Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: http://www.tandfonline.com/loi/uasa20

Functional Data Analysis of Dynamic PET Data

Yakuan* Chen, Jeff Goldsmith & R. Todd Ogden

To cite this article: Yakuan* Chen, Jeff Goldsmith & R. Todd Ogden (2018): Functional Data Analysis of Dynamic PET Data, Journal of the American Statistical Association, DOI: 10.1080/01621459.2018.1497495

To link to this article: https://doi.org/10.1080/01621459.2018.1497495

Functional Data Analysis of Dynamic PET Data

Yakuan Chen[1,*], Jeff Goldsmith[2], and R. Todd Ogden[2]

[1]AT&T Services, Inc.
[2]Department of Biostatistics, Mailman School of Public Health, Columbia University
[*]*yc236k@att.com*

March 30, 2018

Abstract

One application of positron emission tomography (PET), a nuclear imaging technique, in neuroscience involves in vivo estimation of the density of various proteins (often, neuroreceptors) in the brain. PET scanning begins with the injection of a radiolabeled tracer that binds preferentially to the target protein; tracer molecules are then continuously delivered to the brain via the bloodstream. By detecting the radioactive decay of the tracer over time, dynamic PET data are constructed to reflect the concentration of the target protein in the brain at each time. The fundamental problem in the analysis of dynamic PET data involves estimating the impulse response function (IRF), which is necessary for describing the binding behavior of the injected radiotracer. Virtually all existing methods have three common aspects: summarizing the entire IRF with a single scalar measure; modeling each subject separately; and the imposition of parametric restrictions on the IRF. In contrast, we propose a functional data analytic approach that regards each subject's IRF as the basic analysis unit, models multiple subjects simultaneously, and estimates the IRF nonparametrically. We pose our model as a linear mixed effect model in which population level fixed effects and subject-specific random effects are expanded using a B-spline basis. Shrinkage and roughness penalties are incorporated in the model to enforce identifiability and smoothness of the estimated curves, respectively, while monotonicity and non-negativity constraints impose biological information on estimates. We illustrate this approach by applying it to clinical PET data with subjects belonging to three diagnosic groups. We explore differences among groups by means of pointwise confidence intervals of the estimated mean curves based on bootstrap samples.

Key words: Function-on-scalar regression; Constrained estimation; Splines; Nonparametric; Bootstrap

1 Introduction

1.1 Background on PET

Positron emission tomography (PET) is a nuclear imaging technique that allows the study of basic mechanisms of the human body. The application of PET imaging in neuroscience has proven to be a valuable tool to better our current understanding of changes during brain stimulation, cognitive activation, and metabolic processes associated with mental illnesses and neurological disorders. One particular application of PET imaging aims to estimate the density of various proteins throughout the brain. For instance, investigators use PET imaging to study the density of β-amyloid plaque that plays a key role in the pathogenesis of Alzheimer's disease (Zeng and Goodman, 2013); another example is the examination of the serotonin (5-HT) neurotransmitter system in the pathophysiology of depression (Miller *et al.*, 2013) and bipolar disorder (Sullivan *et al.*, 2009), among many others.

The application of PET in such a neuroimaging study begins with the injection of a radiolabeled compound that has affinity for a particular protein in the human brain. This radiolabeled compound, or radiotracer, is designed to bind preferentially to that target protein. Once it is introduced into the bloodstream, the radiotracer is continuously delivered to the brain by the vascular system. Within the brain, each tracer molecule exist in one of three biomedical states: it may be "free" in the synapse, i.e., not bound to any biomolecules; it may be bound "specifically" to the target protein; or it may be "nonspecifically" associated with other macromolecular components. While in the brain, tracer molecules can change from one state to another, potentially making many such transitions during the PET scan. Additionally, because the tracer molecules can cross the blood-brain barriers in both directions, they may also exit the brain and be delivered by the bloodstream to other organs, or back to the brain again.

All radiotracer molecules, no matter their biomedical state, undergo radioactive decay (i.e., emitting positrons) throughout the scan. By detecting the radiation emitted over a given time interval, a three-dimensional image may be obtained via a reconstruction algorithm. Thus, dynamic PET data consist of a sequence of these 3-dimensional images, each voxel of which is a measurement of the concentration of the radiotracer at the correspond-

ing time and location. This concentration depends on the amount of tracer that has been available for delivery in the bloodstream and on the binding behavior of tracer molecules in the brain. Neglecting the noise for the moment, the concentration of the radiotracer may generally be expressed as the convolution between two functions:

$$c(t) = \int_0^t f(s)g(t-s)\,ds \tag{1}$$

The function g is the concentration of the radiotracer in the arterial plasma over time, corrected to account for radioactive metabolites of the tracer; this is termed the "input function" since it represents the amount of tracer available to enter the brain at each time. The function f is the location-specific "impulse response function" (IRF) that represents what the hypothetical concentration of the tracer would be over time if the input function were an instantaneous bolus spike (Dirac delta function). Biologically, if the tracer were to be delivered as an instantaneous bolus spike at time 0, the density of the tracer in the brain would be highest at time 0 and gradually decrease as tracer molecules exit the brain. Therefore, it is expected that the IRF will be non-negative and non-increasing over time.

Because the tracer is designed to bind to the target protein, the IRF is related to the density of that protein in the corresponding location. For instance, in a target-protein-rich region, the IRF decreases slowly because the tracer molecules tend to spend much of the time bound to the target protein. In contrast, in a region with no target proteins, the IRF will decrease at a higher rate. In such a region, because it is completely devoid of the target protein, tracer molecules can only be free or associated with macromolecular components other than the target. If such a region exists, it is termed a "reference region". The binding capacity in the reference region thus represents only "non-specific binding", typically assumed to be uniform throughout the brain; and as a result, specific binding may be estimated based on the difference between IRFs of the region of interest (ROI) and the reference region (Innis *et al.* (2007); Slifstein and Laruelle (2001)).

For a given voxel, the sequence of concentrations across time is termed the time activity curve (TAC). In practice, the observed TAC is contaminated with noise and may be expressed

$$y(t) = \int_0^t f(s)g(t-s)\,ds + \epsilon(t) \tag{2}$$

where f and g are, respectively, the IRF and the input function, and ϵ represents the errors observed due to the Poisson nature of the radioactive decay, the detection process, processing errors, and other sources of error. For each subject, the input function is common across all voxels, but each voxel or region of interest has its specific IRF. In many PET studies, samples of arterial blood are drawn during the scanning; with each sample, the concentration of the tracer is measured and a metabolite analysis is performed. In this way, the input function g can be measured. Although there is some uncertainty in the measured input function, this is generally small relative to the PET noise so it is typically considered "known" in expressions like (2). Hence, any PET modeling technique that involves an input function measured from blood samples must involve deconvolution of the TAC data using this input function to recover the IRF f, which contains information about the density of the target protein.

1.2 Overview of our proposed nonparametric modeling approach

Many approaches for dynamic PET modeling have been proposed. The preponderance of these methods have three characteristics in common. First, once the estimated IRF is obtained, it is summarized using a single scalar measure, and subsequently standard univariate analyses, such as t tests or linear mixed models, are performed on the scalar measure. While the scalar summaries have straightforward biological interpretation, by summarizing the entire IRF using a single scalar, it is possible to lose some important features of this function. Second, all these methods focus on estimating the IRF for one subject and one region at a time. This tends to limit the complexity of models that can be fit to PET data. Third, most of the approaches impose strong parametric assumptions on the model of estimating IRF, and some of these assumptions may not hold in real data applications.

In this paper, we propose an alternative analytic approach to explore the kinetics of the tracer that will improve upon existing techniques in all three of these areas by estimating the IRF nonparametrically using functional data analytic (FDA) techniques. First, in our approach, the entire functions, rather than just the summarized scalars, can be compared across subjects/regions. Comparing the entire estimated IRFs prevents the loss of impor-

tant information, such as local features of the function. Second, in contrast to the current state of the art, our approach models TACs from multiple subjects simultaneously. Modeling multiple subjects simultaneously can help address inherent instability of subject-level parameter estimates and capture patterns for subjects with common characteristics. In this case, population- and subject-level effects can be estimated in a joint model, which increases efficiency. Third, we construct a nonparametric model by applying FDA techniques, which are an important tool set to analyze data that have functional form, such as the TAC in dynamic PET imaging studies. Specifically, the effects of multiple scalar covariates, including continuous covariates, on subject-specific IRFs can be incorporated into our model fitting framework. After incorporating the scalar covariates, which we will discuss in detail in Section 2, Model (2) can be posed as a regression problem with functional responses and scalar predictors, i.e., a function-on-scalar regression (Ramsay and Silverman (2005); Reiss *et al.* (2010); Morris (2015)). Therefore, we can convert this model to a multivariate regression model by extending current FDA techniques.

The advances we propose – to emphasize the IRF as the fundamental unit of interest, rather than a scalar summary; to jointly model IRFs from all subjects; and to model IRFs using a functional data approach in place of the standard parametric model – are a direct response to the shortcomings of existing tools for analyzing dynamic PET data. While our literature review in Section 1.3 will identify some related research, the comprehensive analytic framework developed in this manuscript is a major departure from available tools.

1.3 Brief overview of current estimation methods for dynamic PET data

Traditional approaches for estimating the IRF impose a parametric form on the IRF, which is motivated by a physiological model for tracer distribution. Among these, compartment modeling is the most widely used method to describe the uptake and clearance of a tracer in the tissue (Slifstein and Laruelle, 2001). The compartments of a system can be defined in our application as biomedical states in which each radiotracer molecule can exist: "free", specifically bound to the target protein, nonspecifically associated with other macromolecular components. For many radiotracers in neuroreceptor mapping, the kinetics of the

tracer in the brain can be approximated using a three-tissue compartment (3TC) model (in which compartments are "free" tracer, tracer specifically bound to the target protein, and tracer nonspecifically associated with other macromolecular components) or a more common two-tissue compartment (2TC) model (in which "free" tracer and tracer nonspecifically associated with other macromolecular components are considered as comprising a single compartment). Figure A.1 in the Appendix shows the compartment structures of the two-tissue compartment model and a more simplified one-tissue compartment (1TC) model (in which "free" tracer, tracer specifically bound to the target protein, and tracer nonspecifically associated with other macromolecular components are considered as comprising a single compartment). Basic assumptions of compartment modeling include that all injected tracer molecules will be in exactly one compartment at any given time and that the rates of transfer between compartments are constant over time. These assumptions ensure that the IRF can be expressed as a sum of exponential functions whose time constants and coefficients are functions of the rate parameters. Rate constants involved in the model can be routinely estimated by solving ordinary differential equations and applying nonlinear regression modeling techniques (Cunningham and Jones (1993); Gunn *et al.* (2001)).

Although kinetic models are well-established and almost universally applied, it is generally understood that they are inadequate for modeling many radiotracers, and therefore many alternative modeling strategies have been proposed. For example, "spectral analysis" (Cunningham and Jones, 1993) characterizes the IRF in terms of a set of basis functions and fits the model using non-negative least squares. Gunn *et al.* (2002) extended this basis function framework by imposing an L_1 penalty. Similarly, Jiang and Ogden (2008) and Lin *et al.* (2014) proposed a mixture modeling procedure in which each IRF is represented in terms of a smaller number of basis functions. These approaches have some nonparametric flavor (e.g., the use of basis functions) but they are all derived from the standard compartment model, and so the basis functions are set to be exponentials. In addition, Logan *et al.* (1990) introduced "graphical analysis" which is based on the two tissue compartment model. This approach does not estimate the IRF directly but instead estimates a scalar summary.

As with all parametric models, compartment models rely on assumptions about the data generating process and can perform poorly when these are violated in practice. One key assumption is the assumed compartmental structure itself, which is generally understood to be a simplification of a more realistic (but more complex) model. Additionally, the non-linear least squares methods commonly used to fit compartment models tend to have bias that depends on the parameter values, and these tools can also be somewhat numerically unstable (Peng *et al.*, 2008).

The limitations of parametric methods have helped to motivate the development of non-parametric approaches that allow model-free estimation. O'Sullivan *et al.* (2009) proposed nonparametric "residue analysis" of dynamic PET data based on the indicator dilution theory originally put forth by Meier and Zierler (1954). They base their modeling on Equation (1) but place no parametric restrictions on the IRF f. Instead, they express it as $f(t) = C\left(1 - \int_0^t h(s)ds\right)$, where C is a proportionality constant that is interpretable as an overall flow and h is a probability density function. The term $1 - \int_0^t h(s)ds$ is called the *tissue residue function*, and it reflects the fraction of radiotracer that remains in the system at time t. With this formulation, the IRF is constrained to be non-negative and non-increasing over time. In O'Sullivan *et al.* (2009), the probability density function h is estimated nonparametrically. It is expressed in terms of a natural cubic B-spline basis, and a weighted second derivative penalty is employed to control the roughness of the estimated curve.

Zanderigo *et al.* (2015) proposed a nonparametric method that approximates the problem in terms of a discrete deconvolution operation, which can be solved by using a singular value decomposition (SVD). This method is, however, rather sensitive to noise, potentially causing the estimated curve to oscillate considerably, although these effects can be minimized by eliminating diagonal elements below a certain threshold in the diagonal matrix constructed from SVD. Jiang *et al.* (2015) presented a nonparametric approach for estimating the IRF based on a functional principal component analysis (FPCA). They smoothed the observed PET curves for all voxels using a pre-specified kernel smoother and subsequently applied FPCA on the pre-smoothed curves. Deconvolution was only required on the mean function and the eigenfunctions, rendering it more computationally efficient. The

IRF is then recovered using the functions obtained from the deconvolution operator. Regularization is achieved by selecting the number of components using an *ad hoc* measure of goodness-of-fit. Note that each of these methods estimates the IRFs one at a time.

Whether the IRF is estimated using parametric or nonparametric methods, current practice involves summarizing the estimated IRF using a single scalar measure and then comparing this measure across subjects/regions in subsequent analysis. In the parametric approaches discussed above, these summary measures are related to some aspect of the density of the target protein. For instance, the total volume of distribution of tracer (V_T) is defined as a functional of the IRF; specifically, $V_T = \int_0^\infty f(t)\,dt$. Even with the nonparametric methods for estimating the IRF, interest generally lies in computing some scalar measure that can then be compared across subjects/regions. One option is to calculate $\int_0^{t_{end}} f(t)\,dt$, the area under the IRF until the end of the scan. This is a nonparametric analogue to V_T, although it does not have the same clear biological interpretation. Another option is to calibrate the nonparametric estimator with a specific compartment model so that the resulting summary measure will have the same interpretation in the case that the parametric model holds (O'Sullivan *et al.* (2009); Zanderigo *et al.* (2015)).

In contrast to the preceding, we develop a flexible non-parametric approach to dynamic PET data that 1) focuses on the IRFs, rather than a single summary measure of these functions, as the basis for comparisons; 2) models data from all subjects simultaneously; and 3) estimates the IRF for each subject and includes covariate effects using FDA techniques. The first two of these are novel contributions to the PET literature. The final point required the development of a new model and estimation approach in FDA, which views Model (2) as a special form of the functional historical model with subject- and covariate-dependent coefficients that are estimated subject to biological non-negativity and monotonicity constraints.

The rest of the article is organized as follows. In Section 2, we describe a functional approach to nonparametrically estimate the IRFs of all subjects simultaneously and to compare the entire estimated IRFs across subjects. We conduct a simulation study and present the results in Section 3. An application of our method on some clinical PET data is given in Section 4. Finally, we present some concluding remarks in Section 5.

2 Methodology

2.1 Conceptual model

In this article, we consider only the situation in which the input function g_i is observed for each subject. Following the description given in Section 1, our model for subject i is

$$y_i(t) = \int_0^t f_i(s) g_i(t-s)\, ds + \epsilon_i(t), \qquad i = 1, \ldots, n \tag{3}$$

In principle, both y and f should be indexed by subject and region because the analysis can be performed on any voxel/region. For simplicity, we restrict our attention to a single region and suppress the related index.

For the purpose of estimating f_i in Model (3), we assume f_i can be separated into a population-level fixed effect that may depend on some measured covariates and a subject-level random effect. This formulation allows for the effects of multiple scalar covariates, including continuous covariates, on the IRF to be directly estimated from the model. Then

$$\begin{aligned} f_i(s) &= \beta_0(s) + \sum_{j=1}^{p} x_{ij}\beta_j(s) + \delta_i(s) \\ &= \boldsymbol{x}_i^T \boldsymbol{\beta}(s) + \delta_i(s), \end{aligned} \tag{4}$$

where $\boldsymbol{x}_i = (1, x_{i1}, \cdots, x_{ip})^T$ is the vector of covariate values for subject i and $\boldsymbol{\beta}(s) = (\beta_0(s), \beta_1(s), \cdots, \beta_p(s))^T$. Next, by expanding the functions $\beta_0(\cdot)$, $\beta_1(\cdot)$, $\cdots$, $\beta_p(\cdot)$, and $\delta_i(\cdot)$ in terms of a pre-specified set of basis functions $\phi_1, \cdots, \phi_K$, Model (4) becomes

$$\begin{aligned} f_i(s) &= \boldsymbol{x}_i^T \begin{pmatrix} \sum_{k=1}^K \beta_{0k}\phi_k(s) \\ \vdots \\ \sum_{k=1}^K \beta_{pk}\phi_k(s) \end{pmatrix} + \sum_{k=1}^{K} \delta_{ik}\phi_k(s) \\ &= \boldsymbol{x}_i^T \boldsymbol{B}\boldsymbol{\phi}(s) + \boldsymbol{\delta}_i^T \boldsymbol{\phi}(s) \\ &= \left(\boldsymbol{x}_i^T \boldsymbol{B} + \boldsymbol{\delta}_i^T\right) \boldsymbol{\phi}(s) \end{aligned} \tag{5}$$

where $\boldsymbol{\phi}(s) = (\phi_1(s), \cdots, \phi_K(s))^T$; $\boldsymbol{\delta}_i = (\delta_{i1}, \cdots, \delta_{iK})^T$ are the basis function coefficients of $\delta_i(\cdot)$; and $\boldsymbol{B} = (\boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_p)^T$, where $\boldsymbol{\beta}_j = (\beta_{j1}, \cdots, \beta_{jK})^T$, $j = 0, \cdots, p$ are the basis function coefficients corresponding to $\beta_0(\cdot)$, $\beta_1(\cdot)$, $\cdots$ and $\beta_p(\cdot)$, respectively. We then

replace $f_i(s)$ in Model (3) with the expression in (5), which gives

$$y_i(t) = \int_0^t \left(\boldsymbol{x}_i^T \boldsymbol{B} + \boldsymbol{\delta}_i^T\right) \boldsymbol{\phi}(s) g_i(t-s)\, ds + \epsilon_i(t). \tag{6}$$

2.2 Model for the observed PET data

In practice, the measured concentration values are derived from the decay counts observed over a given time interval across a grid of time points $\{t_{i\ell}\}_{\ell=1,\cdots,L_i}$. By design, the time frames are gradually longer over time during the scan, because of the radioactive decay process and because of decreasing concentration of the tracer. The change in time frames over the scan has two practical implications. First, the discrete grid of time points on which TACs are observed is taken to be the midpoints of the frames, which can be irregular. Second, since data are observed over consecutive time frames of different lengths, the frame duration, the radioactive decay and the overall concentration affect the variability of the response. As a result, weighting schemes that account for these factors are necessary. In the simulation and real data analyses below, weights are set to be the duration of the time-frame corresponding to $t_{i\ell}$ (Zanderigo *et al.*, 2015), although our methodology allows other weighting schemes to be used. Errors are assumed to be uncorrelated over time as they arise originally from decay count data, which are naturally independent, and then are reconstructed and registered separately for each time interval.

With these considerations, the observed data within Model (6) can be expressed in matrix form

$$\begin{aligned} \boldsymbol{y}_i &= \boldsymbol{Z}_i \left(\boldsymbol{B}^T \boldsymbol{x}_i + \boldsymbol{\delta}_i\right) + \boldsymbol{\epsilon}_i \\ &= \boldsymbol{Z}_i \boldsymbol{B}^T \boldsymbol{x}_i + \boldsymbol{Z}_i \boldsymbol{\delta}_i + \boldsymbol{\epsilon}_i \\ &= \boldsymbol{Z}_i \left(\boldsymbol{x}_i^T \otimes \mathbf{I}_K\right) \boldsymbol{\beta} + \boldsymbol{Z}_i \boldsymbol{\delta}_i + \boldsymbol{\epsilon}_i \end{aligned} \tag{7}$$

where $\boldsymbol{y}_i = (y_i(t_{i1}), \cdots, y_i(t_{iL_i}))^T$; $\boldsymbol{Z}_i$ is a $L_i \times K$ matrix with the (ℓ, k)th element $\int_0^{t_{i\ell}} \phi_k(s) g_i(t_{i\ell} - s)\, ds$; $\mathbf{I}_n$ denotes the $n \times n$ identity matrix; $\boldsymbol{\beta} = \text{vec}(\boldsymbol{B}^T)$ is the vector obtained by stacking the rows of $\boldsymbol{B}$, i.e., $\boldsymbol{\beta} = (\boldsymbol{\beta}_0^T, \boldsymbol{\beta}_1^T, \cdots, \boldsymbol{\beta}_p^T)^T$; and $\boldsymbol{\epsilon}_i = (\epsilon_i(t_{i1}), \cdots, \epsilon_i(t_{iL_i}))^T \sim \mathcal{N}(\mathbf{0}, \sigma^2 \boldsymbol{W}_i^{-1})$. $\boldsymbol{W}_i$ is a diagonal matrix with diagonal elements $\{w_{i\ell}\}_{l=1,\cdots,L_i}$, where $\{w_{i\ell}\}$ are as fixed and known observation weights.

Based on our construction, Model (7) can be viewed as a linear mixed effects model where $\boldsymbol{\beta}$ are the population-level fixed effects and $\boldsymbol{\delta}_i \sim \mathcal{N}(\mathbf{0}, \sigma_\delta^2 \mathbf{I}_K)$ are the subject-level random effects. Let $\oplus_{i=1}^n \boldsymbol{D}_i$ denote a block diagonal matrix with diagonal matrix elements $\{\boldsymbol{D}_1, \cdots, \boldsymbol{D}_n\}$, i.e.,

$$\oplus_{i=1}^n \boldsymbol{D}_i = \text{diag}(\boldsymbol{D}_1, \cdots, \boldsymbol{D}_n) = \begin{pmatrix} \boldsymbol{D}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{D}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{D}_n \end{pmatrix}$$

and let $N = \sum_{i=1}^n L_i$. By combining the equations for all subjects in Model (7), we now have

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{Z}\left(\boldsymbol{X} \otimes \mathbf{I}_K\right)\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{\delta} + \boldsymbol{\epsilon} \\ &= \boldsymbol{U}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{\delta} + \boldsymbol{\epsilon}, \end{aligned} \tag{8}$$

where $\boldsymbol{y} = (\boldsymbol{y}_1^T, \cdots, \boldsymbol{y}_n^T)^T$, $\boldsymbol{Z} = \oplus_{i=1}^n \boldsymbol{Z}_i$, $\boldsymbol{X} = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n)^T$, $\boldsymbol{\delta} = (\boldsymbol{\delta}_1^T, \cdots, \boldsymbol{\delta}_n^T)^T$, $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}_1^T, \cdots, \boldsymbol{\epsilon}_n^T)^T$ and $\boldsymbol{U} = \boldsymbol{Z}\left(\boldsymbol{X} \otimes \mathbf{I}_K\right)$.

2.3 Constrained estimation

As discussed in Section 1, the IRF is non-negative and non-increasing over time. Formally,

1. $f_i(s) \geq 0$, $i = 1, \ldots, n$ (non-negativity);
2. $f_i(s) - f_i(t) \geq 0$, $\forall\, s \leq t$, $i = 1, \ldots, n$ (monotonicity).

In addition, we incorporate a roughness penalty on the estimated curves to prevent overfitting.

Based on the likelihood function of Model (8) and by incorporating the preceding constraints and penalties, $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ can be estimated by minimizing

$$\frac{1}{2}(\boldsymbol{y} - \boldsymbol{U}\boldsymbol{\beta} - \boldsymbol{Z}\boldsymbol{\delta})^T \boldsymbol{W} (\boldsymbol{y} - \boldsymbol{U}\boldsymbol{\beta} - \boldsymbol{Z}\boldsymbol{\delta}) + \lambda_1 \boldsymbol{\beta}^T \boldsymbol{P}_\beta^T \boldsymbol{P}_\beta \boldsymbol{\beta} + \lambda_1 \boldsymbol{\delta}^T \boldsymbol{P}_\delta^T \boldsymbol{P}_\delta \boldsymbol{\delta} + \lambda_2 \boldsymbol{\delta}^T \boldsymbol{\delta}, \tag{9}$$

subject to $\boldsymbol{C}(\boldsymbol{X} \otimes \mathbf{I}_K)\boldsymbol{\beta} + \boldsymbol{C}\boldsymbol{\delta} \geq \mathbf{0}$, where $\boldsymbol{W} = \oplus_{i=1}^n \boldsymbol{W}_i$, $\boldsymbol{P}_\beta = \mathbf{I}_{p+1} \otimes \boldsymbol{\Delta}_2 \boldsymbol{\Phi}$ and $\boldsymbol{P}_\delta = \mathbf{I}_n \otimes \boldsymbol{\Delta}_2 \boldsymbol{\Phi}$ and

$$\boldsymbol{C} = \left(\begin{array}{c} \mathbf{I}_n \otimes \boldsymbol{\Phi} \\ \hline \mathbf{I}_n \otimes \boldsymbol{\Delta}_1 \boldsymbol{\Phi} \end{array} \right).$$

Throughout, $\otimes$ represents the Kronecker product of two matrices; $\mathbf{\Phi}$ is the matrix consisting of basis functions evaluated at a pre-specified dense grid $\{\tau_d\}_{d=1,\cdots,D}$, which is equally spaced and lies in the range determined by the irregularly spaced and subject-specific time points at which the data are observed (i.e., the (d,k)th entry of $\mathbf{\Phi}$ is $\phi_k(\tau_d)$ for $d = 1, \ldots, D$ and $k = 1, \ldots, K$); and $\mathbf{\Delta}_1$ and $\mathbf{\Delta_2}$ are the first and second order difference matrices, respectively.

In the loss function (9), the non-negativity and monotonicity constraints are implemented by the inequality $\boldsymbol{C}(\boldsymbol{X} \otimes \mathbf{I}_K)\boldsymbol{\beta} + \boldsymbol{C}\boldsymbol{\delta} \geq \mathbf{0}$, and the upper and lower blocks of $\boldsymbol{C}$ correspond to these two constraints, respectively. The terms $\lambda_1 \boldsymbol{\beta}^T \boldsymbol{P}_\beta^T \boldsymbol{P}_\beta \boldsymbol{\beta}$ and $\lambda_1 \boldsymbol{\delta}^T \boldsymbol{P}_\delta^T \boldsymbol{P}_\delta \boldsymbol{\delta}$ control the smoothness of fixed and random effects, respectively, where $\boldsymbol{P}_\beta = \mathbf{I}_{p+1} \otimes \mathbf{\Delta}_2 \mathbf{\Phi}$ and $\boldsymbol{P}_\delta = \mathbf{I}_n \otimes \mathbf{\Delta}_2 \mathbf{\Phi}$. The magnitudes of both terms are controlled by the same tuning parameter λ_1 since we expect the smoothness of both effects to be similar; separate tuning parameters could also be used, but doing so would be more computationally intensive when choosing the values of the parameters. The term $\lambda_2 \boldsymbol{\delta}^T \boldsymbol{\delta}$ in the loss function (9) controls the magnitude of the variance of the individual-specific random effects and implicitly guarantees that the model is identifiable; this term also reflects the random effects specification in our mixed model representation.

The algorithm that minimizes the loss function (9) is implemented in the `pcls` function in the `mgcv` package (Wood, 2011), which solves least squares problems with quadratic penalties subject to linear equality and/or inequality constraints using quadratic programming.

2.4 Tuning parameter selection

The values of the tuning parameters λ_1 and λ_2 may be chosen by cross validation through a process we describe below. Another parameter to be determined is K, the number of basis functions used in the expansion of the IRF, which could also be determined by cross validation. Provided that the basis set is rich enough to capture all the details of the functions to be estimated, the choice of K is not crucial (Ruppert, 2002). However, the exact choice of K may vary from application to application and some examination of this choice is necessary.

To choose λ_1 and λ_2 we use a bivariate grid search. In each iteration, we generate the test sample by randomly selecting two points from the observed TAC for each subject and treat the unselected data as the training sample. While it is common in some functional data applications to leave out the entire curves (i.e., performing a leave-one-out cross validation at the subject level), our model contains an unobserved subject-level random effect, and therefore randomly leaving out two points from each curve can help assess the performance of the subject-specific effects. For our data, this procedure amounts to leaving out roughly 10% of the observations in each training-test split. The main purpose of performing this "regression-style" cross validation is to strike a good bias/variance tradeoff at the subject and population level. For each split, a full model is fit to the training sample, and prediction error is obtained by applying the fitted model to the test data.

3 Simulation

In this section we conduct a simulation study to assess the quality of our proposed non-parametric method and perform a systematic comparison to existing parametric methods, including one- and two-tissue compartment models. Frankle *et al.* (2006) and Ogden *et al.* (2007) concluded that the one-tissue compartment model has good performance on DASB binding by evaluating common competing modeling techniques using test-retest data, and therefore simulated datasets are generated based on the motivating DASB data. Since the parameter estimates of the two-tissue compartment model on DASB are not identifiable, we use data with [^{11}C]WAY100635 tracer, on which the 2TC model has been proven to be a good fit in a separate analysis, to provide a realistic guide when simulating datasets. In order to evaluate the capability of our method to estimate group differences, we assume subjects in each simulation dataset belong to two groups with different means. The methods are compared in three scenarios: first, when the data follow the one-tissue compartment model; second, when the data follow the two-tissue compartment model; and third, when they follow neither of these compartmental models.

We generate dataset under the first scenario by first fitting the one-tissue compartment model to the motivating DASB data. From the model fit, we extract the mean $\boldsymbol{k}$ and covariance $\boldsymbol{\Sigma}$ of the rate parameters. The "true" mean rate parameters of the two groups,

i.e. $\boldsymbol{k}_0$ and $\boldsymbol{k}_1$, are obtained through small, biologically plausible modifications of $\boldsymbol{k}$, and the difference between the two IRFs determined by $\boldsymbol{k}_0$ and $\boldsymbol{k}_1$ is set to be the "true" group effect for the purposes of this simulation. Individual IRFs are determined by individual rate parameters, which are drawn from a multivariate Gaussian distribution with mean either $\boldsymbol{k}_0$ or $\boldsymbol{k}_1$ and covariance $\boldsymbol{\Sigma}$. A similar process is used to generate data under the second scenario, except that it is based on the model fit of a two-tissue compartment model on WAY tracer data.

Datasets under the third (non-parametric) scenario are generated differently. We begin by fitting our proposed model to the motivating data assuming a group effect. Fitted group mean IRFs and individual deviations from the mean curves are obtained from this model fit. True group mean IRFs in the simulated data are slight modifications of those obtained in the analysis of real data. To generate the individual IRF, we randomly choose two of the individual deviations, take a weighted average, and add it to either of the mean IRFs; weights for the weighted average are α and $1-\alpha$, where α is sampled from a Uniform[0, b] distribution with b chosen to ensure the resulting data has the same mean and variance as the original data. We use this strategy to simulate data in order to guarantee that every simulated individual IRF satisfies both non-negativity and monotonicity constraints.

In each data generating scenario, simulated TACs for each subject are produced using the generated IRF and an input function randomly sampled from the observed data according to Model (2), with errors generated from a mean-zero Gaussian distribution with variance equal to that obtained in real data analysis. We simulate 100 datasets under each scenario, each with 100 subjects split equally into two groups. For each simulated dataset, we fit our proposed model as well as the two parametric compartment models. The values of tuning parameters are determined by a bivariate cross-validation as described in Section 2.4, using a 10×10 grid of possible tuning parameter value and evaluating each combination on 100 training-test splits.

The root integrated mean square errors (RIMSE) of the estimated group effect and of individual IRFs obtained from all three scenarios are presented in Figure 1. Both our proposed method and the 1TC model perform well under the first scenario, although there is a small but expected decrease in performance for the non-parametric approach stemming

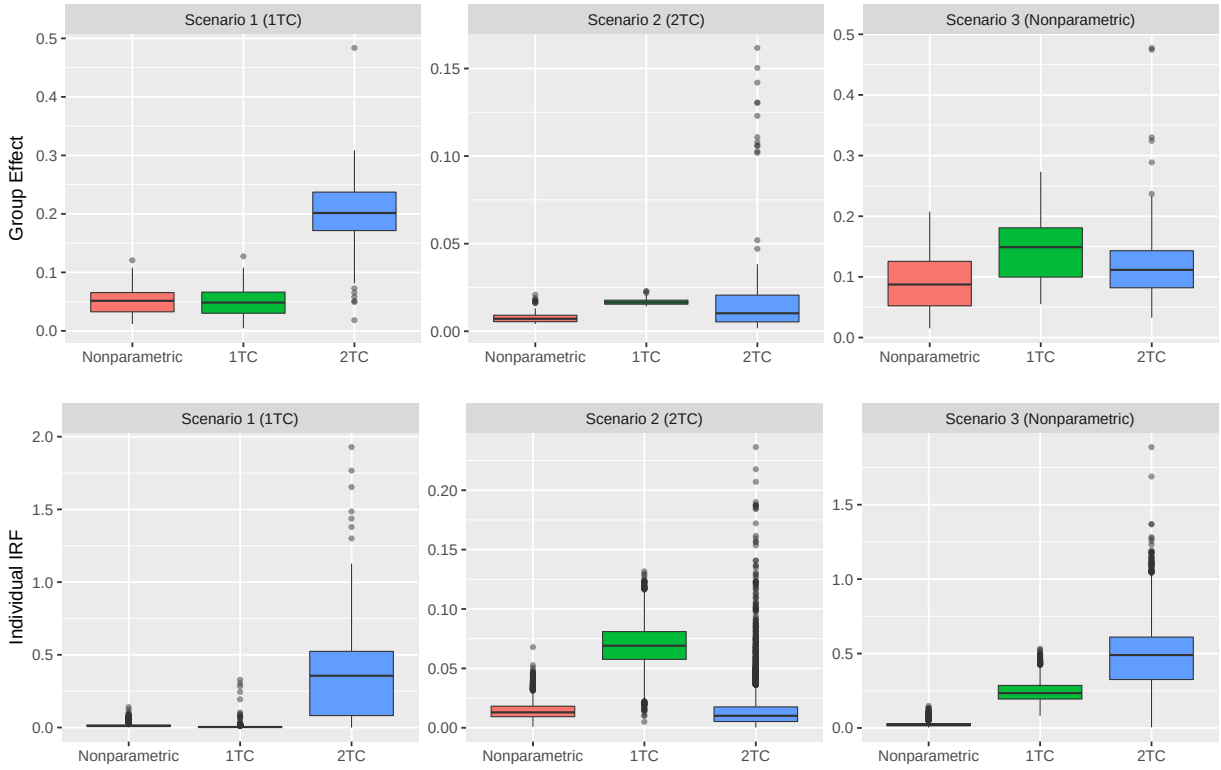

Figure 1: The root integrated mean square errors of the estimated group effect (top) and individual IRFs (bottom) obtained from our proposed method and the parametric compartment models when data follow 1TC model (left), when data follow 2TC model (middle), and when they follow neither of them (right).

from the increase in model complexity. For the second and third scenarios, however, the proposed approach outperforms both parametric methods in terms of estimation of both group effect and individual IRF. Indeed, the performance of the proposed approach is broadly similar across data generating mechanisms, while the performance of the parametric approaches suffer when the assumed model is not true.

4 PET data analysis

Impaired serotonergic function has been implicated in the pathophysiology of major depressive disorder (MDD) and bipolar disorder (BPD). Both have been associated with suicidal behavior and completed suicide. In these studies, the tracer [^{11}C]DASB has been

frequently used to examine the binding capacity of the serotonin transporter in the serotonin (5-HT) neurotransmitter system in the human brain. The one-tissue model generally provides reasonable fit to DASB data and also results in reproducible estimates of binding parameters (Ogden *et al.*, 2007). This model involves a single brain compartment which exchanges tracer molecules with the blood compartment, with tracer particles crossing the blood brain barrier (BBB) into the brain at constant rate K_1 and flowing in the other direction with rate k_2. For this model, the total volume of distribution can be shown to be $V_T = K_1/k_2$ (Innis *et al.*, 2007). Binding potential, a measure related to the density of the target protein in the brain, is typically calculated indirectly with this model, by comparing total distribution volume in a region of interest with that in a reference region.

Our data consist of PET scans using the [^{11}C]DASB tracer of 137 subjects belonging to three diagnostic groups: BPD (20), MDD (83), and normal control (34). Details of the data acquisition are given in Miller *et al.* (2013) and Miller *et al.* (2016). To summarize, injected dose averaged approximately 16mCi for each of the groups with a standard deviation of approximately 2 for all groups. Average injected mass ranged from 4μg to 5μg for the groups with standard deviation approximately 2. PET scanning was done on an Siemens ECAT HR+, and reconstruction was done using the filtered back projection algorithm. Any subject head motion during scanning was corrected by applying the FMRIB linear image registration tool (FLIRT). Frames of data acquisition were generally set to 3 at 20 seconds, 3 at 1 minute, 3 at 2 minutes, 2 at 5 minutes, and either 9 or 10 at 10 minutes, depending on scanning protocol and whether the late frames contain usable data. Concentrations at each time point were determined based on counts while taking into account the decay of the radiotracer (the [^{11}C] isotope has a half-life of 20.3 minutes). Regions of interest were identified on a T1-weighted MRI for each subject and transferred to the PET imaging space by coregistering to the subjects' corresponding sequence of PET images. The regions that we consider in this analysis are relatively large and easy to identify. Arterial samples were drawn every 10 seconds for the first 2 minutes, every 20 seconds for the next two minutes, and then at time points 6, 8, 12, 16, 20, 30, 40, 50, 60, 80, 90, 100, and 120 minutes. Blood samples drawn at 2, 12, 20, 50, 80, and 100 minutes were also analyzed to determine unmetabolized parent compound levels and the arterial data were corrected accordingly

(Parsey *et al.*, 2006b). The model used to describe the arterial concentration data is linear to the peak and a sum of three exponentials after the peak, and the fitted model was used as the input function for each subject.

In this section, we apply our method for estimating IRFs while accounting for covariate effects and constraints, described in Section 2, to this dataset. We first conduct an analysis on the midbrain, an ROI whose importance in the development of depression has been previously demonstrated. Serotonin transporter availability in the midbrain has been shown to be different in depressed subjects by Parsey *et al.* (2006a) and Malison *et al.* (1998) and has been studied in other PET studies of depression (Miller *et al.* (2013); Sullivan *et al.* (2009)). Subsequently, we conduct another analysis that focuses more closely on the binding specific to the target receptor. Raw TACs of the ROI and reference region as well as the input functions for all subjects are shown in Figure 2.

4.1 Analysis of the midbrain data

Focusing first on the midbrain, a model with diagnosis group as the only predictor is fit to the entire dataset. A cubic B-spline basis with 10 basis functions is used to model IRFs; no appreciable difference is observed when we repeat the analysis with either $K = 5$ or 15. Values of the tuning parameters λ_1 and λ_2 are determined by cross validation, as discussed in Section 2.4, using a 20×20 grid of possible tuning parameter value and evaluating each combination on 100 training-test splits. Figure 3 provides the cross validated prediction error over all values of the tuning parameters.

After choosing λ_1 and λ_2, we estimate model parameters using the complete dataset. The fitted group mean and individual IRFs are shown in the left three panels of Figure 4, and mean differences of the three groups are shown in the right panel of Figure 4. These results indicate that the mean IRF of the patients with bipolar disorder tends to be lower than that of the patients with major depression throughout the entire study period. In addition, the mean IRF of the healthy controls is lower than the other two groups at the beginning of the scan, but decreases at a slower rate and is higher than the other groups for most of the scan duration. Such conclusions can only be drawn based on the analysis of the entire IRFs rather than some summarized measures, like V_T. When the IRF is

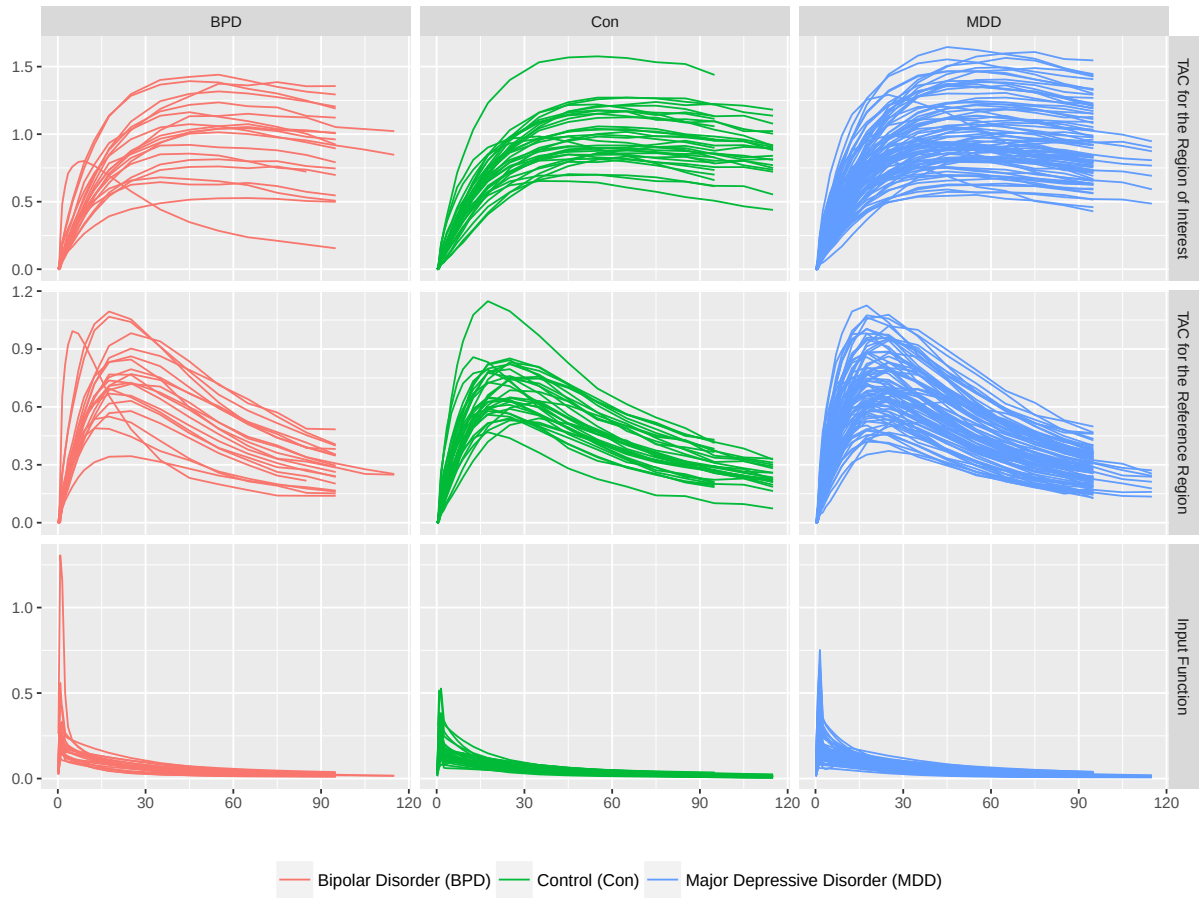

Figure 2: Raw TACs of the ROI (top) and reference region (middle) and the input functions (bottom) by diagnosis group.

estimated according to assumptions required by kinetic models, any difference between two IRFs can only be attributed to differences in the set of rate parameters (which combine to determine the density of target proteins), and that is the extent of the interpretation of such functions. By estimating the IRF nonparametrically, however, a much more flexible interpretation is possible. The IRF reflects the density of the target proteins, to be sure, but going well beyond that it also represents the rate at which the tracer molecules leave the system, which may be time-varying.

We next construct pointwise confidence bands of the group mean differences using a bootstrap algorithm in which 1000 bootstrap samples are generated in the following way. Within each diagnosis group, subjects are chosen with replacement with the sample size of

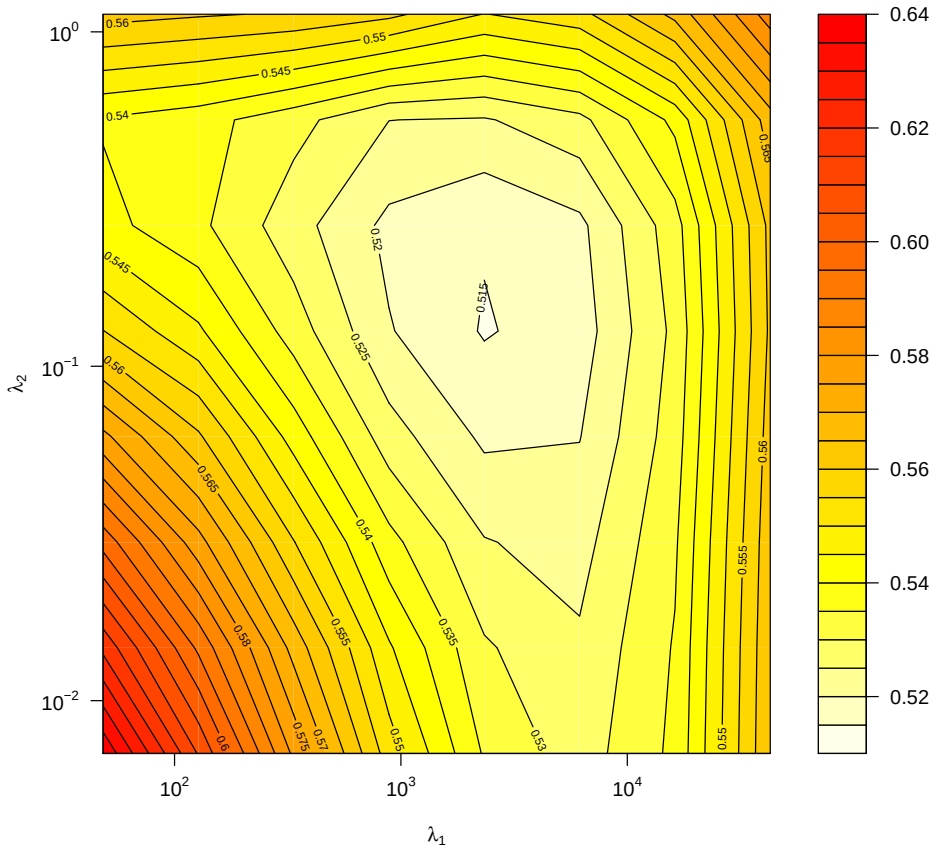

Figure 3: A heat map of cross validated errors, used to determine the tuning parameters λ_1 and λ_2, with contours overlaid.

the bootstrap sample of each group set to be the same as the original sample size. Then we apply our proposed approach on each bootstrap sample to estimate the group mean curves and take the differences among three groups.

The right panel of Figure 4 also shows the 95% pointwise bootstrap confidence bands based on the 1000 bootstrap samples. Due to the relatively large sample sizes of the major depression group and the healthy controls, the confidence band for the difference between major depression patients and healthy controls is narrower than those for the other two comparisons. The confidence bands suggest that the mean IRF of major depression patients is lower than the controls between 75 and 90 minutes and the mean IRF of the bipolar disorder patients is lower than the controls between 75 and 105 minutes.

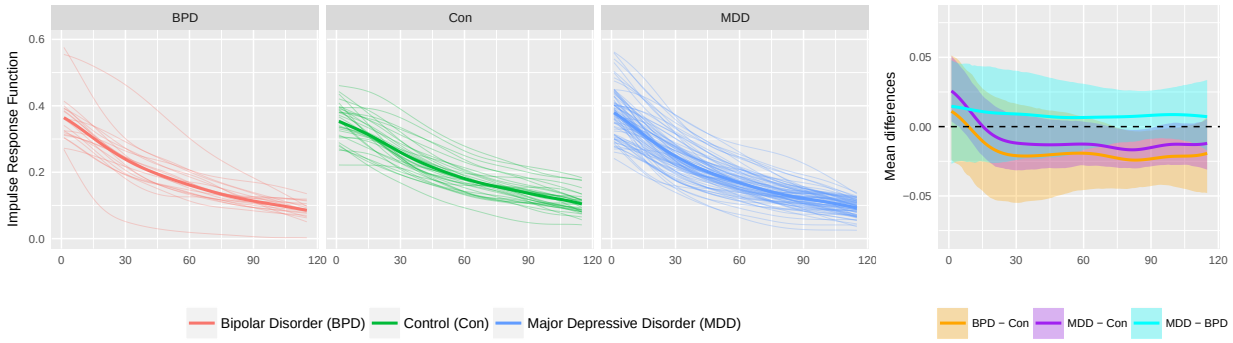

Figure 4: The left three panels show the mean curves of the three groups (thick) and the estimated individual IRFs (thin) obtained using our method. The right panel shows the 95% pointwise bootstrap confidence intervals (shaded areas) with the group mean differences (solid) estimated from the original sample overlaid.

To provide a frame of reference, we also model these data using the one-tissue compartment model, which has been deemed a reasonable compartment model structure for this tracer (Frankle *et al.* (2006); Ogden *et al.* (2007)), and using a two-tissue compartment model. The fitted group mean and individual IRFs obtained using the three methods are presented in Figure A.2 in the appendix.

IRFs for four selected subjects, estimated using our approach and the compartment models, are shown in the top row of Figure 5. The four subjects were selected based on the overall integrated squared difference in the IRFs obtained by the proposed approach and the compartment models. The first subject has the smallest discrepancy between IRF estimates under different approaches; the discrepancies of the second and third subjects are at the 25th and 75th percentiles, respectively; and the fourth subject has the largest disagreement in IRF estimates. For the first and second subjects the estimates are similar, but differences for the remaining subjects can be clearly observed, which suggests that parametric models may be appropriate for some subjects but not others. The bottom row of Figure 5 shows the observed data and the estimated TACs for the same four subjects. Visual inspection of these panels suggest that our method can substantially outperform the parametric approaches in terms of fitted values. Figure A.3 in the appendix shows the distributions of root integrated mean square error of the observed and fitted TACs for all the 137 subjects and, because our approach has smaller RIMSE and fewer extreme values

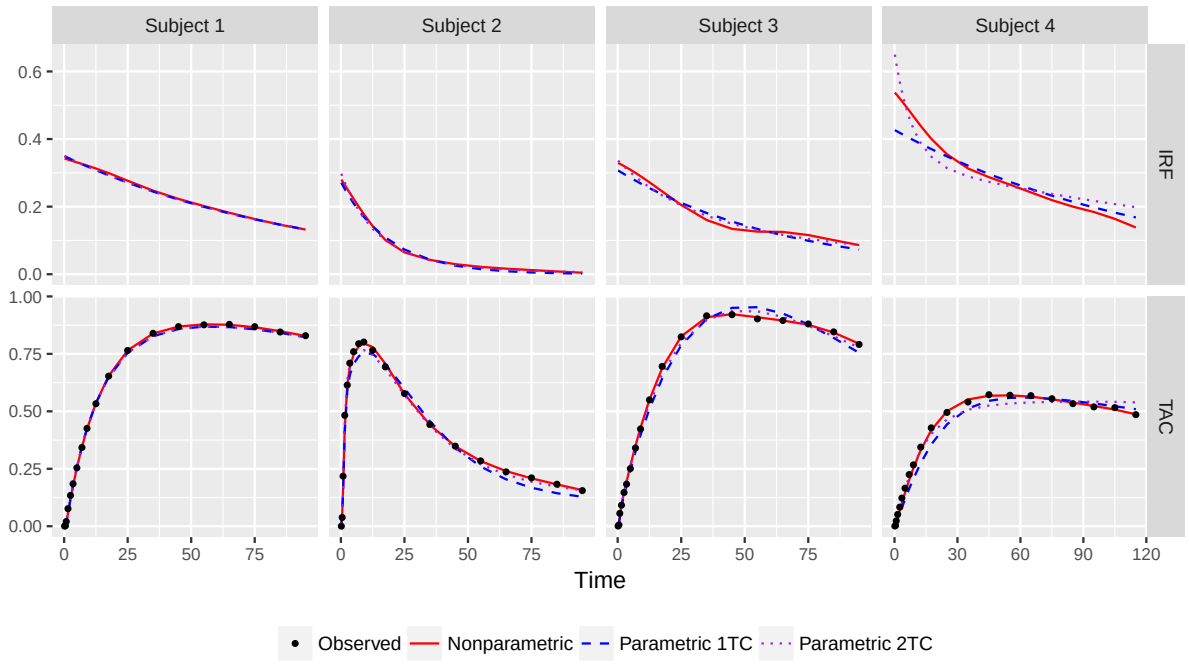

Figure 5: The top row shows comparisons between estimated IRFs using our proposed approach and the 1TC and 2TC models for four selected subjects. The bottom row shows comparisons between estimated TACs using our proposed approach and the 1TC and 2TC models for the same subjects with observed curves overlaid.

than the parametric models, supports the conclusion that nonparametric estimation can produce better fits.

Lastly, Figure 6 shows the residuals obtained using both our nonparametric approach and the parametric methods. For all models, residuals are computed by convolving a subject's estimated IRF and input function to produce the fitted value for the TAC, and subtracting the fitted TAC from the observed TAC at every observation time. Residuals obtained using our approach have roughy constant variance and low within-subject correlation, while the residuals obtained through parametric modeling have larger variability and stronger within-subject correlation. As a supplement to the Figure 6, we provide box plots of the residuals in Figure A.4 in the appendix. The residuals of our proposed method are centered at approximately zero, while the parametric models can miss trends in the data due to relatively inflexible mean structures.

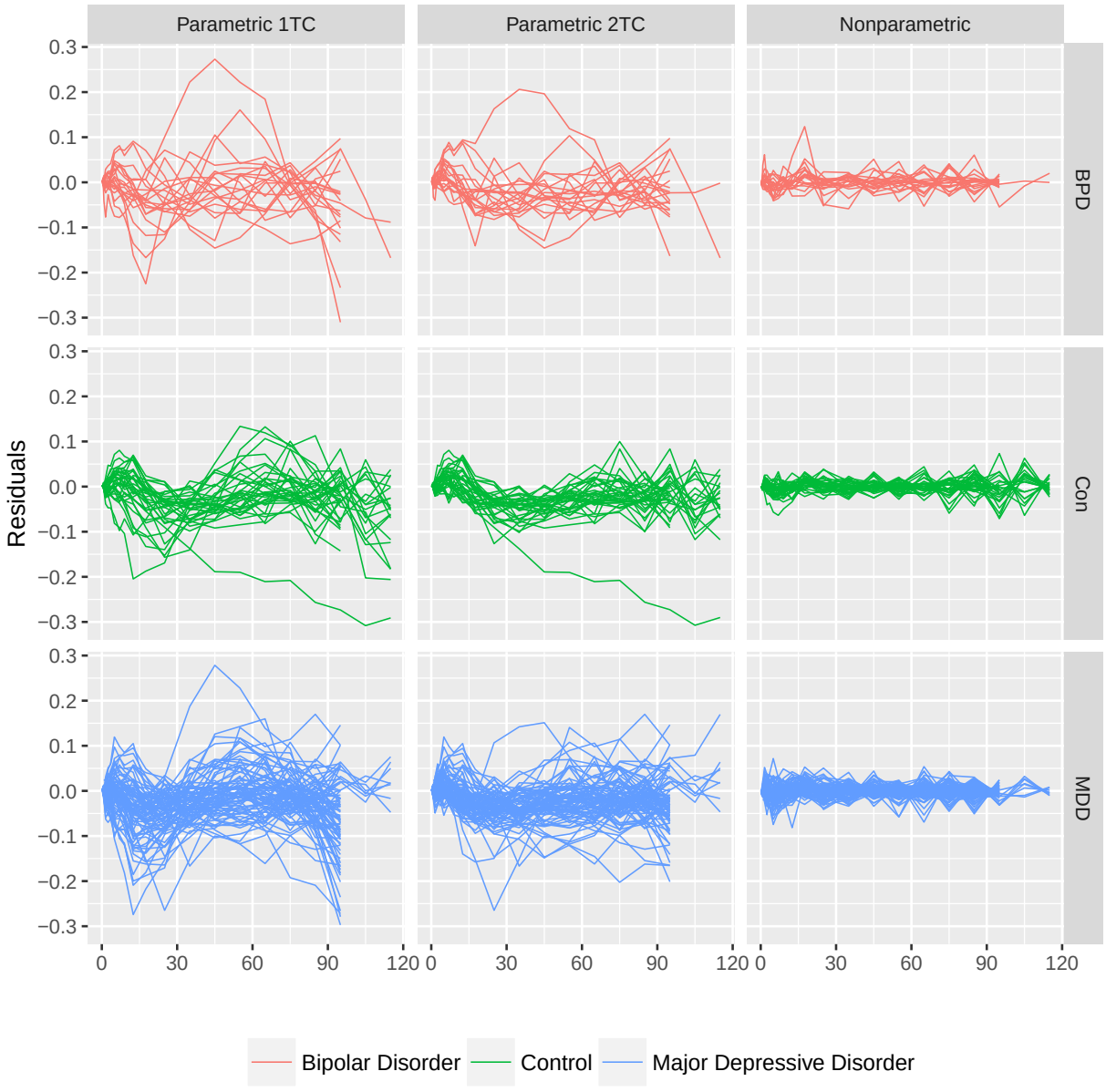

Figure 6: Spaghetti plots with residuals obtained using 1TC (left) and 2TC (middle) models and our proposed approach (right) by diagnosis group.

4.2 Analysis of the difference between the midbrain and the reference region

Next, we focus on isolating the binding capacity that is specific to the target protein. As mentioned in Section 1, tracer molecules may bind to their target protein ("specifi-

cally bound") or they may be associated with other macromolecular components ("non-specifically associated"). However, the observed PET data can measure only total concentration, consisting of unbound molecules as well as those bound to either type of protein, and is unable to discriminate among those states. Thus, as mentioned in Section 1.1, it is common in practice to designate a region that is devoid of the target protein as a "reference region". If non-specific association is uniform across the brain (as is always assumed), a reference region, which will allow only non-specific association, will allow better focus on the binding of the tracer to the specific target protein.

As we discussed in Section 1.3, the total volume of distribution of the tracer (V_T) is a commonly used summary measure for the parametric approaches. It is made up of two components: one that is only involved with specific binding to the target; and the other includes everything else, including volume of unbound tracer and tracer that is associated with other macromolecular components. The V_T of the region of interest represents the total volume of the two components while the V_T of a reference region consists only the second component. Thus, a standard measure of binding is $V_{T\_\text{region}} - V_{T\_\text{ref}} = \int_0^\infty f_{\text{region}}(t)dt - \int_0^\infty f_{\text{ref}}(t)dt = \int_0^\infty \left(f_{\text{region}} - f_{\text{ref}}\right)(t)dt$, i.e., that the binding measure is based on a functional of the difference between two IRFs. Although $V_{T\_\text{region}} - V_{T\_\text{ref}}$ refers only to parametric analysis and the nonparametric analogue $\int_0^{t_{end}} \left(f_{\text{region}} - f_{\text{ref}}\right)(t)dt$ doesn't have the same biological interpretation, it is still reasonable to take the difference between the IRFs obtained from nonparametric approaches because the difference pertains only to specific binding component of the IRF.

In this analysis, we designate the midbrain as the region of interest and the cerebellar gray matter as the reference region (Ichise *et al.*, 2003; Parsey *et al.*, 2005); although in fact it is not a perfect reference region (Parsey *et al.*, 2006a), we use it to illustrate our modeling approach. We modeled IRFs in both the region of interest and the reference region using our approach described in Section 2, with values for the tuning parameters determined separately. The differences between regions within each subject are estimated by subtracting the IRFs of the reference region from those of the region of interest.

The left three panels of Figure 7 display the estimated difference between the IRFs of the region of interest and reference region using this approach. Results indicate that the

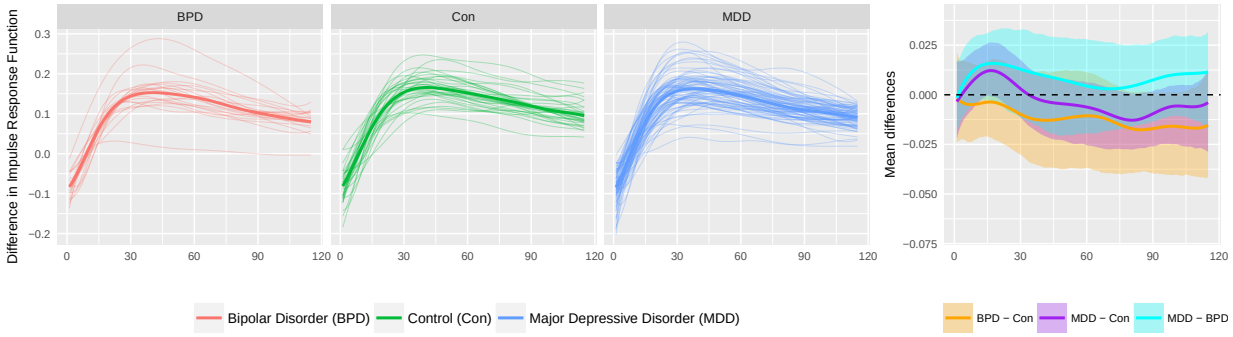

Figure 7: The left three panels shows the estimated differences in IRFs, including both group mean and individual curves, obtained using our approach. The right panel shows the 95% pointwise bootstrap confidence bands (shaded areas) with the group mean differences (solid) estimated from the original sample overlaid.

difference curve starts negative and, as time goes on, reaches a peak and then decreases. The negative difference at the beginning may be due to faster initial uptake in the reference region than in the midbrain region; this rate is unrelated to receptor availability. As time goes on, precisely because of the specific binding of the target protein in the region of interest, tracer molecules in the reference region would tend to be washed out relatively early compared to those in the midbrain. The group differences are shown in the right panel of Figure 7, which indicate that the mean difference IRF of the healthy controls is higher than the patients with major depressive disorder after 30 minutes and that the mean difference IRF of the patients with bipolar disorder is lower than the other two groups for almost the entire time span.

As in Section 4.1, we perform a bootstrap analysis to construct pointwise confidence bands of the group differences of the mean difference IRFs. For each bootstrap sample, we fit the models on the region of interest and on the reference region. Therefore, the group differences of the estimated mean difference IRFs can be obtained for all the bootstrap samples. The 95% pointwise confidence bands shown in the right panel of Figure 7 are constructed based on the bootstrap estimates of the mean difference curves. In contrast to what we observed with the midbrain TACs in Section 4.1, all the confidence bands cover 0 for the entire time range, indicating an insignificant difference among the three diagnosis groups in terms of the mean difference IRFs.

5 Discussion

We proposed a nonparametric model fitting framework that estimates the IRF using functional data analytic techniques. For the first time ever, our method models dynamic PET data from multiple subjects simultaneously. In our approach, IRFs are estimated using a linear mixed effects functional data model with population-level fixed effects and subject-level random effects. In accordance with our biological understanding of the IRF, we imposed appropriate non-negativity and monotonicity constraints on the estimates when fitting the model. Because of its flexibility, our model can be used generally for data with any tracer. The application of our approach to clinical PET data indicates that it successfully captures the structure in IRF, both when we model the region of interest only and when we model the difference between the region of interest and a reference region. Finally, pointwise confidence intervals of the estimated curves were constructed based on bootstrap samples.

In the most general sense, one can view the relationship between the TAC and the input function as a function-on-function regression problem (Scheipl *et al.*, 2015), although it may be more appropriately posed as a model with a historical term: the TAC response at a particular time depends on the history of the (observed) input function up to that time and on the (unknown) subject-specific impulse response function through the convolution of these functions. In contrast to the standard functional historical model, however, the association between the TAC and the input function is not determined by a bivariate coefficient surface, but by an IRF that is assumed to be unchanging over the course of the experiment. Moreover, in PET modeling the IRF is known to be covariate- and subject-specific. The inclusion of factor variables and subject effects in a traditional functional historical models was considered in Rügamer *et al.* (2018). Our proposal similarly includes covariate and subject effects, but does so while accounting for biological constraints that IRFs are constant during the experiment, are monotonically decreasing and are non-negative.

Our proposal focuses on the use of functional data approaches to increase the flexibility in estimating the IRF in comparison to methods that focus the estimation of rate parameters in a compartment model. We do this through the convolution of the IRF and input function which reduces the model to a function-on-scalar regression problem. To this

framework, we add scientifically relevant constraints on monotonicity and non-negativity to the usual estimation process. A careful consideration of the input function and TAC from the perspective of a function-on-function regression model would allow one to study the adequacy of the convolution operator, and is an important direction for future work.

Additional extension of our methodology may take several directions. Because of the way we construct the model, additional covariates, including continuous variables, can be incorporated in the model if the IRF is thought to be associated with those covariates. In addition, it would be useful to develop a goodness-of-fit test based on the estimated curves to evaluate how well the standard parametric models are able to describe the observed data. The development of approaches for TACs in multiple regions is conceptually possible in our modeling framework, but suitable and computationally feasible models for the covariance across regions may be challenging. Lastly, the classification of subjects into diagnostic groups based on PET imaging data is of general interest. However, given the overlap among groups in our data as shown in Figure 4, accurate classification based on PET data alone may not be successful.

6 Acknowledgments

This work was supported in part by Award R01HL123407 from the National Heart, Lung, and Blood Institute, Award R01NS097423 from the National Institute of Neurological Disorders and Stroke, and Award R01EB024526 from the National Institute of Biomedical Imaging and Bioengineering.

References

Cunningham, V. J. and Jones, T. (1993). Spectral analysis of dynamic PET studies. *Journal of Cerebral Blood Flow & Metabolism* **13**:15–23.

Frankle, W. G., Slifstein, M., Gunn, R. N., Huang, Y., Hwang, D.-R., Darr, E. A., Narendran, R., Abi-Dargham, A., and Laruelle, M. (2006). Estimation of serotonin trans-

porter parameters with 11C-DASB in healthy humans: Reproducibility and comparison of methods. *Journal of Nuclear Medicine* **47**:815–826.

Gunn, R. N., Gunn, S. R., and Cunningham, V. J. (2001). Positron emission tomography compartmental models. *Journal of Cerebral Blood Flow & Metabolism* **21**:635–652.

Gunn, R. N., Gunn, S. R., Turkheimer, F. E., Aston, J. A., and Cunningham, V. J. (2002). Positron Emission Tomography compartmental models: A basis pursuit strategy for kinetic modeling. *Journal of Cerebral Blood Flow & Metabolism* **22**:1425–1439.

Ichise, M., Liow, J.-S., Lu, J.-Q., Takano, A., Model, K., Toyama, H., Suhara, T., Suzuki, K., Innis, R. B., and Carson, R. E. (2003). Linearized Reference Tissue Parametric Imaging Methods: Application to [11C]DASB Positron Emission Tomography Studies of the Serotonin Transporter in Human Brain. *Journal of Cerebral Blood Flow & Metabolism* **23**:1096–1112.

Innis, R., Cunningham, V., Delforge, J., Fujita, M., Gjedde, A., Gunn, R., Holden, J., Houle, S., Huang, S.-C., Ichise, M., Iida, H., Ito, H., Kimura, Y., Koeppe, R., Knudsen, G., Knuuti, J., Lammertsma, A., Laruelle, M., Logan, J., Maguire, R., Mintun, M., Morris, E., Parsey, R., Price, J., Slifstein, M., Sossi, V., Suhara, T., Votaw, J., Wong, D., and Carson, R. (2007). Consensus nomenclature for in vivo imaging of reversibly binding radioligands. *Journal of Cerebral Blood Flow & Metabolism* **27**:1533–1539.

Jiang, C.-R., Aston, J. A., and Wang, J.-L. (2015). A functional approach to deconvolve dynamic neuroimaging data. *Journal of the American Statistical Association* (in press).

Jiang, H. and Ogden, R. T. (2008). Mixture modeling for dynamic PET data. *Statistica Sinica* **18**:1341.

Lin, Y., Haldar, J. P., Li, Q., Conti, P. S., and Leahy, R. M. (2014). Sparsity constrained mixture modeling for the estimation of kinetic parameters in dynamic PET. *Medical Imaging, IEEE Transactions on* **33**:173–185.

Logan, J., Fowler, J. S., Volkow, N. D., Wolf, A. P., Dewey, S. L., Schlyer, D. J., MacGregor, R. R., Hitzemann, R., Bendriem, B., Gatley, S. J., and Christman, D. R. (1990).

Graphical analysis of reversible radioligand binding from time-activity measurements applied to [N-11C-methyl]-(-)-cocaine PET studies in human subjects. *Journal of Cerebral Blood Flow & Metabolism* **10**:740–747.

Malison, R. T., Price, L. H., Berman, R., Van Dyck, C. H., Pelton, G. H., Carpenter, L., Sanacora, G., Owens, M. J., Nemeroff, C. B., Rajeevan, N., *et al.* (1998). Reduced brain serotonin transporter availability in major depression as measured by [123 I]-2β-carbomethoxy-3β-(4-iodophenyl) tropane and single photon emission computed tomography. *Biological psychiatry* **44**:1090–1098.

Meier, P. and Zierler, K. L. (1954). On the theory of the indicator-dilution method for measurement of blood flow and volume. *Journal of Applied Physiology* **6**:731–744.

Miller, J. M., Everett, B. A., Oquendo, M. A., Ogden, R. T., Mann, J. J., and Parsey, R. V. (2016). Positron emission tomography quantification of serotonin transporter binding in medication-free bipolar disorder. *Synapse* **70**:24–32.

Miller, J. M., Hesselgrave, N., Ogden, R. T., Sullivan, G. M., Oquendo, M. A., Mann, J. J., and Parsey, R. V. (2013). Positron Emission Tomography quantification of serotonin transporter in suicide attempters with major depressive disorder. *Biological Psychiatry* **74**:287–295.

Morris, J. S. (2015). Functional regression. *Annual Review of Statistics and Its Application* **2**:321–359.

Ogden, R. T., Ojha, A., Erlandsson, K., Oquendo, M. A., Mann, J. J., and Parsey, R. V. (2007). In vivo quantification of serotonin transporters using 11C-DASB and Positron Emission Tomography in humans: Modeling considerations. *Journal of Cerebral Blood Flow & Metabolism* **27**:205–217.

O'Sullivan, F., Muzi, M., Spence, A. M., Mankoff, D. M., O'Sullivan, J. N., Fitzgerald, N., Newman, G. C., and Krohn, K. A. (2009). Nonparametric residue analysis of dynamic PET data with application to cerebral FDG studies in normals. *Journal of the American Statistical Association* **104**:556–571.

Parsey, R. V., Hastings, R. S., Oquendo, M. A., Huang, Y.-y., Simpson, N., Arcement, J., Huang, Y., Ogden, R. T., Van Heertum, R. L., Arango, V., *et al.* (2006a). Lower serotonin transporter binding potential in the human brain during major depressive episodes. *American Journal of Psychiatry* **163**:52–58.

Parsey, R. V., Kent, J. M., Oquendo, M. A., Richards, M. C., Pratap, M., Cooper, T. B., Arango, V., and Mann, J. J. (2005). Acute Occupancy of Brain Serotonin Transporter by Sertraline as Measured by [11C]DASB and Positron Emission Tomography. *Biological Psychiatry* **59**:821–828.

Parsey, R. V., Ojha, A., Ogden, R. T., Erlandsson, K., Kumar, D., Landgrebe, M., Van Heertum, R., and Mann, J. J. (2006b). Metabolite considerations in the in vivo quantification of serotonin transporters using 11C-DASB and PET in humans. *Journal of Nuclear Medicine* **47**:1796–1802.

Peng, J.-Y., Aston, J. A., Gunn, R. N., Liou, C.-Y., and Ashburner, J. (2008). Dynamic positron emission tomography data-driven analysis using sparse Bayesian learning. *IEEE Transactions on Medical Imaging* **27**:1356–1369.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis*. Springer Series in Statistics. Springer, second edition.

Reiss, P., Huang, L., and Mennes, M. (2010). Fast function-on-scalar regression with penalized basis expansions. *International Journal of Biostatistics* **6**:article–28.

Rügamer, D., Brockhaus, S., Gentsch, K., Scherer, K., and Greven, S. (2018). Boosting factor-specific functional historical models for the detection of synchronization in bioelectrical signals. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **67**:621–642.

Ruppert, D. (2002). Selecting the number of knots for penalized splines. *Journal of Computational and Graphical Statistics* **11**:735–757.

Scheipl, F., Staicu, A.-M., and Greven, S. (2015). Functional additive mixed models. *Journal of Computational and Graphical Statistics* **24**:477–501.

Slifstein, M. and Laruelle, M. (2001). Models and methods for derivation of in vivo neuroreceptor parameters with PET and SPECT reversible radiotracers. *Nuclear Medicine and Biology* **28**:595–608.

Sullivan, G. M., Ogden, R. T., Oquendo, M. A., Kumar, J. D., Simpson, N., Huang, Y.-Y., Mann, J. J., and Parsey, R. V. (2009). Positron emission tomography quantification of serotonin-1A receptor binding in medication-free bipolar depression. *Biological Psychiatry* **66**:223–230.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)* **73**:3–36.

Zanderigo, F., Parsey, R. V., and Ogden, R. T. (2015). Model-free quantification of dynamic PET data using nonparametric deconvolution. *Journal of Cerebral Blood Flow & Metabolism* **35**:1368–1379.

Zeng, F. and Goodman, M. M. (2013). Fluorine-18 radiolabeled heterocycles as PET tracers for imaging β-amyloid plaques in Alzheimer's disease. *Current Topics in Medicinal Chemistry* **13**:909–919.