Published in final edited form as:
Stat Methods Med Res. 2013 August ; 22(4): 382–397. doi:10.1177/0962280212448972.

Predicting Brain Activity using a Bayesian Spatial Model

Gordana Derado[*], **F. DuBois Bowman**, and **Lijun Zhang**
Center for Biomedical Imaging Statistics, Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, 1518 Clifton Road, Atlanta, GA 30322, USA

Abstract

Increasing the clinical applicability of functional neuroimaging technology is an emerging objective, e.g., for diagnostic and treatment purposes. We propose a novel Bayesian spatial hierarchical framework for predicting follow-up neural activity based on an individual's baseline functional neuroimaging data. Our approach attempts to overcome some shortcomings of the modeling methods used in other neuroimaging settings, by borrowing strength from the spatial correlations present in the data. Our proposed methodology is applicable to data from various imaging modalities including functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) and we provide an illustration here using PET data from a study of Alzheimer's disease to predict disease progression.

Keywords

Neuroimaging; Bayesian spatial modeling; prediction; Alzheimer's disease

1. Introduction

Functional neuroimaging quantifies noninvasive measures of neurophysiology and neuroreceptor binding and helps to define the neural basis of illnesses and risk factors for major psychiatric disorders such as depression, schizophrenia and Alzheimer's disease (AD). The clinical capabilities of neuroimaging tools for guiding treatment decisions for such disorders, however, have not been fully established. There is emerging interest in using functional neuroimaging to guide treatment selections for individual patients and to predict the progression of disease, prompting the need to develop statistical methodology that would provide clinicians with predictive information about patients' brain activity. Such methods may assist clinicians in making treatment decisions by forecasting post-treatment neural activity. In studying the progression of dementia, they can identify preclinical changes that may predict, for example, the onset of Alzheimer's disease.

Functional brain imaging, such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET), has only recently been used to predict brain and clinical outcomes in individual patients. [1] proposed a predictive statistical model for PET and fMRI data, using a Bayesian hierarchical framework, that uses patient's pretreatment

[*]Address for correspondence: Gordana Derado, Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, 1518 Clifton Road, Atlanta, GA 30322, USA. gderado@emory.edu.

scans, coupled with relevant patient characteristics, to predict brain activity in schizophrenic patients after a specified treatment regimen. [2] used relevance vector regression (RVR) to predict clinical scores from individual scans. In particular, they used individuals' MRI T1 weighted images to predict their performances on established tests used in the evaluation of Alzheimer's disease. Predicted and actual clinical scores were highly correlated. Their analysis uses only structural MRI, rather than neural activity derived from fMRI or PET scans.

Recently, Gaussian processes (GPs), based on Bayesian theory, emerged as an alternative to Support Vector Machines (SVM). A GP is a generalization of the multivariate Gaussian distribution to infinitely many dimensions, with the distribution of the process at any finite number of points having a multivariate normal distribution. [3] evaluate the predictive capability of GP models for two types of quantitative prediction: multivariate regression and probabilistic classification, using whole-brain fMRI volumes from a study investigating subjective responses to thermal pain. They show that GP regression models outperform support vector and relevance vector regression.

Although each of these methods represents an important contribution to predicting treatment outcome or brain activity based on neuroimaging data, none uses the spatial information from the neighboring voxels/regions to improve the prediction accuracy. In addition to background spatial correlations inherent in neuroimaging data, functional neuroimaging data naturally exhibit correlations due to underlying functional connectivity. We propose a model that borrows strength from such correlations, with the goal of improving prediction.

[4] and [5] provide a spatial Bayesian hierarchical model for analyzing functional neuroimaging data that establishes a unified framework to obtain neuroactivation inferences as well as task-related functional connectivity inferences. The model combines whole-brain voxel-by-voxel modeling and ROI analyses. An unstructured variance/-covariance matrix for regional mean parameters allows for the study of inter-regional (long-range) correlations, and the model employs an exchangeable correlation structure to capture intraregional (shorter-range) correlations. A major limitation of this approach for prediction, however, is that it does not capture temporal correlations between the brain activity in repeated scanning sessions.

We propose a novel Bayesian hierarchical framework for predicting follow-up neural activity based on the baseline functional neuroimaging data that attempts to overcome some shortcomings of the modeling methods used in other neuroimaging settings by borrowing strength from the spatial correlations present in the data.

Several types of correlations are addressed at various levels of our hierarchical model. Correlations between repeated scanning sessions are addressed through regional covariance matrices between the baseline and follow-up glucose uptake. Long-range spatial correlations between anatomical brain regions are incorporated through the unstructured covariance matrix of the random effect parameters. Intermediate-range correlations are incorporated through within region (exchangeable) random effects structure, and short-range correlations

between neighbouring voxels are incorporated through a multivariate conditional autoregressive (CAR) structure within first–order neighborhood.

We apply our method to data from a study of Alzheimer's disease and compare it to other competing methods. We also present results from a simulation study to evaluate the accuracy of our estimation procedure.

2. Experimental Data

We use data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) study (http://www.loni.ucla.edu/ADNI/). The goal of this national multi-center project is to develop biomarkers of Alzheimer's Disease (AD) in elderly subjects. We consider [^{18}F]-2-uoro-2-deoxy-2-glucose (FDG) PET scans obtained at baseline (screening) and 6 months. FDG is an analogue of glucose, and in PET it yields concentrations of the injected tracer indicating tissue metabolic activity, in terms of regional glucose uptake. For more details about the ADNI, see [6].

Participants are classified as having mild cognitive impairment (MCI), as Alzheimer's disease (AD) patients, or as healthy controls (HC). The data from 40 AD and 40 HC subjects were used in the training step of the prediction model development. The model is then applied for predicting the month 6 follow-up PET scan, based on the baseline scan, for an additional group of 33 AD and 33 HC subjects.

The processing steps for the PET scans are as follows. (1) Co-registration: In most cases, six five-minute frames are acquired 30 to 60 minutes post-injection. Each extracted frame is co-registered to the first extracted frame of the raw image file. (2) Averaging: the six five-minute frames of the co-registered image set are averaged to create a single 30 min PET image. (3) Standardizing image and voxel size: Each subject's co-registered, averaged image from their baseline PET scan is reoriented into a standard 160 × 160 × 96 voxel image grid, having 1.5 mm cubic voxels. This standardized image then serves as a reference image for all PET scans on that subject. (4) Spatial smoothing: Each image set is filtered with a scanner-specific filter function (can be a non-isotropic filter) to produce images of a uniform isotropic resolution of 8 mm FWHM, the approximate resolution of the lowest resolution scanners used in ADNI. In addition to the above preprocessing steps, we performed a spatial normalization to a standard 91 × 109 × 91 MNI space (see [7]).

The following covariates were included in our analysis: the Alzheimer's Disease Assessment Scale – cognitive subscale [ADAS-Cog] and the subjects' ages (in years). ADAS was designed to measure the severity of the most important symptoms of AD. Its subscale ADAS-cog consists of 11 tasks measuring the disturbances of memory, language, praxis, attention and other cognitive abilities, which are often referred to as the core symptoms of AD.

3. Methodology

We propose a novel Bayesian hierarchical framework for predicting follow-up (or post-treatment) neural activity based on the baseline (or pre-treatment) functional neuroimaging

data. Our prediction model borrows strength from the spatial correlations present in the data (both local, between-voxel, correlations and more long-range, between-region correlations). The model builds on the proper multivariate conditional autoregressive model (MCAR (ρ,Σ)) proposed in [8]. Our algorithm is similar to the predictive method of [1], but our proposed model incorporates spatial information between neighbouring regions, in addition to capturing correlations between repeated scans, thus providing additional information about neural processing. Our model can also be seen as an extension of a hierarchical model for functional neuroimaging data proposed by [4]. Using our proposed method, we analyze the PET data from the study of Alzheimer's disease described in Section 2.

3.1 Model and Estimation

We consider an anatomical parcellation of the brain consisting of $g = 1, \ldots, G$ regions, where we may set G to be as high as 116 ([7]). Alternative anatomical parcellations are also available, such as those based on Brodmann regions ([9]). Let $i = 1, \ldots, n$ denote subjects, $v = 1, \ldots, V$ voxels, and let V_g represent the number of voxels in a particular region indexed by g. We denote the regional glucose use, as a proxy for brain activity at voxel v, by $Y(v)$. Let $\mathbf{Y}_{ig}(v) = (Y_{ig}(v)^{(1)}, Y_{ig}(v)^{(2)})^T$, where superscripts (1) and (2) denote the baseline and follow-up scans, respectively. Before processing, a local neighbourhood of each voxel included in the analysis must be selected, for the voxel-level spatial modeling. We apply a three-dimensional (3D) neighborhood structure containing the 26 immediate neighboring voxels: 9 above, 9 below, and 8 adjacent. In addition, we consider only within-region neighbours. For each voxel, the neighbourhood information is saved in a connectivity matrix W. $W = \{w_{vv'}\}_{(V\times V)}$ is a symmetric matrix (i.e. $w_{vv'} = 1$ if voxel v' is in the defined 3D neighborhood of v), and w_{v+} is the sum of the elements in row v of W.

We propose a multivariate Bayesian hierarchical model that accounts for both spatial correlations between intra-regional voxels, and between regions. The model also accounts for correlations between baseline and follow-up regional glucose use. In particular, our proposed model has the following hierarchical structure:

$$
\begin{gathered}
\boldsymbol{Y}_{ig}(v)|\boldsymbol{\beta}_g, \boldsymbol{\phi}_g, \boldsymbol{\alpha}_{ig}, \boldsymbol{\gamma}_{gv}, \boldsymbol{Z}_g \sim \mathrm{N}(\boldsymbol{\beta}_g(v)+\boldsymbol{\phi}_g(v)+\boldsymbol{\alpha}_{ig}+\mathbf{X}_{ig}\boldsymbol{\gamma}_g, \boldsymbol{Z}_g) \\
\boldsymbol{\phi}_v|\boldsymbol{\phi}_{v'}, v' \neq v, \sum, v{=}1\ldots, V \sim \mathrm{N}(\rho \sum_{w} \tfrac{w_{vv'}}{w_{v+}} \mathbf{I}\boldsymbol{\phi}_{v'}, \tfrac{1}{w_{v+}}\sum) \quad (\mathrm{MCAR}(\rho, \sum)) \\
\boldsymbol{\beta}_{gj}|\lambda^2_{gj} \sim \mathrm{N}(\boldsymbol{\beta}_{0gj}, \lambda^2_{gj}\mathbf{I}) \quad (\lambda_{vgj}{=}\lambda_{gj}, \forall v \in \text{region } g) \\
\boldsymbol{Z}_g^{-1} \sim \mathrm{Wishart}((c_1\boldsymbol{\Omega}_1)^{-1}, c_1) \\
\textstyle\sum^{-1} \sim \mathrm{Wishart}((c_2\boldsymbol{\Omega}_2)^{-1}, c_2) \\
\boldsymbol{\alpha}_{ij}|\boldsymbol{\Gamma}_j \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Gamma}_j) \quad (\boldsymbol{\alpha}_{ij}{=}\boldsymbol{\alpha}_i^{(j)}) \\
(\boldsymbol{\Gamma}_j)^{-1} \sim \mathrm{Wishart}\{(h_j H_j)^{-1}, h_j\} \;\; j{=}1,2 \\
\lambda_{gj}^{-2} \sim \mathrm{Gamma}(a_j, b_j) \\
\gamma_{gjq}|\tau^2_{gjq} \sim \mathrm{N}(0, \tau^2_{gjq}) \quad q{=}1, \ldots, Q \text{ (covariates)} \\
\tau_{gjq}^{-2} \sim \mathrm{Gamma}(e_0, f_0) \\
\rho \sim \mathrm{Uniform}(\{0, 0.05, 0.1, \ldots, 0.8, 0.81, \ldots, 0.9, 0.91, \ldots, 0.99\})
\end{gathered}
\tag{1}
$$

where j denotes the scanning session (i.e., $j = 1$ − baseline and $j = 2$ − follow-up).

For each voxel v, the subject-specific quantities $\mathbf{Y}_{ig}(v)$ are assumed to vary randomly about a mean determined by a population-level mean parameter $\boldsymbol{\beta}_g(v) = (\boldsymbol{\beta}_g(v)^{(1)}, \boldsymbol{\beta}_g(v)^{(2)})^T$, a

population-level spatial dependence parameter $\boldsymbol{\varphi}_g(v) = (\varphi_g(v)^{(1)}, \varphi_g(v)^{(2)})^T$, an individualized random effect component $\boldsymbol{\alpha}_{ig}=(\alpha_{ig}^{(1)}, \alpha_{ig}^{(2)})^T$ and the covariates' parameters $\boldsymbol{\gamma}_g=(\gamma_{g1}^{(1)}, \gamma_{g1}^{(2)}, \ldots, \gamma_{gQ}^{(1)}, \gamma_{gQ}^{(2)})^T$. $\mathbf{X}_{ig}=\begin{bmatrix} X_{ig1}^{(1)} & 0 & \cdots & X_{igQ}^{(1)} & 0 \\ 0 & X_{ig1}^{(2)} & \cdots & 0 & X_{igQ}^{(2)} \end{bmatrix}$ is a $2 \times 2Q$ matrix containing the mean-centered subject-specific covariates, where Q is the number of covariates in the model. We express vectors of random effects across the entire brain as $\boldsymbol{\alpha}_i^{(j)}=(\alpha_{i1}^{(j)}, \ldots, \alpha_{iG}^{(j)})$, where j denotes the baseline and follow-up sessions, and $\boldsymbol{\alpha}_i=(\boldsymbol{\alpha}_{i1}^T, \ldots, \boldsymbol{\alpha}_{iG}^T)^T$. Given $\{\boldsymbol{\beta}_g\}$, $\{\boldsymbol{\varphi}_g\}$, $\{\boldsymbol{\alpha}_{ig}\}$, $\{\boldsymbol{\gamma}_{gv}\}$ and $\mathbf{Z}_g$, the $\{\mathbf{Y}_{ig}\}$ are conditionally independent.

At the second level, the model expresses a prior belief that each voxel's population mean (for the j-th session) arises from a normal distribution with a mean given by the overall region mean $\boldsymbol{\beta}_{0gj}$ and variance λ_{gj}^2. It represents a reasonable starting point to assume that voxels within anatomically defined regions exhibit brain activity that deviates around an overall mean for that region.

Spatial associations are introduced through random effects in the mean structure of the data. Bivariate spatial random effects at the voxel level call for a multivariate CAR (MCAR) specification, where ρ is a scalar parameter representing the overall degree of spatial dependence and Σ is the covariance matrix between $\boldsymbol{\varphi}_v$ and $\boldsymbol{\varphi}_{v'}$. The connectivity (proximity) matrix W used in the MCAR prior for the spatial random effects $\boldsymbol{\varphi}$ is based on the 3D neighbourhood structure described above.

The parameter ρ determines the magnitude of the spatial neighbourhood effect. Following [8], we specify a discrete uniform prior for the spatial autoregression parameter ρ. We assume $\rho < 1$ to ensure propriety, we do not allow $\rho < 0$ since this would violate the similarity of spatial neighbours which we seek, and we place prior mass which favors the upper range of ρ. In particular, we put equal mass on the following 36 values: 0, 0.05, 0.1, …, 0.8, 0.81 0.82, …, 0.90, 0.91, 0.92, …, 0.99. In addition to the local MCAR correlations, we also have exchangeable within-region correlations introduced by the random effects structure.

The model also captures potential functional connections between anatomical brain regions through the covariance matrix $\boldsymbol{\Gamma}_j$. We specify a conjugate inverse-Wishart prior for $\boldsymbol{\Gamma}_j$, allowing a flexible unstructured covariance matrix. The multivariate structure enables us to compute the between-region connectivity matrix, separately for each session. Finally, hyperpriors on the MCAR parameters and other hyperparameters complete the model.

Estimation—Estimation is performed using Markov Chain Monte Carlo (MCMC) techniques implemented via Gibbs sampling. Applying MCMC methods in our context is complicated by the massive amount of data, the large number of spatial locations, and the large number of parameters that need to be estimated. The Gibbs-friendly model specification facilitates estimation by providing substantial reductions in computing time

and memory. We present the full conditionals required to run the Gibbs sampler in the Appendix.

3.2 Prediction

We use the summary statistics (e.g. mean) from the posterior samples to predict the session 2 (month 6) scans for future patients, based on their baseline scans. The prediction proceeds as follows. We first draw the covariance matrix $\boldsymbol{\Gamma}$ from its full conditional distribution. We then draw $\boldsymbol{\alpha}_i$'s, given $\boldsymbol{\Gamma}$, from the prior distribution for $\boldsymbol{\alpha}$. Letting $\mathbf{Y}_i{}^*{}_1$ denote the baseline scan and $\mathbf{Y}_i{}^*{}_2$ the month 6 scan for a future subject, we have the necessary information (parameter estimates for $\boldsymbol{\beta}$, $\boldsymbol{\varphi}$, $\boldsymbol{\Gamma}$, $\boldsymbol{\gamma}$, $\mathbf{Z}$, $\boldsymbol{\alpha}$ and the covariate matrix $\mathbf{X}$) to estimate $\mathbf{Y}_i{}^*{}_2|\mathbf{Y}_i{}^*{}_1$.

At the region level, let $\boldsymbol{Y}=(\boldsymbol{Y}_1^T, \boldsymbol{Y}_2^T)^T$, where $\mathbf{Y}_1$ contains all the session 1 voxel values, and $\mathbf{Y}_2$ all the session 2 voxel values. For region g, then

$\boldsymbol{Y}_g=(\boldsymbol{Y}_{g,1}^T, \boldsymbol{Y}_{g,2}^T)^T \sim \mathrm{N}((\boldsymbol{\mu}_{g,1}^T, \boldsymbol{\mu}_{g,2}^T)^T, \sum_g)$, where $\boldsymbol{\Sigma}_g = \mathbf{Z}_g \otimes \mathbf{I}_{v_g}$.

From model (1), it follows that $\mathbf{Y}_i{}^*{}_{,2}|\mathbf{Y}_i{}^*{}_{,1} \sim \mathrm{N}(\mathbf{b}_i{}^*{}_{g,}\ \mathbf{A}i^*{}_g)$. For region g,

$\mathbf{b}_{i*g}=\boldsymbol{\mu}_{i*g,2}+\sum_{12}^{T}\sum_{11}^{-1}(\boldsymbol{Y}_{i*g,1}-\boldsymbol{\mu}_{i*g,1})$, and $\boldsymbol{\mu}_i{}^*{}_g = \boldsymbol{\beta}_g + \boldsymbol{\varphi}_g + \mathbf{1}_{V_g} \otimes \boldsymbol{\alpha}_i{}^*{}_g + \mathbf{1}_{V_g} \otimes \mathbf{X}_i{}^*{}_{gv}$ $\boldsymbol{\gamma}_{gv}$. $\boldsymbol{\mu}_i{}^*{}_{g,2}$ is a vector consisting of even elements of $\boldsymbol{\mu}_i{}^*{}_g$. Writing $\boldsymbol{Z}_g=\begin{bmatrix} a & c \\ c & b \end{bmatrix}$, we see that $\mathbf{A}_{i*g}=b\mathbf{I}_{V_g}-c\mathbf{I}_{V_g}\cdot\frac{1}{a}\mathbf{I}_{V_g}\cdot c\mathbf{I}_{V_g}=(b-\frac{c^2}{a})\mathbf{I}_{V_g}$.

By inputting the posterior mean of the parameters obtained from the MCMC estimation, we obtain the estimated conditional mean $\hat{\mathbf{b}_i{}^*{}_g}$ and covariance matrix $\hat{\mathbf{A}}_i{}^*{}_g$. The follow-up regional glucose uptake $\mathbf{Y}_{g,2}$ is predicted using the mean of the estimated conditional distribution, i.e. $\hat{\mathbf{b}_i{}^*{}_g}$. Also, a $100(1-\alpha)\%$ prediction interval for $\mathbf{Y}_{g,2}$ can be constructed based on the estimated conditional variance.

3.3 Model Validation: Estimation of the prediction error

For each of the two groups (AD and HC) in the ADNI data, we applied our algorithm to predict the follow-up regional glucose uptake for 33 new (test) subjects using the estimated parameters from the training data. We quantify the prediction error by comparing the observed and predicted follow-up (month 6) brain activity. In neuroimaging, prediction of the brain activity is performed on each voxel, and the prediction error is evaluated across all voxels. Typical squared or absolute error functions are inappropriate since the brain activity measurement (such as regional glucose uptake, or BOLD signal) has different baseline values across the brain. We propose the use of a scale-free loss function so that the prediction error is comparable across all brain voxels. The proposed function is the ratio of the square root of the PMSE and the average effects at voxel v. Specifically, this function (we will refer to it as the "standardized square root of the prediction mean squared error – stPMSE") is defined as

$$\mathrm{stPMSE}(\{Y_i^{(2)}(v)\},\{\hat{Y}_i^{(2)}(v)\})=\frac{\sqrt{(1/N\sum_{i=1}^{N}[\hat{Y}_i^{(2)}(v)-Y_i^{(2)}(v)]^2}}{1/N\sum_{i=1}^{N}Y_i^{(2)}(v)}. \quad (2)$$

This measure of prediction error is similar to one proposed by [1]. We scale by an average of the observed, not the predicted values at voxel *v*, therefore, our measure will be more stable, e.g., against overestimated predicted values.

4. Results

We apply our Bayesian spatial hierarchical model to PET data from the study of Alzheimer's disease (ADNI). We considered 39 regions relevant to AD in our analysis, which were selected from a set of 46 regions of interest considered in the Alzheimer's disease study in [4]. The complete list of the regions with region sizes is given in Table A1, Supplementary Material, section 3. Areas of the temporal and limbic lobes are of particular interest, having been indicated as having either increased or decreased encoding activity between at-risk subjects and controls in the first wave of the Alzheimer's data or being thought to be involved with the (verbal memory) paradigm ([10]). We limit the number of regions in the analysis to be smaller than the number of subjects in the smallest group (in our case, 40) in order to reliably estimate the between-region covariance parameters.

Our model captures spatial correlations at three levels: a) the short-range spatial correlation between voxels within a defined anatomic region, b) exchangeable within-region correlations introduced by the random effects structure and c) the (potentially) long-range inter-regional connectivity, stemming from the covariance matrix $\mathbf{\Gamma}_j$, of the random effect parameters $\boldsymbol{\alpha}_{ij}$. We specify a Wishart prior for $\mathbf{\Gamma}_j^{-1}$.

For the inverse-Wishart prior, the degrees of freedom must satisfy $h_j \geq G$ to yield a proper prior distribution. This prior becomes more diffuse as h_j gets smaller [11]; hence, we set $h_j = G$ to reflect the most diffuse proper prior that our data can support. A seemingly natural choice for H_j is a point estimate of $\mathbf{\Gamma}_j$. We use the sample covariance matrix to obtain H_j, calculated from the subject-specific mean activity levels in each of the anatomical regions. In addition, we examine the sensitivity of our results to the sample covariance matrix by artificially reducing the correlations (covariances) using $H_j^*=(1-w)H_j+w\{diag(H_j)\}$. We also use inverse-Wishart priors for Σ and $\mathbf{Z}_g$, i.e.Σ^{-1} ~ Wishart $((c_2\mathbf{\Omega}_2)^{-1}, c_2)$, $\boldsymbol{Z}_g^{-1}\sim\mathrm{Wishart}((c_1\mathbf{\Omega}_1)^{-1}, c_1)$, where $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$ are 2 × 2 matrices and c_1 and c_2 are shape parameters. Since we have no prior knowledge regarding the nature and extent of dependence, we choose $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$ to be diagonal. The data will inform about dependence a posteriori. We set $c_1 = c_2 = 4$ to provide low precision for these priors.

To complete our Bayesian hierarchical model, we set $a_1 = a_2 = 0.1$, $b1 = 0.005$, $b_2 = 0.001$, and $h_j = G$, resulting in vague or weakly informative priors, to ensure that the information in the data primarily governs the results. However, more informative priors may be employed when fairly precise information is available.

To estimate the model parameters, we performed 3,000 iterations with burn-in of 2,000 iterations (5,000 iterations in total) and thinning of 5 iterations (for storage and computation time). The programming was done in Matlab. The estimation was performed on a Linux cluster. The test/experiment environment consisted of an 8-core system with 16GB of RAM. Each Matlab job was given a dedicated CPU and ran for approximately 26 hours. Future optimizations include taking advantage of the multi-threaded nature of several Matlab image analysis functions or using the Matlab Distributed Computing Environment to use multiple cores per job for speed improvements.

Due to a large number of parameters in our model, it is impractical to monitor the trace plots for each of the parameters. Also, running parallel chains is recommended to monitor how the chains mix. We obtained trace plots for a number of randomly selected voxels in several regions, for each of the voxel-level parameters. Several plots and histograms are given in Supplementary Material, section 3. The chains obtained from the Gibbs sampler (trace plots) are satisfactory, i.e., the generated Markov chains seem to converge, after the burn-in period, to our distributions of interest. We also monitored the trace plots for other (region-specific and scalar) parameters and found them to be satisfactory.

4.1 Prediction of brain activity for PET data from a study of Alzheimer's disease (ADNI)

We apply the proposed prediction algorithm to forecast the follow-up regional glucose uptake for each of the subjects in the test data set, which consists of 33 subjects in each group. Our algorithm provides individualized predictions of the regional glucose uptake, based on the unique information in each subject's baseline scan and relevant personal characteristics (e.g., ADAS-cog score). For prediction, we estimate the covariate effect at the second scanning session by the mean of the second session in the training data. For both groups (AD and HC), the posterior mean of the spatial effect ρ was 0.99, for both scanning sessions, indicating very strong spatial dependence. Also, estimated covariance matrices between the baseline and follow-up regional glucose uptake at a region level, $\mathbf{Z}_g$, yield high between-session correlations (> 0.7 for almost all regions) at a region level.

4.1.1 Individualized prediction maps—Figure 1 shows the individual prediction maps of the regional glucose uptake at 6 month follow-up for four selected AD patients. The prediction maps for several other subjects are given in Supplementary Material, section 3. We see that there are notable differences between subjects in the predicted follow-up activity, indicating that possibly different stages of the disease are present in those individuals. We compare the individual predicted maps (Figure 1 (a)) to the observed maps (Figure 1 (b)) and notice strong agreement between the observed and predicted brain activity. Similar correspondence is observed for most of the subjects. The individual prediction maps, such as those in Figure 1 (a), highlight possible individual differences in the progression of the disease.

We also predict the follow-up brain activity for the healthy controls test data set, which consisted of 33 subjects. The results for four selected subjects are given in Figure 2. Similar satisfactory agreement is observed between the predicted and observed maps. We also notice that this group exhibits smaller between-subject differences than the AD group.

4.2 Comparisons with competing prediction models

We compare our prediction results with the results obtained using three proposed competing methods: the General Linear Model (GLM) [12] based method, the method based on the Bayesian hierarchical model (BHM) proposed in [1], and the method based on the Bayesian spatial hierarchical model for activation and connectivity analysis (BSMac) proposed in [4] and [5].

Comparison with predictions based on the GLM: The GLM models the brain activity for all subjects using common population parameters. Independence and sphericity between scans at baseline and at follow-up scans are assumed. Estimates from the GLM are obtained using OLS. The predicted follow-up brain activity reflects only the population-level expectation and does not take into account the information from the subject's pretreatment scans. That is, $\boldsymbol{Y}^{*}_{g,2}=\mathbf{X}^{*}_{g,2}\cdot\hat{\boldsymbol{\beta}}_{g}$, where subscript 2 denotes session 2 (follow-up) brain activity at region g, superscript * denotes the new subject, and $\hat{\boldsymbol{\beta}_{g}}$ denotes the estimate of the session 2 mean parameter obtained from the training data set. Figure 3(d) displays the square root of the PMSE, relative to the average brain activity, based on the GLM.

Comparison with prediction based on BSMac: The prediction algorithm based on the BSMac is defined in a similar way to the algorithm based on our proposed BSPM. Since BSMac does not estimate the correlations between baseline and follow-up brain activity scans ($\mathbf{Y}_i$'s), i.e., temporal correlations, we estimate those by the sample covariances, at a voxel level. The local spatial parameters are also not estimated, and are dropped from the estimated conditional mean. Figure 3(c) displays the square root of the PMSE, relative to the average brain activity, based on the BSMac.

Comparison with predictions based on the BHM: The prediction based on the Bayesian hierarchical model proposed in [1] enables individualized follow-up predictions of brain response by incorporating the unique information in each subject's baseline functional brain scans and other relevant personal characteristics (covariates). The model also captures the correlations between baseline and follow-up brain activity scans, which enables a similar prediction algorithm to ours. Figure 3(b) displays the square root of the PMSE, relative to the average brain activity, based on the BHM.

A comparison between Figure 3(a) and Figure 3(d) indicates that prediction errors based on our proposed model are lower than those from the GLM, on average. The average error (total sum over all voxels, divided by the number of voxels included in the analysis), for the model based on the GLM is 0.156. Also, the superiority of our prediction model is consistently observed across the brain. A comparison between Figure 3(a) and Figure 3(c) indicates that the prediction errors based on our proposed model are also lower than those from the BSMac, on average. We also notice improvement over the GLM model. The average errors are 0.083 for the BSPM, 0.080 for BHM, 0.104 for the BSMac, and 0.156 for the GLM based prediction method.

Although, for some subjects in this study, there is a strong similarity between the baseline and the month 6 scans, our model draws more predictive power than the prediction based on

the baseline data only. The average error (the whole-brain accuracy) for the 'baseline only prediction' (calculated using the stPMSE (2)) is 0.085, compared to 0.083 for our method. Likely, larger differences would occur for the month 24 follow-up images when compared to baseline. We give more details and show figures with the baseline, the observed and the predicted follow-up (month 6) images, for some selected subjects, in the Supplementary Material.

A comparison between Figure 3(a) and Figure 3(b) indicates that, for these experimental data, the prediction errors based on our proposed model and the model proposed in [1] are very comparable (the difference in average errors is in the third decimal place). However, an additional advantage of our approach is that using our model, for example, inter-regional functional correlations can be estimated through the covariance matrix $\mathbf{\Gamma}_j$. The functional connectivity maps in Figure 4 display selected regions in a brain, with lines connecting regions that have (posterior median) correlations exceeding 0.75. The line thicknesses depends on the strengths of the inter-regional correlations.

5. Simulation

We performed a small simulation study to evaluate the accuracy of our estimation procedure. To generate data for the simulations, we selected 5 regions from the Automated Anatomical Labeling (AAL) map [7] (regions 79, 73, 15, 13, and 85) with sizes ranging from 234 to 4,655 voxels and simulated data for 15 subjects. We used our experimental data described in Section 2 to specify the *true* values for $\boldsymbol{\beta}_g$, spatial parameters $\boldsymbol{\varphi}_g$, random effects parameters $\boldsymbol{\alpha}_i$, and the covariate parameters $\boldsymbol{\gamma}_g$. The rest of the (hyper)parameters were drawn from the prior distributions specified in the model. After all the parameters that define the mean and the covariance parameters in model (1) were specified, 200 simulated data sets were drawn, with the specified mean parameters and the variance-covariance matrix. We then applied our estimation procedure to each of the data sets and analyzed the posterior distributions of the parameters. Table 5 summarizes the results for the parameters of the covariance matrix $\mathbf{Z}_g$, for each of the 5 selected regions. We see that the biases for the covariance matrix parameters $\mathbf{Z}_g$ are small, for each of the regions. Also, the spatial parameter ρ has only a small bias: we set the true value in the simulated data sets to 0.9 and the mean of the estimated ρ, across 200 data sets, was found to be 0.9134 (for each data set, we first find a median of the parameter's posterior distribution, then average those medians). Posterior means of the estimated voxel-level parameters (for randomly selected 25 voxels in each region), as well as the posterior means of the random effects $\boldsymbol{\alpha}_{ij}$ are presented in the Supplementary Material, section 4.

Parameters that are most relevant to the prediction ($\boldsymbol{\beta}$'s, $\boldsymbol{\varphi}$'s, $\boldsymbol{\alpha}$'s, $\mathbf{Z}_g$) have small or relatively small biases, while the parameters that contribute to the prediction less directly, have larger biases.

6. Discussion

In this paper we describe a framework for spatiotemporal modeling of functional neuroimaging data that provides important advantages over some other methods. We propose a novel method for predicting post-baseline (or follow-up) brain scans, based on the

baseline scans. The prediction algorithm is based on a novel Bayesian spatial hierarchical model. Our method is applicable in many clinical situations (e.g., for predicting the progression of a disease, or to predict after treatment brain activity, based on the baseline (pre-treatment) activity). The proposed method may be useful in clinical situations where it is too costly to acquire multiple (repeated) scans on the same subjects.

Our model builds on the predictive model proposed by [1], and a spatial Bayesian hierarchical model proposed by [4], by incorporating a proper multivariate CAR prior for the spatial effect. The proposed model captures the short-range correlations between voxels within a defined anatomical region as well as the (potentially) long-range inter-regional correlations, which provide information about functional connectivity between the brain regions. We consider a 3D neighbourhood structure for estimation of the local spatial associations.

Based on the proposed model, we formulate a prediction algorithm for the follow-up brain activity, based on an individual's baseline functional neuroimaging data and relevant subject characteristics. By borrowing strength from the spatial information, we achieved improved prediction as compared to a couple of competitive prediction models (GLM and BSMac). Our prediction method is comparable to the method by [1], but it offers some additional advantages, e.g., estimation of long-range inter-regional spatial correlations through the covariance matrix $\mathbf{\Gamma}_j$ (Figure 4).

We apply our Bayesian spatial hierarchical model to the PET data from the study of Alzheimer's disease, but the same methodology can easily be applied to data from an fMRI study. In that case, individual summary statistics (i.e., regression coefficients) from a typical first stage GLM based fMRI data analysis would first be obtained (in practice often obtained using software packages such as SPM or FSL), and then used as **Y**'s in our model (1).

One of the limitations of our proposed method is that the computing time is very long (even for relatively small number of iterations of 3,000) for the estimation part of the algorithm. However, this step is done one time, and the prediction step is very fast (it is in order of seconds per one individual). Computing time can be improved by implementing some parallel computing steps (as described in Section 4). Another limitation is that we consider only within-region local 3D neighbourhood in our model, which results in some region border effects in the prediction maps. We will consider across-region local 3D neighbourhood in our future work. Also, we propose a separable bivariate spatial model for computational reasons and simplicity. Using two separate spatial effect parameters could be considered in future analyzes to see whether these differ for different scanning sessions.

Some recent studies ([13], [14], [15]) found that the β-amyloid protein 1–42, total tau protein, and phosphorylated $\text{tau}_{181\text{P}}$ protein concentrations, each derived from cerebral spinal fluid in the brain, may be clinically relevant biological markers for the differential diagnosis of AD. These biological measures may, therefore, serve as potentially useful covariates (predictors) in our model. Some of these proteins were collected in the ADNI study, but the rate of missing data (among subjects with FDG-PET scans) was too substantial for inclusion in our analysis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors are grateful to Dr. Ying Guo from the Department of Biostatistics and Bioinformatics at Emory University for valuable discussions and constructive comments. This research was supported by NIH grants R01-MH079251 (Bowman) and NIH predoctoral training grant T32 GM074909-01 (Derado).

References

1. Guo Y, Bowman FD, Kilts C. Predicting the brain response to treatment using a Bayesian Hierarchical model with application to a study of schizophrenia. Human Brain Mapping. 2008; 29:1092–1109. [PubMed: 17924543]
2. Stonnington CM, Chu C, Klöppel S, Jack CRJ, Ashburner J, Frackowiak RS, et al. Predicting clinical scores from magnetic resonance scans in Alzheimer's disease. NeuroImage. 2010; 51(4): 1405–13. [PubMed: 20347044]
3. Marquand A, MH, Brammer M, Chu C, Coen S, Mourão Miranda J. Quantitative prediction of subjective pain intensity from whole-brain fMRI data using Gaussian processes. NeuroImage. 2010; 49(3):2178–89. [PubMed: 19879364]
4. Bowman FD, Caffo B, Bassett SS, Kilts C. A Bayesian hierarchical framework for spatial modeling of fMRI data. NeuroImage. 2008; 39:146–156. [PubMed: 17936016]
5. Zhang L, Agravat S, Derado G, Chen S, McIntosh B, Bowman F. BSMac: A MATLAB Toolbox Implementing a Bayesian Spatial Model for Brain Activation and Connectivity. Journal of Neuroscience Methods. 2012; 204:133–143. [PubMed: 22101143]
6. Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack C, Jagust W, et al. The Alzheimers Disease Neuroimaging Initiative. Neuroimaging Clinics of North America. 2005:15.
7. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labelling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single subject brain. NeuroImage. 2002; 15:273–289. [PubMed: 11771995]
8. Gelfand AE, Vounatsou P. Proper multivariate conditional autoregressive models for spatial data analysis. Biostatistics. 2003; 4(1):11–25. [PubMed: 12925327]
9. Brodmann, K. Vergleichende Lokalisationlehre der Grosshirnrinde in ihren Prinzipien Dargestellt auf Grund des Zellenbaues. In: Garey, LJ., translator. Brodmann's 'Localisation in the Cerebral Cortex'. Smith-Gordon; London: 1994. original publication date 1909
10. Bassett S, Yousem D, Christinzio C, Kusevic I, Yassa M, Caffo B, et al. Familiar risk for Alzheimers disease alters fMRI activation patterns. Brain. 2006; 129:1229–1239. [PubMed: 16627465]
11. West, M.; Harrison, J. Applied Bayesian Forecasting and Time Series Analysis. Boca Raton: Chapman and Hall/CRC; 1989.
12. Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: A general linear approach. Human Brain Mapping. 1995; 2(4):189–210.
13. De Meyer G, Shapiro F, Vanderstichele H, Vanmechelen E, Engelborghs S, De Deyn PP, et al. Diagnosis-Independent Alzheimer Disease Biomarker Signature in Cognitively Normal Elderly People. Archives of Neurology. 2010; 67(8):949–956. [PubMed: 20697045]
14. Anoop A, Singh PS, Jacob RS, Maji SK. CSF Biomarkers for Alzheimer's Disease Diagnosis. International Journal of Alzheimer's Disease. 2010; (2):377–378.
15. Buerger K, Zinkowski R, Teipel SJ, Tapiola T, Arai H, Blennow K, et al. Differential Diagnosis of Alzheimer Disease With Cerebrospinal Fluid Levels of Tau Protein Phosphorylated at Threonine 231. Archives of Neurology. 2002; 59:1267–1272. [PubMed: 12164722]

Appendix

We give the full conditional distributions of the parameters in the model. We note that the expressions are written at a region level (i.e. for $\boldsymbol{\beta}_g$, instead of $\boldsymbol{\beta}_g(v)$). The details on how the formulas for the full conditionals are derived are given in Supplementary Material, section 2. The full conditionals are given by the following (we omit notation for conditioned variables for simplicity):

1. $\boldsymbol{\beta}_g \sim \mathrm{N}(\boldsymbol{\Omega}_g\mathbf{T}_g, \boldsymbol{\Omega}_g)$, where

   $\boldsymbol{\Omega}_g=\mathbf{I}_{V_g}\otimes(\Lambda_v^{-1}+n\boldsymbol{Z}_g^{-1})^{-1}, \mathbf{T}_g=(\mathbf{I}_{V_g}\otimes\Lambda_v^{-1})\boldsymbol{\beta}_{0g}+\mathbf{V}_g^{-1}\sum_{i=1}^{n}\mathbf{r}_{ig}, \Lambda_v^{-1}=\begin{bmatrix} 1/\lambda_1^2 & 0 \\ 0 & 1/\lambda_2^2 \end{bmatrix}$

   , and $\mathbf{r}_{ig} = \mathbf{Y}_{ig} - \boldsymbol{\varphi}_g - 1 \otimes \boldsymbol{\alpha}_{ig} - 1 \otimes \mathbf{X}_{ig}\,\boldsymbol{\gamma}_g$

2. The full conditional distribution for $\boldsymbol{\varphi}_v$ is a likelihood adjusted version of the conditional distribution for MCAR(ρ, Σ).

   $\boldsymbol{\varphi}_v \sim \mathrm{N}(\mathbf{H}_v\mathbf{P}_v, \mathbf{H}_v)$, where

   $\mathbf{H}_v^{-1}=w_{v+}\sum^{-1}+n\boldsymbol{Z}_g^{-1}, \mathbf{P}_v=w_{v+}\rho\sum^{-1}\sum_{v'\neq v}\frac{w_{vv'}}{w_{v+}}\boldsymbol{\phi}_{v'}+n\boldsymbol{Z}_g^{-1}(\overline{\boldsymbol{Y}}_g(v)-\boldsymbol{\beta}_g(v)-\overline{\boldsymbol{\alpha}}_g-\overline{\mathbf{X}}_g\boldsymbol{\gamma}_g))$

   , and $\overline{\boldsymbol{Y}}_g(v)=\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{Y}_{ig}(v), \overline{\boldsymbol{\alpha}}_g=\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\alpha}_{ig}, \overline{\mathbf{X}}_g=\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_{ig}$.

3. $\boldsymbol{\alpha}_i \sim \mathrm{N}(\boldsymbol{\Psi}\mathbf{P}, \boldsymbol{\Psi})$, where $\boldsymbol{\Psi}=(\boldsymbol{\Gamma}^{-1}+\boldsymbol{\Omega}_\alpha^{-1})^{-1}$,

   $\mathbf{P}=\begin{bmatrix} 1_{V_1}\otimes\boldsymbol{Z}_1^{-1} & \dots & \dots \\ \vdots & \ddots & \vdots \\ \dots & \dots & 1_{V_G}\otimes\boldsymbol{Z}_G^{-1} \end{bmatrix}\cdot\begin{bmatrix} \boldsymbol{Y}_{i1}-\boldsymbol{\beta}_1-\boldsymbol{\phi}_1-1_{V_1}\otimes\mathbf{X}_{i1}\boldsymbol{\gamma}_1 \\ \vdots \\ \boldsymbol{Y}_{iG}-\boldsymbol{\beta}_G-\boldsymbol{\phi}_G-1_{V_G}\otimes\mathbf{X}_{iG}\boldsymbol{\gamma}_G \end{bmatrix}$, and

   $\boldsymbol{\Gamma}^{-1}=\boldsymbol{\Gamma}_1^{-1}\otimes\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}+\boldsymbol{\Gamma}_2^{-1}\otimes\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, while $\boldsymbol{\Omega}_\alpha^{-1}=\begin{bmatrix} V_1\boldsymbol{Z}_1^{-1} & \dots & \dots \\ \vdots & \ddots & \vdots \\ \dots & \dots & V_G\boldsymbol{Z}_G^{-1} \end{bmatrix}$.

4. $\boldsymbol{\Gamma}_j^{-1}\sim\mathrm{Wishart}((h_jH_j+\sum_{i=1}^{n}\boldsymbol{\alpha}_{ij}\boldsymbol{\alpha}_{ij}^T)^{-1}, h_j+n)$.

5. $\boldsymbol{Z}_g^{-1}\propto\mathrm{Wishart}((c_1\Omega_1+\sum_{i=1}^{n}\sum_{v\in g}\mathbf{u}_{igv}\mathbf{u}_{igv}^T))^{-1}, c_1+nV_g$, where $\mathbf{u}_{igv} = \mathbf{Y}_{ig}(v) - \boldsymbol{\beta}_g(v) - \boldsymbol{\varphi}_g(v) - \boldsymbol{\alpha}_{ig} - \mathbf{X}_{ig}\boldsymbol{\gamma}_g$.

6. $\sum^{-1}\sim\mathrm{Wishart}((c_2\Omega_2+\sum_{v=1}^{V}w_{v+}[\boldsymbol{\phi}_v-\rho\sum_{v'\neq v}b_{v'}\boldsymbol{\phi}_{v'}][\boldsymbol{\phi}_v-\rho\sum_{v'\neq v}b_{v'}\boldsymbol{\phi}_{v'}]^T), c_2+V)$

   , where V is the total number of voxels included in the analysis (with at least one within-region neighbour).

7. $\lambda_{gj}^{-2}\sim\mathrm{Gamma}\left(a_j+V_g/4, \left[\frac{1}{b_j}+\frac{(\boldsymbol{\beta}_{gj}-1_{V_g}\beta_{0gj})^T(\boldsymbol{\beta}_{gj}-1_{V_g}\beta_{0gj})}{2}\right]^{-1}\right)$, where $\boldsymbol{\beta}_{gj}=\boldsymbol{\beta}_g^{(j)}$ and $\boldsymbol{\beta}_{0_{gj}}=\boldsymbol{\beta}_{0_g}^{(j)}$, for j = 1,2.

8. For the spatial parameter ρ, we set an array of values and take a uniform prior on that set. The full conditional distribution of the spatial parameter ρ is just the

corresponding set of normalized likelihood weights. More details are given in Supplementary Material, part 2.

9. $\boldsymbol{\gamma}_g \sim \mathrm{N}(\boldsymbol{\Omega}_{\gamma g}\,\mathbf{T}_{\gamma g}, \boldsymbol{\Omega}_{\gamma g})$ where

$$\boldsymbol{\Omega}_{\gamma_g}^{-1}=\boldsymbol{\Lambda}_{\gamma_g}^{-1}+\sum_{i=1}^{n}(\mathbf{1}_{V_g}^{T}\otimes\mathbf{X}_{ig}^{T})\mathbf{V}_g^{-1}(\mathbf{1}_{V_g}\otimes\mathbf{X}_{ig}),\ \mathbf{T}_{\gamma_g}=\sum_{i=1}^{n}(\mathbf{1}_{V_g}^{T}\otimes\mathbf{X}_{ig}^{T})\mathbf{V}_g^{-1}\mathbf{u}_{ig},\ \mathbf{u}_{ig}$$

$$\boldsymbol{\gamma}_g|\Lambda_{\gamma_g}\sim\mathrm{N}(\mathbf{0},\underbrace{\mathrm{diag}(\tau_{g11}^2,\tau_{g21}^2,\ldots,\tau_{g1Q}^2,\tau_{g2Q}^2)}_{\boldsymbol{\Lambda}_{\gamma_g}}).$$

$= \mathbf{Y}_{ig} - \boldsymbol{\beta}_g - \boldsymbol{\varphi}_g - \mathbf{1}_{V_g} \otimes \boldsymbol{\alpha}_{ig}$ and

10. $\tau_{gjq}^2 \sim \mathrm{Gamma}\left(e_0+\frac{1}{2}, \left[\frac{1}{f_0}+\frac{1}{2}\boldsymbol{\gamma}_{gjq}^2\right]^{-1}\right)$.

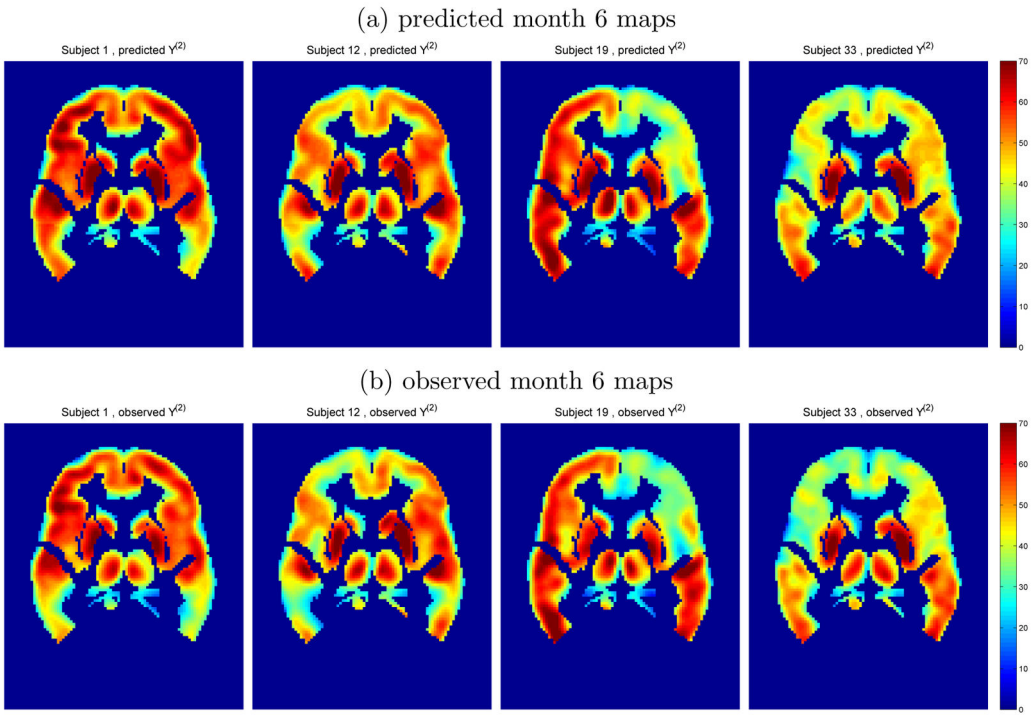

Figure 1.

Individualized predicted and observed month 6 follow-up regional glucose uptake measurements for 4 AD patients from the test data set. Axial slice 40 is shown in radiological view. There is a satisfactory agreement between the observed and predicted post-treatment regional glucose uptake.

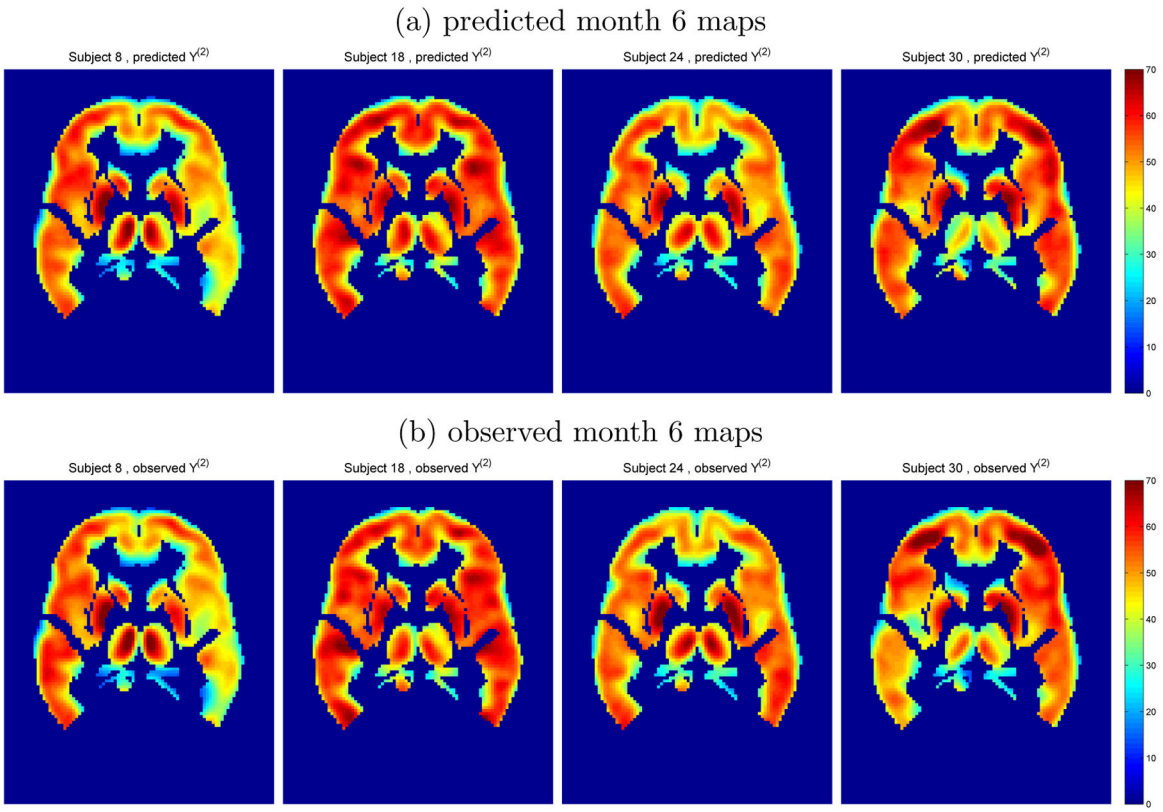

Figure 2.

Individualized predicted and observed month 6 follow-up regional glucose uptake measurements for 4 HC subjects from the test data set; axial slice 40 is shown in radiological view. There is a satisfactory agreement between the observed and predicted post-treatment regional glucose uptake.

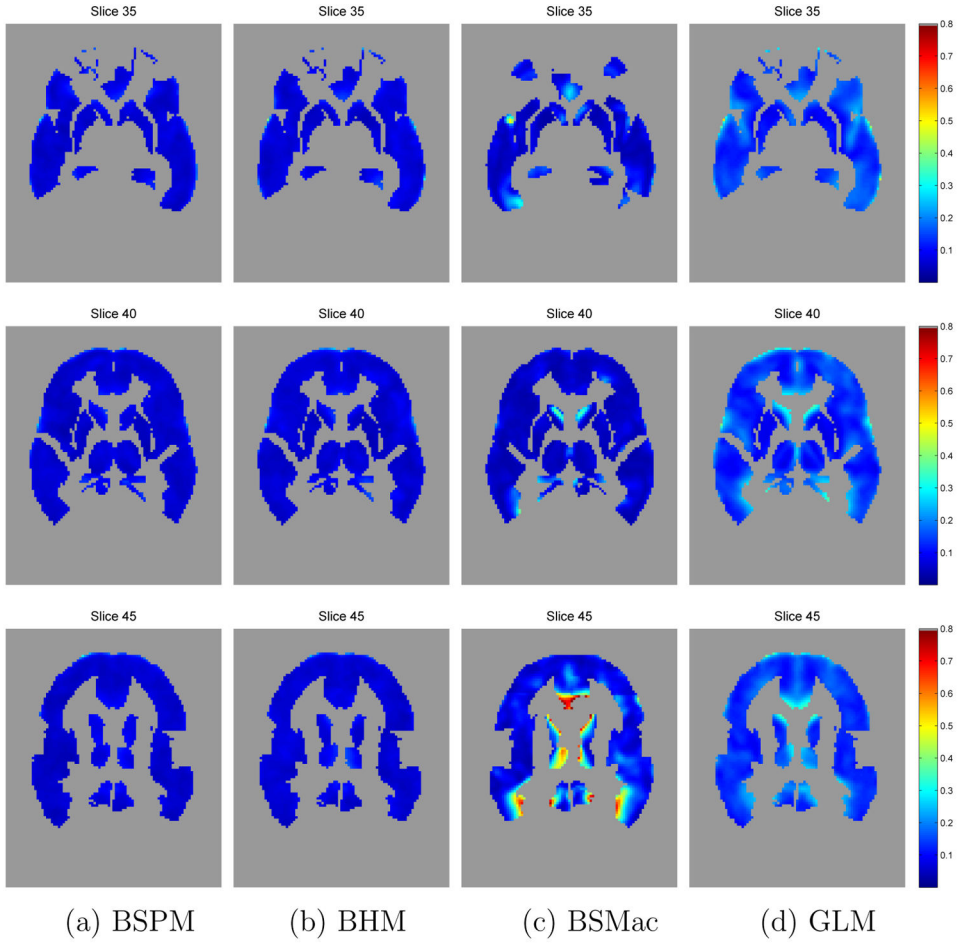

Figure 3.

The images depict the square root of the prediction mean square error, divided by the average observed brain activity (stPMSE, (2)) at each voxel for prediction of the follow-up activity for 33 test subject in the AD group. Axial slices 35, 40 and 45 are shown (in radiological view) of the stPMSE map based on (a) our proposed model, (b) BHM proposed in [1], (c) BSMac proposed in [4], and (c) GLM. Average errors: 0.083 for the BSPM (a), 0.080 for the BHM (b) 0.104 for the BSMac (c), and 0.156 for the GLM (d).

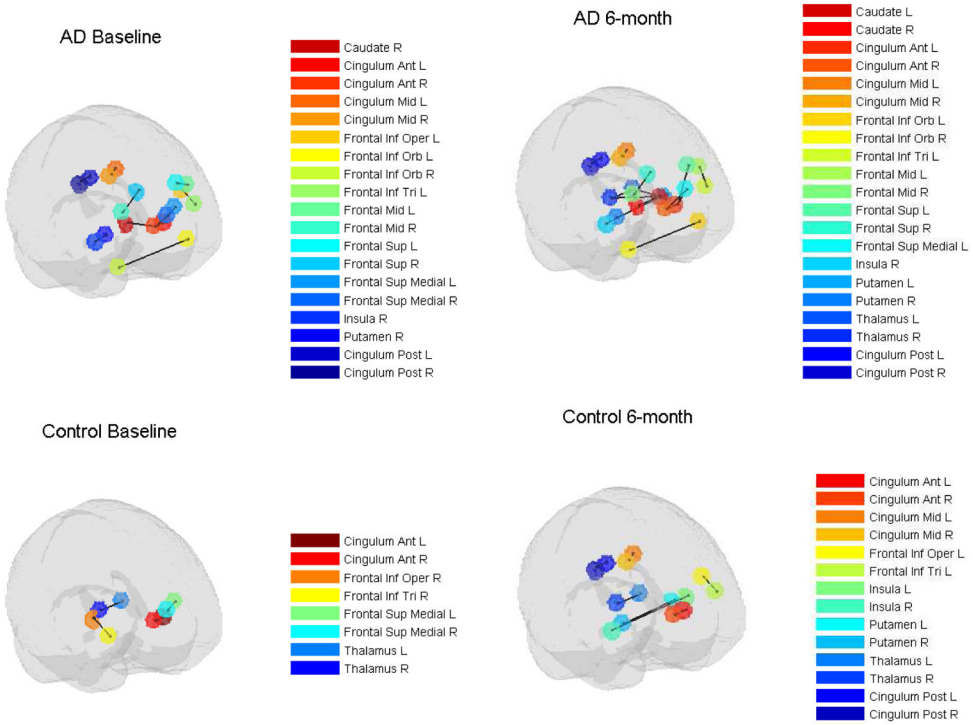

Figure 4.

Region functional connectivity for AD (top) and HC (bottom) subjects. The regions that have (posterior median) correlations exceeding 0.75 are shown. The connecting lines have different thicknesses, corresponding to the strength of the inter-regional correlations.

Table 1

Summary of the simulation results for the parameters in the covariance matrix $\mathbf{Z}_g$, for each region. Z_g^{11} and Z_g^{22} denote the variance, and Z_g^{12} the covariance components.

Param.	Region									
	1		2		3		4		5	
	True	Est.	True	Est.	True	Est.	True	Est.	True	Est.
Z_g^{11}	323.26	321.25	160.61	160.05	11.46	11.46	3.44	3.44	3.57	3.57
Z_g^{22}	120.75	120.50	45.30	45.15	37.22	37.19	10.45	10.45	3.38	3.38
Z_g^{12}	−41.95	−41.40	−33.70	−33.43	16.32	16.30	4.08	4.08	2.40	2.40

Stat Methods Med Res. Author manuscript; available in PMC 2014 September 26.