Patient-Adaptive Population-Based Modeling of Arterial Input Functions

Zhaoyan Xiu, Mark Muzi, Jian Huang, and Eric Wolsztynski

Abstract—**Kinetic modeling of dynamic PET data requires knowledge of tracer concentration in blood plasma, described by the arterial input function (AIF). Arterial blood sampling is the gold standard for AIF measurement, but is invasive and labour intensive. A number of methods have been proposed to accurately estimate the AIF directly from blood sampling and/or imaging data. Here we consider fitting a patient-adaptive mixture of historical population time course profiles to estimate individual AIFs. Travel time of a tracer atom from the injection site to the right ventricle of the heart is modeled as a realization from a Gamma distribution, and the time this atom spends in circulation before being sampled is represented by a subject-specific linear mixture of population profiles. These functions are estimated from independent population data. Individual AIFs are obtained by projection onto this basis of population profile components. The model incorporates knowledge of injection duration into the fit, allowing for varying injection protocols. Analyses of arterial sampling data from ^{18}F-FDG, ^{15}O-H$_2$O and ^{18}F-FLT clinical studies show that the proposed model can outperform reference techniques. The statistically significant gain achieved by using population data to train the basis components, instead of fitting these from the single individual sampling data, is measured on the FDG cohort. Kinetic analyses of simulated data demonstrate the reliability and potential benefit of this approach in estimating physiological parameters. These results are further supported by numerical simulations that demonstrate convergence and stability of the proposed technique under varying training population sizes and noise levels.**

Index Terms—**Arterial input function estimation, kinetic modeling, medical imaging, parametric imaging, PET imaging.**

Zhaoyan Xiu and Jian Huang are with the School of Mathematical Sciences, University College Cork, Cork, T12 CY82 Ireland (e-mail: zhaoyan_xiu@126.com; j.huang@ucc.ie).

Mark Muzi is with the School of Medicine, University of Washington, Seattle, WA 98195 USA (e-mail: muzi@uw.edu).

Eric Wolsztynski is with the School of Mathematical Sciences, University College Cork, Cork, T12 CY82 Ireland, and also with the Insight Centre for Data Analytics, Cork, T12 XF62 Ireland (e-mail: eric.w@ucc.ie).

I. Introduction

A. Motivation

POSITRON Emission Tomography (PET) is widely used in the staging and evaluation of many diseases, and particularly in cancer. From the time of radiotracer injection, a PET scan provides 3D time course imaging of uptake of this radioactive agent in the body. Given the time course of radioactivity concentration in arterial plasma, called the Arterial Input Function (AIF), kinetic modeling of the dynamic imaging data can be used to measure physiologic parameters, such as blood flow and glucose metabolic rate [1], [2], [3], [4]. There exist techniques for estimating physiological parameters that do not require the AIF, such as reference region approaches for receptor-radioligand studies [5]. To be valid, these techniques require the existence of a measurable, reliable reference region that is devoid of the target receptor, and for comparisons to be appropriate, the reference region properties must be independent of treatment effects and groups [5]. A region meeting these criteria has not yet been identified, or might not exist, for many radioligands. AIF estimation therefore constitutes a realistic, intermediate objective for many applications of dynamic data analysis for the determination of metabolic processes.

Determination of the AIF is a challenging task. The gold standard and most accurate approach is to use direct arterial blood sampling, but this requires multiple arterial blood samples over the whole scanning period, and presents several drawbacks including patient discomfort, additional costs, and potential risk due to its invasive nature. A number of less invasive methods have been considered that rely on one of a few strategies for arterial signal characterisation, as follows.

B. Modeling of Arterial Input Data

Image-derived input functions (IDIF) [6], [7], [8], [9], [10], [11], [12], [13] are obtained by fitting an AIF model to arterial blood pool tracer activity extracted from PET image data, rather than obtained from arterial sampling. Such data are usually extracted from a region of interest (ROI) within a dynamic PET image that captures either a large blood vessel or the left cardiac ventricle, and therefore tend to be noisy. Usually, the average or the highest signals in the ROI are extracted to obtain a time course to characterize the concentration of tracer in arterial blood. IDIF approaches may also require at least one physical blood sample to scale the image-derived AIF

curve, and may include tracer uptake information both before and after tracer metabolism. Alternatively, PET signals from multiple regions (not necessarily arterial blood pools) may be leveraged to recover AIF and thereby kinetic parameters, via simultaneous estimation of the input function [14].

Population-based input functions (PBIF) [15], [16], [17], [18], [19], [20] are obtained by estimating a patient's AIF by scaling an overall population arterial input pattern (obtained from historical information on the patient's population or by averaging fits from the overall cohort) based on individual subject information (depending on availability, historical cohorts have 10-30 blood samples through 60-90 minutes of data, and have a combination of patient weight, body surface area, lean body mass or cerebellar FDG activity) [21], [22]. The choice of time points at which this information is extracted is a likely contributor to variability in estimation of the signal scale. With this approach, the individual is assumed to have the same tracer injection protocol and physiological characteristics as the population, which is not always realistic.

Various mathematical constructions such as Feng's model [23], tri-exponential models [24] and convoluted models [25] have been developed to provide a continuous and noise-free description of the AIF obtained from either clinically sampled (as in typical PBIF strategies) or image-derived (as in typical IDIF strategies) signals, as shown with step 3A of the flowchart in Fig. 1. The tri-exponential construction and its convolved alternative rely on a linear mixture of exponential density functions fit to sample arterial data. The tri-exponential model (TE) fits the AIF shape in two parts: a first linearly increasing piece from the time the AIF starts rising at to its peak, and a second piece modeling the AIF from its peak to the end of study using a linear combination of the three exponential curve. In comparison to Feng's model, the TE model provides more flexibility to describe the AIF. However, neither model accounts for the length of the tracer injection and could lead to poor fit to the initial part of the AIF curve. An adaptation of the TE model was proposed [25] that includes the information of injection duration by convolving the injection profile and linear combination of three exponential curves. This convolved tri-exponential (CTE) model addresses some of the issues of the original TE model, but relies on a boxcar function of fixed duration to describe the injection profile, which used on its own may not represent the pattern of tracer arrival into plasma realistically. This particular limitation has been addressed in some works [26], [27], [28], [29], by using a convolution of a double Butterworth function with the TE function [28], or of a square wave with a Gamma function [29]. Final individual AIF estimates are obtained by scaling the estimated AIF pattern to the individual physiological level using either image-derived blood level estimates from a reference region (after partial volume correction) or blood sampling data. (Some alternative scaling methods exist that use routine patient information instead; this is discussed further below.)

An alternative, compartment model based technique was developed by our group to estimate the AIF from a whole-body tracer circulation model (BCM) [29], [30]. In this representation, the fate of each tracer atom is represented by an eight-state Markov chain process discrete by heart beats, where each state represents one important location in the blood circulation system. To use the BCM, theoretically, time course of the tracer concentration from one of the eight states is required. Since the left ventricle is one of those states represented in the stochastic process, the AIF signal can be approximated by the tracer concentration in the left ventricle observed from PET imaging over time. Otherwise clinically sampled arterial data can be used for evaluation of one of these Markov states. Although high dimensionality of the parameters involved in the model constitutes a computational challenge, this physiologically meaningful representation can yield an accurate AIF estimation.

C. Contribution

Here we consider a modified PBIF-type approach by adjusting a population AIF profile to individual characteristics extracted from either arterial sampling or image-derived data, to achieve reliable patient-adaptive AIF estimation throughout the whole time course. In a sense the proposed population-based projection model (PBPM) combines population profiling (as in a PBIF approach) with individual arterial input data modeling (as in an IDIF approach). As such, it differs from both these strategies, in a number of ways. Unlike IDIF strategies, the PBPM leverages population information, and does not depend on the availability of imaging data (although it can also be applied using such data). Unlike PBIF strategies, it does not average population components into a single individual template pattern, but instead produces a different pattern for each individual, by fitting a linear mixture of population profile components to subject-specific arterial sampling data. Fig. 1 describes the overall process and illustrates how it differs from typical PBIF and IDIF processes. To produce individual fits, the PBPM mixture components are fitted to population data, and a combination of these fits is adapted to each subject, unlike with TE-type methods. This approach could also be adapted for use in step 3B in the flowchart, but no such report was found in the literature. The proposed PBPM allows the use of functions other than exponentials, should a particular tracer require it, and can operate with a different number of components depending on the PET tracer used. This number can be automatically selected; two or three such population components are typically sufficient for individual AIF estimation. The proposed PBPM also provides more flexibility with respect to temporal characteristics of the data. It allows for injection duration to be included in the modeling, which makes it applicable to PET studies with different injection durations. Unlike TE-type constructions, it also does not require specification of two key time points (signal start and peak time), and can adapt to different peak times thanks to its parametrization. A model fitting procedure for the PBPM is also proposed hereafter. A penalized positive constrained least squares method is developed and implemented in R [31] for estimation of individual mixing weights. Statistical regularization is built in to adapt relative weight placed on the individual and population information according to the noise level in a given dataset, to facilitate patient-adaptive fitting of cohort characteristics.
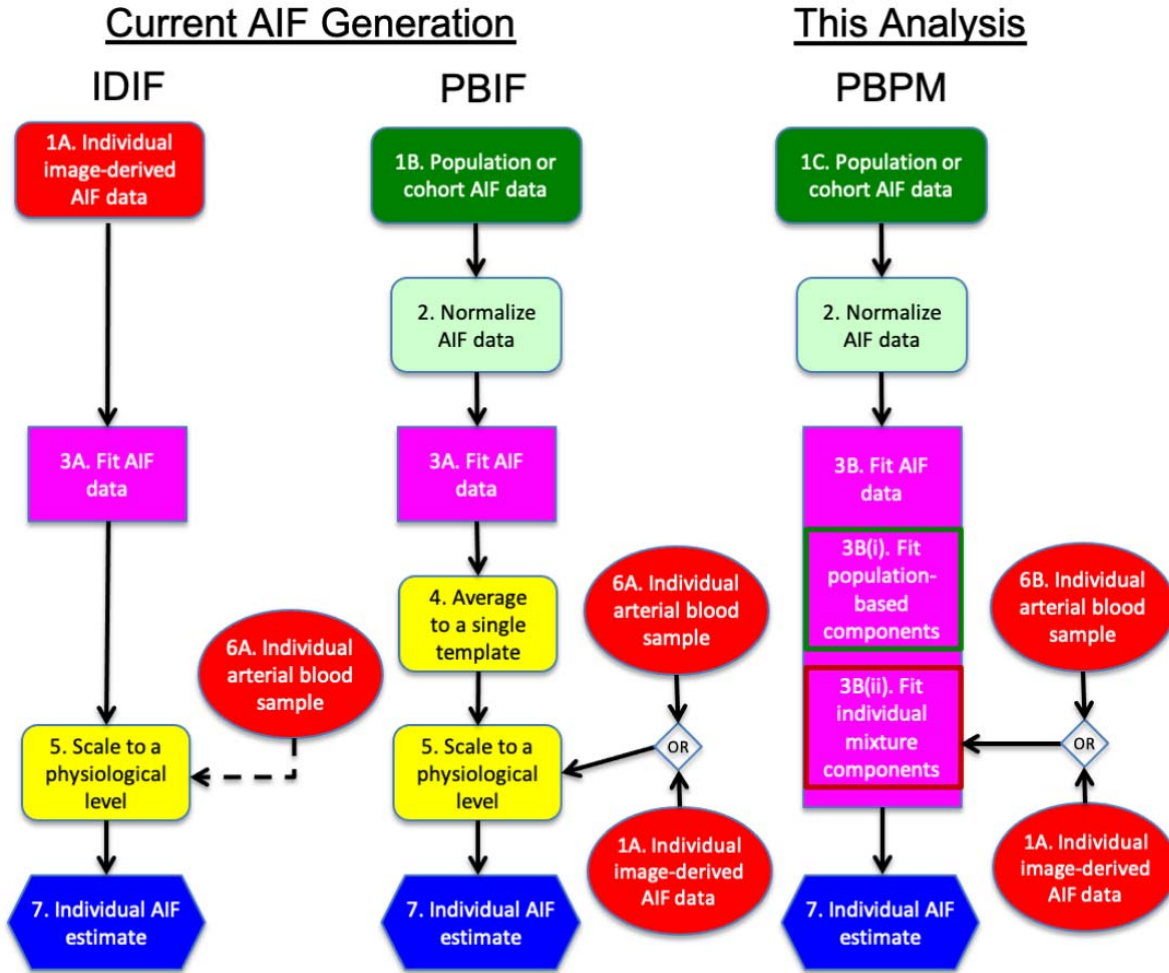
Fig. 1.   Flowchart of typical PBIF and IDIF strategies, along with the proposed PBPM methodology. Red and green blocks indicate use of individual subject and population data respectively. Step 3 of each method is an AIF model fitting step. Blood sampling for step 6A is not used in some IDIF approaches (hence the dashed line); when used, fewer points are typically required compared to blood sampling used in step 6B. Using individual image-derived data (1A) at step 3B(ii) of the PBPM approach (instead of arterial sampling data) would require an extra step in order to scale the output to a physiological level, as in step 5 of an IDIF method.

This paper focuses on describing aspects related to the PBPM (step 3B in the flowchart), and benchmarking this model against reference methods typically used in step 3A of PBIF and IDIF frameworks. Section II develops the proposed model, and provides details on the methodological developments for fitting the model to clinically sampled arterial data, including regularization and cross-validation for selection of the number of components. Section III illustrates model selection and evaluation, convergence of the procedure and other aspects related to its construction, using both simulated and clinical data comprising of three separate arterial sampling datasets of respectively FDG-, H_2O-, and FLT-PET imaging studies. Results from these numerical analyses include validation of an open source implementation of the software in R [31] that we made available [32]. Section IV presents a comparative analysis against another three methods of reference, namely the TE, CTE and BCM models. Results on clinical data demonstrate the gain derived by training the proposed AIF representation on population data, as well as its impact on kinetic analysis is evaluated via resampling of uptake templates in one- and two-tissue compartment models.

II. METHODS

This section presents the modeling strategy used for patient-specific AIF representation based on PET tracer concentration profiling, including aspects of model fitting and calibration.

A. AIF Model Specifications

Let the travel time T of a specific atom at a specific arterial sampling site be modeled as the sum of the time (T_{ir}) for the atom to initially progress from the injection site to the right ventricle (RV) of the heart, and the time (T_c) it spends in circulation, before being sampled at the site:

$$T = T_{ir} + T_c$$

T_{ir} is modeled as a realization from a Gamma distribution (G) with shape and rate parameters α and β, respectively. The circulation time is modeled as a shifted excess circulation time (ξ) with the value of the shift (Δ) being specific to the location of sampled blood-site within the circulatory system

$$T_c = \Delta + \xi$$

If the probability density (f) of ξ is known, it is possible to combine this distribution with the injected dose profile in the right ventricle to evaluate the tracer concentration at any particular time. Under the conventional assumptions of time invariance and linearity of tracer kinetics, a bolus injection at time $t = 0$ will produce a temporal concentration of tracer atoms in the RV proportional to the distribution G. A more complex temporal injection profile, *e.g.* specified by an indicator function $I(\cdot)$, would yield an RV concentration profile C_{RV} characterized by an accumulation over the whole injection time, such that

$$C_{RV}(t) = G \otimes I(t) = \int_0^t G(t-s)I(s)ds$$

The evolution of this profile to the concentration profile at the blood site of interest, C_P, is governed by the circulation time distribution. This is given by a shifted convolution with the excess circulation travel time density

$$C_P(t) = f \otimes C_{RV}(\cdot - \Delta)(t) = \int_0^t f(t-s)G \otimes I(s-\Delta)ds$$

We approximate the distribution f of excess circulation travel time ξ by a linear mixture of J components to obtain an estimate

$$\hat{f}(\xi) = \pi_1 e_1(\xi) + \pi_2 e_2(\xi) + \cdots + \pi_J e_J(\xi)$$

where the $\pi_1, \pi_2, \ldots, \pi_J$ are mixing fractions of the component densities $e_1, e_2, \ldots, e_J$, with constraint $\pi_j > 0, \forall j = 1, \ldots, J$. With this representation, an estimate of the concentration profile becomes

$$\hat{C}_P^J(t) = \sum_{j=1}^J \pi_j \int_0^t e_j(t-s)G \otimes I(s-\Delta)ds \qquad (1)$$

In this expression, the convolution ($\otimes$) of the square wave $I()$ with the Gamma distribution G describes a population injection profile, adjusted to individual subjects by a patient-specific delay Δ. In this approach the population component densities $e_1, e_2, \ldots, e_J$ are assumed to be standard exponentials with characteristic rates $\phi_1 > \phi_2 > \cdots > \phi_J > 0$. While many other possibilities could be considered (*e.g.* Gamma, Weibull, etc.), a mixture of exponential distributions makes for a simple and flexible model. We presume the mixing components are representative of the entire circulatory system, and functional components $e_1, e_2, \ldots, e_J$ and G are held fixed across subjects in the population of study.

Ambiguity may arise in the estimation of the linear coefficients $\hat{\pi}_1$ and $\hat{\pi}_2$, in cases where the corresponding distribution ranges overlap. To remove this ambiguity, a constraint $\pi_1 > \pi_2$ can be imposed when estimating these linear parameters. This constraint ensures that at early time points, the first component contributes mostly to the peak of AIF representation. A similar constraint may be applied to further mixing coefficients, although in our experience on the data of Section IV, it was not necessary as there was no practical ambiguity between π_2 and π_3 in a three-component mixture applied to dynamic FDG-PET data.

B. Model Fitting

Suppose blood time course data from a training population of n subjects

$$\{(t_{ik}, z_{ik}, w_{ik}), i = 1, 2, \ldots, m_k, k = 1, 2, \ldots, n\}$$

are available, where z_{ik} is the concentration of the ith measurement of the kth subject, measured at time t_{ik} with measurement reliability w_{ik} and m_k is the number of time points used in clinical sampling for the kth subject. Each time course z_{ik} is modeled for $i = 1, 2, \ldots, m_k, k = 1, 2, \ldots, n$ as

$$z_{ik} = \hat{C}_P(t_{ik}|\Delta_k, \pi_k, \theta) + w_{ik}^{-1/2}\varepsilon_{ik},$$

where $\hat{C}_P(t|\Delta_k, \pi_k, \theta)$ is defined in the previous subsection with the subject-dependent linear parameter vector $\pi_k = c(\pi_{1k}, \pi_{2k}, \ldots, \pi_{Jk})$, and subject-independent nonlinear parameter vector $\theta = c(\alpha, \beta, \phi_1, \ldots, \phi_J)$. The ε_{ik}'s are assumed to be i.i.d realizations of Gaussian random variables modeling additive noise. We propose to estimate θ by minimizing the objective function defined by

$$l(\theta) = \sum_{k=1}^n \min_{\Delta_k, \pi_k} \sum_{i=1}^{m_k} w_{ik}(z_{ik} - \hat{C}_P(t_{ik}|\Delta_k, \pi_k, \theta))^2$$

An algorithm for this optimization is implemented in R [31]. Minimization of $l(\theta)$ requires two-step estimation of π_k and Δ_k, which is described in the following two subsections, respectively.

1) Estimation of π_k Given Δ_k and θ: The linear mixture is fitted by minimizing, for $k = 1, \ldots, n$, criterion

$$l_1(\pi_k|\Delta_k, \theta) = \sum_{i=1}^{m_k} w_{ik}(z_{ik} - \hat{C}_P(t_{ik}|\Delta_k, \pi_k, \theta))^2$$

$$= \sum_{i=1}^{m_k} w_{ik}\left(z_{ik} - \sum_{j=1}^J \pi_{jk}A_{ijk}\right)^2$$

with respect to π_k given Δ_k and θ, where the jth component of the linear mixture for the kth patient, hereafter denoted by A_{ijk} for convenience, can be calculated as

$$A_{ijk} = A_j(t_{ik}|\Delta_k, \theta) = \int_0^{t_{ik}} e_j(t_{ik}-s)G \otimes I(s-\Delta_k)ds$$

for $i = 1, \ldots, m_k, j = 1, \ldots, J$, where the mixing parameters $\{\pi_{jk}\}_{j=1}^J$ are constrained positive. An R implementation of the Goldfarb-Idnami method [33] is employed to solve this positivity-constrained least-squares problem. We denote the resulting minimizer of $l_1(\pi_k|\Delta_k, \theta)$ as $\hat{\pi}_k = \hat{\pi}(\Delta_k, \theta)$.

2) Estimation of Delay Δ_k Given θ: Inserting $\hat{\pi}_k(\Delta_k, \theta)$ in the objective function l_1, we have the reduced objective function

$$l_2(\Delta_k|\theta) = l_1(\hat{\pi}_k(\Delta_k, \theta)|\theta)$$

Given population characteristics θ, a grid search may be carried out to minimize $l_2(\Delta_k|\theta)$ with respect to Δ_k. The measured data may not be coincident with the timing of the injection, and a parameter Δ_k, positive or negative, is introduced to account for this. In this study, 21 evenly distributed grid points in the range of $[-60\text{ s}, +60\text{ s}]$ were used. The resulting minimizer of l_2 is denoted as $\hat{\Delta}_k$.

3) Estimation of Population Parameter θ: Finally, the population-specific parameters $\theta = (\alpha, \beta, \phi_1, \ldots, \phi_J)$ are estimated by minimizing

$$l(\theta) = \sum_{k=1}^{n} \sum_{i=1}^{m_k} w_{ik}(z_{ik} - \hat{C}_P(t_{ik}|\hat{\Delta}_k, \hat{\pi}_k, \theta))^2$$

Initial values for $\theta^0 = (\alpha^0, \beta^0, \phi_1^0, \ldots, \phi_J^0)$ may be selected based on experience from previous experiments and some physiologic reasoning about the average travel time of each type of tracer in the blood circulation system. Lower and upper bounds $\theta^L = (\alpha^L, \beta^L, \phi_1^L, \ldots, \phi_J^L)$ and $\theta^H = (\alpha^H, \beta^H, \phi_1^H, \ldots, \phi_J^H)$, respectively, can be set based on experience from the population of patients receiving the same tracer injection, and physiologic reasoning. Since the scales of the parameters differ considerably, and also to avoid the complexity of constrained minimization, the following parameter transformation may be performed. Let $p = \frac{\theta - \theta^L}{\theta^H - \theta^L}$ be transformed via a logit transform into

$$\varsigma = \log \frac{p}{1-p}$$

Inversely,

$$\theta = \theta^L + (\theta^H - \theta^L) \frac{e^\varsigma}{1 + e^\varsigma}$$

Thus, the constrained minimization with respect to θ can be performed as unconstrained minimization of $l(\varsigma)$ with respect to ς. A Gauss-Newton type method was implemented in R to optimize $l(\varsigma)$.

4) Fitting Individual Curves: The population parameter estimates $\hat{\theta} = \arg\min_\theta l(\theta)$ defined in Step (II-B.3) can be used instead of being re-evaluated when fitting time activity curves for "new" patients (i.e. patients outside of the training population of n studies used to obtain $\hat{\theta}$), thus reducing the number of parameters required to obtain an individual fit. In such cases, simultaneous estimation of π_k and Δ_k can be achieved by grid search over G values $\{\hat{\Delta}_k^1, \ldots, \hat{\Delta}_k^G\}$, corresponding estimates $\{\hat{\pi}_k^1, \ldots, \hat{\pi}_k^G\}$ being obtained instantaneously via linear regression.

5) Model Regularization: In cases of low signal-to-noise ratio,[1] regularization of the patient-specific mixing weights π can be applied to control variability of these estimates. Once the number and parameters of the population components have been determined, a penalized nonlinear least squares approach using the historical linear parameters as a Bayesian penalty for individual fit [34] is used for this approach. Regularization on $\pi = \pi_k$ parameters for the kth patient is achieved using the following individual objective function:

$$\begin{aligned} f_\lambda(\pi|\Delta, \theta) &= W(\mathbf{z}^T - \pi^T A_\Delta^T)^2 \\ &\quad + \lambda(\pi^T - \pi_0^T)\Omega^{-1}(\pi^T - \pi_0^T)^T \\ &= -[(A_\Delta^T(\sqrt{W}\mathbf{z}))^T + \lambda\pi_0^T\Omega^{-1}]\pi \\ &\quad + \pi^T(A_\Delta^T W A_\Delta + \lambda\Omega^{-1})\pi \end{aligned} \quad (2)$$

where $\mathbf{z} = \{z_{ik}\}_{i=1}^{m_k}$, $A_\Delta = [A_{ijk}]_{i,j}$ is a matrix of dimension $m_k \times J$, and $W = W_k$ are defined in Section II-B.1. All components in the vectorized mixing weights π are constrained

[1]e.g. if using small and/or noisy PET imaging data for AIF estimation

to be nonnegative, and π_0 and Ω respectively denote the mean and covariance matrix of linear parameters obtained from population-based estimation. regularization parameter λ controls the trade-off between individual and population data contributions, larger values of λ constraining individual estimates to be closer to the population profile π_0, which suits higher-noise scenarios. The optimization is a quadratic problem for which we used the Goldfarb-Idnami method [33].

Parameter λ is set by minimizing the generalized cross-validation (GCV) score which is defined as the mean squared error adjusted by the effective degree of freedom:

$$GCV(\lambda) = \frac{\frac{1}{m}\sum_{i=1}^{m}(z_{ik} - \hat{z}_{ik}(\lambda))^2}{\left(\frac{1}{m}\text{Trace}(I - A_\Delta(A_\Delta^T A_\Delta + m\lambda I)^{-1}A_\Delta^T)\right)^2},$$

where $\hat{z}_{ik}(\lambda)$ is $\hat{C}_P(t_{ik}|\Delta_k, \pi_k, \theta)$ based on minimization of the objective function defined in Eqn. (2), for a given λ.

C. Model Selection

Similarly to compartmental analysis of dynamic PET data, the number of components needed in the linear mixture defined in Eqn. (1) may depend on the tracer used in the studies. More components may be included if the tracer molecules can participate in metabolic activity. For example, in water studies, tracers do not participate in any metabolic activity, and a one-tissue compartmental model is normally used for water dynamic studies. Models using more than one compartment are applied to studies involving other tracers for more complicated circulatory or metabolic systems [4].

Here a leave-one-out approach is developed for selection of the number of components to be used in the proposed linear mixture defined in Eqn. (1) so as to minimize model complexity defined as

$$Err(J) = \sum_{k=1}^{n} \min_{\Delta_k, \pi_k} \sum_{i=1}^{m_k} \frac{w_{ik}}{m_k}(z_{ik} - C_P^J(t_{ik}, \Delta_k, \pi_k, \hat{\theta}_k))^2, \quad (3)$$

where $J = 2, 3 \ldots$ is the number of components, $C_P^J(\cdot)$ defines the predicted AIF based on PBPM with J components given linear and non-linear parameters, and

$$\hat{\theta}_k = \arg\min_{\theta_k} \sum_{j \neq k}^{n} \min_{\Delta_k, \pi_k} \sum_{i=1}^{m_k} \frac{w_{i,k}}{m_k}\left(z_{i,k} - C_P^J(t_{i,k}, \Delta_k, \pi_k, \theta_k)\right)^2$$

A cross-validated error ($\hat{Err}$) may not necessarily be the minimum model complexity Err for the chosen J, but should be reasonably small for all values $Err(J)$ in the range of J.

D. Scale Factor Estimation

An overarching objective of AIF estimation is to design methods that can rely on image-derived input data instead of invasive sampling. In this context, scaling of PBPM estimates of the AIF obtained from image-derived arterial signals is required to obtain final parametric imaging estimates. This evaluation typically requires at least one blood sample point (from either arterial or venous blood) [15], [35], [36], or an adequately calibrated image-derived measurement from reference tissue (e.g. cerebral tissue) [22], [37], which is not

always feasible. Routine individual subject information may alternatively be used to this end. In [37] the authors use FDG uptake measured in cerebellum ROI from a single scan acquisition to normalize a PBIF estimate. In [38] the authors show that body surface area can be used as an alternative to blood sampling in PBIF normalization in brain ^{11}C-TMSX PET studies. In [39], the authors use injection dose and weight information to scale PBIF estimates in ^{11}C-DPA-713 PET studies. Here we propose a regression model for the scale factor that does not require arterial blood sampling or a large blood-pool in the field of view. This model uses injection dose X_{ID}, blood volume X_{BV} (estimated from non-invasive patient information on body weight, height and total blood volume [40]) and tail height of the AIF curve ($\hat{\pi}_3$, the linear coefficient for the third component in the PBPM) as independent variables:

$$log(S) = \beta_0 + \beta_1 X_{ID} + \beta_2 X_{BV} + \beta_3 \hat{\pi}_3 + \text{noise} \qquad (4)$$

III. EVALUATION

This section introduces the three clinical datasets used for evaluation of the proposed PBPM model. Model selection, evaluation of the impact of training population size on model fitting, and scaling of the final AIF estimates are covered in these analyses. Complementary simulations based on the sampling data also illustrate convergence of the procedure at varying noise levels.

A. Data

The proposed model was applied to three separate datasets, comprising blood sampling data obtained during the acquisition of respectively 105 FDG-PET studies, 39 H$_2$O-PET studies, and 32 FLT-PET studies, all acquired at the University of Washington [9], [29], [41], [42], [43], [44]. The data were from both healthy subjects and patients with cancer conditions. All FLT patients had cancer (brain, breast, lung, sarcoma). Of the 105 FDG blood curves used in the analysis, 12 were from normal human subjects with the rest from patients with cancer, including brain, lung, liver and colon cancer patients. Eighteen of the 39 water curves came from normal subjects. Durations of FDG injections were either 1 or 2 minutes (58 and 33 cases respectively), bolus injection for H$_2$O was assumed to be 5 seconds, and duration of FLT injection was 1 minute. The datasets are summarised in Table I. The data included sample AIFs of the subjects along with gender, weight, height and injection dose. Median sampling durations and numbers of sampling time points per tracer are also summarised in that table. The most common sampling frame for FDG was as follows (min): 0, 0.25, 0.5, 0.75, 1, 1.5, 2, 2.5, 3, 4, 5, 6, 7, 10, 13, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90. For H$_2$O, it was 0, 0.067, 0.13, 0.20, 0.27, 0.33, 0.40, 0.47, 0.53, 0.60, 0.67, 0.73, 0.80, 0.87, 0.93, 1.08, 1.27, 1.42, 1.58, 1.75, 1.92, 2.25, 2.58, 2.92, 3.25, 3.58, 3.92, 4.25, 4.58, 4.92. For FLT, it was 0, 0.25, 0.5, 0.75, 1.00, 1.25, 1.5, 1.75, 2, 2.5, 3, 4, 5, 6, 7, 8, 10, 15, 20, 30, 40, 50, 60, 90, 110. All water studies were manual bolus injections. Only the initial 7 FDG studies were manual injections over 1 min. The other FDG studies had either a 1 min or 2 min programmed infusion using a syringe

TABLE I

TOP SECTION: PATIENT INFORMATION IN MEAN (SD) OF 105 FDG, 39 H$_2$O AND 32 FLT STUDIES (ONLY 101 WEIGHT AND HEIGHT FOR FDG STUDIES). BOTTOM SECTION: MEDIAN SAMPLING DURATIONS AND MEDIAN NUMBERS OF TIME POINTS IN THE SAMPLING FRAMES PER TRACER

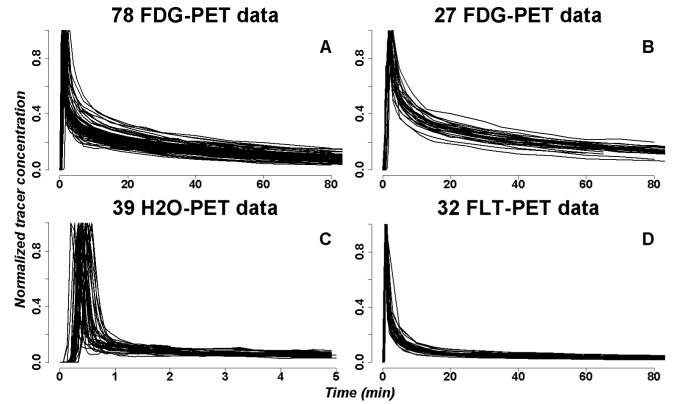| Tracer	FDG	H$_2$O	FLT
Injection Dose (mCi)	8.92 (1.48)	22.72 (20.92)	4.43 (0.59)
Gender	37F; 68M	15F; 24M	7F; 25M
Weight (kg)	79.1 (15.6)	76.1 (14.4)	82.4 (17.3)
Height (cm)	174.7 (9.6)	175.8 (8.3)	175.6 (9.5)
Peak (mCi)	1.87 (0.57)	13.57 (15.03)	0.77 (0.47)
Number of sampling points	26	30	27
Sampling duration (min)	90	4.917	110

Fig. 2. Normalised kinetic PET time activity curves acquired at the University of Washington [9], [29], [41], [43], [44]: (A) 78 FDG-PET curves with 1-minute injection duration, (B) 27 FDG-PET curves with 2-minute injection duration; (C) 39 H$_2$O-PET curves with bolus injection; (D) 32 FLT-PET curves with 1-minute injection.

pump. All FLT injections were a 1 min injection from an infusion pump. A saline flush was used after every injection. Ten studies from each of these 3 sets were used for PBPM fitting of the population components θ. The rest of the data were used to obtain individual fits from each of the models used in the following comparative analyses. The normalized AIFs are presented in Fig. 2. Each curve typically includes three parts: a plateau of 0 concentration at the begin, followed by a rapid increase after injection, and an exponential decay. No imaging data were used in the following analyses.

B. Model Selection

The clinical datasets were used to estimate the population-based parameters in the proposed PBPM model for varying numbers of components. Figure 3 depicts the cross-validated error $\hat{Err}$ defined from $Err(J)$, for the FDG- and H$_2$O-PET cohorts respectively, after carrying out the model selection procedure described in Section II-C. For the FDG set an elbow was clearly identified at 3 components, indicating a 3-component mixture to be optimal in the sense of loss $Err(J)$. Three exponentials with different rates convolved with a Gamma distribution were thus estimated from the population data to approximate the travel time distribution of the FDG and FLT tracers, yielding a reasonable representation in PBPM (which aligns with the traditional two-tissue compartment modeling of FDG molecules in the blood circulation system). For the H$_2$O set,
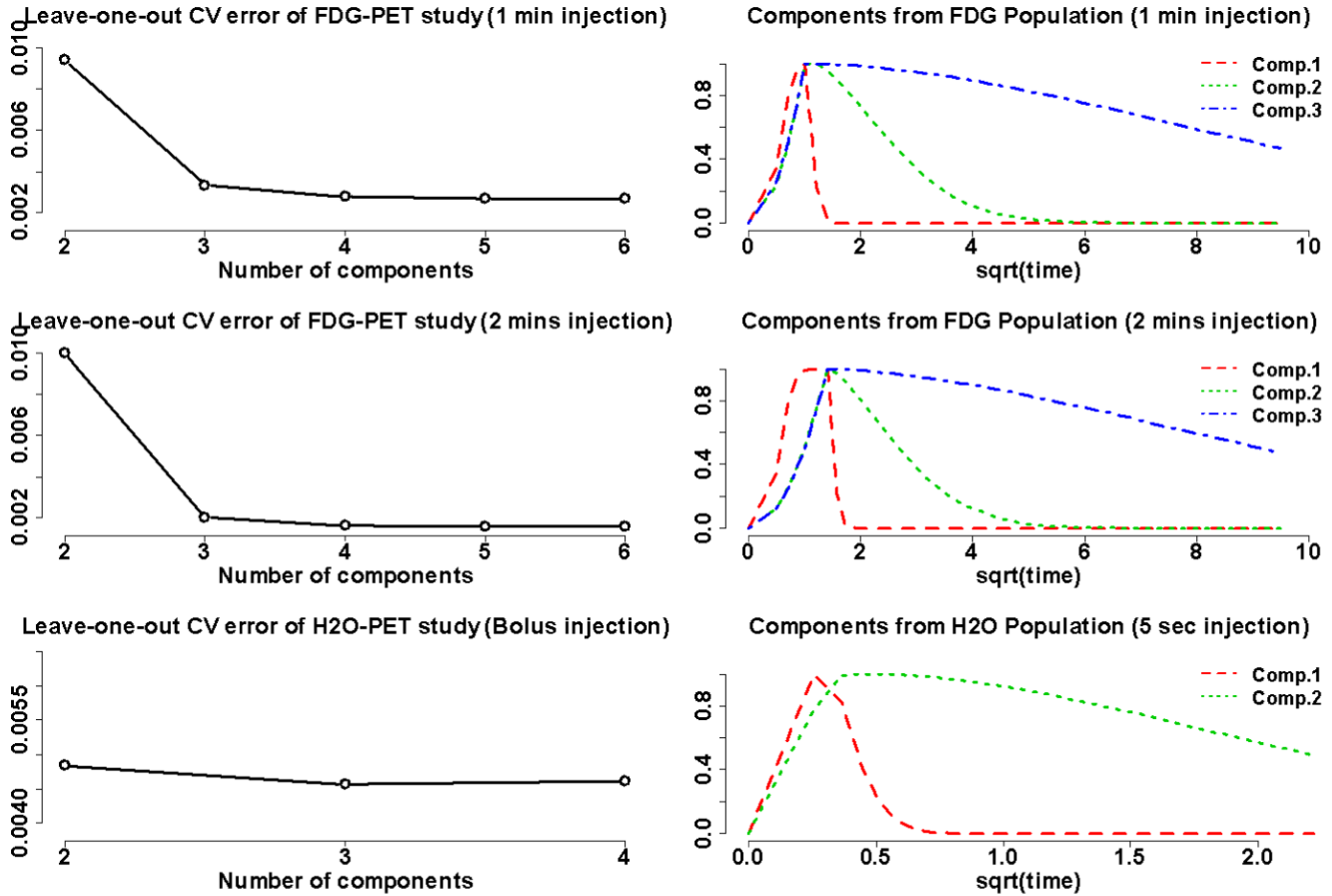
Fig. 3. Left: model selection for FDG (top for one-minute injections; centre for two-minute injections) and H_2O (bottom) studies on the basis of the leave-one-out cross-validated error defined in Section III.C. Right: corresponding basis component fits.

there was no clear elbow and a 2-component mixture seemed the optimal choice according to this selection method. Since, unlike FDG, H_2O is not involved in the metabolic process, it seems reasonable that the number of components used to describe H_2O AIF might be one less than FDG. Note that this model selection rule is only indicative, and using an over-specified mixture (i.e. using more components) may also mechanically impact the bias-variance tradeoff achieved by the model.

C. Model Fitting

Model fitting was analysed on both clinical data (Section III-C.1) and simulated data created using the clinical data as templates (Section III-C.2), to illustrate the procedure and demonstrate its convergence and reliability at varying noise levels.

1) Model Fitting: Figure 4 shows examples of AIF fits obtained from PBPM for FDG-, H_2O-, and FLT-PET clinical studies based on the above model selection. These examples of individual PBPM fits illustrate how the third component in the FDG and FLT PBPM fits, and the second component in the H_2O PBPM fit, could adequately represent the sample AIF tail patterns.

2) Model Validation and Assessment: A set of complementary simulation-based analyses were conducted to validate the

model and its implementation. A first analysis was carried out on AIF templates generated by adding random noise to the PBPM model curve, to validate the implementation and evaluate model performance at varying noise levels. Details on the simulation settings used for these analyses are provided in Appendix A. Figure 5 presents one example simulated curve and its fit for each noise level.

Figure 6 presents the performance of PBPM on 400 simulated AIFs where medium noise level is as above when the noise pattern is set to be similar to the residuals obtained from fitting the PBPM model to the directly sampled AIFs. The RMSE distributions showed overall improvement in model fit as noise level decreases, as expected. The parameter estimation error distributions in Fig. 6 showed an increase in estimation efficiency as noise level decreases. The figure also illustrates that model performance using nonlinear parameters estimates for $\theta = (\alpha, \beta, \phi_1, \phi_2, \phi_3)$ was comparable to that obtained when using the true population parameter values. These results demonstrate that the PBPM approach is consistent and accurate, and validate its implementation [32].

D. Sensitivity to Training Population Size

As for the training population size required to fit the mixture components, behaviour of the PBPM model may vary with the PET tracer used. An experiment on the FDG cohort was
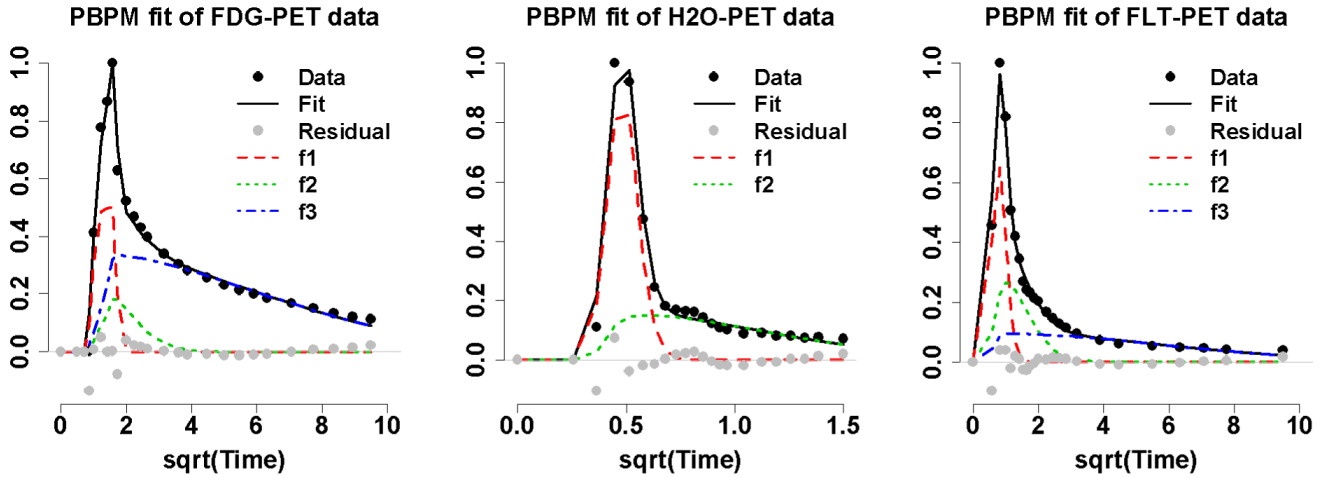
Fig. 4. Examples of individual PBPM fits to FDG (left), H_2O (middle) and FLT (right) arterial data, where the population-fitting components f_j shown in each plot, defined as the convolutions between the exponential components e_j with the population Gamma profile G in Eqn. (1), are scaled by individual mixing weights π_j (with $j = 1, 2, 3$ for FDG and FLT data and $j = 1, 2$ for H_2O data).
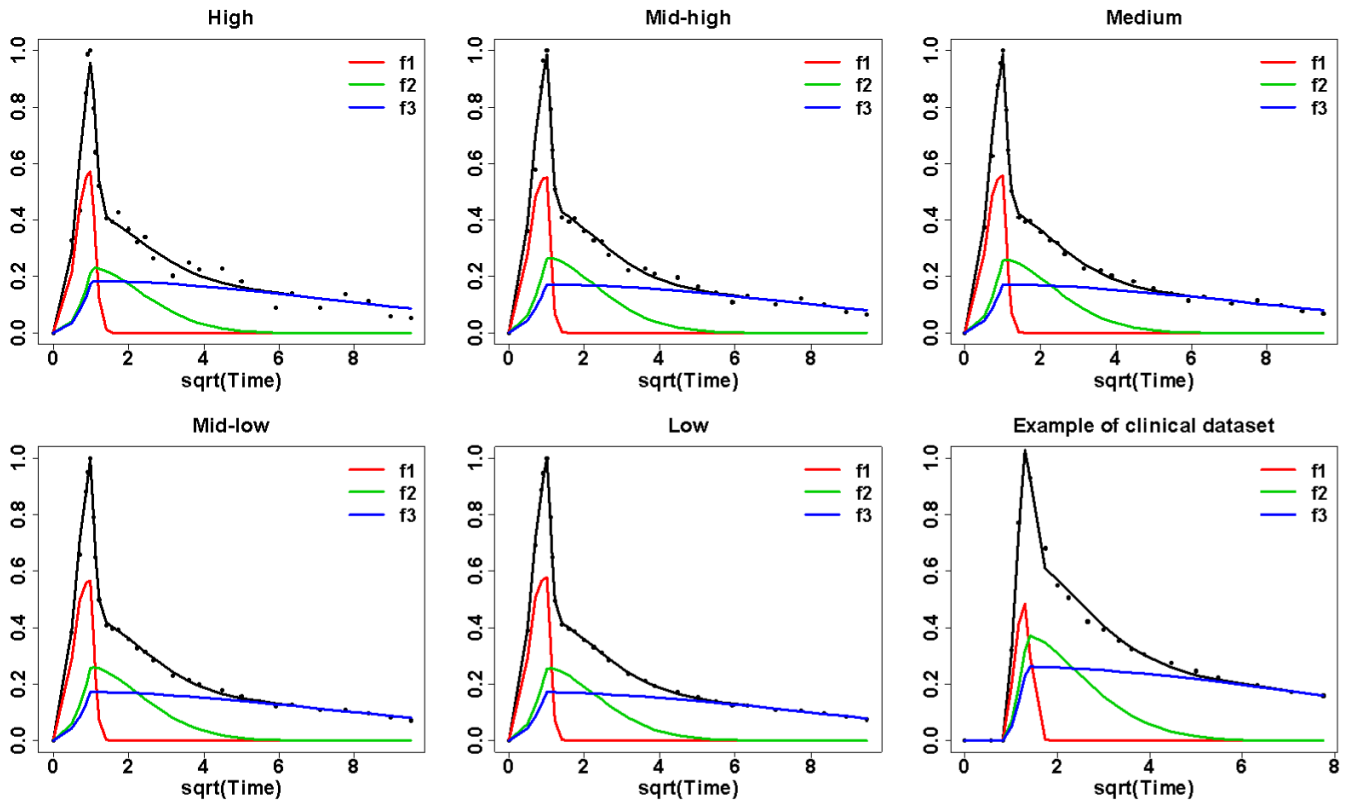

Fig. 5. Example of PBPM fit on a simulated curve for each noise level.

carried out, using random subsets ranging from $n = 5$ to $n = 60$ studies taken from the overall set as training population studies, and monitoring cross-validated model fit errors. The results from this analysis, shown in Fig. 7, indicated that model fitting was relatively insensitive to this parameter, with comparable error distributions obtained when using population parameter estimated across the range of population sizes.

E. Scale Factor Estimation

Results on 91 FDG AIF studies from the scale factor estimation procedure described in Section II-D are depicted

in Fig. 8 (4 patients were removed from the set of 95 curves used in other analyses, due to missing information). They show significant alignment between estimated scale factors (Eqn. (4)) and actual scale factor values obtained directly from arterial sampling data (Pearson correlation $\rho = 0.83$, two-sided test with $p < 1e^{-10}$ under the null hypothesis $H_0 : \rho = 0$).

IV. RESULTS

This section provides output of comparative analyses between the proposed PBPM and reference approaches, using
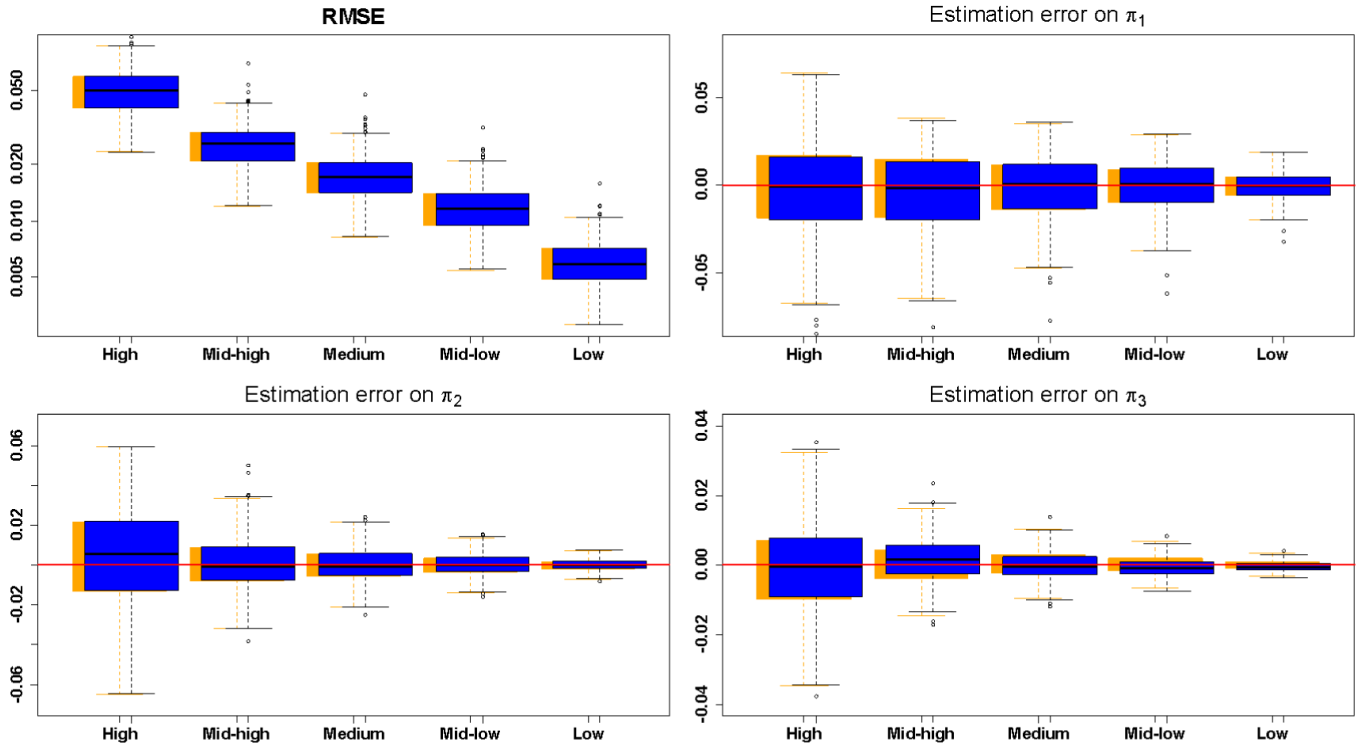
Fig. 6.   Top-left: root mean square error (RMSE) distributions for the 400 AIF fits obtained from PBPM at each noise level. Top-right and bottom: distributions of estimation errors $\pi_i - \hat{\pi}_i, i = 1, 2, 3$. Blue boxplots represent the distributions of parameter estimates obtained using estimated nonlinear parameters for $\theta = (\alpha, \beta, \phi_1, \phi_2, \phi_3)$, whilst the orange (shadow) boxplots indicate those same distributions when the true values for the nonlinear parameters were used.
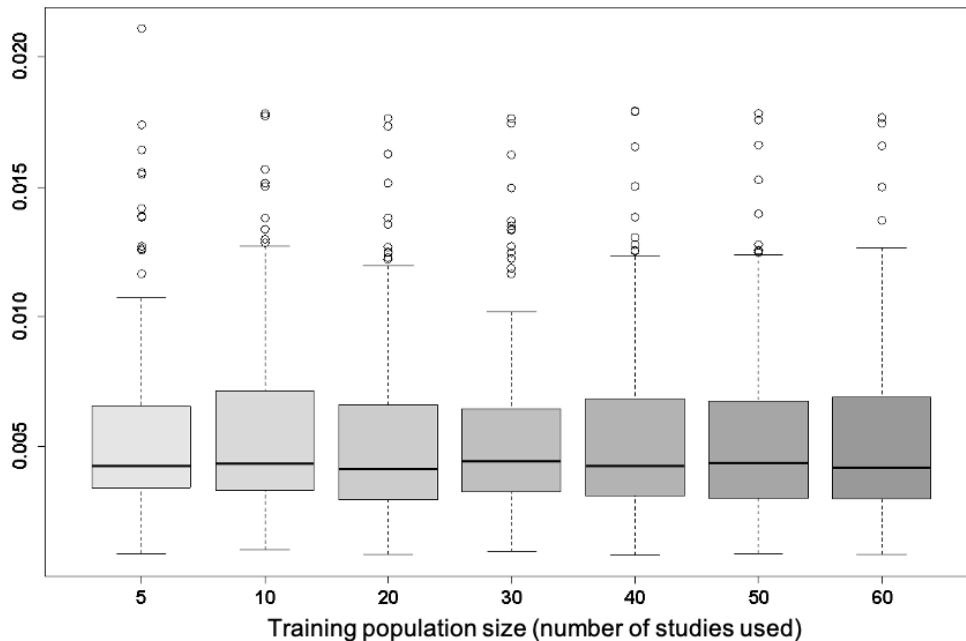

Fig. 7.   Distribution of 3-fold cross-validated AIF PBPM fit errors for varying training population sizes n.

the three clinical datasets introduced in Section III-A. The benefit of using basis components fitted to separate population data is also evaluated hereafter.

A. Comparative Analysis

PBPM-derived AIF estimates were compared to alternatives obtained from another three models described in Section I-B,

namely a whole-body blood circulation model (BCM) [29], [30], a straightforward tri-exponential model (TE) [5], [23], and a convolved tri-exponential model (CTE) [25]. For each tracer, the PBPM was fitted using a random sample of 10 arterial blood sampling curves for estimation of the population parameters $\theta = (\alpha, \beta, \phi_1, \ldots, \phi_J)$ (and not for any other step), and the rest of the data for estimation of individual AIF signals,

TABLE II

ONE-SIDED MANN-WHITNEY TEST P-VALUES FOR THE CROSS-VALIDATED BCM, TE, CTE AND PBPM MSEs OF FIG. 9. THE ALTERNATIVE HYPOTHESES H_A FOR THESE TESTS ARE PROVIDED TO ASSIST IN INTERPRETING THE TEST OUTPUTS. HERE H_A IS ACCEPTED AT THE 5% SIGNIFICANCE LEVEL FOR P-VALUES LOWER THAN 0.05 (SIGNIFICANT RESULTS IN BOLD)

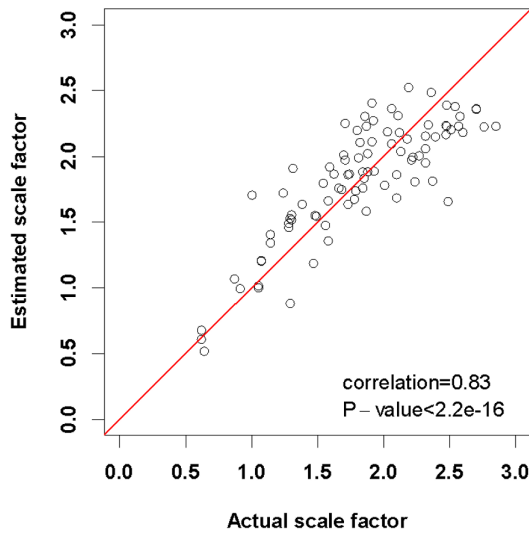| Model Reference	FDG	H_2O	FLT
MSE: mean (sd)			
BCM	0.0116 (0.0094)	0.0086 (0.0069)	0.0127 (0.0072)
TE	0.0073 (0.0047)	0.0099 (0.0040)	0.0066 (0.0035)
CTE	0.0073 (0.0047)	0.0169 (0.0080)	0.0112 (0.0055)
PBPM	0.0063 (0.0045)	0.0073 (0.0039)	0.0070 (0.0041)
Mann-Whitney p-value			
H_A: $MSE_{BCM} > MSE_{PBPM}$	**< 0.001**	0.312	**0.002**
H_A: $MSE_{TE} > MSE_{PBPM}$	**0.002**	**0.004**	0.512
H_A: $MSE_{CTE} > MSE_{PBPM}$	**< 0.001**	**< 0.001**	**0.010**

Fig. 8. Scatterplot of estimated scale factors defined in Eqn. (4) against actual values obtained directly from arterial sampling data from 91 FDG AIF studies, showing a significant association between the two samples.

i.e. of $\{(\pi_k, \Delta_k), k = 1, \ldots, n\}$ (and not for any other step). In other words, the 10 randomly selected "population curves" were used as "historical data" in this framework. These were only used for the purpose of fitting the PBPM, and not the BCM, TE and CTE models. A comparative analysis between these four models was carried out in terms of the sum of mean squared errors obtained for each of the n scaled AIF fits $\{\hat{z}_{i,k} = \hat{C}_P(t_{i,k}|\hat{\pi}_k)\}_{k=1}^n$, defined by

$$E = \sum_{k=1}^{n} \left(\frac{\sum_{i=1}^{m_k}(z_{i,k} - \hat{z}_{i,k})^2}{m_k} \right)$$

This fit error was cross-validated in the comparative analysis, using

$$\hat{E} = \sum_{k=1}^{n} \left(\frac{\sum_{i=2}^{m_k-1}(z_{i,k} - \hat{z}_{i,k}^*)^2}{m_k - 2} \right) \quad (5)$$

for comparison, where $\hat{z}_{i,k}^* = \eta^*(t_{i,k}|\hat{\pi}_k^*)$ and $\eta^*(t_{i,k}|\hat{\pi}_k^*)$ is the set of optimal parameter estimates for the data $\{(t_{j,k}, z_{j,k})\}_{j=1, j\neq i}^m$.

The boxplots of cross-validated error distributions of Fig. 9 (bottom) illustrate that PBPM is competitive for all three

tracers. For the FDG set, the PBPM model yields comparable performance and improved accuracy over TE, and outperforms BCM both in terms of median error and standard error. For the H_2O set, the BCM and PBPM models show greater flexibility in fitting the AIF data compared to the tri-exponential alternatives, and PBPM also yields a lower median error compared to BCM. For the FLT set, TE and PBPM perform comparably and both models outperform BCM and CTE in terms of variability and median errors. Nonparametric, one-sided (Mann-Whitney) tests under the null hypothesis of no difference in model error distribution locations (Table II) indicated a significantly lower PBPM error compared to TE and CTE for both the FDG and H_2O sets at the 1% significance level.

B. Gain From a Population-Based Approach

An initial numerical experiment was carried out on simulated data to assess the potential gain offered by employing a population-based strategy over a fully individual approach to estimate the basis components when fitting the PBPM to sampling data. The same simulation process defined by Eqn. (11) was applied to generate 210 simulated curves from the PBPM templates. A random set of 10 of these curves were used as a training population to obtain estimates $\hat{\theta}$ of the nonlinear parameters θ. The remaining 200 curves were fitted once following the proposed methodology (using this estimate $\hat{\theta}$ to generate PBPM fits), and independently a second time by re-estimating the nonlinear parameters θ individually (i.e. without using the estimate derived from the training population). We refer to the former AIF estimate as the PBPM fit, and the latter AIF estimate as the individual fit. Fit errors were calculated for each of the 200 curves and cross-validated using leave-one-out cross-validation. Figure 10 shows the distributions of cross-validated fit errors obtained from each of the PBPM and individual fits, at varying noise levels. The results show a clear gain resulting from the population-based methodology.

A second analysis was carried out on the clinical data to confirm these findings. Figure 11 illustrates the gain obtained by using a population-based strategy to estimate the nonlinear parameters θ for the PBPM basis component profiles, over a fully individual fitting process where all model parameters (π, Δ, θ) would be estimated from the single patient sampling
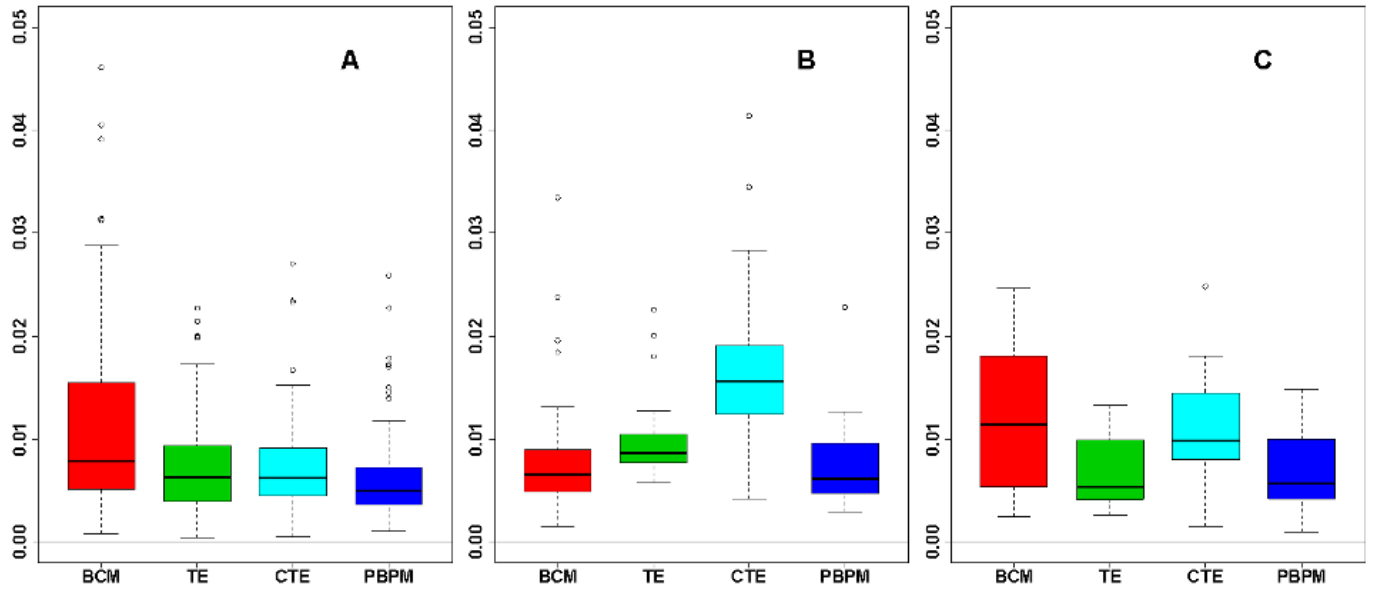
Fig. 9. Model performance comparison for BCM, TE, CTE and PBPM on the FDG- (A), H_2O- (B) and FLT- (C) PET studies. Distributions of cross-validated errors (Eqn. (5)) depict the relative fitting performance of the four models, and indicate that PBPM remains competitive for different tracers.
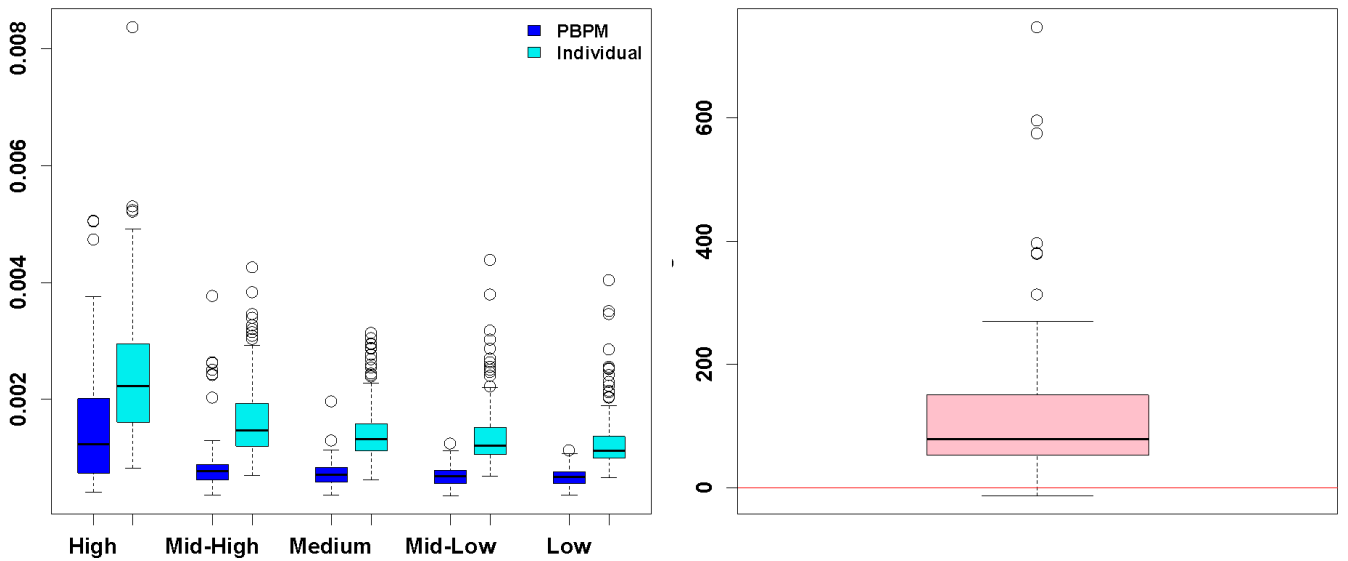

Fig. 10. Left: distributions of PBPM and individual leave-one-out cross-validated fit errors (respectively darker and lighter coloured) at varying noise levels. Right: distribution of percentage differences between these PBPM and individual fit errors (($\hat{E}_{indiv} - \hat{E}_{PBPM})/\hat{E}_{PBPM} \times 100$).

data. A one-sided paired Wilcoxon test, under the null hypothesis that the average cross-validated PBPM error is not less than the average cross-validated individual fit error, confirmed the gain achieved when using estimates $\hat{\theta}$ obtained from a separate training population ($p < 0.0001$).

C. Impact on Kinetic Modeling

AIF estimation is of interest in the context where there is calculation of kinetic parameters. Impact of the proposed PBPM approach on parametric imaging was evaluated in that sense via Monte Carlo simulation of H_2O- and FDG-PET time activity curves generated from one-tissue compartment (1C) and two-tissue compartment

(2C) models respectively [4], [45], [46], [47], [48], where

$$\frac{dC_1(t)}{dt} = K_1 C_p(t) - (k_2 + k_3)C_1(t) + k_4 C_2(t) \quad (6)$$

$$\frac{dC_2(t)}{dt} = k_3 C_1(t) - k_4 C_2(t) \quad (7)$$

(with $k_3 = k_4 = 0$ in the 1C model) describe exchanges between arterial plasma C_p and the first and second tissue compartments C_1 and C_2. Details on the modeling and implementation of this system for this analysis are provided in Appendix B. A set of template H_2O time activity curves $\{z_{t,\text{true}}\}_{i=1}^N$ were created by evaluating the 1C model with each of the $N = 39$ clinical arterial samples from the H_2O dataset described in Section III-A and with true parameter values $K_1 = 0.2$ mL $\cdot$ cm$^{-3} \cdot$ min^{-1}, $k_2 = 0.8$ min^{-1},
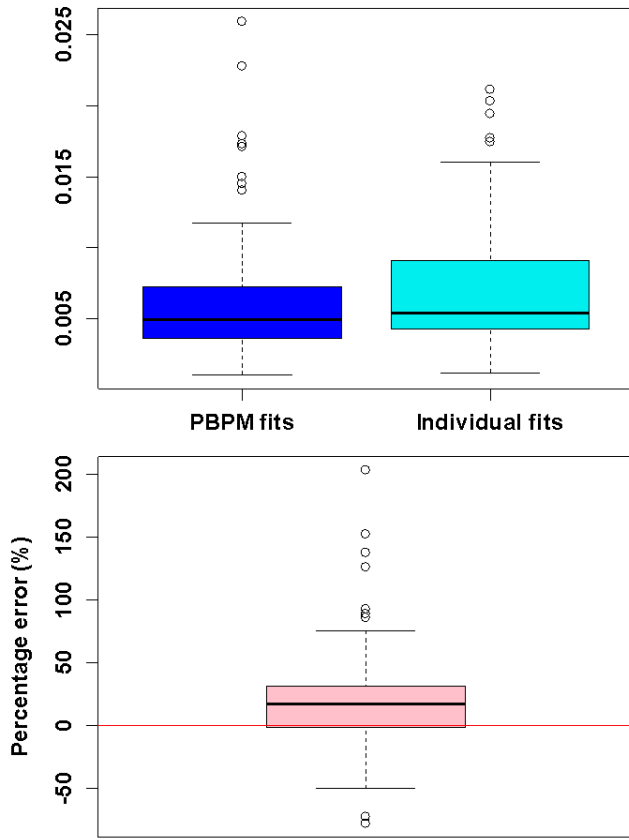
Fig. 11. Top: comparison of cross-validated errors obtained when using a training population strategy to estimate the nonlinear parameters θ that determine the structure of the PBPM basis components (left), over a fully individual approach where the parametric vector θ is estimated directly from the same individual sampling data (right), showing a lower error tends to be achieved from a population-based strategy (one-sided paired Wilcoxon test $p < 0.0001$). Bottom: direct comparison based on percent differences between these PBPM and individual fit errors $((\hat{E}_{indiv} - \hat{E}_{PBPM})/\hat{E}_{PBPM} \times 100)$.

$V_B = 0.05 \text{ mL} \cdot \text{cm}^{-3}$, $\Delta = 0.25$ s (found relevant in other studies [49] in modeling tissue activity around the aorta in tumor blood flow studies). Similarly, a set of template FDG time activity curves $\{z_{t,\text{true}}\}_{i=1}^N$ were obtained by evaluating the 2C model with each of the $N = 91$ clinical arterial samples from the FDG dataset of Section III-A and with true parameter values $K_1 = 0.102 \text{ mL} \cdot \text{cm}^{-3} \cdot \text{min}^{-1}$; $k_2 = 0.13 \text{ min}^{-1}$; $k_3 = 0.062 \text{ min}^{-1}$; $k_4 = 0.0068 \text{ min}^{-1}$; $V_B = 0.04 \text{ mL} \cdot \text{cm}^{-3}$; $\Delta = 0.2667$ s (found relevant in other studies [50] for modeling of grey matter activity in cerebral MRglu studies). This process yielded a set of respectively 39 and 91 unique template uptake curves. These templates were normalised to 1 so estimation of a scale factor was not required in this experiment. Finally, simulated uptake curves were obtained for each scenario by adding Gaussian noise with standard deviation $\sigma = \phi \sqrt{z_{t,\text{true}}}$ with $\phi = 0.001$ for H_2O and $\phi = 0.04$ for FDG, to 10 replicates of each of these template curves. This yielded respectively 390 and 910 simulated time activity curves for these two settings.

Estimates of flow (K_1), rate k_2, and distribution volume $V_D = K_1/k_2$ for the H_2O scenario, and of flux ($K_i = K_1 k_3/(k_2 + k_3)$), distribution volume ($V_D = K_i/k_3$), and rates

	K_i (FDG)	K_1 (FDG)	K_1 (H_2O)
BCM	**0.0317**	< 0.0001	**0.0001**
TE	< 0.0001	0.7392	< 0.0001
CTE	0.1417	0.5621	< 0.0001

	V_D (FDG)	k_2 (FDG)	k_2 (H_2O)
BCM	< 0.0001	< 0.0001	0.4665
TE	**0.0263**	**0.0220**	**0.0074**
CTE	< 0.0001	0.5170	< 0.0001

	V_{nd} (FDG)	k_3 (FDG)	V_t (H_2O)
BCM	0.9997	0.9954	**0.0006**
TE	**0.0002**	< 0.0001	< 0.0001
CTE	0.4531	0.4974	0.2311

K_1, k_2 and k_3 for the FDG scenario, were compared across the range of 1C and 2C model fits obtained using AIF estimates from BCM, TE, CTE and PBPM for the H_2O- and FDG-PET cohorts. Note that in the FDG case here, quantity V_D [51], [52] differs from the distribution volume $V_t = K_1/k_2(1 + k_3/k_4)$ of total ligand uptake in tissue relative to total concentration of ligand in plasma typically of interest for 2-TC models, since we assume $k_4 = 0$ [53]. For FDG, the distribution of non-displaceable compartment relative to total concentration of ligand in plasma $V_{nd} = K_1/k_2$ is also reported on. (In the case of water, these quantities are equivalent.) Delay was held constant for the purpose of this analysis, and the compartment models were optimized with respect to (K_1, k_2, k_3, V_B) (with $k_3 = k_4 = 0$ in the 1C model, and $k_4 = 0$ in the 2C model) via nonlinear estimation.

Figure 12 shows the distributions of percentage errors in estimating these parameters using varying AIF estimates, and Table III provides the p-values of corresponding one-tailed Wilcoxon (Mann-Whitney) tests to assess bias from these distributions, under the null hypothesis of a strictly lower bias for the alternative model, compared to PBPM (i.e. $H_0 : Err_{PBPM} \geq Err_{other}$, after adjustment of error signs where appropriate for the comparison). In complement, "mirror" tests assessing $H_0 : Err_{PBPM} \leq Err_{other}$ indicated that PBPM bias was significantly greater than BCM bias for estimation of V_{nd} and rate k_3 in the FDG setting ($p < 0.005$ for both). No other such tests yielded significance. In particular, TE and CTE did not yield significantly lower bias compared to PBPM in any scenario.

In summary, the PBPM model is a viable option for evaluation of physiological parameters in the FDG and H_2O settings, where it led to either statistically comparable or significantly reduced bias in every scenario for all kinetic parameters, with the only exception of BCM for estimation of V_{nd} and k_3 in FDG. These results illustrate its competitiveness and reliability for kinetic analysis across different tracers, against the most effective methods of reference.
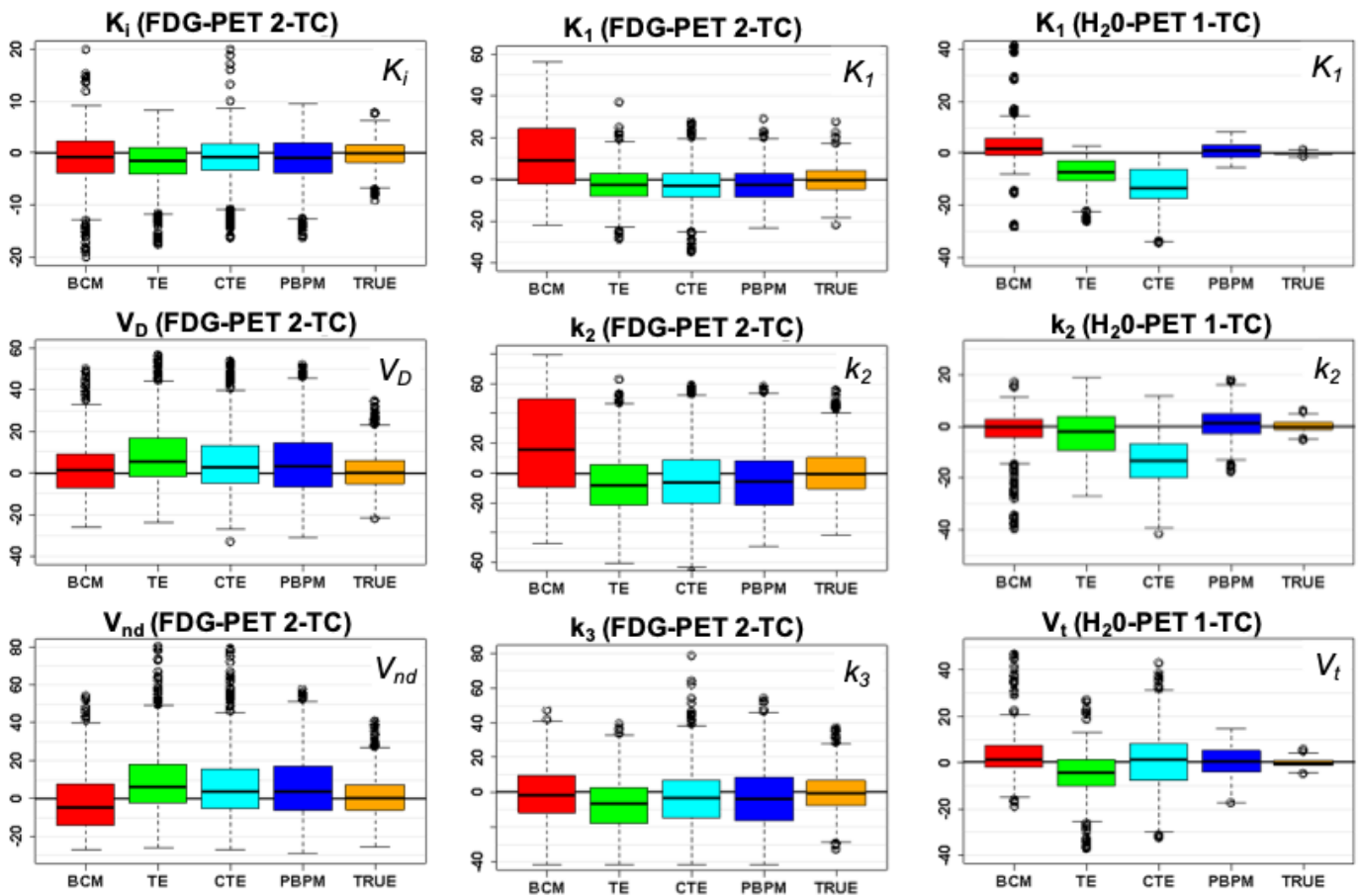
Fig. 12. Distributions of percentage errors in estimates of flux K_i (top left), distribution volume V_D (centre left), V_{nd} (bottom left), K_1 (centre, top), k_2 (center), and k_3 (centre, bottom) parameters from 2C modeling of simulated FDG uptake data, and of flow K_1 (top right), k_2 (center right) and V_t (bottom right) parameters from 1C modeling of simulated H_2O uptake data, using respectively BCM, TE, CTE, PBPM AIF estimates, or the true AIF used in simulating the template curves.

V. DISCUSSION

This paper introduced a novel population-based projection model (PBPM) that combines strategies used in both image-derived input function (IDIF) and population-based input function (PBIF) approaches in order to achieve flexible and reliable AIF estimation throughout the whole time course, by fitting a mixture of population-informed uptake patterns to individual patient imaging data. In this sense this methodology can be seen as the projection of individual patient PET tracer kinetics characteristics onto a basis of population uptake Using a few population-based components, the proposed method can incorporate knowledge of the injection duration into the model fit, thus not requiring the injection protocol to be the same in all PET studies, unlike common PBIF methods.

A comparative analysis was carried out that included two reference models, namely the tri-exponential (TE) and convolved tri-exponential (TCE) models, and a whole-body blood circulation model (BCM) as references for FDG-, H_2O- and FLT-PET datasets, as well as simulated datasets obtained from resampling of the clinical information. Results demonstrated that the proposed PBPM has overall competitive or lower AIF estimation error and greater adaptability to different tracers compared to the other three methods, and illustrated

its consistency at varying noise levels. Application of the proposed approach to parametric imaging, using one- and two-tissue compartment kinetic modeling scenarios, was also evaluated. In this experiment, the PBPM method overall yielded appropriate characterisation of kinetic rates as well as flow, flux and distribution volume, and yielded either comparable or significantly improved bias over the other models (except for two cases where alternate BCM yielded statistically significant improvement over PBPM), showing improved adaptivity to different tracers.

A natural end-goal is to apply AIF modeling to image-derived data, either because clinical sampling is not available or to facilitate non-invasive evaluation. As part of the technical contributions of this paper, regularization of the model should confer robustness against increased noise levels typically found in imaged blood-pool data. A region of interest from PET imaging of the left ventricle could be used in this sense, as the contrast of concentration of tracer from ventricles and their surrounding muscle is usually distinct. Follow-on work is planned around future access to datasets that combine sampling and imaging data to carry out such an evaluation. This will also yield the opportunity to evaluate the proposed PBPM method compared to IDIF techniques used commonly with major PET tracers.

Where only individual venous blood samples and no reliable arterial blood information are available, the proposed technique may still be considered, for example after conversion of the venous signals into arterial scales. Some venous input function (VIF) approaches were proposed that apply an arterio-venous transform scheme to estimate AIF, for example in ^{11}C-labelled studies [54].

The requirement for scale factor estimation, discussed in Section II-D, impacts all modeling techniques, including the proposed PBPM strategy, but it is not yet clear how. Analysing this impact will be the scope of follow-on work.

A critical minimum number of PET studies required to use the proposed PBPM technique with future studies remains to be determined and could vary with the PET radiotracer used. Numerical experiments demonstrated robustness of this model in this aspect, with stable model fit errors for even very few population studies.

This study has not discriminated the data in terms of healthy and cancer subjects, nor in terms of disease types. Although we do not anticipate significant differences in AIF profiles across such strata, using imaged data may introduce differences in PBPM estimates of AIF profiles. Further analyses would be required to explore this.

VI. CONCLUSION

We proposed a statistical model of the arterial input function of a patient that leverages arterial uptake profile information at both the cohort and subject levels. Results from simulations demonstrated numerical stability and consistency of the approach at varying noise levels. Improvement in accuracy of the AIF representation over reference AIF models was assessed on three clinical cohorts, each imaged with one of three routinely used PET tracers (namely, FDG, H_2O and FLT). The proposed model either outperformed or remained competitive against three reference techniques, in terms of AIF estimation RMSE, for these tracers. Experiments further demonstrated that this representation of the AIF is viable for kinetic analysis of physiological parameters such as flow, flux and distribution volume from one- and two-tissue compartment models. Overall the proposed model yielded statistically comparable or significantly lower bias against the three alternate models for determination of kinetics from typical FDG template data in almost all comparisons. The proposed methodology can be incorporated in either population-based or image-derived input function estimation frameworks; the latter will be the scope of follow-on work assessing this model for non-invasive AIF estimation.

APPENDIX

Here we provide technical details on the analyses carried out in the main paper for model selection, simulation-based validation studies, and implementation of the one- and two-tissue compartment models used for kinetic analysis.

A. Validation Study on Model-Derived Template

In the first analysis of Section III-C.2, three mixture components were fit to the FDG-PET imaging cohort to obtain estimates of the Gamma and exponential parameters with parameters set as follows:

$$\hat{\alpha} = 2.1784, \hat{\beta} = 7.8741,$$
$$\hat{\phi}_1 = 88.3992, \hat{\phi}_2 = 0.1488, \hat{\phi}_3 = 0.0088.$$

To simulate individual AIF curves, the linear mixing parameters were generated randomly via an acceptance-rejection algorithm, drawing values from a Normal distribution so that $\pi_i \sim N(\bar{\pi}_i, \sigma(\pi_i)^2)$, $i = 1, 2$, and enforcing some *ad hoc* constraints. A value for π_3 was then set given π_1 and π_2 to ensure that $\sum_{i=1}^{3} \pi_i \geq 1$, and the simulated time courses were defined as

$$C_p^{tem}(t_i) = \pi_1 C_1(t_i|\theta) + \pi_2 C_2(t_i|\theta) + \pi_3 C_3(t_i|\theta) \quad (8)$$

The following steps were taken to simulate the AIF curves:

Step 1 - Generate random component curves $C_1(t_i|\theta)$, $C_2(t_i|\theta)$ and $C_3(t_i|\theta)$ using the nonlinear parameter values given above;

Step 2 - Generate $0.4 < \pi_1 < 0.8$ and $0.08 < \pi_2 < 0.35$ via acceptance-rejection;

Step 3 - Generate $\pi_3 = \frac{1 - max\{\pi_1 C_1(t_i|\theta) + \pi_2 C_2(t_i|\theta)\}}{max\{C_3(t_i|\theta)\}}$;

Step 4 - Test that π_3 satisfies

$$0.17 < \pi_3 < 0.37 \quad (9)$$
$$0.99 < \pi_1 + \pi_1 + \pi_3 < 1.15 \quad (10)$$

and repeat Steps 2–4 until the above constraints are satisfied, to ensure adequate mixing coefficient values.

In the second analysis of Section III-C.2, sets of 400 simulated curves were generated using the same process at varying noise levels termed *high*, *mid-high*, *medium*, *mid-low* and *low*

$$C_P^{sim}(t_i) = C_p^{tem}(t_i) + \epsilon(\cdot, t_i) \quad (11)$$

where $\epsilon(medium) \sim N(0, .05)$ when $t_i < 1.5$ min and $\epsilon(medium) \sim N(0, .01)$ when $t_i \geq 1.5$ min for level *medium*. $\epsilon(high)$, $\epsilon(mid-high)$, $\epsilon(mid-low)$ and $\epsilon(low)$ are 3, $\frac{3}{2}$, $\frac{2}{3}$ and $\frac{1}{3}$ times of $\epsilon(medium)$, respectively.

B. Evaluation of Compartment Models

Here we provide details on the parametric modeling approach used for kinetic analysis in Section IV-C. The arterial plasma concentration of the tracer is the input for kinetic models. Direct measurement of plasma concentration may not always be feasible, but if the arterial whole blood activity is known (or estimated), it may be used to infer the plasma concentration [29], [55]. Implementation of the one- and two-tissue compartment models given concentration in plasma $C_P(t)$ involves computation of the concentration in tissue $C_T(t)$, represented as the following convolution:

$$C_T(t) = (\varphi_1 e^{-\theta_1 t} + \varphi_2 e^{-\theta_2 t}) \otimes C_P(t)$$

This concentration $C_T(t)$ can be seen as the sum of concentrations in compartments 1 and 2, i.e. $C_T(t) = C_1(t) + C_2(t)$. Exchanges between arterial plasma C_p and the first and second tissue compartments C_1 and C_2 are described

respectively as follows (with $k_3 = k_4 = 0$ in the 1C model) [4], [9], [48], [56]:

$$\frac{dC_1(t)}{dt} = K_1 C_p(t) - (k_2 + k_3)C_1(t) + k_4 C_2(t) \quad (12)$$

$$\frac{dC_2(t)}{dt} = k_3 C_1(t) - k_4 C_2(t). \quad (13)$$

The compartment-specific contributions are derived from Eqn. (12) and Eqn. (13) respectively as follows ([9], [56]):

$$C_1(t) = \frac{K_1}{\theta_2 - \theta_1} \left((k_4 - \theta_1)e^{-\theta_1 t} + (\theta_2 - k_4)e^{-\theta_2 t} \right) \otimes C_P(t)$$

$$C_2(t) = \frac{K_1 k_3}{\theta_2 - \theta_1} \left(e^{-\theta_1 t} - e^{-\theta_2 t} \right) \otimes C_P(t)$$

Specifically, we can use

$$\Delta = \sqrt{(k_2 + k_3 + k_4)^2 - 4k_2 k_4}$$

$$\theta_1 = \frac{k_2 + k_3 + k_4 + \Delta}{2}$$

$$\theta_2 = \frac{k_2 + k_3 + k_4 - \Delta}{2}$$

$$\varphi_1 = K_1 \frac{\theta_1 - k_3 - k_4}{\Delta}$$

$$\varphi_2 = -K_1 \frac{\theta_2 - k_3 - k_4}{\Delta}$$

to evaluate the model for given (K_1, k_2, k_3, k_4) values. For the one-tissue compartment model, we simply have $k_3 = k_4 = 0$.

ACKNOWLEDGMENT

REFERENCES

[1] S. C. Huang and M. E. Phelps, "Principles of tracer kinetic modeling in positron emission tomography and autoradiography," in *Positron Emission Tomography and Autoradiography: Principles and Applications for the Brain and Heart*, M. Phelps, J. Mazziotta, and H. Schelbert, Eds. New York, NY, USA: Raven Press, 1986, pp. 287–346.

[2] P. Herscovitch, J. Markham, and M. E. Raichle, "Brain blood flow measured with intravenous $H_2^{15}O$. I. Theory and error analysis," *J. Nucl. Med.*, vol. 24, pp. 782–789, Sep. 1983.

[3] M. E. Raichle, W. R. Martin, P. Herscovitch, M. A. Mintun, and J. Markham, "Brain blood flow measured with intravenous $H_2^{15}O$. II. Implementation and validation," *J. Nucl. Med.*, vol. 24, pp. 790–798, Sep. 1983.

[4] L. Sokoloff et al., "The [^{14}C]deoxyglucose method for the measurement of local cerebral glucose utilization: Theory, procedure, and normal values in the conscious and anesthetized albino rat," *J. Neurochem.*, vol. 28, no. 5, pp. 897–916, 1977.

[5] F. Zanderigo, R. T. Ogden, and R. V. Parsey, "Noninvasive blood-free full quantification of positron emission tomography radioligand binding," *J. Cereb. Blood Flow Metab.*, vol. 35, no. 1, pp. 148–156, 2014.

[6] K. Chen et al., "Noninvasive quantitation of the cerebral metabolic rate for glucose using positron emission tomography, ^{18}F-fluorodeoxyglucose, the Patlak method, and an imaged-derived input function," *J. Cerebral Blood Flow Metabolism*, vol. 18, pp. 716–723, Jul. 1998.

[7] R. Laforest et al., "Measurement of input functions in rodents: Challenges and solutions," *Nucl. Med. Biol.*, vol. 32, pp. 679–685, Oct. 2005.

[8] J. Kim et al., "Minimally invasive method of determining blood input function from PET images in rodents," *J. Nucl. Med.*, vol. 47, no. 2, pp. 330–336, 2006.

[9] F. O'Sullivan et al., "Nonparametric residue analysis of dynamic PET data with application to cerebral FDG studies in normals," *J. Amer. Stat. Assoc.*, vol. 104, pp. 556–571, Jun. 2009.

[10] E. Croteau et al., "Image-derived input function in dynamic human PET/CT: Methodology and validation with ^{11}C-acetate and ^{18}F-fluorothioheptadecanoic acid in muscle and ^{18}F-fluorodeoxyglucose in brain," *Eur. J. Nucl. Med. Mol. Imag.*, vol. 37, pp. 1539–1550, Aug. 2010.

[11] P. Zanotti-Fregonara, K. Chen, J.-S. Liow, M. Fujita, and R. B. Innis, "Image-derived input function for brain PET studies: Many challenges and few opportunities," *J. Cerebral Blood Flow Metabolism*, vol. 31, pp. 1986–1998, Aug. 2011.

[12] E. K. Fung and R. E. Carson, "Cerebral blood flow with [^{15}O] water PET studies using an imaged-derived input function and MR-defined carotid centerlines," *Phys. Med. Biol.*, vol. 58, pp. 1903–1923, Mar. 2013.

[13] H. Sari et al., "Estimation of an image derived input function with MR-defined carotid arteries in FDG-PET human studies using a novel partial volume correction method," *J. Cerebral Blood Flow Metabolism*, vol. 37, no. 4, pp. 1398–1409, 2017.

[14] R. T. Ogden, F. Zanderigo, S. Choy, J. J. Mann, and R. V. Parsey, "Simultaneous estimation of input functions: An empirical study," *J. Cerebral Blood Flow Metabolism*, vol. 30, no. 4, pp. 816–826, 2010.

[15] S. Takikawa et al., "Noninvasive quantitative fluorodeoxyglucose PET studies with an estimated input function derived from a population-based arterial blood curve," *Radiology*, vol. 188, no. 1, pp. 131–136, 1993.

[16] S. Eberl, A. R. Anayat, R. R. Fulton, P. K. Hooper, and M. J. Fulham, "Evaluation of two population-based input functions for quantitative neurological FDG PET studies," *Eur. J. Nucl. Med. Mol. Imag.*, vol. 24, pp. 299–304, Mar. 1997.

[17] T. Shiozaki et al., "Noninvasive estimation of FDG input function for quantification of cerebral metabolic rate of glucose: Optimization and multicenter evaluation," *Nucl. Med. Commun.*, vol. 41, pp. 1612–1618, Oct. 2000.

[18] P. Meyer, V. Circiumaru, C. A. Cardi, D. A. Thomas, H. Bal, and P. D. Acton, "Simplified quantification of small animal [^{18}F]FDG PET studies using a standard arterial input function," *Eur. J. Nucl. Med. Mol. Imag.*, vol. 33, no. 8, pp. 948–954, 2006.

[19] P. Zanotti-Fregonara et al., "Population-based input function and image-derived input function for [^{11}C](R)-rolipram PET imaging: Methodology, validation and application to the study of major depressive disorder," *NeuroImage*, vol. 63, no. 3, pp. 1532–1541, 2012.

[20] F. O'Sullivan et al., "Kinetic quantitation of cerebral PET-FDG studies without concurrent blood sampling: Statistical recovery of the arterial input function," *IEEE IEEE Trans. Med. Imag.*, vol. 29, no. 3, pp. 610–624, Mar. 2010.

[21] F. Zanderigo, R. T. Ogden, and R. V. Parsey, "Reference region approaches in PET: A comparative study on multiple radioligands," *J. Cerebral Blood Flow Metabolism*, vol. 33, no. 6, pp. 888–897, 2013.

[22] M. Bentourkia, "Kinetic modeling of PET-FDG in the brain without blood sampling," *Computerized Med. Imag. Graph.*, vol. 30, no. 8, pp. 447–451, 2006.

[23] D. Feng, S.-C. Huang, and X. Wang, "Models for computer simulation studies of input functions for tracer kinetic modeling with positron emission tomography," *Int. J. Bio-Med. Comput.*, vol. 32, pp. 95–110, Mar. 1993.

[24] R. V. Parsey et al., "Validation and reproducibility of measurement of 5-HT$_{1A}$ receptor parameters with [carbony1-^{11}C]WAY-100635 in humans: Comparison of arterial and reference tissue input functions," *J. Cereb. Blood Flow Metabolism*, vol. 20, pp. 1111–1133, Jul. 2000.

[25] M. Tonietto, G. Rizzo, M. Veronese, and A. Bertoldo, "Modelling arterial input functions in positron emission tomography dynamic studies," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 2247–2250.

[26] A. Holt, E. Pasca, S. W. Heijmink, J. Teertstra, S. H. Müller, and U. A. van der Heide, "The impact of overall injection time on the arterial input function and pharmaco-kinetic analysis using the Tofts model in DCE-MRI for prostate cancer patients," in *Proc. Int. Soc. Mag. Reson. Med.*, vol. 20, 2012, p. 238.

[27] E. Poulin et al., "Conversion of arterial input functions for dual pharmacokinetic modeling using Gd-DTPA/MRI and ^{18}F-FDG/PET," *Magn. Reson. Med.*, vol. 69, no. 3, pp. 781–792, 2013.

[28] H. Sari *et al.*, "Modelling the impact of injection time on the bolus shapes in PET-MRI AIF conversion," *EJNMMI Phys.*, vol. 1, pp. 1–3, Dec. 2014.

[29] J. Huang and F. O'Sullivan, "An analysis of whole body tracer kinetics in dynamic PET studies with application to image-based blood input function extraction," *IEEE Trans. Med. Imag.*, vol. 33, no. 5, pp. 1093–1108, May 2014.

[30] F. O'Sullivan *et al.*, "Assessment of a statistical AIF extraction method for dynamic PET studies with ^{15}O water and ^{18}F fluorodeoxyglucose in locally advanced breast cancer patients," *J. Med. Imag.*, vol. 5, no. 1, 2017, Art. no. 011010.

[31] *R: A Language and Environment for Statistical Computing*, R Core Team, R Foundation for Statistical Computing, Vienna, Austria, 2016.

[32] Z. Xiu, J. Huang, and E. Wolsztynski. (2020). *PBCMM*. [Online]. Available: https://github.com/ryanxiu/pbcmm.git

[33] B. A. Turlach and A. Weingessel, *Quadprog: Functions to Solve Quadratic Programming Problems*, document R Package Version 1.5-4, 2011.

[34] F. O'Sullivan and G. Wahba, "A cross validated Bayesian retrieval algorithm for nonlinear remote sensing experiments," *J. Comput. Phys.*, vol. 59, pp. 441–455, Jul. 1985.

[35] C. S. Brock, H. Young, S. Osman, S. K. Luthra, T. Jones, and P. M. Price, "Glucose metabolism in brain tumours can be estimated using [^{18}F]2-fluorodeoxyglucose positron emission tomography and a population-derived input function scaled using a single arterialised venous blood sample," *Int. J. Oncol.*, vol. 26, no. 5, p. 1377, 2005.

[36] M. Tonietto, F. Zanderigo, A. Bertoldo, D. P. Devanand, and B. Stankoff, "Multicenter validation of population-based input function with nonlinear mixed effect modeling for voxel-wise quantification of [^{18}f]FDG metabolic rate," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 376–379.

[37] M. Bentourkia, "Kinetic modeling of PET data without blood sampling," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 3, pp. 697–702, Jun. 2005.

[38] E. Rissanen *et al.*, "Automated reference region extraction and population-based input function for brain [^{11}C]TMSX pet image analyses," *J. Cerebral Blood Flow Metabolism*, vol. 35, no. 1, pp. 157–165, 2015.

[39] M. I. Akerele, S. A. Zein, S. Pandya, A. Nikolopoulou, and S. A. Nehmeh, "Population-based input function for TSPO quantification and kinetic modeling with [^{11}C]-DPA-713," *EJNMMI Phys.*, vol. 8, no. 1, pp. 1–17, 2021.

[40] G. M. March, S. Elfatori, and P. E. Beaulé, "Clinical experience with tranexamic acid during primary total hip arthroplasty," *Hip Int.*, vol. 23, no. 1, pp. 72–79, 2013.

[41] M. M. Graham *et al.*, "The fluorodeoxyglucose lumped constant in normal human brain," *J. Nucl. Med.*, vol. 43, pp. 1157–1166, 2002.

[42] M. Muzi *et al.*, "Kinetic analysis of 3'-deoxy-3'-^{18}F-fluorothymidine in patients with gliomas," *J. Nucl. Med.*, vol. 47, no. 10, pp. 1612–1621, 2006.

[43] M. Muzi *et al.*, "Imaging of cyclosporine inhibition of P-glycoprotein activity using ^{11}C-verapamil in the brain: Studies of heathy humans," *J. Nucl. Med.*, vol. 50, no. 8, pp. 1267–1275, 2009.

[44] S. Eyal *et al.*, "Regional P-glycoprotein activity and inhibition at the human blood-brain barrier as imaged by positron emission tomography," *Clin. Pharmacol. Therapeut.*, vol. 87, no. 5, pp. 579–585, 2010.

[45] E. D. Morris, C. J. Endres, K. C. Schmidt, B. T. Christian, R. F. Muzic, and R. E. Fisher, "Kinetic modeling in positron emission tomography," *Emission Tomogr., Fundam. PET SPECT*, vol. 46, no. 1, pp. 499–540, 2004.

[46] D. Hawe, F. R. H. Fernández, L. O'Suilleabháin, J. Huang, E. Wolsztynski, and F. O'Sullivan, "Kinetic analysis of dynamic positron emission tomography data using open-source image processing and statistical inference tools," *Wiley Interdiscipl. Rev. Comput. Statist.*, vol. 4, no. 3, pp. 316–322, 2012.

[47] A. Bertoldo, G. Rizzo, and M. Veronese, "Deriving physiological information from PET images: From SUV to compartmental modelling," *Clin. Transl. Imag.*, vol. 2, no. 3, pp. 239–251, 2014.

[48] A. Dimitrakopoulou-Strauss, L. Pan, and C. Sachpekidis, "Kinetic modeling and parametric imaging with dynamic PET for oncological applications: General considerations, current clinical applications, and future perspectives," *Eur. J. Nucl. Med. Mol. Imag.*, vol. 48, pp. 21–39, Jan. 2021.

[49] W. Watabe *et al.*, "Noninvasive estimation of the aorta input function for measurement of tumor blood flow with [^{15}/O] water," *IEEE Trans. Med. Imag.*, vol. 20, no. 3, pp. 164–174, Mar. 2001.

[50] D. Feng *et al.*, "An evaluation of the algorithms for determining local cerebral metabolic rates of glucose using positron emission tomography dynamic data," *IEEE Trans. Med. Imag.*, vol. 14, no. 4, pp. 697–710, Dec. 1995.

[51] Y. Choi *et al.*, "Evaluation of the effect of glucose ingestion and kinetic model configurations of FDG in the normal liver," *J. Nucl. Med.*, vol. 35, no. 5, pp. 818–823, 1994.

[52] E. Taylor *et al.*, "Impact of tissue transport on pet hypoxia quantification in pancreatic tumours," *EJNMMI Res.*, vol. 7, no. 1, pp. 1–11, 2017.

[53] R. B. Innis *et al.*, "Consensus nomenclature for *in vivo* imaging of reversibly binding radioligands," *J. Cerebral Blood Flow Metabolism*, vol. 27, no. 9, pp. 1533–1539, 2007.

[54] S. Syvanen, G. Blomquist, L. Appel, M. Hammarlund-Udenaes, B. Langstrom, and M. Bergstrom, "Predicting brain concentrations of drug using positron emission tomography and venous input: Modeling of arterial-venous concentration differences," *Eur. J. Clin. Pharmacol.*, vol. 62, no. 10, pp. 839–848, 2006.

[55] P. H. Hinderling, "Red blood cells: A neglected compartment in pharmacokinetics and pharmacodynamics," *Pharmacolog. Rev.*, vol. 49, no. 3, pp. 279–295, 1997.

[56] R. N. Gunn, S. R. Gunn, and V. J. Cunningham, "Positron emission tomography compartmental models," *J. Cerebral Blood Flow Metabolism*, vol. 21, no. 6, pp. 635–652, 2001.