EJNMMI Research

ORIGINAL RESEARCH

Open Access

Non-invasive arterial input function estimation using an MRA atlas and machine learning

Rajat Vashistha[1,2], Hamed Moradi[1,2,3], Amanda Hammond[3], Kieran O'Brien[3], Axel Rominger[5], Hasan Sari[4,5], Kuangyu Shi[5], Viktor Vegh[1,2*] and David Reutens[1,2]

Abstract

Background Quantifying biological parameters of interest through dynamic positron emission tomography (PET) requires an arterial input function (AIF) conventionally obtained from arterial blood samples. The AIF can also be non-invasively estimated from blood pools in PET images, often identified using co-registered MRI images. Deploying methods without blood sampling or the use of MRI generally requires total body PET systems with a long axial field-of-view (LAFOV) that includes a large cardiovascular blood pool. However, the number of such systems in clinical use is currently much smaller than that of short axial field-of-view (SAFOV) scanners. We propose a data-driven approach for AIF estimation for SAFOV PET scanners, which is non-invasive and does not require MRI or blood sampling using brain PET scans. The proposed method was validated using dynamic ^{18}F-fluorodeoxyglucose [^{18}F]FDG total body PET data from 10 subjects. A variational inference-based machine learning approach was employed to correct for peak activity. The prior was estimated using a probabilistic vascular MRI atlas, registered to each subject's PET image to identify cerebral arteries in the brain.

Results The estimated AIF using brain PET images (IDIF-Brain) was compared to that obtained using data from the descending aorta of the heart (IDIF-DA). Kinetic rate constants (K_1, k_2, k_3) and net radiotracer influx (K_i) for both cases were computed and compared. Qualitatively, the shape of IDIF-Brain matched that of IDIF-DA, capturing information on both the peak and tail of the AIF. The area under the curve (AUC) of IDIF-Brain and IDIF-DA were similar, with an average relative error of 9%. The mean Pearson correlations between kinetic parameters (K_1, k_2, k_3) estimated with IDIF-DA and IDIF-Brain for each voxel were between 0.92 and 0.99 in all subjects, and for K_i, it was above 0.97.

Conclusion This study introduces a new approach for AIF estimation in dynamic PET using brain PET images, a probabilistic vascular atlas, and machine learning techniques. The findings demonstrate the feasibility of non-invasive and subject-specific AIF estimation for SAFOV scanners.

Keywords PET, Dynamic scanning, Kinetic modelling, Arterial input function, Variational inference

*Correspondence:
Viktor Vegh
viktor.vegh@cai.uq.edu.au
[1]Centre for Advanced Imaging, University of Queensland, Brisbane, Australia
[2]ARC Training Centre for Innovation in Biomedical Imaging Technology, University of Queensland, Brisbane, Australia
[3]Siemens Healthcare Pty Ltd, Melbourne, Australia
[4]Advanced Clinical Imaging Technology, Siemens Healthcare AG, Lausanne, Switzerland
[5]Department of Nuclear Medicine, Bern University Hospital, Inselspital, University of Bern, Freiburgstrasse 18, Bern 3010, Switzerland

Introduction

Dynamic PET images capture the time course of radiopharmaceutical uptake within specific organs or tissues. Kinetic modelling maps these images to images of kinetic parameters, which reflect physiological and biochemical processes within tissue or regions of interest analysis [1, 2]. Parametric imaging is increasingly used in research with a growing interest in its applications in clinical diagnosis, therapy monitoring and treatment planning [3], as disease states may perturb tissue kinetic parameters [4–7]. For instance, the kinetic analysis of ^{18}F-fluorodeoxyglucose ([^{18}F]FDG) uptake can differentiate lung cancer patients into those with either short or long survivals [8]. In contrast, the kinetic constants K_1 and K_i for ^{18}F-fluorocholine can distinguish between malignant and healthy prostate tissue [9]. Analysis of radiotracer kinetics can also increase diagnostic accuracy in oncological applications [3, 10].

Estimation of kinetic parameters requires knowledge of the arterial input function (AIF) [11]. An ideal input function estimation method should be non-invasive, accurate, specific to the individual and amenable to wide implementation. Traditionally, the input function has been measured using arterial blood sampling, which is invasive and labor-intensive [12]. A non-invasive alternative is to use an image-derived input function, estimated from blood pools in large vascular structures within the PET image. Imaging of such vascular structures is easily achieved with long axial field-of-view (LAFOV) total body PET systems (with axial length greater than 70 cm), in which the heart and proximal aorta are present in the field-of-view (FOV). For brain scans performed with short axial field-of-view (SAFOV) scanners, the largest vessels imaged, the carotid arteries, are of smaller diameter (~5 mm) in comparison to the aorta, and additional co-registered magnetic resonance scans are often utilised for their segmentation [12–14]. Another consequence of the small diameter of the carotid arteries is that the estimated image derived input function is affected by radiotracer spill-over and spill-in due to partial volume effects. Following a bolus injection of radiotracer, the input function radiotracer concentration rises rapidly to a peak which precedes a tail in which activity gradually decreases over time. Counts originating from intra-arterial radiotracer that spill over into adjacent voxels lower the estimated input function peak, while counts from adjacent voxels that spill into arterial voxels elevate the tail of the input function. Deconvolution to correct for these partial volume effects requires knowledge of the scanner point spread function and of patient-specific arterial anatomy [14–16].

A different approach is to use a population-based input function, generally scaled using at least one blood sample from the subject [17]. Here, inaccuracies, particularly in microparameter estimates, can result when the shape of the actual input function differs from the population average, for example, when radiotracer injection protocols differ. Simultaneous estimation of the input function and kinetic parameters has also been performed using methods incorporating a mathematical model for the input function, additionally requiring at least one venous blood sample [18].

Machine learning (ML) methods have also been applied to non-invasive AIF estimation in the individual subject without blood sampling. Varnyu et al. employed a deep neural network implementation to estimate arterial input function (AIF) utilizing sinogram data [19]. Kuttner et al. showcased that long short-term memory (LSTM) recurrent neural network models yield reduced error rates compared to Gaussian process regression in estimating the input function from tissue time activity curves. However, most methods are supervised learning methods requiring paired training data [20, 21]. This is a major drawback because the amount of paired data available for training is generally small, resulting in poor generalization. This motivated us to develop a machine learning method that does not require arterial blood sampling data for training.

We propose a novel method for non-invasive AIF estimation in individual subjects from head images that can be used with SAFOV scanners. The method incorporates information from a probabilistic vascular atlas [22], and applies variational inference, a probabilistic machine learning technique. This approach enables the estimation of peak activity corrected input function from limited data. The prior for Bayesian variational inference is the mean input function calculated from the subject specific brain PET images using vascular atlas. To our knowledge, this is the first non-invasive method for input function estimation using a machine learning approach that does not require training with previous example data.

Materials and methods

Human data and acquisition protocol

Total body PET data were acquired using a Biograph Vision Quadra in list mode started 15 s prior to the intravenous injection of [^{18}F]FDG (mean activity: 235 ± 51 MBq; i.e., approximately 3 MBq/kg). Emission data were acquired for 65 min. The data were binned into 62 frames using the following frame durations: 2 × 10 s, 30 × 2 s, 4 × 10 s, 8 × 30 s, 4 × 60 s, 5 × 120 s, and 9 × 300 s. Framing windows are fixed as this work is based on the use of retrospective data. Images were reconstructed using the Siemens PSF + TOF algorithm with conventional attenuation, scatter and random count correction strategies [23]. Of the 24 datasets from oncological patients, 10 were selected randomly to test the developed method. The Sample size of 10 detects a large (Cohen's d = 0.8)

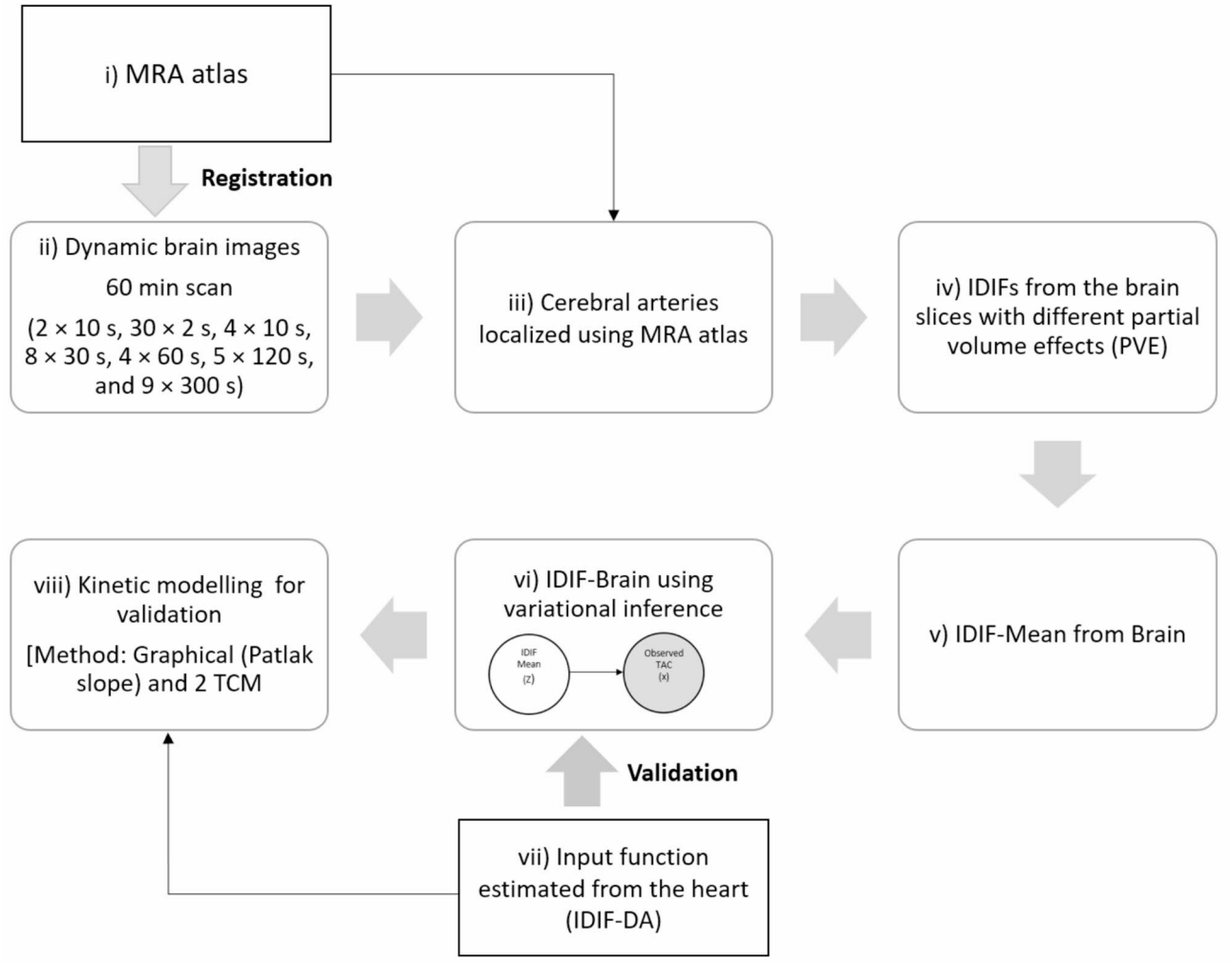


Fig. 1 The pipeline for estimating the input function using brain images is illustrated. The first five steps (i-v) involve the estimation of the IDIF-Mean from dynamic brain PET images using a probabilistic atlas. Step vi) involves correction for peak activity effects using variational inference (IDIF-Brain). Step vii) involves obtaining a reference AIF from the heart (IDIF-DA). Step viii) involves validating the estimated input function using kinetic modelling

difference in image-defined input function between the proposed method (IDIF-Brain) and a currently used (IDIF-DA) method with a power of 80% and $\alpha = 0.05$ (paired t-test), when calculated using G*Power software package [24].

Benchmark AIF

Image derived input functions were obtained from circular regions of interest (ROI), 10 mm in diameter, over the descending aorta (IDIF-DA). This input function was used as the reference standard because it exhibited less rhythmic motion than input functions derived from cardiac images. In addition, as the mean delay observed in peaks from the carotid arteries (mean delay: 3.2 ± 1.7 s), and descending aorta (mean delay: 3.8 ± 1.8 s) were found to be negligible [23]. It is due to this reason we have not considered the delay correction. Previous studies have validated the equivalence of IDIF-DA to image-derived input functions via blood sampling [25], and in kinetic parameters have been estimated accurately using aorta derived input functions [26].

Identification in brain arteries

Voxels in the brain image with a high probability of containing an artery were identified using a previously-described arterial occurrence probability atlas that was created from 603 high-resolution T1-weighted MRI and time-of-flight magnetic resonance angiography (MRA) images in healthy subjects (refer to Supplementary Fig. 1s) [22]. The atlas is a spatially resolved probability map of arterial locations in the brain.

For each subject, the MRA atlas was registered to the weighted average brain PET images using FSL (version 6.0.5.1). Here, brain PET images used for registration were extracted manually from the total body PET dataset. To select larger arteries (mean radius above ~2 mm)

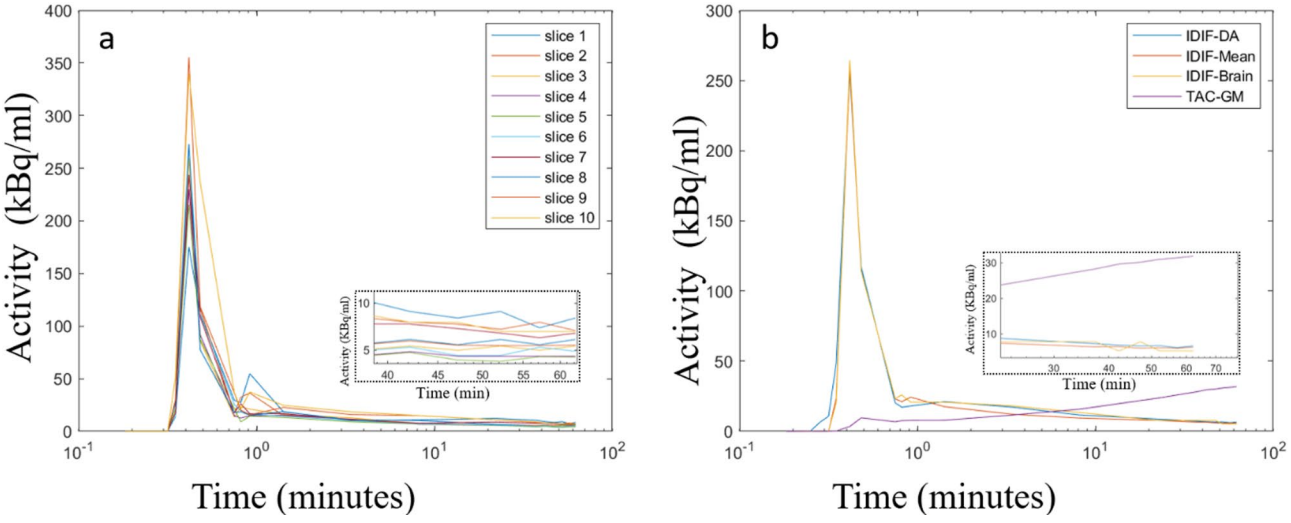


Fig. 2 Comparison between the estimated input function from the brain and the reference input function from the descending aorta (DA) for Subject 1 (S1). **a**. input functions from different brain slices. **b**. Comparison of AIF estimated from descending aorta (IDIF-DA, blue), mean atlas-based AIF estimate (IDIF-Mean, red), input function estimate corrected using variational inference (IDIF-Brain, yellow), and gray matter tissue time activity curve (TAC-GM, purple). The dashed boxes compare the tail of input function with log scale on the both axes

in the brain PET images, only voxels with a probability > 80% were selected. Reducing the probability threshold below 80% resulted in a lower estimated peak, most likely because of the peak effect in smaller arteries. Note, at a threshold of > 70% contribution arises mostly from the internal carotid artery and basilar artery segments outside the brain, which are large structures, reducing the likelihood of post-registration misalignment at the 80% threshold level. A minimum of 5 and a maximum of 10 axial slices were used for input function estimation in each subject with the choice of axial slices for input function extraction reflecting the correspondence between individual anatomy and the probabilistic atlas.

To extract input function curves from the brain PET images, highest voxel intensity in the selected vascular voxels in each axial slice was used at each time step to minimise the influence of peak activity, as depicted in Fig. 1A. The input function in each subject was obtained by averaging the curves across all slices, as illustrated in Fig. 2B (IDIF-Mean). Figure 2A illustrates in an example subject that IDIF-Mean is still influenced by spill over and spill in that vary between slices, necessitating correction using variational inference.

Variational inference for correcting IDIF-Mean

Variational inference is used in Bayesian analysis to approximate the posterior distributions for the latent variables [27]. We used this approach to correct for peak activity effects in the IDIF-Mean by formulating a probabilistic graphical model linking the IDIF-Mean (known as the prior or latent variable or z) and the average tissue time activity curve (TAC, observation or x) from the gray matter of the brain (see, Fig. 1-vi). The proposed graphical model aimed to correct the IDIF-Mean by learning the relationship between observed tissue time activity curve and the IDIF-Mean based on z. The problem of posterior estimation is formulated as the conditional probability density of z given x [28], denoted as P(z = IDIF-Mean | x = TAC):

$$\mathrm{P}\left(\mathrm{Z}\,|\,\mathrm{X}\right) = \frac{\mathrm{P}\left(\mathrm{z},\ \mathrm{x}\right)}{\mathrm{P}\ \left(\mathrm{x}\right),} \tag{1}$$

where $\mathrm{P}\ (\mathrm{z},\ \mathrm{x})$ is the joint probability density, and $\mathrm{P}\ (\mathrm{x})$, known as evidence, contains the marginal density of the observations such that P(x) = ∫ P(z|x) * P(z) dz. A closed form solution for P(x) does not exist, hence a surrogate posterior, q, is used as an approximation for the true posterior. The value of q is typically chosen from a simpler family of variational distributions (each distribution has its own set of variational parameters) and optimised to closely match the target posterior.

The true distribution of arterial input function is a complex problem to capture mathematically, but it is valid to assume it follows a sum of exponential distributions [29]. Hence, the surrogate posterior was formulated such that the latent variables (z) followed an exponential distribution with a mean ($1/\lambda\ z$) and the observed variables (x) followed a Gaussian distribution with a mean and standard deviation ($\mu_{\ x,}\sigma_{\ x}$) at each time step. The values of $\lambda\ z,\ \mu_{\ x,}\sigma_{\ x}$ were obtained from the IDIF-Mean (mean of the distribution) and the tissue time activity curve of gray matter (mean and deviation of the distribution) in the brain PET images, respectively. The surrogate posterior was then calculated using the evidence lower bound:

$$ELBO_q = E\ \{\log p\,(\mathrm{x}\mid \mathrm{z})\} - KL\,(q(\,z)|\,|\mathrm{p}\,(\mathrm{z}))\,, \quad (2)$$

where $E\ \{\log p\,(\mathrm{x}\mid \mathrm{z})\}$, is the expected log-likelihood of the observed TAC of gray matter and Kullback-Leibler divergence $KL\,(q\,(z)|\,|\mathrm{p}\,(\mathrm{z}))$, measures the closeness between the true ($\mathrm{p}\,(\mathrm{z})$) and surrogate ($q\,(z)$) posterior distributions. The objective is to identify parameter values ($\lambda z,\ \mu_{x,}\sigma_x$) that maximise the evidence lower bound (or minimize the Kullback-Leibler divergence, measuring the difference in information represented by two distributions) and result in an accurate approximation of the target posterior distribution. The TensorFlow probability package (Version 2.4) was used. In the supplementary material, we have explained the variational inference process in detail. The code has been provided in the supplementary material.

AIF validation

To evaluate the accuracy of IDIF-Brain, we compared their area under the curves (AUC) for whole curve, obtained using trapezoidal integration, with those of the reference input functions, IDIF-DA. In addition, we separately compared the AUC to the first peak of the curves (AUC_P) and the tail region of the curves (AUC_T). Here, peak and tail regions are based on the early 30×2 s and later 9×300 s frames, respectively.

Statistical analysis

The statistical significance of differences in mean AUC was evaluated using the paired *t*-test with $p < 0.05$ (two-tailed test) being used as the threshold for statistical significance. Normality assumptions were confirmed using the Shapiro-Wilks test [30] using a significance level of $p < 0.05$.

The reliability of IDIF-Brain for parameter estimation was assessed by calculating the Pearson correlation coefficient (r) between kinetic parameters (K_1, k_2, $k_{3,}$ K_i) estimated using IDIF-DA and IDIF-Brain [31]. Further, the statistical significance of the correlation was tested using Fisher's z-transformation to test the null hypothesis $r = 1$. Welch's unequal variance t-test was used to compare the estimated kinetic parameters with those previously reported in literature [23]. Single-factor ANOVA was performed to compare the kinetic parameters estimated using IDIF-DA, IDIF-Brain and IDIF-Mean. Homoscedasticity assumptions were confirmed using the Breusch-Pagan Test [32]. In addition, to assess the reliability and consistency intraclass correlation coefficient (ICC) are computed between K_i values.

Kinetic modelling and parameter Estimation

Kinetic parameters were estimated using the irreversible two tissue compartment model described by [33]. The rate constant representing the dephosphorylation of metabolically trapped FDG was assumed to be zero:

$$C_{Tn}\,(t) = (1 - V_b)\,.\left[\left(\frac{K_1\,k_2}{k_2 + k_3}.e^{-(k2+k3)t} + \frac{K_1\,k_3}{k_2 + k_3}\right) \otimes\ C'_b\,(t)\right] + C'_b\,(t)\,.V_b, \quad (3)$$

$C_{Tn}\,(t)$ is the tissue time activity curve obtained by convolving arterial input function, $C'_b\,(t)$, and tissue impulse response function at time t and V_b is the vascular blood volume. The micro kinetic parameters, K_1*(ml/cm*^{3}*/min)*, k_2 *(1/min)* and k_3 (1*/min*), are the rate constants for the unidirectional exchange of radiotracer between compartments (see Eq. 3). We used blood activity directly since K_1, k_2 and k_3 were compared across the two IDIFs. Non-linear least squares (NLS) fitting of Eq. (3) to the tissue time activity curve for each pixel was used to estimate K_1, k_2, and k_3. Using the built-in function lsqcurvefit available in MATLAB 2019b, optimization was performed employing the Levenberg-Marquardt (LM) algorithm. Initial values for fitting the parameters were set to 0.01/min. Lower bounds for all parameters (K_1, k_2, k_3 *and* V_b) were set to zero, while upper bounds were set to 1/min. Net radiotracer influx, K_i (ml/*cm*^{3}/min), was calculated according to $Ki = K_1 k_3 / (k_2 + k_3)$. The Patlak graphical method [34] was also used to estimate K_i for the purpose of validation.

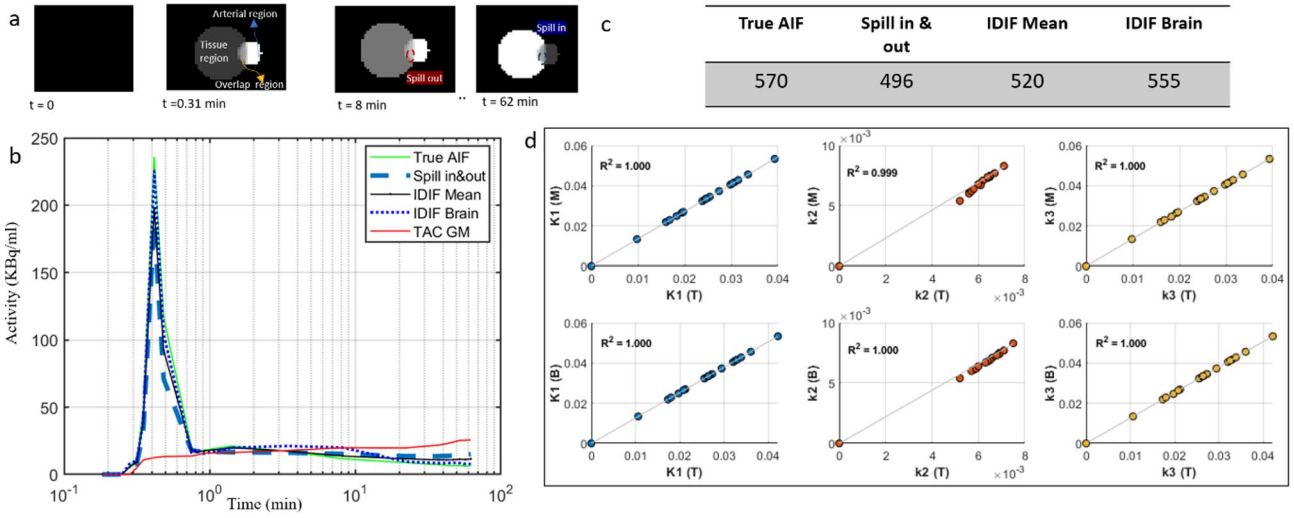
Simulation

We performed a simulation study to validate the application of variational inference. In Fig. 3A the simulation setup is shown, which we used to distinguish between voxels corresponding to an arterial region and those corresponding to a blood-free tissue region. An overlapping pixel region was introduced to simulate varying degrees of spill-in and spill-out (by the amount of pixel overlap) between the artery and tissue at their interfaces. The input function for the arterial region was modelled to represent the true AIF at time points that are kept similar to the acquired data.

Results

Simulation study

In Fig. 3B, the simulated True AIF was compared against four different scenarios: (i) the IDIF Mean, calculated following the procedures outlined in method Sect. 2.3; (ii) the tissue time activity curve (TAC GM), derived using the reference kinetic parameters and Eq. 3; (iii) the IDIF-brain, calculated using the IDIF-Mean and IDIF-Brain; and (iv) spill in and out due to the overlapping region at early and late timepoints. In Fig. 3c, the area under the curve comparison demonstrates that the IDIF estimated from the brain (AUC = 555) closely matches the true AIF (AUC = 570), In Fig. 3D, the resulting kinetic parameters $K_{1,}$ $k_{2,}$ and k_3 were compared between the applied



True AIF	Spill in & out	IDIF Mean	IDIF Brain
570	496	520	555

Fig. 3 **a**) simulation was conducted to differentiate voxels in arterial and blood-free tissue regions. An overlapping region simulating spill-in and spill-out between artery and tissue. **b**) compares the True AIF with the IDIF Mean, tissue time activity curve, IDIF brain (using variational inference). **c**) area under the curve estimates. **d**) shows a comparison of kinetic parameters between the True values (T) and those obtained from the IDIF Mean (M) and IDIF-Brain (**b**) for the simulations

simulation values (T) and those derived from the IDIF Mean (M) and IDIF Brain (B) for the simulated scenarios. The proposed method's effectiveness is highlighted by the medium to high correlation observed between kinetic parameters estimated using input function estimated from the brain and the simulated true input function.

AUC analysis

Figure 2 represents the qualitative comparison of the input function curves extracted from the brain regions and DA. In plot A, input functions from different axial brain slices are shown for Subject 1. Plot B compares IDIF-DA and IDIF-Brain and illustrates the concordance between the peaks of the two AIFs and the good correspondence between AIF tails.

Table 1 presents a quantitative comparison of the AUC values (in kBq.min.ml^{-1}) for descending aorta and brain AIF whole curves based on 10 subjects (refer to Supplementary Fig. 2s to compare input function for 10 subjects). The mean whole curve AUC for IDIF-Brain (721 ± 105; mean ± standard deviation, $n = 10$) was similar to that for IDIF-DA (741 ± 114; $p = 0.46$, paired t-test). Our proposed method resulted in an overestimation of the AUC in four subjects (underlined in Table 1) and an underestimation in six. The underestimation and overestimation are mainly due to differences in the peak values. This observation was validated by comparing the mean AUC_P for the IDIF-Brain (370 ± 180), which was lower than for IDIF-DA (422 ± 190, $p = 0.16$). However, when analysing the mean AUC_T, the IDIF-Brain (49 ± 11) was comparable to IDIF-DA (46 ± 8, $p = 0.11$). Subject 8 had the lowest absolute difference between IDIF-DA and IDIF-Brain (2 in kBq.min.ml^{-1}; 0.2%), while Subject 3 had the highest absolute difference (148 in kBq.min.ml^{-1}; 22.2%). The mean AUC percentage error between IDIF-Brain and IDIF-DA for the whole curve was 9%.

In contrast, the mean whole curve AUC for IDIF-Mean (643 ± 110) was lower than that for IDIF-DA (741 ± 114: $p = 0.02$) with an error percentage of 16%. However, no statistically significant difference was found between IDIF-Brain and IDIF-DA ($p = 0.46$).

Comparison based on kinetic modelling

Mean kinetic parameters in gray matter and white matter voxels estimated using NLS and Patlak methods for IDIF-DA, IDIF-Brain and IDIF-Mean are summarised in Table 2. For each AIF, the estimated kinetic parameters are in agreement with previously reported values with statistically insignificant differences ($p > 0.1$, Welch's t-test) [23, 35]. In addition, no statistically significant difference in estimated K_i values computed with kinetic parameters (K_1, k_2, k_3) using IDIF-DA, IDIF-Brain and IDIF-Mean was found on ANOVA ($F\ [2, 27] = 0.14$, $p = 0.86$).

In gray matter, when compared against the reference (IDIF-DA), the mean percentage errors for K_i obtained using NLS and Patlak methods with IDIF-Mean, were higher (by 8% and 7%, respectively) than estimates using IDIF-Brain ($p = 0.17$). Similarly, for white matter, the use of IDIF-Mean resulted in larger mean error percentages (by 3% and 10%) than the use of IDIF-Brain ($p = 0.14$). In addition, the mean Pearson correlation coefficient between IDIF-DA and IDIF-Brain for K_i was significantly higher ($r = 0.97$) than that between IDIF-DA and IDIF-Mean ($r = 0.89$, $p < 0.0001$). However, mean K_i values estimated using IDIF-DA did not differ significantly

Table 1 Comparison of the AUC for different AIF Estimation methods in the 10 subjects. AIF estimated from the descending aorta (IDIF-DA), the mean atlas-based AIF estimate (IDIF-Mean), and the AIF estimate corrected using variational inference (IDIF-Brain) are compared. AUC difference denotes the mean percentage absolute error against the reference (IDIF-DA). The underline in the table represents the overestimation of AUC using IDIF-Brain in comparison to the IDIF-DA

Subject	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Mean AUC	AUC difference (%)
IDIF-DA	655	799	666	878	678	807	737	922	508	767	741 ± 114	REF
IDIF-Brain	633	703	*814*	803	*710*	688	623	920	*536*	*784*	721 ± 105	9
IDIF-Mean	541	656	805	767	616	568	486	728	510	758	643 ± 110	16

from either IDIF-Brain ($p = 0.27$) or IDIF-Mean ($p = 0.12$). For the K_i values the intraclass correlation coefficient between IDIF-Brain and IDIF-DA was found to be higher (ICC = 0.79) than IDIF-DA and IDIF-Mean (ICC = 0.22). Notably, intraclass correlation coefficients between 0.6 and 0.8 indicate the estimated K_i values are fairly consistent and provide a good level (assumed between 0.75 and 0.90) of reliability [36].

Figure 4 depicts the regression between the individual kinetic parameters (K_1, k_2, k_3, and K_i) estimated using IDIF-Brain and IDIF-DA for a region of interest in the cortical region of Subject 3. The correlation coefficient (r) between estimates for all subjects are provided in Table 3. The mean correlation coefficients for K_1, k_2, k_3 and K_i ($p = 0.01$, 0.01, 0.03, 0.19, Fisher transformation) were found not to be significantly different from 1, indicating a strong correlation.

Voxel probability selection

The selection of voxel probabilities for the MRA atlas was identified as a critical factor. In Fig. 5, the IDIF-Mean was computed using two distinct voxel probability thresholds for MRA and compared against the IDIF-DA. In Fig. 5B and C, for the five different regions of interest in the brain, the kinetic parameters derived from the IDIF-Mean at 70% and 80% voxel probabilities were compared to those obtained from the IDIF-DA. R^2 values for the kinetic parameters were evaluated for both conditions. Figure 5D compares the input function after applying variational inference to IDIF-Mean at 70% and 80% voxel probabilities with the IDIF-DA. Figure 5E compares the individual kinetic parameters after applying the variational inference on IDIF-Mean at 70% and 80%. Figure 5F compares the parameters between corrected IDIF and the IDIF using ground truth from the heart. The IDIF peak calculated with the 80% voxel probability threshold (depicted in yellow) is more closely aligned with the IDIF-DA. Similarly, individual parameters are closely aligned with IDIF-Mean with a higher threshold (compare Fig. 5B and C). Nonetheless, after applying the variational inference for two different threshold voxel probabilities, kinetic parameters were found to be closely related to IDIF from the heart.

Discussion

The accurate estimation of the AIF for a subject is essential for robust radiotracer kinetic modelling. In this study, we propose a non-invasive input function estimation method only using brain PET images. Validation employed metrics such as the AUC and linear regression parameters. To evaluate the impact of using IDIF-Brain in parametric mapping, we compared kinetic parameters estimated using the [^{18}F]FDG two-tissue compartment

Table 2 Mean and standard deviation of the kinetic parameter, K_1 (ml/cm^3/min), k_2 (1/min) and k_3 (1/min), estimates for Gray matter (top) and white matter (bottom) were obtained using IDIF-DA (used as reference value, REF), IDIF-Brain and IDIF-Mean. The mean percentage absolute errors between reference and IDIF-Brain/Mean for Ki (ml/cm^3/min) obtained using NLS and Patlak are shown

Gray matter	**K_1**	**k_2**	**k_3**	**Ki-NLS**	**Ki_Patlak**	**% Error, Ki**	
Method						**NLS**	**Patlak**
IDIF-DA	0.160 ± 0.075	0.147 ± 0.147	0.066 ± 0.056	0.052 ± 0.020	0.067 ± 0.026	REF	
IDIF-Brain	0.160 ± 0.071	0.108 ± 0.107	0.038 ± 0.050	0.047 ± 0.022	0.062 ± 0.014	10	8
IDIF-Mean	0.170 ± 0.059	0.062 ± 0.072	0.035 ± 0.049	0.062 ± 0.023	0.078 ± 0.021	18	15
White matter	***K_1***	***k_2***	***k_3***	***Ki-NLS***	***Ki_Patlak***	**% Error, *Ki***	
Method						**NLS**	**Patlak**
IDIF-DA	0.064 ± 0.030	0.105 ± 0.113	0.025 ± 0.031	0.012 ± 0.006	0.016 ± 0.009	REF	
IDIF-Brain	0.072 ± 0.033	0.093 ± 0.085	0.017 ± 0.028	0.012 ± 0.008	0.015 ± 0.004	6	6
IDIF-Mean	0.079 ± 0.034	0.073 ± 0.066	0.013 ± 0.023	0.012 ± 0.008	0.018 ± 0.011	3	16

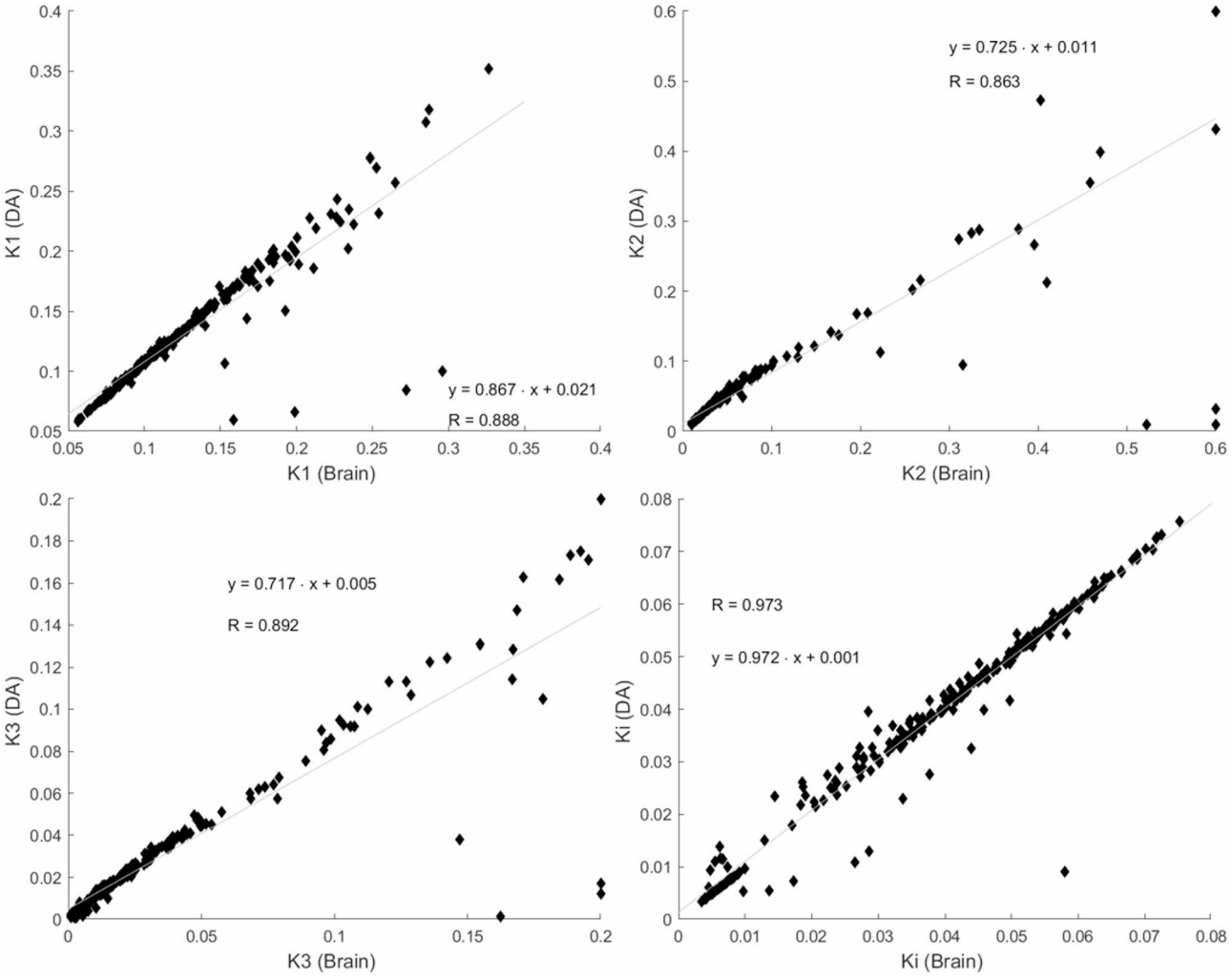


Fig. 4 Voxel-wise comparison between the kinetic parameters calculated using IDIF-Brain and IDIF-DA in the region of interest of cortical region of Gray matter in Subject 3. Linear regression indicates a strong linear relationship

model utilising IDIF-Brain and IDIF-DA, and found them to be highly correlated.

The proposed method can be used with SAFOV PET-CT scanners to estimate the input function from small (~ 5 mm diameter) brain vessels. Other methods require the individual's MRI scan to be co-registered with brain PET images to localise arterial voxels. In addition to requiring another set of images, registration errors arise when PET and MRI images are acquired during separate sessions [13]. The use of variational inference for peak

Table 3 Correlation coefficient (r) for individual kinetic parameters (K_1, k_2, k_3, and K_i) calculated using IDIF-Brain and IDIF-DA for the 10 subjects

Subject	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
K_1	0.85	0.96	0.88	0.98	0.98	0.91	0.93	0.93	0.96	0.79
k_2	0.82	0.97	0.86	0.99	0.97	0.92	0.86	0.94	0.97	0.76
k_3	0.88	0.95	0.89	0.98	0.99	0.89	0.93	0.96	0.98	0.89
Ki-NLS	0.95	0.99	0.97	0.99	0.99	0.93	0.98	0.91	0.97	0.97

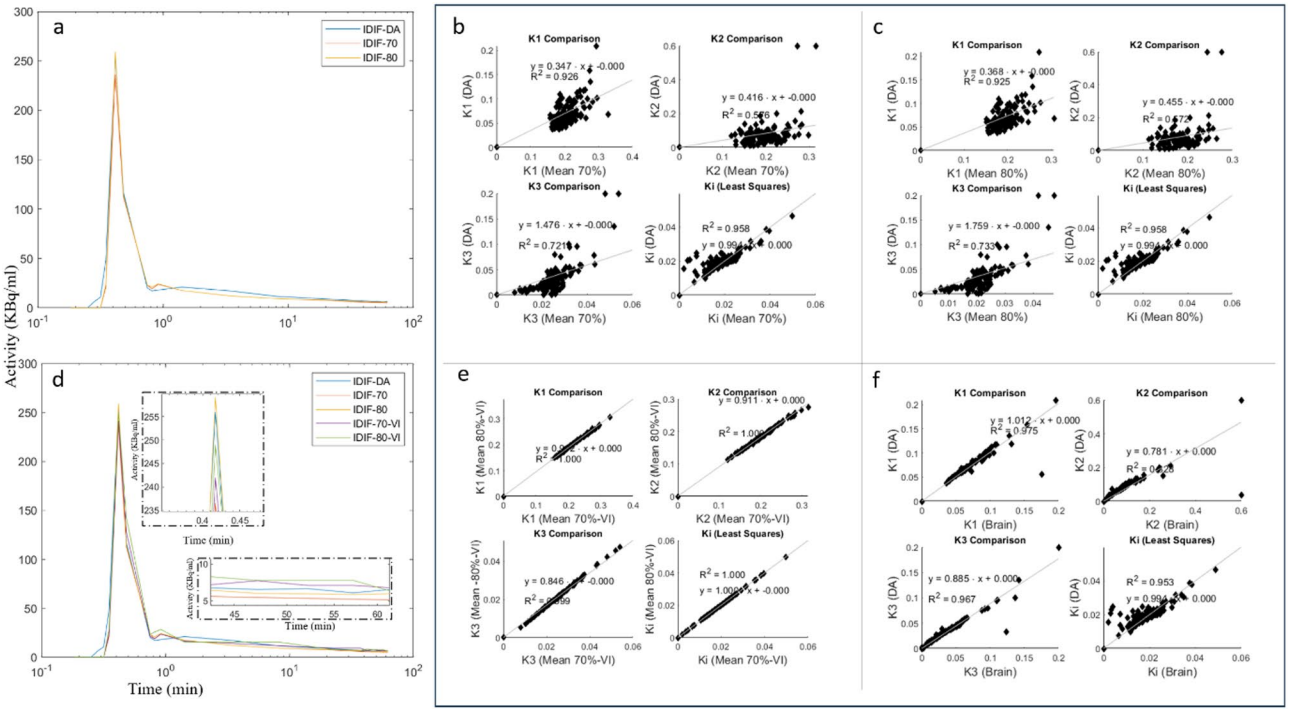


Fig. 5 In **a**, the IDIF-Mean was calculated using two different voxel probability thresholds for MRA and compared to the IDIF-DA. **b** and **c** compares kinetic parameters derived from the IDIF-Mean at 70% and 80% voxel probabilities against the IDIF-DA for five distinct regions of interest in the brain. R^2 values for the kinetic parameters were also evaluated for both conditions. In **d**, the input function after applying variational inference on the IDIF-DA at 70% and 80% voxel probabilities is compared with the IDIF-DA. **e** compares individual kinetic parameters after applying variational inference (VI) to the IDIF-DA at 70% and 80%. Lastly, **f** illustrate the comparison of kinetic parameters between the corrected IDIF and the IDIF based on ground truth from the heart

activity correction is advantageous compared to existing methods. It uses input function estimated from different brain slices as mean, and the tissue time activity of the gray matter as a prior. Previous studies have relied on the use of the point spread function, but did not consider its variation across the FOV [12]. Another method which does not use the point spread function involves the manual selection of regions of interest along the periphery of arterial brain vessels [13]. The method did not consider partial volume effects at the centre of the vessel, leading to underestimated input functions, mainly around the peak [13].

We have utilised a Bayesian deep learning approach incorporating information from the subject's PET images alone. AUC difference in Table 1 shows the notable improvement in peak activity correction achieved through variational inference. Specifically, the IDIF-Mean (without correction) exhibits a higher mean absolute percentage error (by 98 kBq.min.ml-1) compared to IDIF-Brain (with correction, by 20 kBq.min.ml^{-1}), primarily due to underestimation of peak values. Furthermore, the estimated kinetic parameters demonstrate a higher percentage error with IDIF-Mean relative to IDIF-DA, as detailed in Table 2. In addition, the use of gradient-based variational inference offers faster convergence compared to traditional Bayesian methods, such as the Markov Chain Monte Carlo technique. Another advantage compared to existing supervised learning techniques for input function estimation is that it does not require paired training data. It is widely acknowledged that supervised machine learning methods generally perform optimally within the distribution of data they are trained on, but when tested with data from different distributions, their performance deteriorates due to the lack of generalisation [20, 21]. Nevertheless, variational inference requires assumptions, including factorising the

variables to simplify computation and selecting a family of distributions which can be restrictive, potentially failing to capture the true complexity of the posterior and resulting in biased inferences. The limitation of the factorization assumption was addressed using the TensorFlow Probability package by using joint distributions to capture variable dependencies [37]. The second assumption was mitigated by employing informative priors incorporating domain knowledge, such as those derived from the MRA vascular atlas.

Nonetheless, variational inference is a probabilistic method whose accuracy in correcting for partial volume effects depends on the estimation of the posterior and prior. Unlike deterministic AI-based methods pre-trained on large datasets to handle varying amounts of partial volume effect, our approach is subject-specific and inherently influenced by the choice of prior and the evidence or posterior (TAC-GM). Although the proposed method has the potential to estimate the input function using the MRA atlas and corrects for the peak activity, it has limitations around correcting the partial volume effect at the tail. For instance, at late time points in Fig. 5 and in Supplementary Fig. 2s, subjects 4, 5, 7, and 9 exhibited IDIF Brain values slightly higher than IDIF-DA, indicating an incomplete correction for spill-in effects. We understand this variability to arise from subject-specific differences in TAC-GM. Additionally, selecting priors with lower voxel probabilities on the MRA atlas leads to inferior area under the curve and kinetic parameter estimates compared to the true AIF. This is likely due to priors with low probability failing to adequately capture the peak of the input function, as illustrated in Fig. 5, and by comparing the 70% and 80% thresholds for a subject.

In our method, the MRA angiogram atlas was registered to late PET frames. Precise registration is a challenging task as the patient positioning can differ across the dynamic PET scans, affecting the localisation of arterial voxels. By estimating IDIF-Mean from different brain slices with a high probability of containing arterial vasculature, the method tries to effectively reduce mis-registration errors, if not completely eliminate them, and also address the shortcomings of previous studies that heavily relied on identifying the carotid arteries [38]. However, artefacts such as scatter, motion and noise in early time frames can still introduce misalignment. Subsequently, this can affect the accuracy of input function peak estimation. We observed an overestimation of the AUC for four subjects (underlined in Table 1) and an underestimation for others due to differences between the peak values of IDIF-DA and IDIF-Brain. We attribute these variations to the inherently low number of counts in the short acquisition time frames (30×2 s) at early time points.

In variational inference, surrogate posterior estimation is employed because a closed-form solution is not available for calculating the true posterior (refer to Eq. 1). Estimating the surrogate posterior relies on the selection of specific parameters that capture the characteristics of the true posterior distribution (refer to Eq. 2). At each time point, the values of these parameters (λz, μ_x, and σ_x) were dependent on the selected window frame. A shorter framing window is typically employed to achieve higher precision in AIF estimation [12]. However, sudden and erratic increases in frame activity have been reported for low net true counts in the previous study based on Siemens mMR scanner (for shorter time windows less than 9×5 s, 3×10 s) [13]. A phantom study performed subsequently attributed these increases to inaccurate scatter correction [39]. It will be interesting to investigate if in the Vision Quadra total body scanner scatter-based increases in shorter frames can be found. An alternative approach may also be to use the sinogram or listmode data [1, 2]. Nonetheless, kinetic modelling ideally uses metabolite-corrected arterial plasma from whole blood activity. The whole blood-to-plasma ratio in FDG studies follows a mono-exponential function due to the lack of blood-borne metabolites and quick exchange between plasma and red-blood cells [40]. Hence, an arterial-plasma correction was not applied in our study. IDIF-Brain and IDIF-DA were measured in whole blood, which makes them directly comparable, whereas kinetic parameters based on whole blood activity can be corrected by a small factor to account for plasma AUC being about 3% higher than whole blood AUC [41].

The frame architecture used in this study was 30×2 s. Longer frames, typically 5 to 10 s, which are more commonly used in dynamic studies performed using SAFOV scanners with less sensitivity, would be expected to reduce the impact of noise on input function estimation, particularly around peak values. While the proposed methodology demonstrates promising outcomes, it requires further validation on SAFOV scanners when the carotid arteries are typically located near the edge of the FOV (or beyond). Additionally, in this work, we only had access to reconstructed total body PET therefore modifications of frame length were not feasible. Lastly the work could be extended by implementing automated registration techniques to enhance the accuracy and efficiency of PET data registration to the MRA atlas. An important advantage of using the MRA atlas in our method is that an additional MRI scan does not have to be acquired for the AIF estimation. Future studies may consider the individual specific MRA scan for IDIF-Brain estimation where available. Additionally, exploring the applicability of this approach to a broader range of PET radiotracers beyond FDG and validating against arterial blood sampling could broaden its utility and impact in clinical practice. Another limitation of our study is the absence of frame-by-frame motion correction algorithm.

However, to mitigate this we carefully examined dynamic PET data for each subject, focusing on the last 10 min to exclude data affected by significant movement. Future research should evaluate the impact of motion correction on kinetic parameter estimation using MRA to improve accuracy and robustness of results.

Conclusion

We outline a novel machine learning approach for input function estimation from dynamic brain PET images. The results highlight the feasibility of non-invasively estimating a subject-specific input function, providing a alternative to invasive blood sampling, or additional MRI acquisition. This approach can simplify parametric PET mapping and be adopted in available clinical installations, particularly with SAFOV PET scanners.

Abbreviations

SAFOV	Short axial field-of-view
LAFOV	Long axial field-of-view
AIF	Arterial input function
AUC	Area under the Curve
FOV	Field-of-view
PET	Positron emission tomography
DA	Descending Aorta
NLS	Non-linear least squares
[^{18}F]FDG	^{18}F-fluorodeoxyglucose

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13550-025-01253-3.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4: Figure 1s: Representation of the MRA brain images with arterial probability, input function extracted from brain slices and location for estimation of IDIF-DA in heart. Figure 2s: Comparisons between IDIF-DA, IDIF-Mean and IDIF-Brain for the 10 subjects.

Acknowledgements

Not applicable.

Author contributions

RV: Designed the study, Processed and analysed the data, generated all figures, Drafted the manuscript. HM: Performed data processing and contributed in designing of study. AR, HS, KS: Acquired the Total body PET data and involved in manuscript revision. AH, KB, VV, DR: Supervised the study, provided scientific direction and manuscript revision.

Funding

This research was conducted by the Australian Research Council Training Centre for Innovation in Biomedical Imaging Technology (project number IC170100035) and funded by the Australian Government through the Australian Research Council. No other potential conflicts of interest relevant to this article exist.

Data availability

The data that support the findings of this study are available from [23] but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. The code for variation inference implemented using Python, and extraction of the IDIF-Mean using Matlab, is available upon request.

Declarations

Ethics approval and consent to participate

The data that support the findings of this study was previously approved by the local institutional review board of Department of Nuclear Medicine, Inselspital, Bern University Hospital, University of Bern [23] (KEK 2019–02193), and written informed consent was obtained from all patients. The study was performed in accordance with the declaration of Helsinki.

Consent for publication

Written informed consent for publication was obtained.

Competing interests

The authors declare that they have no competing interests.

Received: 14 March 2024 / Accepted: 3 May 2025
Published online: 23 May 2025

References

1. Wang G, Qi J. Direct Estimation of kinetic parametric images for dynamic PET. Theranostics. 2013;3(10):802–15.
2. Gunn RN, Gunn SR, Cunningham VJ. Positron emission tomography compartmental models. J Cereb Blood Flow Metab. 2001;21(6):635–52.
3. Dimitrakopoulou-Strauss A, Pan L, Sachpekidis C. Kinetic modeling and parametric imaging with dynamic PET for oncological applications: general considerations, current clinical applications, and future perspectives. Eur J Nucl Med Mol Imaging. 2021;48:21–39.
4. Boutin H, Chauveau F, Thominiaux C, Grégoire M-C, James ML, Trebossen R, et al. 11 C-DPA-713: a novel peripheral benzodiazepine receptor PET ligand for in vivo imaging of neuroinflammation. J Nucl Med. 2007;48(4):573–81.
5. Bartels A, Willemsen A, Doorduin J, De Vries E, Dierckx R, Leenders K. [11 C]-PK11195 PET: quantification of neuroinflammation and a monitor of anti-inflammatory treatment in Parkinson's disease? Parkinsonism Relat Disord. 2010;16(1):57–9.
6. Hahn A, Schain M, Erlandsson M, Sjölin P, James GM, Strandberg OT, et al. Modeling strategies for quantification of in vivo 18F-AV-1451 binding in patients with Tau pathology. J Nucl Med. 2017;58(4):623–31.
7. Vashistha R, Moradi H, Hammond A, O'Brien K, Rominger A, Sari H, et al. ParaPET: non-invasive deep learning method for direct parametric brain PET reconstruction using histoimages. EJNMMI Res. 2024;14(1):10.
8. Dimitrakopoulou-Strauss A, Hoffmann M, Bergner R, Uppenkamp M, Eisenhut M, Pan L, et al. Prediction of short-term survival in patients with advanced nonsmall cell lung cancer following chemotherapy based on 2-deoxy-2-[F-18] fluoro-D-glucose-positron emission tomography: a feasibility study. Mol Imaging Biology. 2007;9:308–17.
9. Schaefferkoetter JD, Wang Z, Stephenson MC, Roy S, Conti M, Eriksson L, et al. Quantitative 18 F-fluorocholine positron emission tomography for prostate cancer: correlation between kinetic parameters and Gleason scoring. EJNMMI Res. 2017;7:1–13.
10. Takesh M. The potential benefit by application of kinetic analysis of PET in the clinical oncology. International Scholarly Research Notices. 2012.
11. Graham MM, Muzi M, Spence AM, O'Sullivan F, Lewellen TK, Link JM, et al. The FDG lumped constant in normal human brain. J Nucl Med. 2002;43(9):1157–66.
12. Khalighi MM, Deller TW, Fan AP, Gulaka PK, Shen B, Singh P, et al. Image-derived input function Estimation on a TOF-enabled PET/MR for cerebral blood flow mapping. J Cereb Blood Flow Metabolism. 2018;38(1):126–35.
13. Vestergaard MB, Calvo OP, Hansen AE, Rosenbaum S, Larsson HB, Henriksen OM, et al. Validation of kinetic modeling of [15O] H2O PET using an image derived input function on hybrid PET/MRI. NeuroImage. 2021;233:117950.
14. Zanotti-Fregonara P, Chen K, Liow J-S, Fujita M, Innis RB. Image-derived input function for brain PET studies: many challenges and few opportunities. J Cereb Blood Flow Metabolism. 2011;31(10):1986–98.
15. Sari H, Erlandsson K, Law I, Larsson HB, Ourselin S, Arridge S, 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 Cereb Blood Flow Metabolism. 2017;37(4):1398–409.
16. Kamenskiy AV, Pipinos II, Carson JS, MacTaggart JN, Baxter BT. Age and disease-related geometric and structural remodeling of the carotid artery. J Vasc Surg. 2015;62(6):1521–8.
17. Zanotti-Fregonara P, Hines CS, Zoghbi SS, Liow JS, Zhang Y, Pike VW, et al. Population-based input function and image-derived input function for [(1)(1)C](R)-rolipram PET imaging: methodology, validation and application to the study of major depressive disorder. NeuroImage. 2012;63(3):1532–41.
18. Ogden RT, Zanderigo F, Choy S, Mann JJ, Parsey RV. Simultaneous Estimation of input functions: an empirical study. J Cereb Blood Flow Metab. 2010;30(4):816–26.
19. Varnyú D, Szirmay-Kalos L, editors. Blood Input Function Estimation in Positron Emission Tomography with Deep Learning. 2021 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC); 2021:IEEE.
20. Kuttner S, Wickstrøm KK, Kalda G, Dorraji SE, Martin-Armas M, Oteiza A et al. Machine learning derived input-function in a dynamic 18F-FDG PET study of mice. Biomedical Phys Eng Express. 2020;6(1).
21. Kuttner S, Wickstrøm KK, Lubberink M, Tolf A, Burman J, Sundset R, et al. Cerebral blood flow measurements with 15O-water PET using a non-invasive machine-learning-derived arterial input function. J Cereb Blood Flow Metabolism. 2021;41(9):2229–41.
22. Mouches P, Forkert ND. A statistical atlas of cerebral arteries generated using multi-center MRA datasets from healthy subjects. Sci Data. 2019;6(1):29.
23. Sari H, Mingels C, Alberts I, Hu J, Buesser D, Shah V, et al. First results on kinetic modelling and parametric imaging of dynamic (18)F-FDG datasets from a long axial FOV PET scanner in oncological patients. Eur J Nucl Med Mol Imaging. 2022;49(6):1997–2009.
24. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G*power 3.1: tests for correlation and regression analyses. Behav Res Methods. 2009;41(4):1149–60.
25. Reed MB, Godbersen GM, Vraka C, Rausch I, Ponce de León M, Popper V, et al. Comparison of cardiac image-derived input functions for quantitative whole body [18F] FDG imaging with arterial blood sampling. Front Physiol. 2023;14:1074052.
26. Sari H, Mingels C, Alberts I, Hu J, Buesser D, Shah V et al. First results on kinetic modelling and parametric imaging of dynamic 18 F-FDG datasets from a long axial FOV PET scanner in oncological patients. Eur J Nucl Med Mol Imaging. 2022;1–13.
27. Jospin LV, Laga H, Boussaid F, Buntine W, Bennamoun M. Hands-on bayesian neural networks—A tutorial for deep learning users. IEEE Comput Intell Mag. 2022;17(2):29–48.
28. Neiswanger W. Probabilistic Graphical Models 2015. Available from: https://www.cs.cmu.edu/~epxing/Class/10708-17/slides/lecture13-VI.pdf
29. Feng D, Wong K-P, Wu C-M, Siu W-C. A technique for extracting physiological parameters and the required input function simultaneously from PET image measurements: theory and simulation study. IEEE Trans Inf Technol Biomed. 1997;1(4):243–54.
30. Mudholkar GS, Srivastava DK, Thomas Lin C. Some p-variate adaptations of the Shapiro-Wilk test of normality. Commun Statistics-Theory Methods. 1995;24(4):953–85.
31. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. Anesth Analgesia. 2018;126(5):1763–8.
32. Halunga AG, Orme CD, Yamagata T. A heteroskedasticity robust Breusch–Pagan test for contemporaneous correlation in dynamic panel data models. J Econ. 2017;198(2):209–30.
33. Gunn RN, Gunn SR, Turkheimer FE, Aston JA, Cunningham VJ. Positron emission tomography compartmental models: a basis pursuit strategy for kinetic modeling. J Cereb Blood Flow Metab. 2002;22(12):1425–39.
34. Patlak CS, Blasberg RG, Fenstermacher JD. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. J Cereb Blood Flow Metabolism. 1983;3(1):1–7.
35. Moradi H, Vashistha R, Ghosh S, O'Brien K, Hammond A, Rominger A, et al. Automated extraction of the arterial input function from brain images for parametric PET studies. EJNMMI Res. 2024;14(1):33.
36. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Conditioning Res. 2005;19(1):231–40.
37. Shih A, Ermon S. Probabilistic circuits for variational inference in discrete graphical models. Adv Neural Inf Process Syst. 2020;33:4635–46.
38. Fung EK, Carson RE. Cerebral blood flow with [15O] water PET studies using an image-derived input function and MR-defined carotid centerlines. Phys Med Biol. 2013;58(6):1903.
39. Ottilia Dalgaard M, Nøhr Ladefoged C, Law I, Littrup Andersen F, editors. Predicting the lower count limit for adequate scatter correction using dynamic data with low count frames on Siemens mMR. 7th Conference on PET/MR and SPECT/MR; 2018; La Biodola, Isola d'Elba.
40. Naganawa M, Gallezot J-D, Shah V, Mulnix T, Young C, Dias M, et al. Assessment of population-based input functions for Patlak imaging of whole body dynamic 18 F-FDG PET. EJNMMI Phys. 2020;7:1–15.
41. Dias AH, Smith AM, Shah V, Pigg D, Gormsen LC, Munk OL. Clinical validation of a population-based input function for 20-min dynamic whole-body 18F-FDG multiparametric PET imaging. EJNMMI Phys. 2022;9(1):60.

Publisher's note