Quality and Precision of Parametric Images Created From PET Sinogram Data by Direct Reconstruction: Proof of Concept

Mustafa E. Kamasak*, Bradley T. Christian, Charles A. Bouman, and Evan D. Morris

***Abstract*—We have previously implemented the direct reconstruction of dense kinetic model parameter images ("parametric images") from sinogram data, and compared it to conventional image domain kinetic parameter estimation methods [1]–[3]. Although it has been shown that the direct reconstruction algorithm estimates the kinetic model parameters with lower root mean squared error than the conventional image domain techniques, some theoretical obstacles remain. These obstacles include the difficulty of evaluating the accuracy and precision of the estimated parameters. In image domain techniques, the reconstructed time activity curve (TAC) and the model predicted TAC are compared, and the goodness-of-fit is evaluated as a measure of the accuracy and precision of the estimated parameters. This approach cannot be applied to the direct reconstruction technique as there are no reconstructed TACs. In this paper, we propose ways of evaluating the precision and goodness-of-fit of the kinetic model parameters estimated by the direct reconstruction algorithm. Specifically, precision of the estimates requires the calculation of variance images for each parameter, and goodness-of-fit is addressed by reconstructing the difference between the measured and the fitted sinograms. We demonstrate that backprojecting the difference from sinogram space to image space creates error images that can be examined for goodness-of-fit and model selection purposes. The presence of nonrandom structures in the error images may indicate an inadequacy of the kinetic model that has been incorporated into the direct reconstruction algorithm. We introduce three types of goodness-of-fit images. We propose and demonstrate a number-of-runs image as a means of quantifying the adequacy or deficiency of the model. We further propose and demonstrate images of the F statistic and the change in the Akaike Information Criterion as devices for identifying the statistical advantage of one model over another at each voxel. As direct reconstruction to parametric images proliferates, it will be essential for imagers to adopt methods such as those proposed herein to assess the accuracy and precision of their parametric images.**

***Index Terms*—Akaike information criterion, compartmental models, error images, F statistic, four-dimensional (4-D) positron emission tomography (PET) reconstruction, goodness of fit, residuals, runs test, variance images.**

Manuscript received August 19, 2013; revised November 01, 2013; accepted November 14, 2013. Date of publication December 12, 2013; date of current version February 27, 2014. *Asterisk indicates corresponding author.*

*M. E. Kamasak is with the Department of Computer Engineering, Istanbul Technical University, 38030 Istanbul, Turkey (e-mail: kamasak@itu.edu.tr).

B. T. Christian is with Waisman Brain Imaging Laboratory, University of Wisconsin, Madison, WI 53705 USA (e-mail: bchristian@wisc.edu).

C. A. Bouman is with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA (e-mail: bouman@ecn.purdue.edu).

E. D. Morris is with Yale PET Center, Yale University, New Haven, CT 06520 USA (e-mail: evan.morris@yale.edu).

Digital Object Identifier 10.1109/TMI.2013.2294627

I. Introduction

POSITRON emission tomography (PET) imaging is unique among medical imaging modalities because of its molecular specificity. To realize this unique power of PET often requires the application of sophisticated kinetic modelling. Dynamic PET data represent a combination of emissions from all states of the radioactive tracer. Modelling and parameter estimation facilitate extracting the contribution to the data attributable solely to a particular state of the tracer. Depending on the tracer and kinetic model parameters, the contribution of the tracer in a certain state can be interpreted physiologically. Conventional kinetic model parameter estimation techniques require the reconstruction of the acquired sinogram data into emission-time images. Voxel-based time activity curves (TAC) are then extracted from the reconstructed images, and the model parameters are estimated using linear [4]–[9] and nonlinear methods [10], [11]. Linear techniques transform the data, so that the parameters of interest can be estimated by linear regression methods, while nonlinear techniques generally estimate the kinetic parameters by iteratively minimizing a properly weighted distance metric between the extracted TACs and the model. These image domain methods are common because they are simple and intuitive. However, the signal-to-noise (SNR) ratio of the reconstructed PET images is usually not high enough for voxel-by-voxel-based kinetic model parameter estimation except in the case of certain simplified models which yield compound parameters. In other words, if it is desired to estimate every model parameter at each voxel, image domain methods may suffer high spatial variance. Therefore, in the conventional image domain techniques, the TAC is not extracted from individual voxels but from a region of interest made up of multiple voxels that are assumed to be physiologically similar. High noise at individual voxels in the emission-time images had led to another approach for the solution of the problem: direct reconstruction of kinetic model parameters from the sinograms.

Direct reconstruction was proposed earlier for single positron emission tomography (SPECT) [12], [13] and PET [14]–[17]. Computational complexity and lack of other technologies (i.e., inter-frame registration methods) limited the use of direct reconstruction. After computationally efficient algorithms were developed, direct reconstruction was applied to different imaging modalities [18], [19], to different compartment models [1], [20], [21]. The direct reconstruction of kinetic model parameters were

later improved by many studies [22]–[28]. Refer to [24], [29], [30] for surveys of the direct kinetic model parameter estimation techniques.

There are many factors that may affect the reliability of the estimated kinetic model parameters. Among them are low signal-to-noise ratio (SNR) of the measured data, convergence to a local minima, and an inadequate kinetic model. Therefore, we seek ways to assess the reliability of the estimations. In image domain parameter estimation techniques, the reconstructed and model-predicted TACs are compared. This comparison often yields multiple measures of the precision of the estimated parameters. During the direct reconstruction procedure, there are neither reconstructed TACs nor explicit model-predicted TACs. The estimation is performed on the sinogram data without the reconstruction of images over time (and hence without any voxelwise TACs). Therefore, we cannot assess the model fit to the data in the conventional ways.

In this paper, we analyze the estimation reliability of the direct reconstruction technique. That is, how precise are the estimated parameters at each voxel? How can we know that a selected kinetic model is appropriate? How can we know that a particular model is better than an alternative model? To answer these questions, we have adapted some standard practices from image domain parameter estimation to image reconstruction. First, we present variance images for each parameter based on Monte Carlo simulations. The images of the variance—or a related error index—provide an estimate of the precision of the voxelwise parameter estimates. Without such estimates, it is impossible to say definitively that one image is statistically different from another. Second, we introduce a new type of image upon which assessments of precision and goodness-of-fit can be made. These new images are based on filtered backprojections of the residual error from the sinogram domain into the image domain. Similar to our work, Votaw *et al.* [15] backprojected the residual sinograms, however they did not compare model fits to data and did not analyze the goodness-of-fit. We evaluate the potential of error images for visualizing the goodness-of-fit of the model to the data over time and space, simultaneously. With the images of the filter backprojected residuals, an adequate kinetic model can be distinguished from a deficient one. That is, the residual images are proposed as a tool for proper model selection in direct reconstruction. Three established indexes, the number-of-runs, the F statistic, and the Akaike Criterion, are extended from standard 1-D parameter estimation to evaluate the error images that we create.

II. Methods

Fig. 1 shows the three domains of parametric image generation. The parameter domain refers to the kinetic model parameters that we seek to estimate, the image domain refers to the emission values of the voxels, and the sinogram domain signifies the projections of the emission values. These domains are connected to each other via the compartment and scanner models, respectively.

The compartment model uses the parameters at each voxel to compute the value of the emissions (radioactivity concentration) from each respective voxel at a given time. The computed emission images are the inputs to the scanner model, which computes their projections according to the characteristics of the scanner. In the parameter estimation problem (to produce parametric images), sinograms are the measured data acquired by the PET scanner. The scanner model and compartment model must be chosen carefully to yield physiologically meaningful estimates. The scanner model is chosen based on the geometry and response characteristics of the PET scanner. An appropriate kinetic model is selected by considering the underlying physiology and properties of the tracer.

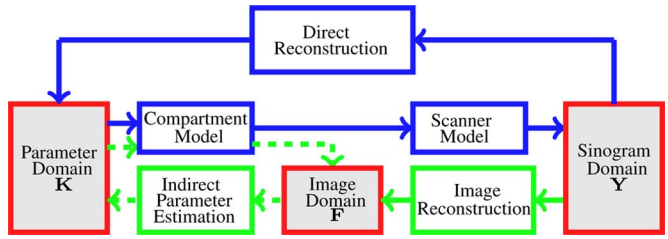


Fig. 1. Domains of the kinetic model parameter estimation problem—alternative algorithmic paths of direct versus indirect estimations. Conventionally, parameters are estimated indirectly (green paths). The indirect estimation has two steps: first, emission images are estimated from the sinogram data (solid green) and second, kinetic parameters are estimated from the emission images. The latter process is iterative (dashed green path) if the model is nonlinear. Alternatively (Kamasak, 2005), kinetic parameters can be estimated directly from the sinogram data (blue path) via an iterative procedure. The image domain (emission image) is skipped in the direct estimation.

A. Experimental Data Acquisition

A healthy male rhesus monkey (Macaca mulatta; 7 kg.) was scanned using a Siemens EXACT HR+ scanner. The animal was anaesthetized using ketamine (10 mg/kg) and subsequently maintained on 0.5%–1.5% isoflurane. A 5-min transmission scan using ^{68}Ge/^{68}Ga rod sources was acquired prior to administration of the radiopharmaceutical. ^{18}F-fallypride was administrated as a bolus injection (90 MBq) through an i.v. catheter placed in the saphanous vein. The data were collected into 40 time frames consisting of 6×0.5 min, 7×1 min, 5×2 min, 4×5 min, and 18×10 min frames for a total acquisition time of 220 min. The acquisition of the data was performed in 3-D mode (septa retracted) using $2\times$ angular compression, span of 9 and a maximum ring difference of 22, resulting in a 3-D sinogram with five segments $(-2, -1, 0, +1, +2)$.

During the acquisition, arterial blood samples were collected from an arterial catheter placed in the femoral artery. Commencing with injection of ^{18}F-fallypride, 1 mL samples were initially drawn at 10-s intervals for the first 3 min, 5 min-intervals for 15 min, and 30-min intervals until the end of the study. The arterial plasma input function of parent radioligand was obtained using the method of ethyl acetate extraction as previously reported [31].

After data acquisition, the subject was removed from anaesthesia and continuously monitored until fully alert. The experimental procedures were in accordance with institutional guidelines.

The 3-D sinogram was corrected for random coincidences, deadtime losses, scatter, attenuation, and scanner normalization

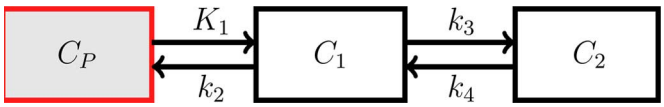


Fig. 2. Schematic diagram of two-tissue compartment, four parameter model. Plasma radioactivity (red box) is measured and therefore not a compartment in the mathematical sense.

[32]. The sinogram was not decay corrected as it was handled in the kinetic model (see Section II-B). To facilitate the implementation of our algorithm, the raw 3-D sinogram was rebinned into a 2-D matrix using the Fourier rebinning algorithm (FORE) supplied by the vendor (ECAT ver. 7.2) [33].

B. Kinetic Model

The kinetic model that we have used in our estimation framework is the standard four parameter, two-tissue compartment model used commonly to describe the uptake and retention of high specific activity receptor-ligand tracers in PET. This model, shown in Fig. 2, has four parameters representing the rate of tracer exchange between possible states. Depending on the nature of the injected tracer, these kinetic parameters would represent the rate constants of meaningful physiological processes such as metabolism or receptor-ligand association. The model requires a measured plasma input function and can be solved via numerical convolution of the input function with the impulse response function.[1] The predicted emissions, f, at voxel s, over time-frame $[t_i, t_j]$ are a function of kinetic parameters $K_s = [k_{1s}, k_{2s}, \ldots, k_{Ps}]^t$ (where P is the number of parameters in the model) that are unique to voxel s, and are given generically as

$$f(K_s)_{t_i}^{t_j} = \frac{1}{(t_j - t_i)} \times \int_{t_i}^{t_j} \left\{ V_B C_{WB}(s,t) + (1 - V_B) SA(t) \sum_{l=1}^{M} C_l(s,t) \right\} dt \tag{1}$$

where $SA(t)$ is the time-varying specific activity of the tracer, C_{WB} is activity concentration in the blood, V_B is the blood volume fraction of the tissue, M is the number of tissue compartments, and C_l is the molar concentration of the lth compartment of the kinetic model (depicted in Fig. 2). [2]

C. Direct Reconstruction of Parametric Images

Classical kinetic parameter estimation techniques in PET use a two-step approach. First the measured sinograms are reconstructed to emission images using a tomographic reconstruction technique. The compartment model parameters are then estimated at the region or voxel level by comparing the compartment model output to the TACs. In the direct reconstruction technique, the compartment model and the scanner model can be combined in order to estimate kinetic model parameters directly from the measured sinograms [1].

Direct reconstruction of parametric images from sinogram data has certain advantages over indirect estimation methods as all the available data are used simultaneously and spatial regularization is performed in the kinetic parameter domain.

Within the iterative direct reconstruction algorithm, the emissions map is projected to sinograms via the scanner model

$$Y_p(\mathbf{K}) = AF(\mathbf{K}) \tag{2}$$

where $Y_p(\mathbf{K})$ are the predicted sinograms, F are the emission images, $\mathbf{K}$ are the parametric images, and A is the projection matrix (also known as the scanner model). The parametric images are reconstructed by minimizing the cost function given as

$$\mathbf{K} = \arg\min_{\mathbf{K}} \{ -\mathrm{LL}(Y|\mathbf{K}) + \beta S(\mathbf{K}) \} \tag{3}$$

where LL is the log likelihood, Y are the measured sinograms, S represents the regularization kernels that are applied locally to each of the parametric images on each iteration of the minimization procedure, and β is a constant that determines the amount of regularization. At the end of each iteration, the parametric images are updated until the convergence criteria are met.

The minimization of (3) was implemented using the "parametric iterative coordinate descent" (PICD) algorithm that was introduced in [1]. PICD is an adaptation of the iterative coordinate descent (ICD) algorithm where each voxel's parameters are updated sequentially (while other voxels' parameters are kept constant) in order to minimize (3). In PICD, the kinetic parameters are transformed into linear and nonlinear parameters that are not coupled with each other. These nonlinear and linear parameters are then updated using a nested optimization approach which is computationally efficient.

Direct reconstruction was compared extensively to indirect estimation methods in our previous work [1]. More recently, direct reconstruction has been applied to different imaging modalities [18], [19], using different compartment models [20], [21]. In addition, a number of similar studies confirmed our results and demonstrated the advantages of the direct reconstruction approach [22]–[24], [26]–[28], [34], [35]. These studies indicate that the direct parametric reconstruction can substantially reduce estimation error in kinetic parameters compared to the indirect estimation methods. Direct reconstruction methods were reviewed in a survey paper by Tsoumpas *et al.* [29].

Choice of Spatial Regularization: As indicated by (3), direct reconstruction includes spatial regularization in one or more of the parameters. The degree of spatial regularization can be different for each parameter and has been initially computed according to the method of Saquib *et al.* [36]. In brief, this method includes maximum likelihood (ML) estimation of Markov random field (MRF) parameters from incomplete observations. The amount of regularization for all kinetic parameters is manually adjusted by scaling the β in (3) with a single scale factor.

[1]Some kinetic models (i.e., ones with nonlinear components) may not have closed form solutions and may require other forms of solution.

[2]This equation assumes that the tracer is not restricted from any part of the voxel. This may not be correct for all tracers and tissue types. In those cases, $(1 - V_B)$ can be changed by a new weighting coefficient. However, parameter identifiability should be considered whenever new parameters are added to the model.

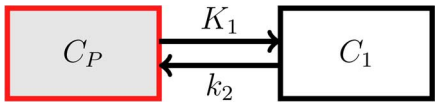


Fig. 3. Schematic diagram of one-tissue compartment, two parameter model. Plasma radioactivity (red box) is a measured function. This model was incorporated into direct reconstruction in order to test the use of backprojected residuals for kinetic model selection.

Volume of Distribution: Images of the distribution volume of the tracer can be calculated directly from the individual parametric images. The distribution volume for a two-tissue compartment model is given by

$$V_T = \frac{K_1}{k_2}\left(1 + \frac{k_3}{k_4}\right). \tag{4}$$

The distribution volume of the tracer is expected to be the highest in the striatum coinciding with the highest density of D2 receptors (largest k_3).

D. Backprojection of Residuals From the Sinogram Domain

In direct reconstruction, the kinetic model parameter estimation is performed by minimization of the objective function in the sinogram domain. Creation of emission images and TACs are bypassed. For direct reconstruction, a conventional goodness-of-fit analysis on each of the projections is not practical as it would not be possible to visualize the result of the analysis in the image domain. Therefore, we adapted the goodness-of-fit concept.

In sinogram space, we fit the model prediction to the acquired data for every time-frame and angle as a single fitted (4D) curve of events. However, to visualize all the fits collectively in image space, we have chosen to backproject the residuals between the fitted (predicted) and measured projections for all angles using filtered backprojection. Thus, let $\mathrm{Res}(\mathbf{K}, t_n)$ denote the residual sinogram

$$\mathrm{Res}(\mathbf{K}, t) = Y_p(\mathbf{K}, t) - Y(t_n) \tag{5}$$

where $Y(t_n)$ is the measured sinogram and $Y_p(\mathbf{K}, t_n)$ is the fitted sinogram at time frame n. Furthermore, let $\mathrm{ImRes}(\mathbf{K}, t_n)$ be the filtered backprojection of $\mathrm{Res}(\mathbf{K}, t)$ for time frame n. In the case of an ideal model fit to the data, the $\mathrm{Res}(\mathbf{K}, t_n)$ and $\mathrm{ImRes}(\mathbf{K}, t_n)$ would be zero mean Gaussian distributed for all time frames.

Utility of ImRes Images: In order to evaluate the utility of the backprojected residuals as a means of model selection, we reconstructed the parametric images using two different compartment models. First, we performed a direct reconstruction that incorporated the four parameter, two-tissue compartment model (shown in Fig. 2) and second, we performed a direct reconstruction with a two parameter, one-tissue compartment model (shown in Fig. 3). The two parameter model is likely inadequate (or at least inferior to the four parameter model) in describing the uptake of a tracer that binds to a receptor site (requiring a binding compartment), therefore we expected some visual indication of model deficiency in the backprojected residual images, $\mathrm{ImRes}(\mathbf{K})$.

E. Residual Image and Quantitative Measures of Goodness-of-Fit

1) Number-of-Runs Images and Runs Histograms: The adequacy of any fit of a model to data can be tested by examining the nonrandomness of the residuals using a runs test [37]. Assuming that residuals have zero mean, a "run" is defined as a subsequence of residuals having the same sign (the number-of-runs is equal to one plus the number of zero crossings of the residuals plotted versus an independent variable such as time). A very small number-of-runs is an indication of nonrandomness in the residuals and thus an inability of the model to describe fully all the components of the data. While a large number-of-runs is generally desirable, an extremely large number may be an indication of periodicity in the data that is not accounted for in the model. In the present work, we used distribution of the number-of-runs in image space, (i.e., an image of runs in time at every voxel in $\mathrm{ImRes}(\mathbf{K})$) as an image of the goodness-of-fit which can be used to assess the local performance of a given model. We also examined the histogram of the number-of-runs across all voxels as an indicator of the relative goodness-of-fit of two candidate models, globally. While the number-of-runs will vary from voxel to voxel for any model, a histogram of runs values with a higher mean for model A as compared to model B is an image-wide indication of a better fit with model A.

2) F Statistic: The F statistic can be used to compare nested compartment models [38], [39]. Let P_1 and P_2 denote the number of parameters in one and two-tissue compartment models, respectively, and T denotes the number of time frames. The F statistic for voxel s is computed as

$$F(s) = \frac{\frac{\mathrm{RSS}_1(s) - \mathrm{RSS}_2(s)}{P_2 - P_1}}{\frac{\mathrm{RSS}_2(s)}{T - P_2}} \tag{6}$$

where $\mathrm{RSS}_1(s)$ and $\mathrm{RSS}_2(s)$ are the sum of squares of the residual values of voxel s that are computed from ImRes images corresponding to one- and two-tissue compartment models. The residual sum of squares for voxel s is computed as

$$\mathrm{RSS}_l(s) = \sum_n w_n \mathrm{ImRes}_l(K_s, t_n)^2 \tag{7}$$

where w_n is a weight constant for time frame n and $\mathrm{ImRes}_l(K_s, t_n)$ is the residue value of voxel s obtained from the filter backprojected error image for compartment model l at time frame n. These coefficients (w_n) for time frames can be used to weight the image value according to the variance of the measurement error. The weights are typically chosen as the inverses of the noise variances. Although the noise variance cannot be known, it is commonly assumed to be inversely proportional to frame duration. Other weighting choices are available in the literature [40], [41]. In this study, no weighting was applied in the computation of RSS_i ($w_n = 1$ for $\forall n$).

From the F-distribution with $(P_2 - P_1, T - P_2)$ degrees-of-freedom, it is possible to find the ratio that corresponds to a particular significance level. For a two-tissue compartment model $P_2 = 4$, whereas for a one-tissue compartment model $P_1 = 2$. With 40 time frames, the F distribution with (2; 36) degrees-of-

freedom, an F-ratio that is higher than 5.3 corresponds to a significance level of $p < 0.01$, and a ratio higher than 3.3 corresponds to $p < 0.05$. Therefore, for the voxels whose F-ratios are higher than 5.3, there is a probability of 99% or greater $(p < 0.01)$ that the improved fit thanks to the two-tissue compartment model over the one-tissue compartment is not a chance happening.

3) Akaike Information Criteria: The Akaike information criteria (AIC) can be used to compare both nested and nonnested models [38], [42], [43]. The AIC for voxel s can be computed as follows:

$$\mathrm{AIC}(s) = T\ln\left(\frac{\mathrm{RSS}_l(s)}{T}\right) + 2(P_i + 1) \tag{8}$$

where $\ln(\cdot)$ denotes natural logarithm and P_i is the number of parameters (degrees-of-freedom) for compartment model i. Let $\mathrm{AIC}_1(s)$ and $\mathrm{AIC}_2(s)$ denote the AIC for one and two-tissue compartment models respectively for voxel s. The AIC difference for voxel s can be computed as

$$\begin{aligned}\Delta\mathrm{AIC}(s) &\triangleq \mathrm{AIC}_2(s) - \mathrm{AIC}_1(s) \\ &= T\ln\left(\frac{\mathrm{RSS}_2(s)}{\mathrm{RSS}_1(s)}\right) + 2(P_2 - P_1).\end{aligned} \tag{9}$$

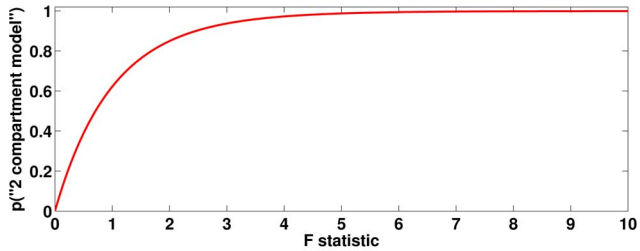
For the voxels whose ΔAIC values are negative, the two-tissue compartment model is better compared to the one-tissue compartment model. Conversely, the one-tissue compartment model is better for voxels with positive ΔAIC values. The probability of the TAC being consistent with a two-tissue compartment model is given as [38]

$$\begin{aligned}&p(\text{“two}-\text{tissue compartment model”}|\Delta\mathrm{AIC}) \\ &\qquad = \frac{e^{-0.5\Delta\mathrm{AIC}}}{1 + e^{-0.5\Delta\mathrm{AIC}}}.\end{aligned} \tag{10}$$

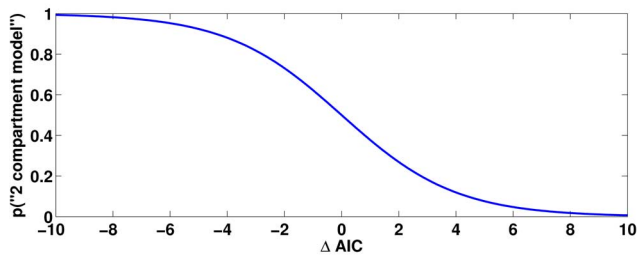
The probability of a voxel TAC being consistent with a two-tissue compartment model is shown in Fig. 4(b). ΔAIC values that are smaller than -5.9 correspond to a significance level of $p < 0.05$, and values smaller than -9.2 correspond to $p < 0.01$.

F. Monte Carlo Simulations

Parametric images of the uptake and retention of ^{18}F-fallypride into a monkey brain were reconstructed using the parametric iterative coordinate descent (PICD) algorithm [1]. For the purpose of calculating the variance images, the parametric images were assumed to be ground truth. Using the four parameter two-tissue compartment model that is described in Section II-B, the voxel emissions at different times were computed. Using the Siemens HR+ EXACT scanner model [44], the emission images were forward projected into 25 different sets of sinograms to which distinct instantiations of identically distributed Poisson noise were added. The number of noise realizations was chosen such that the final change in mean parametric images was below a certain percentage. Next, parametric images were directly reconstructed from these simulated sinograms. The absolute bias and coefficient of variance (CV) of each parametric image was calculated voxel by voxel



(a)



(b)

Fig. 4. (a) Probability of a voxel TAC being better modelled with two-tissue compartment model than one-tissue compartment model. The F ratio that is higher than 5.3 corresponds to $p < 0.01$, higher than 3.3 corresponds to $p < 0.05$ based on an F distribution with (2, 36) degrees-of-freedom. (b) Probability of a voxel TAC better modelled with two-tissue compartment model than one-tissue compartment model. Two-tissue compartment model is more suitable than one-tissue compartment model for voxels with negative AIC values. ΔAIC that is smaller than -9.2 corresponds to level $p < 0.01$, and smaller than -5.9 corresponds to a significance level of $p < 0.05$.

from these 25 sets of reconstructed parametric images. The absolute bias of parameter i for voxel s is computed as

$$\text{absolute bias} = |\bar{k}_{\mathrm{is}} - k_{\mathrm{is}}^{\mathrm{correct}}| \tag{11}$$

where $\bar{k}_{\mathrm{is}}$ is the mean value and $k_{\mathrm{is}}^{\mathrm{correct}}$ is the correct value of parameter i of voxel s. Furthermore, the coefficient of variation of parameter i for voxel s is computed as

$$\mathrm{CV}(k_{\mathrm{is}}) = \frac{\sigma(k_{\mathrm{is}})}{\bar{k}_{\mathrm{is}}} \tag{12}$$

where $\sigma(k_{\mathrm{is}})$ is the standard deviation of parameter i of voxel s. The mean $(\bar{k}_{\mathrm{is}})$ and standard deviation $(\sigma(k_{\mathrm{is}}))$ of parameter i of voxel s are computed from multiple noise realizations as follows:

$$\bar{k}_{\mathrm{is}} = \frac{1}{J}\sum_{j=1}^{J} k_{\mathrm{is}}^{j} \tag{13}$$

and

$$\sigma(k_{\mathrm{is}}) = \sqrt{\frac{\sum_{j=1}^{J}\left(k_{\mathrm{is}}^{j} - \bar{k}_{\mathrm{is}}\right)^2}{J-1}} \tag{14}$$

where J is the number of noise realizations and k_{is}^{j} is the estimated value of parameter i for voxel s of noise realization j.

III. Results

Parametric images resulting from application of the direct reconstruction approach with a two-tissue compartment model to

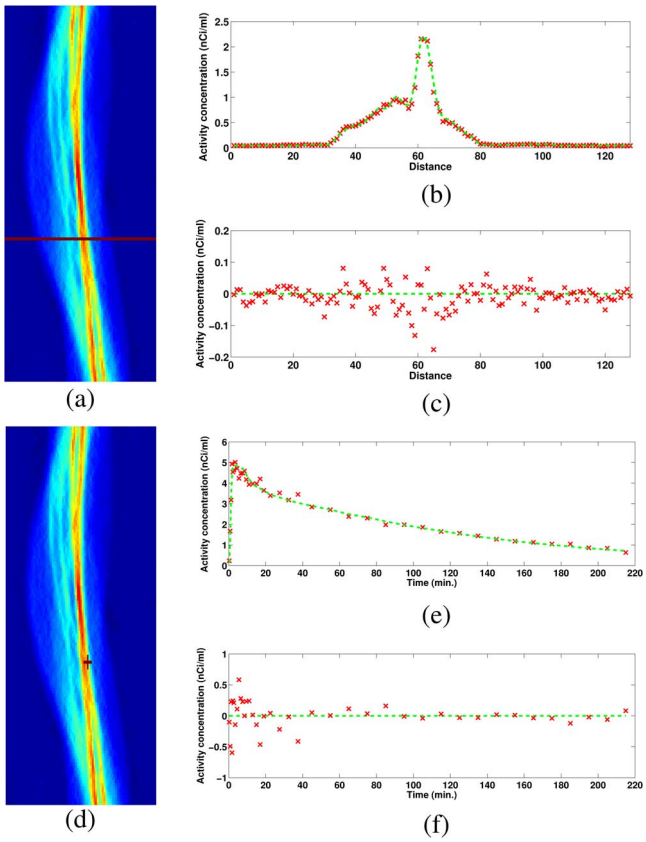


Fig. 5. (a) Sinogram data; time frame 30, angle 80 is identified by horizontal line through all distances. (b) Measured sinogram projection versus distance (red) is overlaid on the predicted sinogram projection (green). (c) Residuals of fit in (b). Residuals shown in (c) are one line of the residual sinogram that is filter backprojected to create error images. (d) Sinogram data; time frame 30, angle 80, radial distance 64 is marked by a plus sign. (e) Measured sinogram projection versus time (in minutes) is overlaid on the predicted sinogram projection (green). (f) Residuals of fit in (e).

^{18}F-fallypride data acquired in a monkey were created. The images were estimated directly from measured sinograms via (3); no intermediate TACs or emission images were created. Spatial regularization for each parameter [shown as $S(\mathbf{K})$ in (3)] was incorporated into reconstruction as described in Section II-C.

A. Residuals in Sinogram Domain

An example fit to the recorded counts in the sinogram domain (at a single projection angle and time) versus radial distance is shown in Fig. 5(a)–(c). The selected projection angle is indicated in Fig. 5(a) by a horizontal line across the sinogram. The measured and fitted curves at this particular time and projection angle are shown in Fig. 5(b). The selected residual plot is displayed in Fig. 5(c) as the error between the measured and fitted curves. The error shown in Fig. 5(c) appears uncorrelated in radial distance which suggests that the kinetic model (in this case a two-tissue compartment model) was adequate to fit the data. Had the residuals been distributed nonrandomly, we would have taken that as one indication of a deficiency of the kinetic model.

Similarly, it is possible to investigate the fit of a certain voxel in the sinogram over time frames as shown in Fig. 5(d)–(f). The fit of a certain voxel in the sinogram [marked with a plus sign in Fig. 5(d)] at different time frames is shown in Fig. 5(e). The residual in time of this sinogram voxel is shown in Fig. 5(f); again there is no apparent correlation in time.

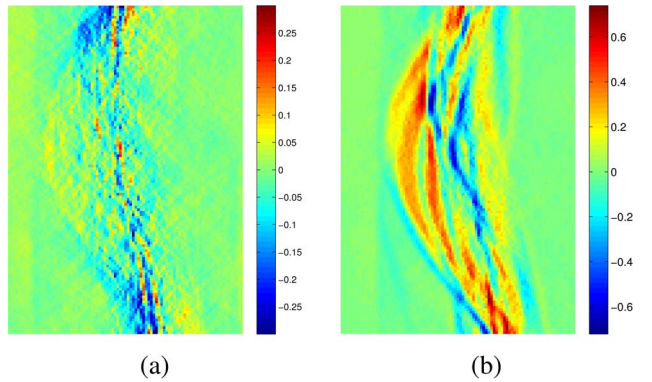


Fig. 6. Residual sinograms of (a) two-tissue (b) one-tissue compartment model for time frame 25.

The residual sinograms for time frame 25 corresponding to two-tissue and one-tissue compartment models are shown in Fig. 6(a) and (b), respectively. The range of residuals is larger for the one-compartment model than for the two-tissue compartment model. The residual sinogram for the one-tissue compartment model appears to have a more definite form as well.

B. Error Images

The error images, $\mathrm{ImRes}(\mathbf{K})$ reconstructed via filtered backprojection from the residual sinograms (see Fig. 6) are shown in Fig. 7. The left-hand column of Fig. 7 shows the emission image at various time frames from the same slice. The middle and right-hand columns show $\mathrm{ImRes}(\mathbf{K})$ for the models shown in Fig. 2 (two-tissue compartment, four parameter) and Fig. 3 (one-tissue compartment, two parameter), respectively. In a given row, $\mathrm{ImRes}(\mathbf{K})$ corresponds to a single time-frame, indicated at the left. All error images, $\mathrm{ImRes}(\mathbf{K})$, are displayed using the same grayscale. Purely randomly distributed residuals in the sinogram domain should correspond to filter backprojected residual images with no visible structure (i.e., "white").[3]

In the backprojected residual images $\mathrm{ImRes}(\mathbf{K})$, objects for the one-tissue compartment model are more apparent than for the two-tissue compartment model. These visually apparent structures are the results of large residuals clusters that are visible in Fig. 6(b). We interpret the residual images of the one-tissue compartment model as being less random than the corresponding residual images of the two-tissue compartment model.

C. Quantitative Analysis of Error Images

1) Runs Histogram: We can compare the behavior of the two models, globally, by examining a histogram of all the runs for all the voxels inside the brain (i.e., over multiple slices). The histograms in Fig. 8 corresponding to the application of the two models are clearly different. The shift of the mean number-of-runs from a lower value (for the one-tissue compartment model) to a higher value for the two-tissue compartment model [from Fig. 8(a) to (b)] is another indication of the superior description

[3]Due to varying variance in the residuals, some structure may be visible even in the case of purely randomly distributed residuals.

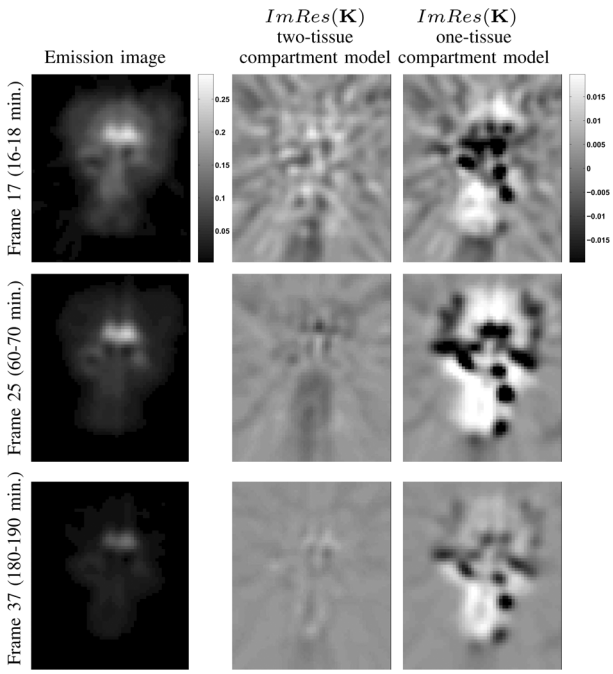


Fig. 7. Emission images and error images that are obtained from filter backprojected sinogram residuals at different time frames. The second and third columns show the error images that are corresponding two-tissue and one-tissue compartment models, respectively.

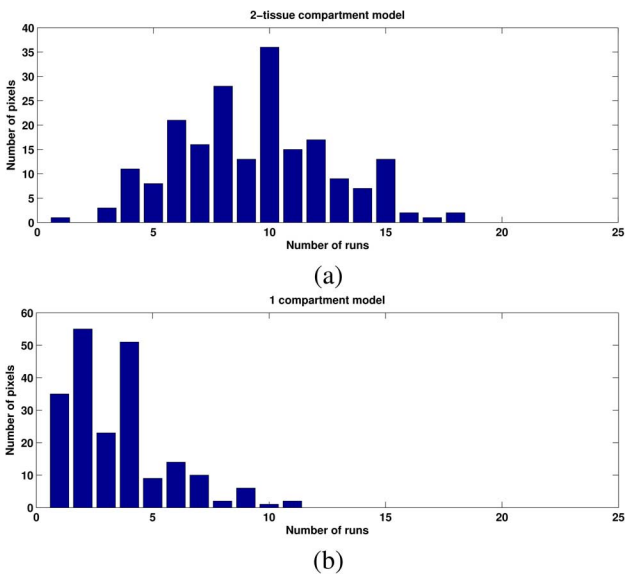


Fig. 8. Histogram of runs for the voxels of (a) two-tissue compartment model (b) one-tissue compartment model.

of the data by the four parameter model and thus greater randomness of the residuals at all voxels. The minimum number-of-runs is 1 and highest possible run value is 40 because there were 40 time frames in the reconstruction. Although the runs histogram does not retain spatial information, it depicts the overall behavior of the model across multiple voxels, simultaneously.

2) Number-of-Runs Images: Because the $\mathrm{Res}(\mathbf{K})$ for all time frames are backprojected at once, the number-of-runs in $\mathrm{ImRes}(\mathbf{K})$ at any given voxel is calculated from the number

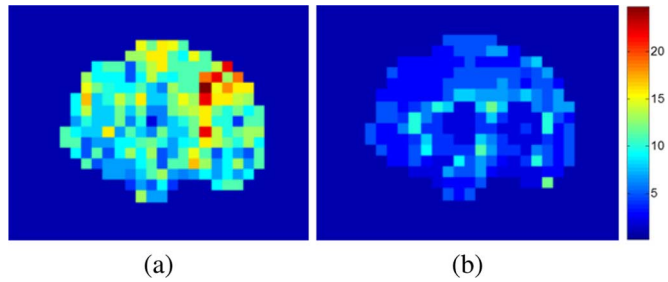


Fig. 9. Number-of-runs images obtained from $\mathrm{ImRes}(\mathbf{K})$ for (a) two-tissue compartment model (b) one-tissue compartment model. Images correspond to the same slice as shown in previous figures. Only the brain voxels are displayed.

of subsequences of positive and negative values over time at that voxel. Fig. 9(a), and (b) shows the number-of-runs images for all the voxels within the brain obtained from $\mathrm{ImRes}(\mathbf{K})$ for both the two and one-tissue compartment models, respectively. A low number-of-runs at a given voxel indicates a possible deficiency of the model for describing the temporal behavior (uptake, retention, efflux) of the tracer at the corresponding voxel. The "holes" in the runs image for the one-tissue compartment model [Fig. 9(b)] at the locations of the striata are confirmation of the maximal deficiency of the two parameter model in those regions where both fast and slow components are most needed to describe the data (i.e., in areas of very high binding). Generally higher numbers of runs in Fig. 9(a) than (b), indicates more randomness in the error, and therefore better fits to the data in most regions of the brain when more parameters are added to the kinetic model.

3) F Statistic and $\Delta\mathrm{AIC}$ *Images:* Calculation of the F statistic at every voxel, as described in Section II-E2, creates an image of the F-ratio values to determine the statistical improvement achieved by a two-tissue compartment model over a one-tissue compartment model. As demonstrated in Fig. 10, the F statistic image can be thresholded [Fig. 10(b)] at the appropriate level to identify those voxels whose improved fit satisfies the desired probability level [Fig. 4(a)].

One can use another popular index of goodness-of-fit, AIC, for a similar purpose. AIC is a residual sum of squares index that is adjusted for number of parameters in a given model. As described in Section II-E3, the difference in AIC ($\Delta\mathrm{AIC}$) between models can be related to a probability of improved fit. The $\Delta\mathrm{AIC}$ image [Fig. 10(c)], created (voxel by voxel over time) from the $\mathrm{ImRes}(\mathbf{K})$ images is shown for the same data as used elsewhere in the paper. Fig. 4(b) shows the corresponding probability plot for different values of $\Delta\mathrm{AIC}$ from which a desired probability threshold can be determined. Fig. 10(d) shows the results of thresholding the $\Delta\mathrm{AIC}$ image at the $p < 0.05$ level.

Comparison of Fig. 10(b) and (d) shows good agreement. Both images indicate that statistically significant improvement in the fits of the dynamic data via direct reconstruction occurs primarily in the striatum and in the cortex, the areas of high and moderate binding of ^{18}F-fallypride to the dopamine D2 receptor. Looking at the particular slice shown in these images, the F statistic appears to produce a slightly more permissive threshold (recall both images are thresholded at $\mathrm{p} < 0.05$).

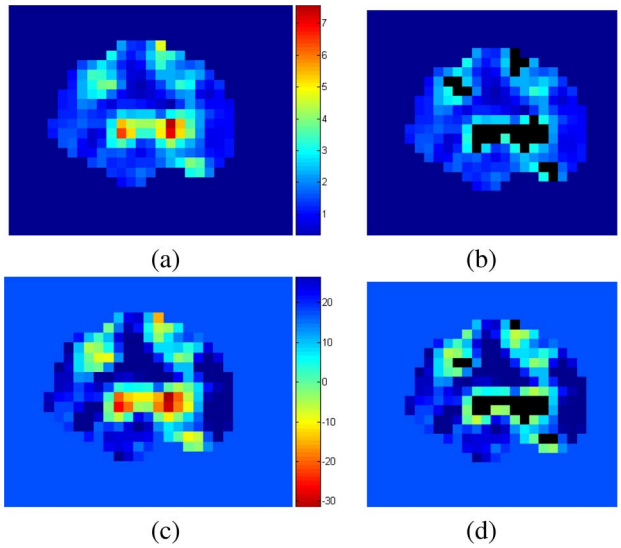


Fig. 10. (a) F statistic image computed from the residual ImRes images of one and two-tissue compartment models. (b) Voxels that exceed the 95% confidence level (with F statistics) for use of the two-tissue compartment model are shown in black. (c) ΔAIC image computed from the residual ImRes images of one and two-tissue compartment models. two-tissue compartment model is more suitable for voxels with negative ΔAIC values. (d) Voxels that exceed the 95% confidence level (with ΔAIC) for use of the two-tissue compartment model are shown in black.

D. Monte Carlo Simulations

The parametric images shown in the first column of Fig. 11 were defined as "true" in our simulations. Monte Carlo simulations were performed as described in Section II-F. The mean of each parametric image obtained from multiple noise realizations is shown in the second column of Fig. 11.

One iteration of direct reconstruction takes about 3 min on a computer with an Intel i7 processor and 8 GB memory. The number of iterations required for convergence depends on the regularization level. Typically it takes 5–15 iterations for convergence. This means 15-45 min per noise realization, and approximately 6–18 h for the Monte Carlo results with 25 noise realizations.[4] With the Intel i7 processor, eight reconstructions (one on each core) can be executed in parallel. Therefore, total computation time for 25 noise realization is reduced to 50–140 min.

1) Precision: The coefficient of variation images corresponding to each parameter are presented in Fig. 11. The coefficient of variation (CV) for each voxel is computed by dividing the standard deviation of voxel value (based on multiple reconstructions) by its own mean (averaged over reconstructions of multiple noise realizations). The coefficient of variation images provide a means of comparing the precision of parametric images even if the images or their variances are on different scales. The mean, absolute bias, and CV (computed from multiple reconstructions) are shown in Fig. 11. We can see from this figure that k_3 is greatest in the striatum, where there is the most D2/D3 binding. From the CV images (in the last column) in Fig. 11, it is also apparent that k_3 is estimated most precisely in that same area of the brain (a region of very low variation). By contrast, it might seem as if we have estimated the blood flow parameters (K_1, k_2) in the muscle and skull outside the brain but the CV images for these parameters indicate that the estimates for voxels outside the brain are much less reliable than for those inside the brain (see especially the $\mathrm{CV}(K_1)$ image inside and outside the brain).

2) Low and High Regularization Levels: In order to provide insight into the trade-off between accuracy and precision, the bias and CV images are generated for low and high regularization levels. In order to obtain different regularization levels, the manually adjusted regularization parameter [β in (3)] was scaled by a factor of 100 and 0.01 for high and low regularization levels, respectively.

Fig. 12 shows the mean parametric images of direct reconstruction with high and low regularization on the first and second rows, respectively. It is quite apparent that the parametric images on the first row (high regularization level) have little spatial variance, and most of the details and edges are over-smoothed. On the other hand, parametric images on the second row (low regularization level) are under-regularized and have high spatial variance.

In Fig. 13 the absolute bias and CV of the parametric images reconstructed with high and low regularization levels are shown. This figure demonstrates that the parametric images obtained using a low regularization level have lower bias but higher spatial variance compared to the parametric images obtained using a high regularization level.

IV. Discussion

Parametric images are of only limited usefulness if their bias and precision are not known. Without a measure of precision, we cannot say definitively that two images are different. Because direct reconstruction produces multidimensional images of each kinetic parameter, we claim that new evaluation and visualization techniques are needed for representing the bias and precision of estimated images in a convenient and digestible format. We note that, this is a proof-of-concept study in which we proposed and demonstrated the applicability of multiple novel approaches to visualize the bias and the variance in kinetic parameter estimates using a single dataset and only two different compartment models. Therefore, more rigorous tests should be performed with more datasets and with different (and possibly more than 2) compartment models.

A. Accuracy Versus Precision

We calculated the bias and CV images via Monte Carlo simulations. In Fig. 11, low variability is observed in the parameter estimates in the striatal and cortical regions that are known to have significant amounts of ^{18}F-fallypride binding to D2/D3 receptors and where we would expect reliable parameter estimates. On the other hand, high variation in the k_3 image correspond to the ventricles, which consist only of cerebrospinal fluid (no brain tissue), thus estimated k_3 values in those voxels are expected to be unreliable (effectively, meaningless). Figs. 11–13 demonstrate that the well-known bias-variance tradeoff (of spatial regularization levels) can be visualized using the Monte Carlo method.

[4]Overhead such as sinogram creation, noise generation, etc., can be ignored as they take considerably less time compared to the direct reconstruction of parametric images.

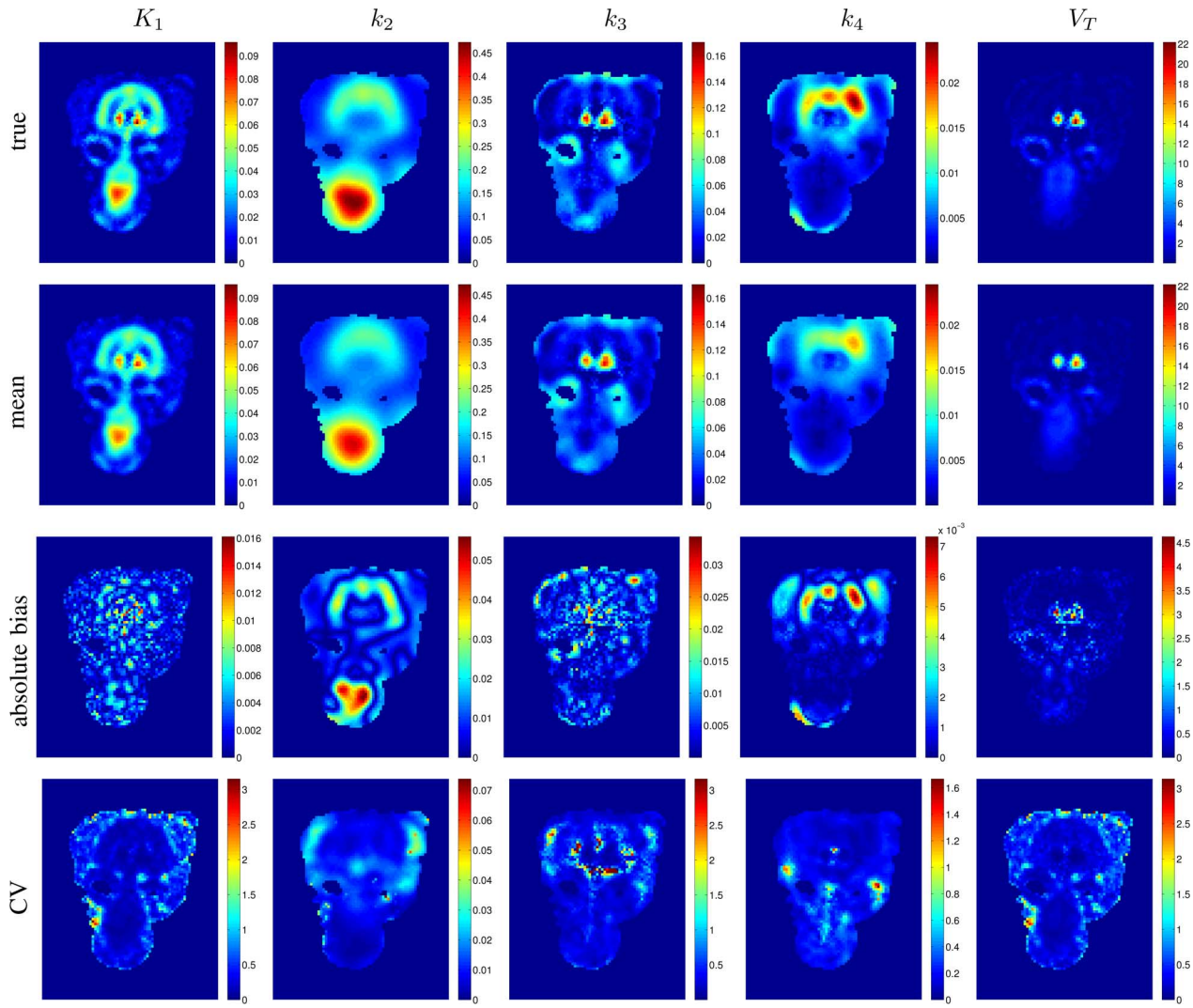


Fig. 11. Parametric images for K_1, k_2, k_3, k_4, and V_T. Images on column one are the true images, on column two are the recontructed mean images, on column three are the absolute bias of parametric images, column four are the CV of parametric images.

Besides Monte Carlo methods, analytical frameworks for the computation of bias and variance have been developed for static PET [45] and ROI analysis of dynamic PET [46]. A similar formulation was then adapted for dynamic PET reconstruction with different amounts of spatial regularization [34]. The analytical computation of bias and variance for kinetic parameter estimations was introduced by Wang *et al.* [34], and it was validated by others [47], [48]. It may be possible to develop a similar analytical framework for the computation of bias and variance in direct reconstruction of parametric images.

B. Goodness-of-Fit of a Model

We proposed a new goodness-of-fit metric for the direct reconstruction algorithm whose fits are carried out in the sinogram domain. To allow an imager to inspect local goodness-of-fit, a number-of-runs technique has been adapted from conventional 1-D parameter estimation resulting in a new type of image. In our adaptation, the error between the measured and the forward projected model sinograms are filter backprojected into error images, $\mathrm{ImRes}(\mathbf{K})$, in the spatial domain. We interpret these filter backprojected error images as the remaining activity that cannot be explained by the chosen kinetic model. If such a distribution is highly nonrandom, then the kinetic model that has been incorporated into the reconstruction must be deemed inadequate. The structure and low number-of-runs in the filter backrojected error images arise from nonrandom residuals in sinogram space. In our initial test with ^{18}F-fallypride, the greater nonrandomness in the $\mathrm{ImRes}(\mathbf{K})$ made from a one-tissue compartment model confirms that a one-tissue compartment model is inadequate to describe the uptake and retention of ^{18}F-fallypride in all regions of the monkey brain. Our technique is a novel way to view the local goodness-of-fit, to alert consumers of parametric images to possible limitations in the images, and/or to select the least complicated model necessary (model parsimony).

C. Model Selection

Model selection is merely an extension of goodness-of-fit. A deficient model cannot adequately describe the data (i.e., the estimated parameters will be biased) and the error images will reflect unwanted structure. On the other hand, an over-parame-

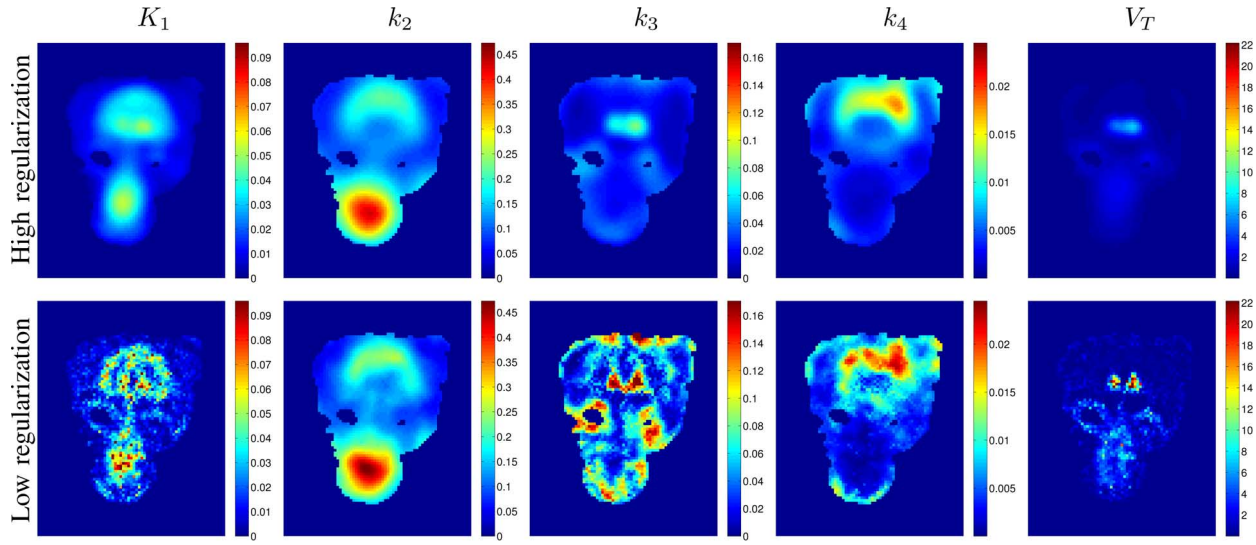


Fig. 12. Mean parametric images for K_1, k_2, k_3, k_4, and V_T in coronal perspective reconstructed high and low regularization levels on the first and second row, respectively.

terized model will produce random error images but the kinetic parameters will suffer high variance and poor identifiability.

1) Sinogram Domain Metrics: The residuals between the measured and fitted sinograms can be analyzed to compare the fits of different compartment models. This analysis, however, will not yield any information on a particular voxel.

2) Image Domain Metrics: Another way to examine model adequacy, at the ROI level, can be achieved by the histogram of runs as shown in Fig. 8 (in this case the ROI is the whole brain). If the number-of-runs for each voxel within a ROI is computed and binned into a histogram, model sufficiency for the ROI can be compared across models. In a direct comparison of models, the one with uniformly higher runs is the preferred one for the ROI or the entire image. It may even be possible to see multiple populations of voxels with different peaks in the runs histogram. This may be interpreted to indicate that a ROI is kinetically heterogeneous (i.e., time activity curves of the voxels within the ROI are dissimilar) and additionally one model cannot adequately describe the behavior of all the voxels in the ROI. Better fits (more runs) may be achieved at the expense of too great an increase in parameter variance, in which case a balance must be struck between describing all the data and estimating parameters precisely.

3) Quantitative Analysis on Image Domain: The probability that one model fits the data better than another can be quantified and used for model selection through the use of the F statistic or $\Delta\mathrm{AIC}$.[5] These indexes are ways of comparing the residual sum of squares between model fits while accounting for the decreased degrees-of-freedom imposed by additional model parameters. Fig. 10 showed—as expected—that high binding regions containing appreciable quantities of dopamine receptors were statistically better fit when the direct reconstruction algorithm incorporated the two-tissue compartment model as opposed to the one-tissue compartment. By extending the F statistic and $\Delta\mathrm{AIC}$ indexes to 3-D images, we have introduced a theoretically based way of selecting the proper model for incorporation into a direct reconstruction algorithm.

Some caveats in the quantitative analysis of error images some of which are listed below.

1) The voxels in a dataset may be modelled with multiple (more than 2) compartment models. In these cases, the idea presented in this work should be extended to handle statistical testing to find the best model amongst multiple compartment models.
2) Some correction may be required to the confidence level in Fig. 10(b) and (d) as multiple comparisons (for all voxels) are performed to generate these statistical significance images. With the increasing number of comparisons, there may be many voxels that appear to violate the null hypothesis (i.e., supporting the two-tissue compartment model) simply by chance.
3) RSS_l is comprised of two types of error in ImRes: Error from the model mismatch and error from measurements. The weight coefficients (w_n) in (7) should be used to normalize the effects of the variance of the measurement error to RSS_i. In this work, uniform weights (i.e., $w_n = 1 \forall n$) were used. Besides the uniform weights, weights obtained from inverse frame durations $(w_n = 1/(t_n - t_{n-1}))$ were also tried. The results were similar. We believe for the data used in this work the error from the model mismatch dominated the error from the measurements in RSS_l. Subtraction of RSS_l in F-statistic and calculating the division of RSS_l in $\Delta\mathrm{AIC}$ may partially cancel out the contribution from the measurement error. If TACs of the tested models are similar or if the data are extremely noisy, error from the model mismatch may become comparable to the measurement error term. The importance of weight coefficients in these cases should be investigated.

D. Limitations

There are certain limitations of the direct reconstruction algorithm and methods/data used in this study. In this study, it is assumed that all voxels can be described by the selected kinetic

[5]It is possible to generate p-value images using runs, $\Delta\mathrm{AIC}$, or F statistics, where these p-value images can be thresholded to produce masks that adhere to or violate a given model.

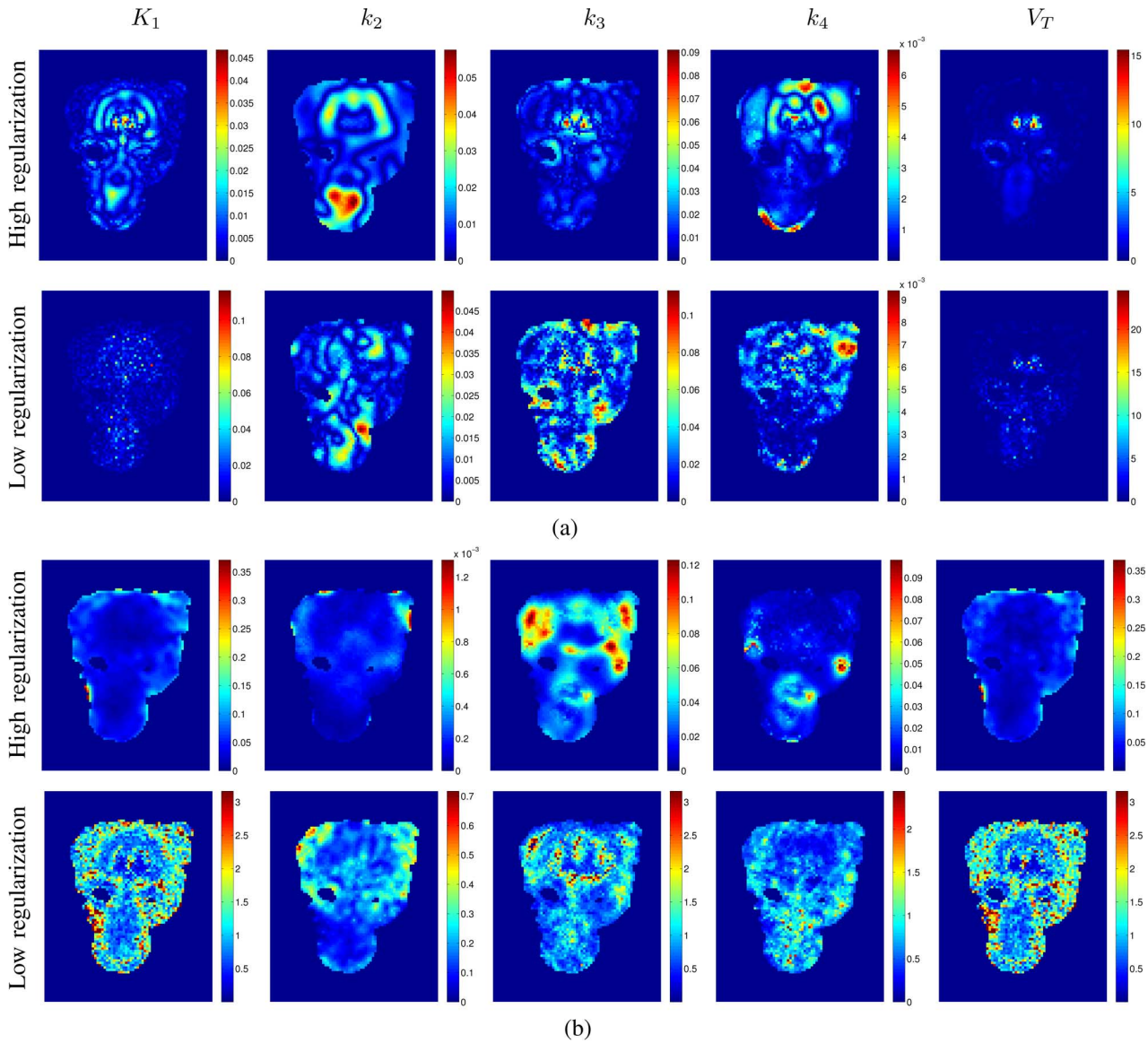


Fig. 13. (a) Absolute bias and (b) CV of parametric images for K_1, k_2, k_3, k_4, and V_T in coronal perspective reconstructed at high and low regularization levels.

model. However, there may exist some voxels which cannot be described with the selected or any other kinetic models. It may be possible to detect such voxels using a preliminary segmentation in the image space and to allow for different models (or no model at all) in different areas of the image.

A possible limitation of the direct reconstruction is its susceptibility to motion artifacts. In conventional image domain estimation methods, it is possible to register the emission images at each time point for motion compensation. However, it may be possible to record motion during acquisition and correct the data in an automated way using external measurement devices [49], [50]. In addition, there are direct reconstruction algorithms that apply to listmode data [29], [51]. The listmode data can be corrected event-by-event with the aid of high frequency motion detectors [52], [53].

V. Conclusion

In this paper, we have introduced three means of verifying the precision and validity of the parametric images created by direct reconstruction from sinograms [1]. The coefficient of variation images are an essential element for assessing the accuracy of any parametric image. As in any experiment, an acceptable coefficient of variation can only be determined in the context of the expected effect size, number of subjects in a study, etc., Monte Carlo simulations are probably too cumbersome for use as a routine means of calculating the variability in the images, but they will serve as a check of other methods for estimating variance. Besides the parametric images produced by direct reconstruction, we have introduced a multi-dimensional framework for exploring the reliability and adequacy of the kinetic models that are incorporated into the reconstruction. This required a combination of innovations that extended concepts in parameter estimation from one to three dimensions. To help interpret the results, we have introduced new ways of visualizing the metrics of goodness-of-fit by backprojecting the errors from sinogram space to image space, and creating new numbers of runs images. The best model produces the most valid (precise, unbiased) parameters. Selection of the best model can be di-

rected through the use of the F statistic and ΔAIC images introduced here. While our immediate focus is on the validity of the direct reconstruction algorithm that we introduced previously [1], we believe that the concepts and the framework we have introduced in this paper apply to the characterization and evaluation of other algorithms that produce parametric images from dynamic imaging data.

REFERENCES

[1] M. E. Kamasak, C. A. Bouman, E. D. Morris, and K. Sauer, "Direct reconstruction of kinetic parameter images from dynamic PET data," *IEEE Trans. Med. Imag.*, vol. 24, no. 5, pp. 636–650, May 2005.

[2] M. E. Kamasak, E. D. Morris, C. A. Bouman, and B. T. Christian, "Direct reconstruction of sinograms to parametric images of 18f-fallypride binding in monkey brain," presented at the Int. Symp. Funct. NeuroReceptor Mapp. Living Brain, Copenhagen, Denmark, Jul. 6–8, 2006.

[3] E. D. Morris, M. E. Kamasak, B. T. Christian, T. E. Cheng, and C. A. Bouman, "Visualizing all the fits: Evaluating the quality and precision of parametric images created from direct reconstruction of PET sinogram data," in *Proc. Int. Symp. Biomed. Imag.*, Arlington, VA, Apr. 6–9, 2006, pp. 291–294.

[4] C. S. Patlak and R. G. Blasberg, "Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data: generalizations," *J. Cereb. Blood Flow Metab.*, vol. 5, pp. 584–590, 1985.

[5] R. E. Carson, S. C. Huang, and M. V. Green, "Weighted integration method for local cerebral blood flow measurements with positron emission tomography," *J. Cereb. Blood Flow Metab.*, vol. 6, pp. 245–258, 1986.

[6] J. Logan, J. S. Fowler, N. D. Volkow, A. P. Wolf, D. J. S. S. L. Dewey, R. R. MacGregor, R. Hitzemann, S. J. G. B. Bendriem, and D. R. Christman, "Graphical analysis of reversible radioligand binding from time-activity measurements applied to [n-11c-methyl]-(-)-cocaine PET studies in human subjects," *J. Cereb. Blood Flow Metab.*, vol. 10, no. 5, pp. 740–747, 1990.

[7] D. Fang, Z. Wang, and S. C. Huang, "A study on statistically reliable and computationally efficient algorithms for generating local cerebral blood flow parametric images with positron emission tomography," *IEEE Trans. Med. Imag.*, vol. 12, no. 2, pp. 182–188, Jun. 1993.

[8] J. A. Thie, G. T. Smith, and K. F. Hubner, "Linear least squares compartmental-model-independent parameter identification in PET," *IEEE Trans. Med. Imag.*, vol. 16, no. 1, pp. 11–16, Feb. 1997.

[9] K. Chen, M. Lawson, E. Reiman, A. Cooper, D. Feng, S. C. Huang, D. Bandy, and D. Ho, "Generalized linear least squares method for fast generation of myocardial blood flow parametric images with n-13 ammonia PET," *IEEE Trans. Med. Imag.*, vol. 17, no. 2, pp. 236–243, Apr. 1998.

[10] R. E. Carson, "Tracer kinetic modeling in PET," in *Positron Emission Tomography, Basic Science and Clinical Practice*, P. E. Valk, D. L. Bailey, D. W. Townsend, and M. N. Maisey, Eds. London, U.K.: Springer, 2002.

[11] E. D. Morris, C. J. Endres, K. C. Schmidt, B. T. Christian, R. F. Muzic, Jr., and R. E. Fisher, "Kinetic modeling in PET," in *Emission Tomography: The Fundamentals of PET and SPECT*, M. Wernick and J. Aarsvold, Eds. San Diego, CA: Academic, 2004, ch. 23.

[12] M. A. Limber, M. N. Limber, A. Cellar, J. S. Barney, and J. M. Borwein, "Direct reconstruction of functional parameters for dynamic SPECT," *IEEE Trans. Nucl. Sci.*, vol. 42, no. 4, pp. 1249–1256, Aug. 1995.

[13] B. W. Reutter, G. T. Gullberg, and R. H. Huesman, "Accuracy and precision of compartmental model parameters obtained from directly estimated dynamic SPECT time-activity curves," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 1, pp. 170–176, Feb. 2004.

[14] R. H. Huesman, "A new fast algorithm for the evaluation of regions of interest and statistical uncertainty in computed tomography," *Phys. Med. Biol.*, vol. 29, no. 5, pp. 543–552, 1984.

[15] J. R. Votaw and H. H. Li, "Analysis of PET neurofunctional mapping studies," *J. Cereb. Blood Flow Metab.*, vol. 15, no. 3, pp. 492–504, 1995.

[16] G. L. Zeng, G. T. Gullberg, and R. H. Huesman, "Using linear time-invariant system theory to estimate kinetic parameters directly from projection measurements," *IEEE Trans. Nucl. Sci.*, vol. 42, no. 6, pp. 2339–2346, Dec. 1995.

[17] J. Matthews, D. Bailey, P. Price, and V. Cunningham, "The direct calculation of parametric images from dynamic PET data using maximum-likelihood iterative reconstruction," *Phys. Med. Biol.*, vol. 42, pp. 1155–1173, 1997.

[18] A. B. Milstein, K. J. Webb, and C. A. Bouman, "Estimation of kinetic model parameters in fluorescence optical diffusion tomography," *J. Opt. Soc. Am.*, vol. 22, no. 7, pp. 1357–1368, 2005.

[19] B. Alacam and B. Yazici, "Direct reconstruction of pharmacokinetic-rate images of optical fluorophores from NIR measurements," *IEEE Trans. Med. Imag.*, vol. 28, no. 9, pp. 1337–1353, Sep. 2009.

[20] G. Wang, L. Fu, and J. Qi, "Maximum a posteriori reconstruction of the Patlak parametric image from sinograms in dynamic PET," *Phys. Med. Biol.*, vol. 53, no. 3, pp. 593–604, 2008.

[21] G. Wang and J. Qi, "Direct reconstruction of PET receptor binding parametric images using a simplified reference tissue model," in *Proc. SPIE Med. Imag.*, 2009, p. 72580V.

[22] F. Pektas and I. S. Yetik, "Direct estimation of kinetic parameters from the sinogram with an unknown blood function," in *Proc. 3rd IEEE Int. Symp. Biomed. Imag.*, 2006, pp. 295–298.

[23] G. Wang and J. Qi, "Iterative nonlinear least squares algorithms for direct reconstruction of parametric images from dynamic PET," in *Proc. 5th IEEE Int. Symp. Biomed. Imag.*, 2008, pp. 1031–1034.

[24] C. Tsoumpas, F. E. Turkheimer, and K. Thielemans, "Study of direct and indirect parametric estimation methods of linear models in dynamic positron emission tomography," *Med. Phys.*, vol. 35, no. 4, pp. 1299–1309, 2008.

[25] G. Wang and J. Qi, "Generalized algorithms for direct reconstruction of parametric images from dynamic PET data," *IEEE Trans. Med. Imag.*, vol. 28, no. 11, pp. 1717–1726, Nov. 2008.

[26] J. Yan, B. Planeta-Wilson, J. Gallezot, and R. E. Carson, "Initial evaluation of direct 4d parametric reconstruction with human PET data," in *Proc. IEEE Nucl. Sci. Symp. Conf. Rec.*, 2009, pp. 2503–2506.

[27] G. Wang and J. Qi, "Acceleration of the direct reconstruction of linear parametric images using nested algorithms," *Phys. Med. Biol.*, vol. 55, no. 5, pp. 1505–1518, 2010.

[28] J. Matthews, G. I. Angelis, F. A. Kotasidis, P. J. Markiewicz, and A. J. Reader, "Direct reconstruction of parametric images using any spatiotemporal 4d image based model and maximum likelihood expectation maximisation," in *Proc. IEEE Nucl. Sci. Symp. Conf. Rec.*, 2010, pp. 2435–2441.

[29] C. Tsoumpas, F. E. Turkheimer, and K. Thielemans, "A survey of approaches for direct parametric image reconstruction in emission tomography," *Med. Phys.*, vol. 35, no. 9, pp. 3963–3971, 2008.

[30] A. Rahmim, J. Tang, and H. Zaidi, "Four-dimensional (4D) image reconstruction strategies in dynamic PET: Beyond conventional independent frame reconstruction," *Med. Phys.*, vol. 36, no. 8, 2009.

[31] B. T. Christian, T. K. Narayanan, B. Shi, and J. Mukherjee, "Quantitation of striatal and extrastriatal D-2 dopamine receptors using PET imaging of ^{18}F fallypride in nonhuman primates," *Synapse*, vol. 38, no. 1, pp. 71–79, Oct. 2000.

[32] M. E. Casey and E. J. Hoffman, "Quantitation in positron emission computed tomography. 7. A technique to reduce noise in accidental coincidence measurements and coincidence efficiency calibration," *J. Comput. Assist. Tomogr.*, vol. 10, pp. 845–850, 1986.

[33] M. Defrise, P. Kinahan, D. Townsend, C. Michel, M. Sibomana, and D. Newport, "Exact and approximate rebinning algorithms for 3-D PET data," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 145–158, Apr. 1997.

[34] G. Wang and J. Qi, "Analysis of penalized likelihood image reconstruction for dynamic PET quantification," *IEEE Trans. Med. Imag.*, vol. 28, no. 4, pp. 608–620, Apr. 2009.

[35] T.-C. Y. K.-H. Su and Y.-H. D. Fang, "A novel approach for direct reconstruction of parametric images for myocardial blood flow from PET imaging," *Med. Phys.*, vol. 40, no. 10, p. 102505, 2013.

[36] S. S. Saquib, C. A. Bouman, and K. Sauer, "ML parameter estimation for Markov random fields, with applications to Bayesian tomography," *IEEE Trans. Med. Imag.*, vol. 7, no. 7, pp. 1029–1044, Jul. 1998.

[37] C. Cobelli, D. Foster, and G. Toffolo, *Tracer Kinetics in Biomedical Research: From Data to Model*. New York: Kluwer, 2000.

[38] H. Motulsky and A. Christopoulos, *Fitting Models to Biological Data Using Linear and Nonlinear Regression: A Practical Guide to Curve Fitting*, 1st ed. New York: Oxford Univ. Press, 2004.

[39] S. M. Ross, *Introductory Statistics*, 2nd ed. New York: Elsevier Academic, 2005.

[40] R. F. Muzic, Jr. and B. T. Christian, "Evaluation of objective functions for estimation of kinetic parameters," *Med. Phys.*, vol. 33, no. 2, pp. 342–353, Feb. 2006.

[41] M. A. K. M. Yaqub, R. Boellaard, and A. A. Lammertsma, "Optimization algorithms and weighting factors for analysis of dynamic PET studies," *Phys. Med. Biol.*, vol. 51, pp. 4217–4232, 2006.

[42] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. 19, no. 6, pp. 716–723, Dec. 1974.

[43] K. P. Burnham and D. Anderson, *Model Selection and Multi-Model Inference*, 2nd ed. New York: Springer, 2002.

[44] J. Qi, R. M. Leahy, C. Hsu, T. H. Farquar, and S. R. Cherry, "Fully 3-D Bayesian image reconstruction for the ECAT EXACT HR+," *IEEE Trans. Nucl. Sci.*, vol. 95, pp. 1096–1103, 1998.

[45] J. A. Fessler, "Mean and variance of implicitly defined bias estimators (such as penalized maximum likelihood): Applications to tomography," *IEEE Trans. Image Process.*, vol. 5, no. 3, pp. 493–506, Mar. 1996.

[46] J. Qi and R. Huessman, "Theoretical study of penalized-likelihood image reconstruction for region of interest quantification," *IEEE Trans. Med. Imag.*, vol. 25, no. 5, pp. 640–648, May 2006.

[47] F. Pektas and I. S. Yetik, "Covariance expressions for implicit estimators of kinetic parameters with dynamic PET," in *Proc. IEEE Int. Conf. Electro/Inf. Technol.*, 2008, pp. 316–320.

[48] M. E. Kamasak, "Computation of variance in compartment model parameter estimates from dynamic PET data," *Med. Phys.*, vol. 39, no. 5, pp. 2638–2648, May 2012.

[49] P. M. Bloomfield, T. J. Spinks, J. Reed, L. Schnorr, A. M. Westrip, L. Livieratos, R. Fulton, and T. Jones, "The design and implementation of a motion correction scheme for neurological PET," *Phys. Med. Biol.*, vol. 48, no. 8, pp. 959–978, 2003.

[50] A. J. Montgomery, K. Thielemans, M. A. Mehta, F. Turkheimer, S. Mustafovic, and P. M. Grasby, "Correction of head movement on PET studies: Comparison of methods," *J. Nucl. Med.*, vol. 47, no. 12, pp. 1936–1944, 2006.

[51] J. Yan, B. Planeta-Wilson, and R. E. Carson, "Direct 4-D PET list mode parametric reconstruction with a novel EM algorithm," *IEEE Trans. Med. Imag.*, vol. 31, no. 12, pp. 2213–2223, Dec. 2012.

[52] R. E. Carson, W. C. Barker, J. S. Liow, and C. A. Johnson, "Design of a motion compensation OSEM list-mode algorithm for resolution-recovery reconstruction for the HRRT," in *IEEE Nucl. Sci. Symp. Conf. Rec.*, 2003, vol. 5, pp. 3281–3285.

[53] S.-K. Woo, H. Watabe, Y. Choi, K. M. Kim, C. C. Park, R. R. Fulton, P. M. Bloomfield, and H. Iida, "Development of event-based motion correction technique for PET study using list-mode acquisition and optical motion tracking system," in *Proc. SPIE*, M. Sonka and J. M. Fitzpatrick, Eds., 2003, vol. 5032, pp. 1300–1307.