang X, Hu X, Chen H, Yang N, Lv J, Han J, Guo L, Liu T. Fusing DTI and fMRI data: a survey of methods and applications. *Neuroimage*. In press.
125. Bowman FD, Zhang L, Derado G, Chen S. Determining functional connectivity using fMRI data with diffusion-based anatomical weighting. *Neuroimage* 2012, 62:1769–1779.
126. Damoiseaux JS, Greicius MD. Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity. *Brain Struct Funct* 2009, 213:525–533.
127. Greicius MD, Supekar K, Menon V, Dougherty RF. Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cereb Cortex* 2009, 19:72–78.
128. Iyer SP, Shafran I, Grayson D, Gates K, Nigg JT, Fair DA. Inferring functional connectivity in MRI using Bayesian network structure learning with a modified PC algorithm. *Neuroimage* 2013, 75: 165–175.
129. Chen J, Calhoun V, Pearlson G, Ehrlich S, Turner J, Ho B, Wassink T, Michael A, Liu J. Multifaceted genomic risk for brain function in schizophrenia. *Neuroimage* 2012, 610:866–875.
130. Liu J, Pearlson G, Windemuth A, Ruano G, Perrone-Bizzozero N, Calhoun V. Combining fMRI and SNP data to investigate connections between brain function and genetics using parallel ICA. *Hum Brain Mapp* 2009, 300:241–255.
131. Vounou M, Nichols TE, Montana G. Discovering genetic associations with high-dimensional neuroimaging phenotypes: a sparse reduced-rank regression approach. *Neuroimage* 2010, 530:1147–1159.
132. Stingo F, Guindani M, Vannucci M, Calhoun V. An integrative Bayesian modeling approach to imaging genetics. *J Am Stat Assoc* 2013, 1080:876–891.
133. Salazar E, Nikolova Y, Hariri AR, Carin L. A Bayesian framework for joint analysis of heterogeneous neuroscience data. Submitted for publication.
134. Glahn DC, Winkler AM, Kochunov P, Almasy L, Duggirala R, Carless MA, Curran JC, Olvera RL, Laird AR, Smith SM. Genetic control over the resting brain. *Proc Natl Acad Sci USA* 2010, 1070: 1223–1228.
135. van den Berg SM, Setiawan A, Bartels M, Polderman T, van der Vaart AW, Boomsma DI. Individual differences in puberty onset in girls: Bayesian estimation of heritabilities and genetic correlations. *Behav Genet* 2006, 360:261–270.
136. Bishop CM. *Pattern Recognition and Machine Learning*. New York: Springer; 2006.
137. Rue H, Maritno S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrate nested laplace approximations. *J R Stat Soc Ser B* 2009, 71:319–392.
138. Woolrich MW, Jbabdi S, Patenaude B, Chappell M, Makni S, Behrens T, Beckmann C, Jenkinson M, Smith S. Bayesian analysis of neuroimaging data in fsl. *Neuroimage* 2009, 450:S173–186.
139. Kang J, Johnson TD, Nichols TE, Wager TD. Meta analysis of functional neuroimaging data via Bayesian spatial point processes. *J Am Stat Assoc* 2011, 1060: 124–134.
140. Kang J, Nichols TE, Wager TD, Johnson TD. A Bayesian hierarchical spatial point process model for multi-type neuroimaging meta analysis. *Ann Appl Stat*. In press.
141. Yue YR, Lindquist MA, Loh J. Meta-analysis of functional neuroimaging data using Bayesian nonparametric binary regression. *Ann Appl Stat* 2012, 6: 697–718.