



Original paper

Robust nonlinear parameter estimation in tracer kinetic analysis using infinity norm regularization and particle swarm optimization

Seung Kwan Kang^{a,b}, Seongho Seo^{c,d,*}, Chul-Hee Lee^{a,e}, Mi Jeong Kim^b, Su Jin Kim^f,
Jae Sung Lee^{a,b,g,*}

^a Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Republic of Korea

^b Department of Nuclear Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

^c Department of Neuroscience, Gachon University College of Medicine, Incheon, Republic of Korea

^d Department of Electronic Engineering, Pai Chai University, Daejeon, Republic of Korea

^e Department of Nuclear Medicine, Korea Institute of Radiological and Medical Sciences, Seoul, Republic of Korea

^f Department of Nuclear Medicine, Seoul National University Bundang Hospital, Seongnam, Republic of Korea

^g Institute of Radiation Medicine, Medical Research Center, Seoul National University, Seoul, Republic of Korea



ARTICLE INFO

Keywords:

Tracer kinetic analysis
Infinity-norm regularization
Particle swarm optimization
Non-convex optimization
Positron emission tomography

ABSTRACT

In positron emission tomography (PET) studies, the voxel-wise calculation of individual rate constants describing the tracer kinetics is quite challenging because of the nonlinear relationship between the rate constants and PET data and the high noise level in voxel data. Based on preliminary simulations using a standard two-tissue compartment model, we can hypothesize that it is possible to reduce errors in the rate constant estimates when constraining the overestimation of the larger of two exponents in the model equation. We thus propose a novel approach based on infinity-norm regularization for limiting this exponent. Owing to the non-smooth cost function of this regularization scheme, which prevents the use of conventional Jacobian-based optimization methods, we examined a proximal gradient algorithm and the particle swarm optimization (PSO) through a simulation study. Because it exploits multiple initial values, the PSO method shows much better convergence than the proximal gradient algorithm, which is susceptible to the initial values. In the implementation of PSO, the use of a Gamma distribution to govern random movements was shown to improve the convergence rate and stability compared to a uniform distribution. Consequently, Gamma-based PSO with regularization was shown to outperform all other methods tested, including the conventional basis function method and Levenberg–Marquardt algorithm, in terms of its statistical properties.

1. Introduction

Dynamic positron emission tomography (PET) scanning provides quantitative and accurate information regarding the amount of radiotracers moving in a living body with a sufficiently fine temporal resolution for visualizing their spatiotemporal distribution. In addition to visualization, a kinetic analysis of the dynamic PET data based on compartmental modeling allows us to quantify the rate constants of the radiotracer exchange between compartments and the physiological parameters associated with these rate constants [1–9]. Because compartment models bring about nonlinear least squares (NLS) problems [10–12] in estimating the rate constants, investigators have commonly used iterative algorithms employing Jacobians of the cost function, such as the Levenberg–Marquardt algorithm (LMA), to solve such

problems. However, one of the main limitations of an iterative NLS algorithm is that its solution may converge to the sub-optimal local minima if an improper initial value is selected. Another limitation is that often there are effectively an infinite number of solutions that yield nearly the same value of the cost function. Given a high noise level, these cannot be distinguished mathematically or statistically. In practice, the parameters are highly correlated, leading to a parameter space with long shallow valleys in which there is only minor change in the cost function. Owing to these limitations and the high noise levels in the dynamic PET voxels, iterative NLS algorithms usually produce noisy and biased parametric images in a voxel-wise full kinetic analysis. On the other hand, regularizing cost function provides additional criteria for selecting the solution and can reduce the noise at the potential expense of higher bias. Another widely used fitting method for PET

* Corresponding authors at: Department of Nuclear Medicine, Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea (Jae Sung Lee); Department of Electronic Engineering, Pai Chai University, 155-40, Baejae-ro, Seo-gu, Daejeon 35345, Republic of Korea (Seongho Seo).

E-mail addresses: seos@pcu.ac.kr (S. Seo), jaes@snu.ac.kr (J.S. Lee).

<https://doi.org/10.1016/j.ejmp.2020.03.013>

Received 12 November 2019; Received in revised form 6 March 2020; Accepted 8 March 2020

Available online 20 March 2020

1120-1797/ © 2020 Associazione Italiana di Fisica Medica. Published by Elsevier Ltd. All rights reserved.

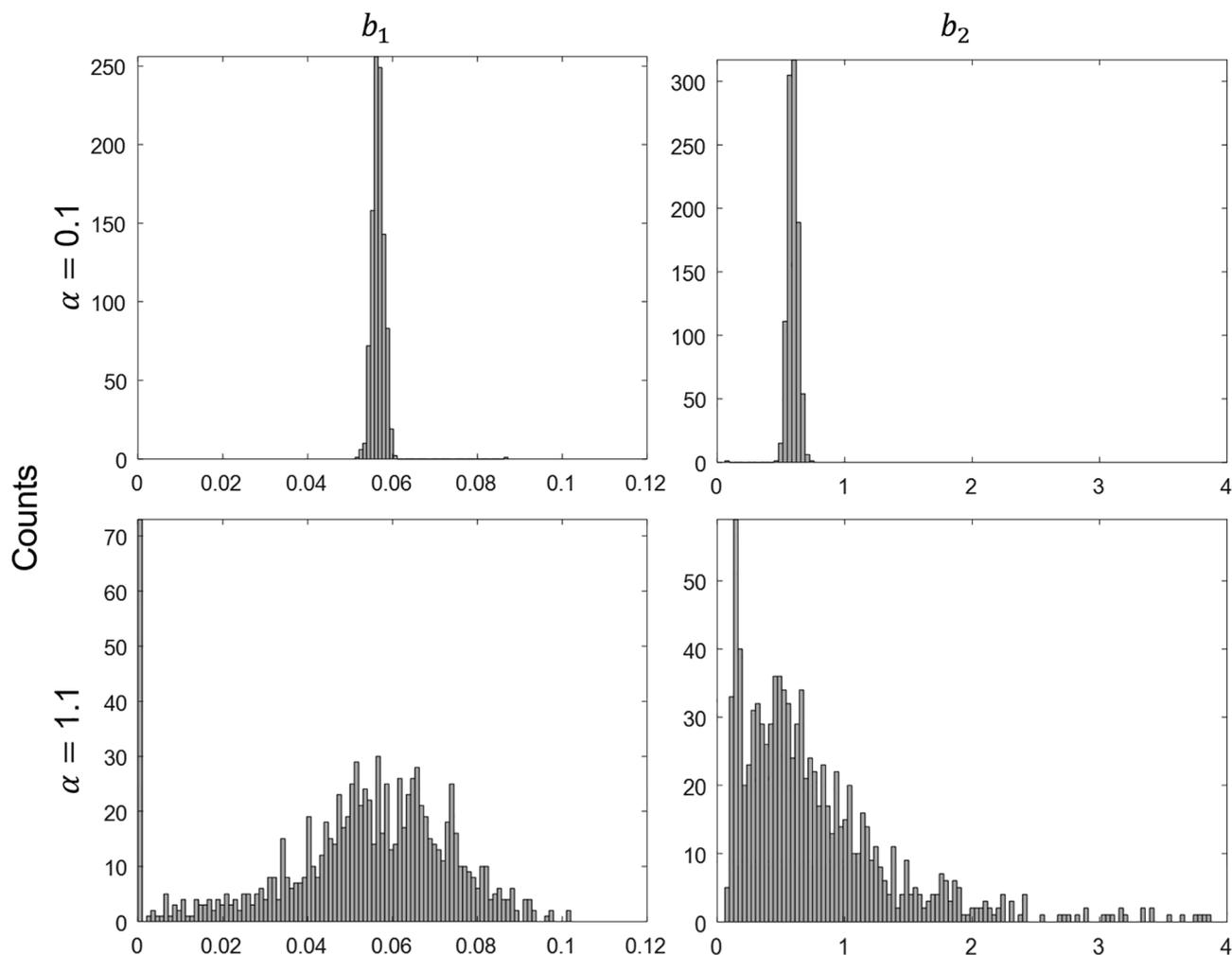


Fig. 1. Histograms of exponent parameters b_1 and b_2 in equation (4) estimated using LMA with two different noise levels: $\alpha = 0.1$ and $\alpha = 1.1$ (refer to Fig. 2 and section *Simulations* for the noise level).

kinetics is the basis function method (BF) [13–15]. Based on a pre-defined set of basis-function exponents, the BF method estimates the corresponding coefficients using linear least square method. Then, the best parameter set is determined by comparing the weighted least square values of all combinations of exponents and coefficients. The BF method's performance depends heavily on the initialization of the exponents set, especially the set's size and range.

In a PET kinetic analysis, a two-tissue compartment model with four kinetic parameters is most commonly used for describing the kinetics of [^{18}F]fluorodeoxyglucose ([^{18}F]FDG) and various other radiotracers with reversible uptake or binding [16,17]. In this study, from a careful preliminary analysis using the aforementioned model, we found that the estimate of the larger of two exponents in the model equation is exceptionally more sensitive to noise than that of the smaller exponent, yielding a positively skewed distribution under highly noisy circumstances, as shown in Fig. 1. To mitigate this severe bias problem, this study adopted an l-infinity (l_∞)-norm ($\|\cdot\|_\infty$) regularization strategy based on the assumption that errors in the rate constant estimates can be reduced when constraining the overestimation of the larger exponent.

Because the introduction of l_∞ norm makes the regularized cost function non-smooth, conventional Jacobian-based algorithms are no longer relevant to the suggested optimization problem. Therefore, we examined two representative algorithms that do not require smoothness of the cost function: a proximal gradient algorithm [18] and particle swarm optimization (PSO) [19–23].

There have been limited number of studies that applied the PSO to the PET tracer kinetics and biochemical process analysis [24,25]. The PSO is an iterative global optimization algorithm based on a random search of the solution space at each iteration. The previous researches handled the large search space using this random search and predefined moving rules. In this study, we also propose the use of Gamma distribution and partial linearization in the PSO for achieving better convergence in kinetic parameter estimation. Detailed description will be provided in Section 2.5.

There also has been studies using the randomness in the Bayesian framework [26–28]. The posterior distribution was calculated from the prior information, and kinetic parameters were inferred from the posterior. On the other hand, PSO is more heuristic algorithm focusing on the optimization problem. It facilitates the optimization process by computing cost function directly and comparing between randomized candidates rather than finding complex gradient of cost function. Accordingly, incorporating various regularization, such as infinity norm, is possible into the original problem.

In the following sections, we describe a novel l_∞ -norm regularization scheme for the two-tissue compartment model in dynamic PET studies and the optimization algorithms we examined for the regularization problem. The results of a simulation study conducted to evaluate the performance of these approaches are then discussed. We also describe how we optimized the l_∞ -norm regularization.

2. Methods

2.1. PET compartment modeling

In the two-tissue compartment model with four kinetic parameters, the compartments represent the radiotracer concentration in arterial plasma ($C_a(t)$), a free or nonspecifically bound radiotracer ($C_f(t)$), and a specifically bound radiotracer ($C_b(t)$) at time t , respectively. The radiotracer exchanges between compartments are described using four rate constants, K_1 ($\text{ml g}^{-1} \text{min}^{-1}$), k_2 , k_3 , and k_4 (min^{-1}), as follows:

$$\frac{dC_f(t)}{dt} = K_1 C_a(t) - (k_2 + k_3)C_f(t) + k_4 C_b(t), \quad (1)$$

$$\frac{dC_b(t)}{dt} = k_3 C_f(t) - k_4 C_b(t). \quad (2)$$

The time-activity curve (TAC) in a tissue ($C_T(t)$) equals the summation of $C_f(t)$ and $C_b(t)$ if we pre-correct the blood volume component in the tissue to simplify the problem:

$$C_T(t) = C_f(t) + C_b(t). \quad (3)$$

By solving Eqs. (1) and (2), the analytic solution of $C_T(t)$ is therefore derived as

$$C_T(t) = (a_1 e^{-b_1 t} + a_2 e^{-b_2 t}) \otimes C_a(t), \quad (4)$$

where

$$a_1 = \frac{K_1}{(b_2 - b_1)} (k_3 + k_4 - b_1),$$

$$a_2 = \frac{K_1}{(b_2 - b_1)} (b_2 - k_3 - k_4),$$

$$b_{1,2} = \frac{(k_2 + k_3 + k_4) \mp \sqrt{(k_2 + k_3 + k_4)^2 - 4k_2 k_4}}{2}.$$

Our goal with compartmental modeling is to estimate the four rate constants by fitting Eq. (4) to the noisy measurements of $C_T(t)$ obtained from dynamic PET frames. This is equivalent to finding an optimal solution ($\hat{\theta}$) to the following optimization problem regarding $\theta = [a_1 \ b_1 \ a_2 \ b_2]$, and subsequently retrieving the estimates of the rate constants from $\hat{\theta}$ [29], i.e.,

$$\hat{\theta} = \underset{\theta}{\text{argmin}} \|\tilde{C}_T - C_T(\theta)\|_{\mathbf{W}}^2 \quad (5)$$

where C_T is a vector of the noisy measurement of tissue TAC at PET frames; $C_T(\theta)$ is the TAC constructed using θ and Eq. (4); and $\|\cdot\|_{\mathbf{W}}^2$ is the weighted sum of squares for a given diagonal matrix of weights \mathbf{W} . The weights usually consider the difference in variance of the measurement errors; for the weights, we used the frame durations divided by the decay factor to concern the data statistics (Eq. (24)). Because solution (4) is nonlinear in b_1 and b_2 , problem (5) is an NLS problem and should thus be solved using an iterative algorithm such as LMA.

2.2. Partial linearization

For fixed b_1 and b_2 , the analytic solution to $C_T(t)$ in (4) can be regarded as a linear combination of two curves, $e^{-b_1 t} \otimes C_a = C_1(t)$ and $e^{-b_2 t} \otimes C_a = C_2(t)$. Thus, assuming that b_1 and b_2 are already known or estimated, the estimation problem in (5) can be simplified as follows:

$$\min \|C_T - \mathbf{A}\phi_a\|_{\mathbf{W}}^2, \quad (6)$$

where $\mathbf{A} = e^{-\phi_b t} \otimes C_a$ ($\phi_b = [b_1 \ b_2]^T$) and $\phi_a = [a_1 \ a_2]^T$. Its weighted least squares solution can then be obtained as

$$\hat{\phi}_a = [\hat{a}_1 \ \hat{a}_2]^T = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{W} \mathbf{A}^T C_T. \quad (7)$$

Finally, we can obtain $\hat{\theta}$ by applying iterative algorithms that update $\hat{\phi}_b = [\hat{b}_1 \ \hat{b}_2]^T$ using nonlinear estimation methods and subsequently

compute $\hat{\phi}_a$ using (7) for the updated $\hat{\phi}_b$ during every iteration.

2.3. Infinity norm regularization

Although Eq. (5) is a simple NLS curve-fitting problem, it usually yields biased estimation results from the highly noisy PET TACs. Between the estimates of the two nonlinear parameters b_1 and b_2 ($b_1 < b_2$) in Eq. (4), the larger one (b_2) is likely to be more sensitive to noise in $C_T(t)$ than the smaller one. This is because the larger exponent in Eq. (4) is associated more with the early fast rising part of $C_T(t)$ rather than the late relatively slow clearance part; indeed, we observed a larger overestimation of b_2 at a high noise level, as shown in Fig. 1. The figure shows the histograms of b_1 and b_2 estimated using the LMA under two different noise levels of a tissue TAC (Section 2.6 Simulations to see how the noisy TACs were generated). At a low noise level ($\alpha = 0.1$), b_1 and b_2 show similar symmetric Gaussian distributions. At a high noise level ($\alpha = 1.1$), b_1 still shows a Gaussian distribution except for the peak at around zero. However, the distribution of b_2 is skewed and has a long tail toward the positive direction.

Therefore, to alleviate such positive bias of b_2 , which is observed only for the larger exponent, we can consider the regularization of the exponent parameters using a special form of infinity norm, $\|\mathbf{M}\theta\|_{\infty}$, ($\mathbf{M} = [0e_2 0e_4]$), which selects the element with the absolute maximum between two exponents by definition. Thus, the cost function can be modified as follows:

$$\|C_T - C_T(\theta)\|_{\mathbf{W}}^2 + \gamma \|\mathbf{M}\theta\|_{\infty}, \quad (8)$$

where γ is a regularization parameter controlling the influence of $\|\mathbf{M}\theta\|_{\infty}$. Because $\|\mathbf{M}\theta\|_{\infty}$ is non-differentiable, conventional Jacobian-based optimization algorithms such as LMA cannot properly handle it. In Sections 2.4 and 2.5, we describe how we utilized two representative algorithms for a non-smooth function minimization: proximal gradient and PSO.

2.4. Optimization algorithm 1: Proximal gradient

The proximal gradient algorithm solves the following unconstrained optimization problem:

$$\text{ming}(\theta) + h(\theta)$$

where g is differentiable and h is closed, convex, and possibly non-smooth. In our problem, functions g and h are given as

$$g(\theta) = \|C_T - C_T(\theta)\|_{\mathbf{W}}^2 \text{ and } h(\theta) = \gamma \|\mathbf{M}\theta\|_{\infty}. \quad (9)$$

The update equations of the proximal gradient algorithm is as follows:

$$\theta^{(k)} = \mathbf{prox}_{t_k h}(\theta^{(k-1)} - t_k \nabla g(\theta^{(k-1)})) \quad (10)$$

where k is the iteration and t_k is the step size at k . The parameters are updated using the gradient descent for function g , followed by a proximal mapping for function h . The proximal operator \mathbf{prox} is defined using the following minimizer:

$$\mathbf{prox}_{th}(\theta) = \underset{u}{\text{argmin}} h(u) + \frac{1}{2t} \|u - \theta\|_2^2 \quad (11)$$

where t is a constant. Through a Moreau decomposition [18], we have

$$\mathbf{prox}_h(\theta) = \mathbf{M}\theta - \mathbf{prox}_{h^*}(\mathbf{M}\theta) \quad (12)$$

The dual norm of $\|\cdot\|_{\infty}$ (h) is given through the indication function on l_1 ball, and the proximal mapping of the indication function is simply given using a projection to a convex set. Therefore, the final proximal mapping of our function h is

$$\mathbf{prox}_{th}(\theta) = \mathbf{M}\theta - t\gamma \mathbf{proj}_{B_1}(\mathbf{M}\theta/t\gamma) \quad (13)$$

where $\mathbf{proj}_{B_1}(\cdot)$ describes the projection to l_1 ball. We used an accelerated and descent proximal gradient algorithm shown in [18].

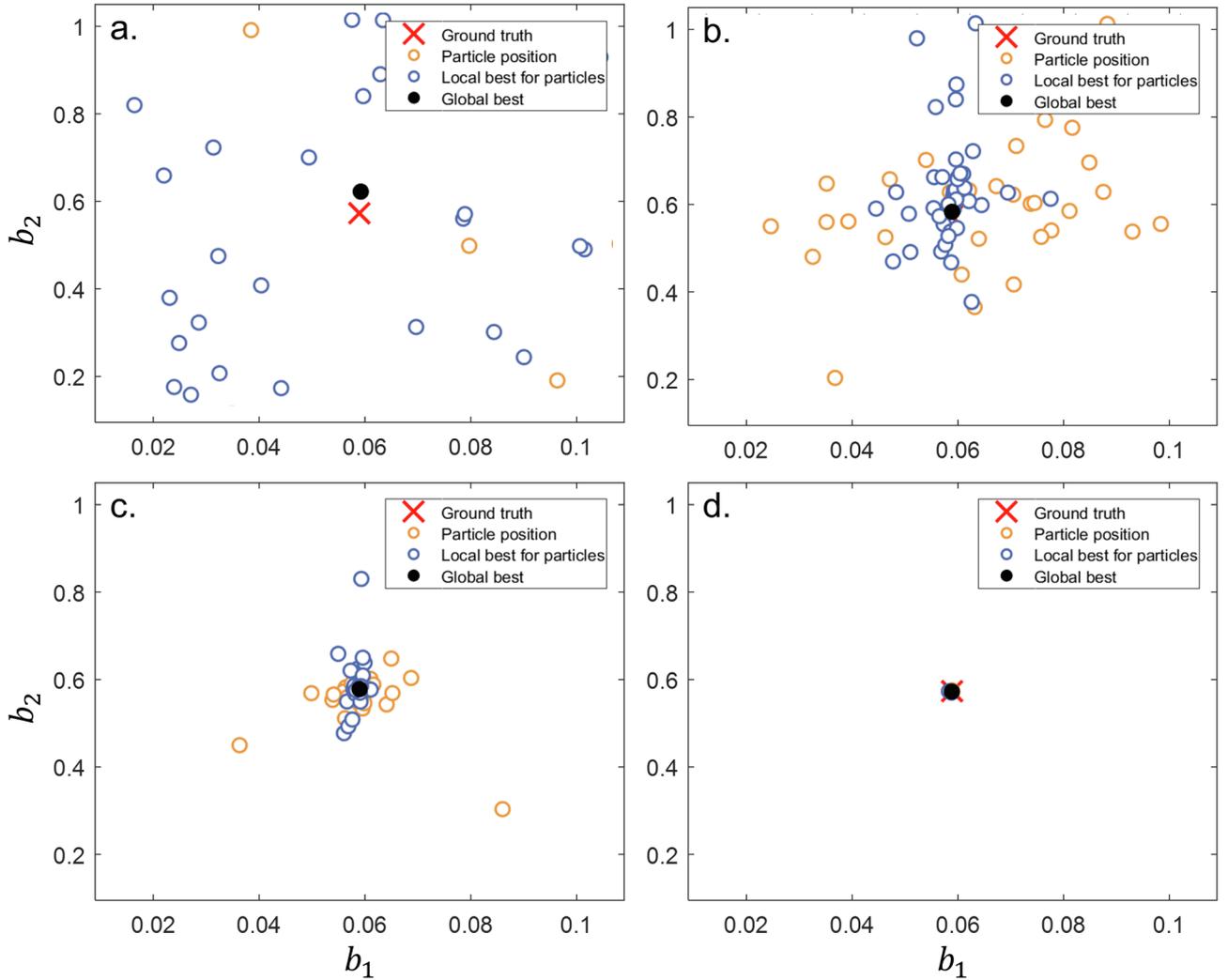


Fig. 2. Movements of particles in PSO by the iterations, which are represented in parametric space. (a) Initial values. (b) 10th iteration. (c) 20th iteration. (d) 40th iteration.

$$\omega^{(k)} = \text{prox}_{t_k h}(\theta^{(k-1)} - t_k \nabla g(\theta^{(k-1)})) \quad (14)$$

$$\kappa^{(k)} = \begin{cases} \omega^{(k)} f(\omega^{(k)} \leq f(\kappa^{(k-1)})) \\ \kappa^{(k-1)} \text{ otherwise} \end{cases} \quad (15)$$

$$\tau^{(k)} = \kappa^{(k-1)} + \frac{1}{\rho_k} (\omega^{(k)} - \kappa^{(k-1)}) \quad (16)$$

$$\theta^{(k)} = (1 - \rho_{k+1}) \kappa^{(k)} + \rho_{k+1} \tau^{(k)} \quad (17)$$

$$\rho_k = \frac{2}{k+1} \quad (18)$$

The step size, t_k , was selected as 10^{-4} .

2.5. Optimization algorithm 2: particle swarm optimization (PSO)

2.5.1. Standard PSO

Standard PSO minimizes the cost function through the random movement of particles (candidate solutions) (Fig. 2; for further details, see Appendix A1). After the initial guess of the particle positions, their velocities (\mathbf{v}) and positions (\mathbf{x}) in the next iterations are determined using the following update equations:

$$\mathbf{v}_i(m+1) = w\mathbf{v}_i(m) + c_1\varphi_1(\mathbf{p}_i^L(m) - \mathbf{x}_i(m)) + c_2\varphi_2(\mathbf{p}^G(m) - \mathbf{x}_i(m)), \quad (19)$$

$$\mathbf{x}_i(m+1) = \mathbf{x}_i(m) + \mathbf{v}_i(m+1), \quad (20)$$

where the constant w is the inertia weight affecting the convergence of particles within the range $[0, 1]$; φ_1 and φ_2 are independent random numbers uniformly distributed between $[0, 1]$ for regulating the randomness of particle movement; and the two constants c_1 and c_2 are the acceleration coefficients that control the broadness of particle movement. At each iteration, i th particle's best position \mathbf{p}_i^L is also updated if the i th particle's new position $\mathbf{x}_i(m+1)$ yields a lower cost function (f) value than its current best position, i.e.,

$$\mathbf{p}_i^L(m+1) = \begin{cases} \mathbf{x}_i(m+1), & \text{iff } (f(\mathbf{x}_i(m+1)) \leq f(\mathbf{p}_i^L(m))), \\ \mathbf{p}_i^L(m), & \text{otherwise.} \end{cases} \quad (21)$$

Among the updated $\mathbf{p}_i^L(m+1)$, the best one with the lowest cost function is selected as the population's best position, i.e.,

$$\mathbf{p}^G(m+1) = \text{argmin}_{\mathbf{p}_i^L} f(\mathbf{p}_i^L(m+1)). \quad (22)$$

If a stopping criterion is satisfied, the PSO algorithm finally returns \mathbf{p}^G as an optimal solution to the problem.

2.5.2. Gamma distribution for particle movement

The PSO relies on several control parameters, i.e., w , c_1 , and c_2 , which heavily affect the performance of the PSO and whose selection is radically heuristic. Therefore, various efforts have been made to

Table 1
Simulated rate constants for digital phantom.

	K_1	k_2	k_3	k_4
Region 1	0.16	0.4	0.15	0.08
Region 2	0.08	0.2	0.09	0.1
Region 3	0.08	0.2	0.225	0.09
Region 4	0.4	0.7	0.3	0.2
Region 5	0.52	1.2	0.3	0.05
	a_1	b_1	a_2	b_2
Region 1	0.0538	0.0557	0.1062	0.5743
Region 2	0.0385	0.0607	0.0415	0.3293
Region 3	0.0505	0.0377	0.0295	0.4773
Region 4	0.1574	0.1310	0.2426	1.0690
Region 5	0.1097	0.0397	0.4103	1.5103

achieve a good convergence rate and stability of the algorithm by investigating a strategy for selecting the control parameters [30–33], or the use of a non-uniform distribution for φ_1 and φ_2 to change the randomness characteristics [34,35].

In this study, we adopted the control parameters optimized in a previous study on the PSO in a multidimensional complex space. A total of $N = 20$ particles, $w = 0.7298$, $c_1 = 1.825$, and $c_2 = 1.168$ were used [31]. In addition, we employed a Gamma distribution for random movements of particles to assure the positivity of the estimated parameters and long tail of the distribution. The Gamma distribution will enable a faster search of the minimum position than a uniform distribution because particles sometimes move farther while preserving their random movement characteristics (for further details, see Appendix A2).

By integrating (19) and (20), we can express the position of each particle at the next iteration $\mathbf{x}_i(m+1)$ in terms of only the position information as follows:

$$\begin{aligned} \mathbf{x}_i(m+1) = & \mathbf{x}_i(m) + w(\mathbf{x}_i(m) - \mathbf{x}_i(m-1)) + c_1\varphi_1(\mathbf{p}_i^L(m) - \mathbf{x}_i(m)) \\ & + c_2\varphi_2(\mathbf{p}^G(m) - \mathbf{x}_i(m)), \end{aligned} \quad (23)$$

which indicates that the next position of each particle is influenced by the weighted sum of three different distances from the current position ($\mathbf{x}_i(m)$) to the previous, local best, and global best positions ($\mathbf{x}_i(m-1)$, $\mathbf{p}_i^L(m)$, and $\mathbf{p}^G(m)$), respectively.

2.6. Simulations

To evaluate the performance levels of the algorithms, a total of 1000 realizations of a single TAC containing noise were generated using a real arterial input function obtained from a [^{11}C]raclopride human PET study. As the rate constants, 0.17, 0.42, 0.15, and 0.08 were selected for K_1 ($\text{mlg}^{-1}\text{min}^{-1}$), k_2 , k_3 , and k_4 (min^{-1}), respectively. To explore the effects of the noise levels, noises with various levels were added to noiseless TACs using the following noise model [36]:

$$\text{Noise } N(0, 1) \times \alpha \sqrt{\frac{C_T(t)e^{\lambda t_i}}{\Delta t_i(\text{second})}}, \quad (24)$$

where the value of α represents the noise level, which is set at 0.1, 0.3, 0.5, 0.7, 0.9, 1.1, or 1.3; t_i and Δt_i are the mid-point and the scan duration of i th frame in seconds; and λ is the decay constant of the radioisotope used. We assumed C-11 in this simulation while other radioisotopes could be considered, such as F-18 in our real animal data used for further evaluation (Section 2.7). In the case of F-18, the simulated noise could be less than the C-11.

Using this set of simulation data, we first examined the non-smooth function minimization algorithms we described in Sections 2.4 and 2.5 in terms of convergence rate and stability; for the PSO, the new formulation based on Gamma distribution for random particle movements were evaluated. Then, the statistical properties of the most promising one, among the algorithms considered, were compared against the most

widely used nonlinear least square algorithm, LMA as well as BF in terms of the bias and coefficient of variance (CV) (see Appendix A3). For the sake of fairness during the initial-value selection, we repeated the LMA for all initial PSO particles, which were randomly selected to be between 0 and 1.5, and chose the best solution among the results of repetition. Finally, we explored the effect of choosing the regularization parameter on the PSO performance. For the BF method, we used 100 candidates logarithmically spaced within [0.001, 0.1] for b_1 and [0.1, 10] for b_2 to construct the basis functions.

To test whether the proposed methods can achieve the same performance irrespective of the simulation settings, we have conducted additional simulation study using a mathematical phantom with five different regions. We assigned different rate constants to each region in the phantom as summarized in Table 1. The dynamic phantom data was projected to produce sinograms to we added noise. We then reconstructed the noisy sinogram using maximum-likelihood expectation–maximization algorithm with 100 iterations. Then, the rate constants were estimated pixel-by-pixel, and normalized root mean squared error (NRMSE) was calculated after smoothing the estimated and ground truth images using 3-mm Gaussian filters.

2.7. Application to real data

We retrospectively analyzed the dynamic [^{18}F]FDG PET data obtained from two BALB/c nude mice (male, 6 weeks old) in our previous investigation [37]. All the animal studies were approved by the Institutional Animal Care and Use Committee (IACUC) of Seoul National University Hospital. Two different tumors cells, MDA-MB-231 (five million cells) and HepG2 (ten million), were implanted subcutaneously into the left front thigh of each mouse. The mice were fasted for at least six hours and 16.6 MBq of [^{18}F]FDG was injected into the tail vein. Dynamic PET images were obtained for two hours (4×3 s, 8×6 s, 8×30 s, 1×300 s, 11×600 s) using an animal PET/CT scanner (eXplore Vista CT; GE Healthcare, Waukesha, USA). Arterial input function was extracted from the left ventricle by manually drawing volume of interest (VOI). BF, LMA, Unregularized PSO, and Regularized PSO ($\gamma = 10^{-2.75}$) were applied to generate parametric images of individual kinetic parameters in two-tissue compartment model. Initial values for LMA estimation and initial mean for PSO analysis were determined by applying VOI-based analysis to the whole tumor regions (supplementary Fig. 1). To consider the intravascular activity in tumors, we modified each algorithm by introducing the blood volume fraction, v_a , so that $(1 - v_a)C_T(t) + v_a C_a(t)$ was fitted to the noisy measurement of [^{18}F]FDG tissue TAC $C_T(t)$ using the image-derived input function for $C_a(t)$ [38]. All the estimation algorithms were implemented in the MATLAB 2018b (<http://www.mathworks.com>) with the Intel Core i7-7700k processor.

3. Results

3.1. Comparison of non-smooth function minimization algorithms

Fig. 3(a) shows the shape of cost function (8) in the parameter space for b_1 and b_2 and trajectories of parameter estimates using proximal gradient and Gamma-distribution-based PSO algorithms in a single simulation. Noise level (α) and regularization parameter (γ) were 0.3 and $10^{-2.75}$, respectively. The maximum steps of the proposed PSO and proximal gradient were 120 and 30,000, respectively. The proximal gradient algorithm was extremely sensitive to the initial value and inner parameters (e.g., step size). The proximal gradient was unable to reach the ground truth even with the good initial parameters ($[0.1; 0.1; 0.05; 0.05]^T$). Although the initial values of the proposed PSO were randomly selected, it searched for the solution much better than the proximal operator. This is likely because our problem is highly non-convex, particularly under highly noisy circumstances.

The proposed PSO with a Gamma distribution for random particle

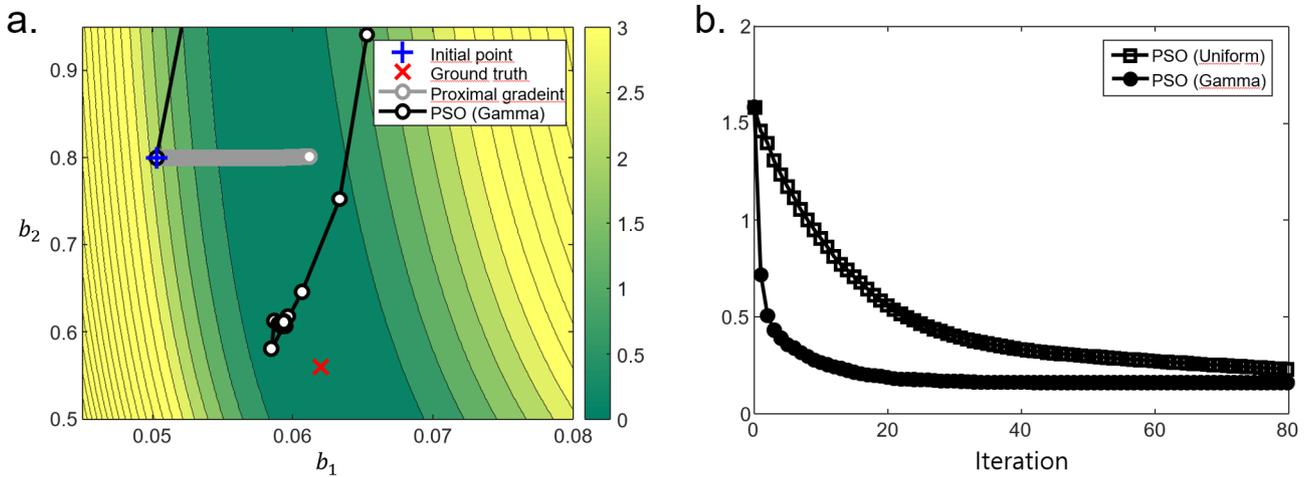


Fig. 3. Cost function evaluations of PSO and proximal gradient: (a) Trajectories of estimates using proximal gradient and Gamma-based PSO with l_∞ -norm regularization and (b) Convergence speed of PSO algorithms.

Table 2

Comparison of the standard PSO with a uniform distribution for random particle movement, and the proposed PSO with a Gamma distribution (mean and standard deviation of the residual value of the cost function).

Noise level	PSO (Uniform)	PSO (Gamma)
0.1	0.165 ± 0.050	0.099 ± 0.015
0.3	0.312 ± 0.500	0.278 ± 0.043
0.5	0.486 ± 0.073	0.458 ± 0.071
0.7	0.675 ± 0.101	0.651 ± 0.103
0.9	0.850 ± 0.131	0.829 ± 0.133
1.1	1.041 ± 0.166	1.021 ± 0.166
1.3	1.218 ± 0.192	1.195 ± 0.191

movement showed a better performance than the standard PSO with a uniform distribution (Table 2). In addition, the proposed PSO with a Gamma distribution showed a much faster convergence rate than the standard PSO with a uniform distribution (Fig. 3(b)): Accordingly, in the following sub-sections, we present only the results from the Gamma-distribution-based PSO algorithm.

3.2. Comparison to conventional approach

This sub-section presents the results of a comparative evaluation of the proposed method with the LMA-based methods. The examples of simulated TACs for the various noise levels is given in Fig. 4. Fig. 5 shows the bias and CV of the parameter estimation using five different approaches: BF, LMA, the repetition of LMA (Repeated LMA), and Gamma-based PSO without and with l_∞ -norm regularization (Unregularized PSO and Regularized PSO, respectively). Although they showed similar bias at low noise levels, the difference in bias was substantial at high noise levels, particularly in the k_2 and k_3 estimations. BF method yielded better performances than LMA in bias. The repetition of the LMA algorithm for diverse initial guess of parameters improved the bias property; however, it still yielded a higher bias than the Regularized PSO, and required a much higher computing time owing to the repetition. By combining Gamma-based PSO and l_∞ -norm regularization ($\gamma = 10^{-2.75}$), the bias was substantially reduced in the K_1 , k_2 , and k_3 estimations.

The results of the CV analysis are also shown in Fig. 5, which also indicates that the Regularized PSO yielded the best performance. The BF method showed comparable performances with regularized PSO except for k_4 . The Repeated LMA substantially improve the CV properties except for k_4 . A more detailed analysis of regularization is given in the section below. The average computation time of BF, LMA, and PSO for 1000 noisy simulations was 0.1158, 0.0658, and 0.0932 s,

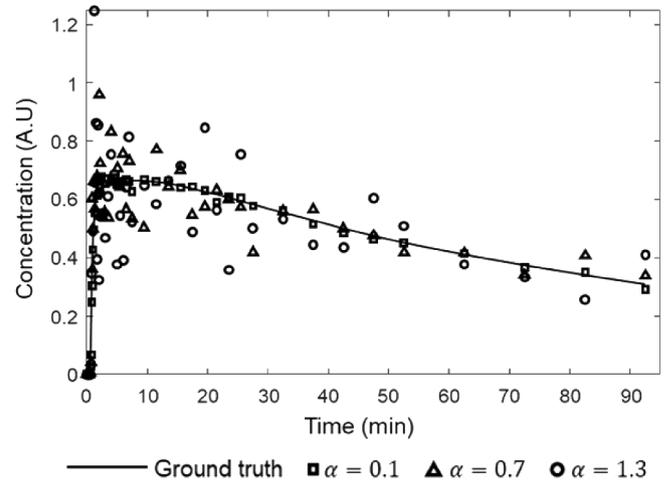


Fig. 4. Simulated TACs with different noise levels (a).

respectively. However, these times could depend significantly on algorithmic settings such as the number of bases, number of iterations, and number of particles.

3.3. Optimization of regularization parameter

The bias and CV property in parameter estimation at a high noise level were improved by applying moderate l_∞ regularization, as shown in Fig. 6. Too much regularization (e.g. $\gamma = 10^{-1}$) was shown to yield a poor performance. The CV for K_1 , k_2 , and k_3 decreased as the regularization parameter increased. The best performance was obtained with $\gamma = 10^{-2.75}$ through $10^{-2.25}$. However, this optimal regularization parameter may not be valid for other parameter sets. The optimal regularization parameter depends on the larger exponent, b_2 , and it would be challenging to find the global optimal rule for individual b_2 . Our results show that too high or too low regularization parameter causes poor performance. Based on the expectation that the optimal range derived from the simulation will work well for a similar b_2 by avoiding extreme cases, we continued to use the range in later experiments.

3.4. Parametric images

Fig. 7 shows the digital phantom used for the simulation. Five circular regions were realized and ground truth sinogram was

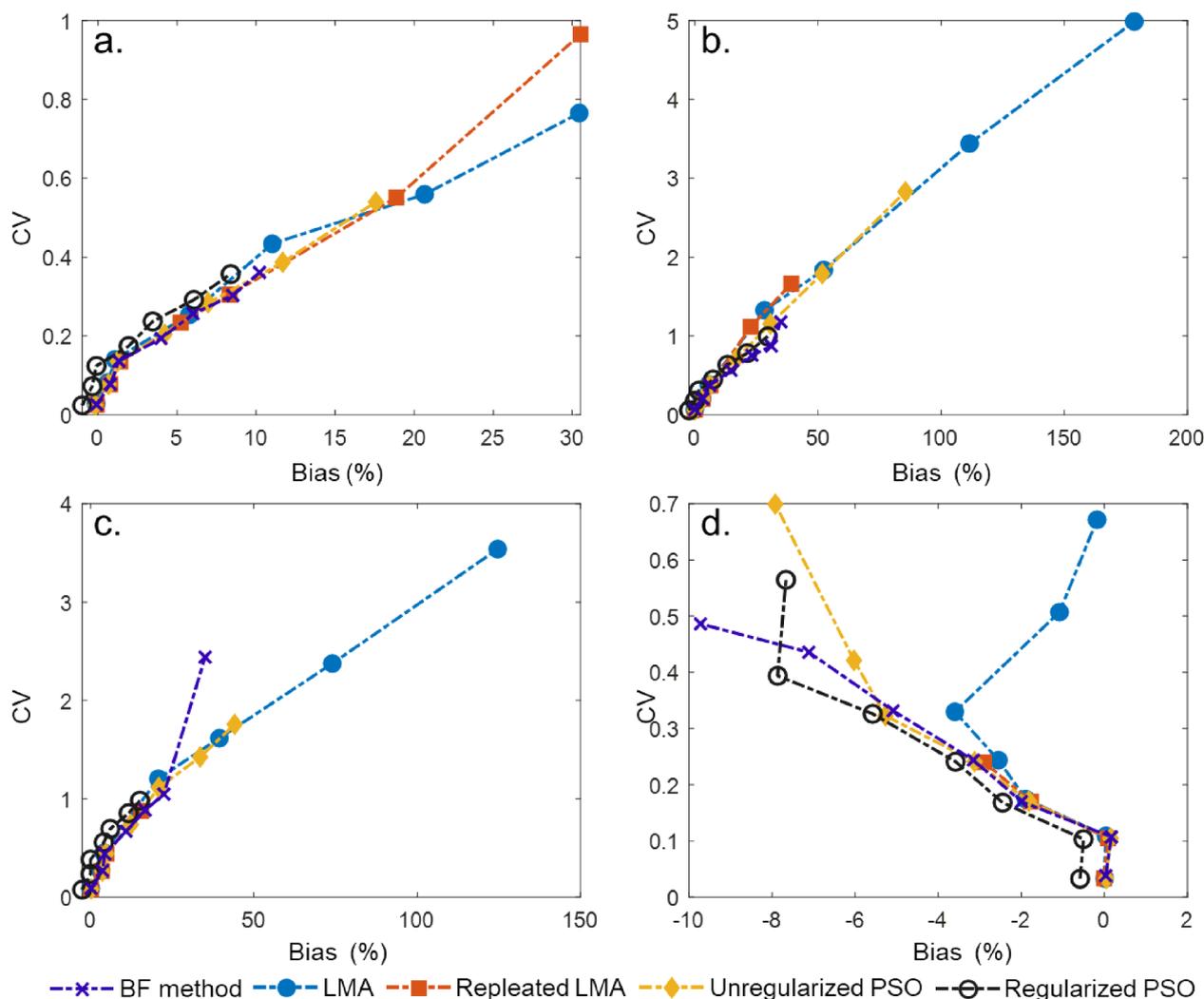


Fig. 5. CV versus Bias plots for parameter estimation using four different methods. (a) K_1 . (b) k_2 . (c) k_3 . (d) k_4 .

contaminated by Poisson noise. The parametric images of each rate constant (Fig. 8) for the simulation data show that the l^∞ regularization is efficient for reducing bias and variation through wide range of rate constants. NRMSE values of each region demonstrate that the Regularized PSO is more accurate than any other compared methods especially for k_3 and k_4 (Table 3). Fig. 9 shows the static (left) and parametric (right) images of $[^{18}\text{F}]$ FDG PET studies on BALB/c nude mice. The parameter values estimated from VOI-based analysis (initial values for LMA and PSO) are indicated by yellow arrows under the colorbars. The LMA parametric images show much higher or lower pixel values relative to the results of VOI analysis. Although the Unregularized and Regularized PSO methods yielded visually similar parametric images, the Unregularized PSO resulted in the over-estimation of k_2 and k_3 in HepG2 tumor model. These findings in the evaluations using real data are consistent with those in the simulation study. Although the heterogenous distribution of kinetic parameters was observed in the parametric images generated using the proposed methods, we could not confirm their validity because of no availability of histochemical staining data of tumors.

4. Discussions

This study aimed to robustly estimate the individual rate constants from the highly noisy TACs of a two-tissue compartment model. The noise in measured TACs led to the bias in this parameter estimation. In particular, the high noise level in the TACs yielded considerably skewed

distribution of the larger exponents in Eq. (4) with a long tail (Fig. 1). The method proposed in this study to reduce the bias due to the skewed parameter distribution is the minimization of infinity norm that takes the largest elements in a given vector. By minimizing the infinity norm, we can regularize the larger element not to have extremely large values.

The LMA, the most widely used nonlinear estimation algorithm for least squares curve fitting, cannot deal with the proposed infinity norm regularizer that is non-differentiable. This is because the LMA, in each iteration step to find solution, should use Jacobian and locally linearized Hessian information respectively consisted of the first and second order partial derivatives. In the LMA, a damping parameter to mix these two pieces of information is adjusted through the iterations. If the cost function is rapidly reduced, a small value of damping parameter is used, and the LMA behaves like Gauss-Newton algorithm that uses both Jacobian and Hessian information. Otherwise, the damping parameter is adjusted so that the LMA operates like a gradient-descent algorithm that only uses Jacobian, yielding slower searching speed than Gauss-Newton.

Between the alternative algorithms to handle the non-differentiable optimization problem on account of the infinity norm regularization, the PSO outperformed the proximal gradient algorithm. The proximal gradient algorithm that uses only the gradient information of the differentiable part of cost function showed slow convergence and was easily trapped in local minima as shown in Fig. 5a. It would hardly proceed in the direction of less descent evaluated for nonlinear parameters. On the contrary, the PSO based on the random search in the

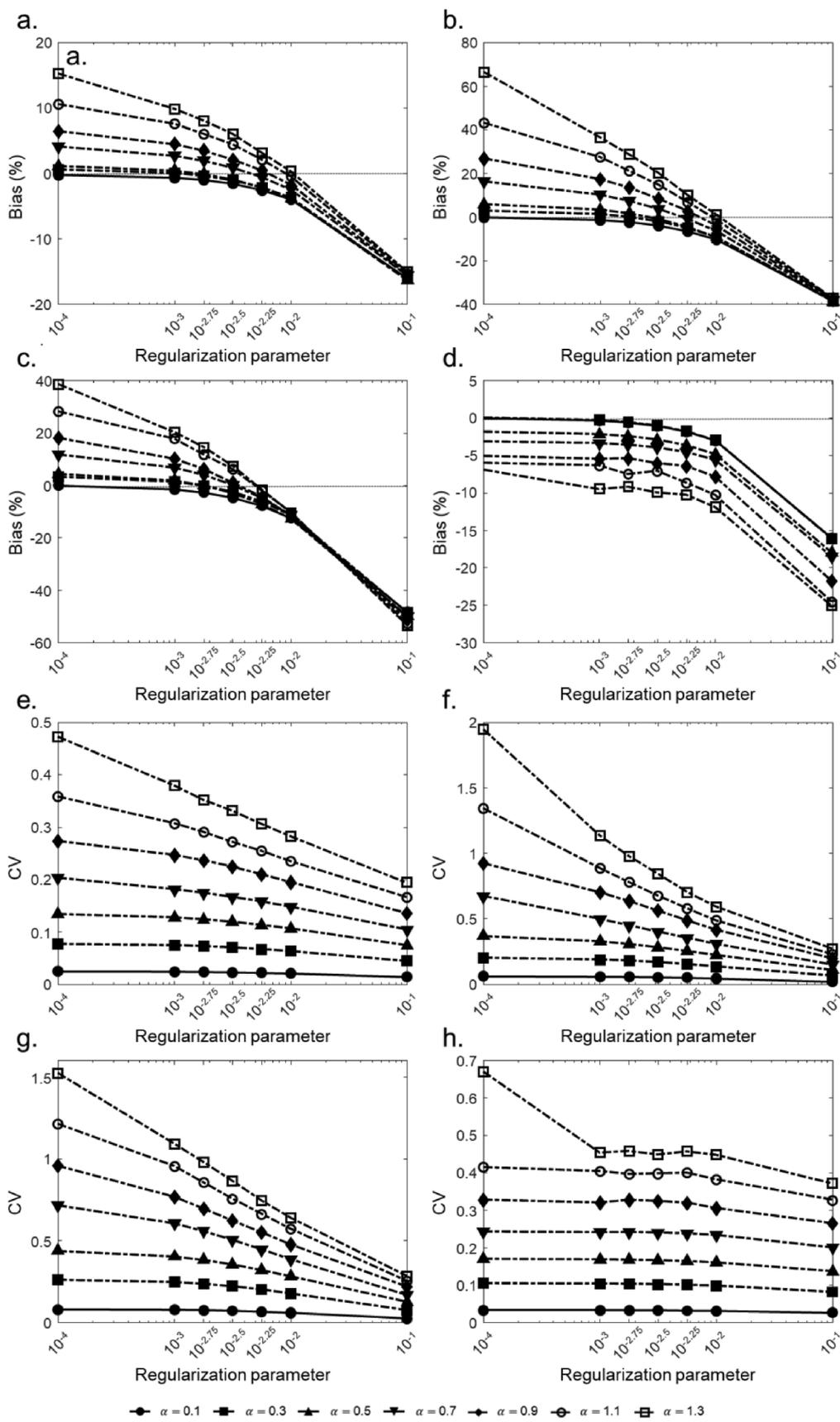


Fig. 6. Parameter optimization for l_∞ -norm regularization: (a) bias of K_1 , (b) bias of k_2 , (c) bias of k_3 , (d) bias of k_4 , (e) CV of K_1 , (f) CV of k_2 , (g) CV of k_3 and (h) CV of k_4 ,

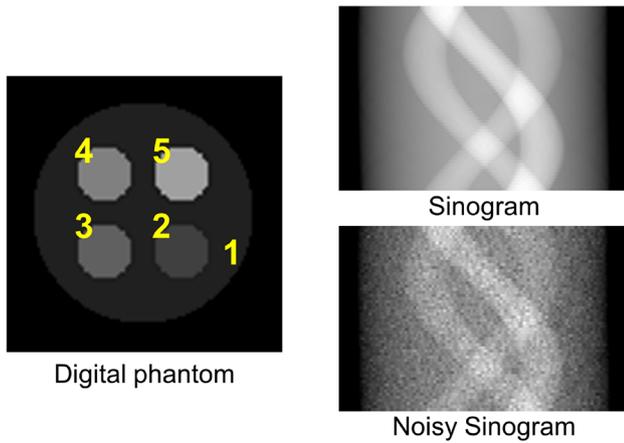


Fig. 7. Simulated digital phantom and sinogram.

parameter space showed better convergence properties than the proximal gradient, indicating this search strategy is more effective than gradient-based one in this specific optimization problem.

The partial linearization of parameter estimation and the use of gamma distribution for random particle movement improved the performance of the PSO algorithm. The dimension of nonlinear parameter space was reduced by the partial linearization that decomposes the parameter estimation into linear and nonlinear problems. Consequently, we could reduce the uncertainty and bias caused by the iterative search of many parameters with the limited number of noisy data samples. In the original PSO, the randomness of particle movement is constrained by the uniform random distribution: The particles move to the far and near points in parameter search space with the same probability. However, the gamma distribution has more centralized distribution, allowing for faster convergence, and the extended tail allows for a sudden departure from the local minimum.

Table 3

Normalized root mean square error (NRMSE) for parametric images and various regions in digital phantom. Bold numbers show the lowest NRMSE values for each region and kinetic parameters.

		k_1	k_2	k_3	k_4
Region 1	BF	0.71	0.791	3.814	3.935
	LMA	0.743	0.796	3.874	3.952
Region 2	Unregularized PSO	0.704	0.788	3.859	3.969
	Regularized PSO	0.630	0.506	2.629	3.875
	BF	0.855	0.837	4.474	2.630
Region 3	LMA	1.276	1.066	7.276	3.642
	Unregularized PSO	0.868	1.031	7.238	3.597
	Regularized PSO	0.794	0.731	4.919	3.323
Region 4	BF	1.046	1.125	1.539	2.204
	LMA	1.033	1.119	1.528	2.136
	Unregularized PSO	1.046	1.122	1.538	2.191
Region 5	Regularized PSO	1.052	0.853	1.251	2.188
	BF	0.431	0.545	2.075	1.921
	LMA	0.232	0.333	1.456	1.015
Region 5	Unregularized PSO	0.228	0.330	1.471	1.038
	Regularized PSO	0.258	0.254	0.982	1.016
	BF	0.157	0.110	0.584	3.338
Region 5	LMA	0.161	0.116	0.601	3.324
	Unregularized PSO	0.160	0.115	0.617	3.351
Region 5	Regularized PSO	0.193	0.094	0.428	3.467

The same linearization strategy underlies the BF method, but the BF and the PSO are different in how to prepare the candidates for exponents. The BF method picks out the best set of exponents from a pre-defined bounded and discrete pool of candidates. In contrast, the PSO searches for the best set by moving particles randomly over unbounded and continuous parameter space. Therefore, the BF method's performance depends on the size and range of the pool while the PSO is robust to the initialization of particle positions. For example, the BF method showed poor performance for region 4 in Fig. 8, but the Unregularized or Regularized PSO showed better results.

The simulation study with dynamic mathematical phantom showed

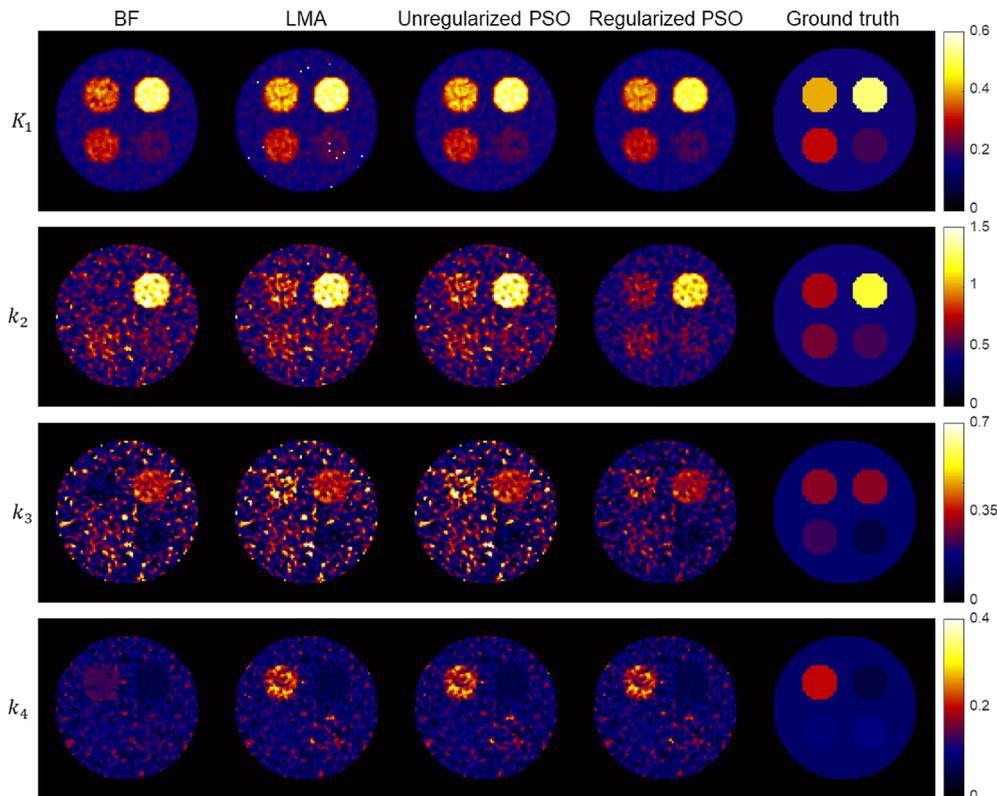


Fig. 8. Parametric images for digital phantom.

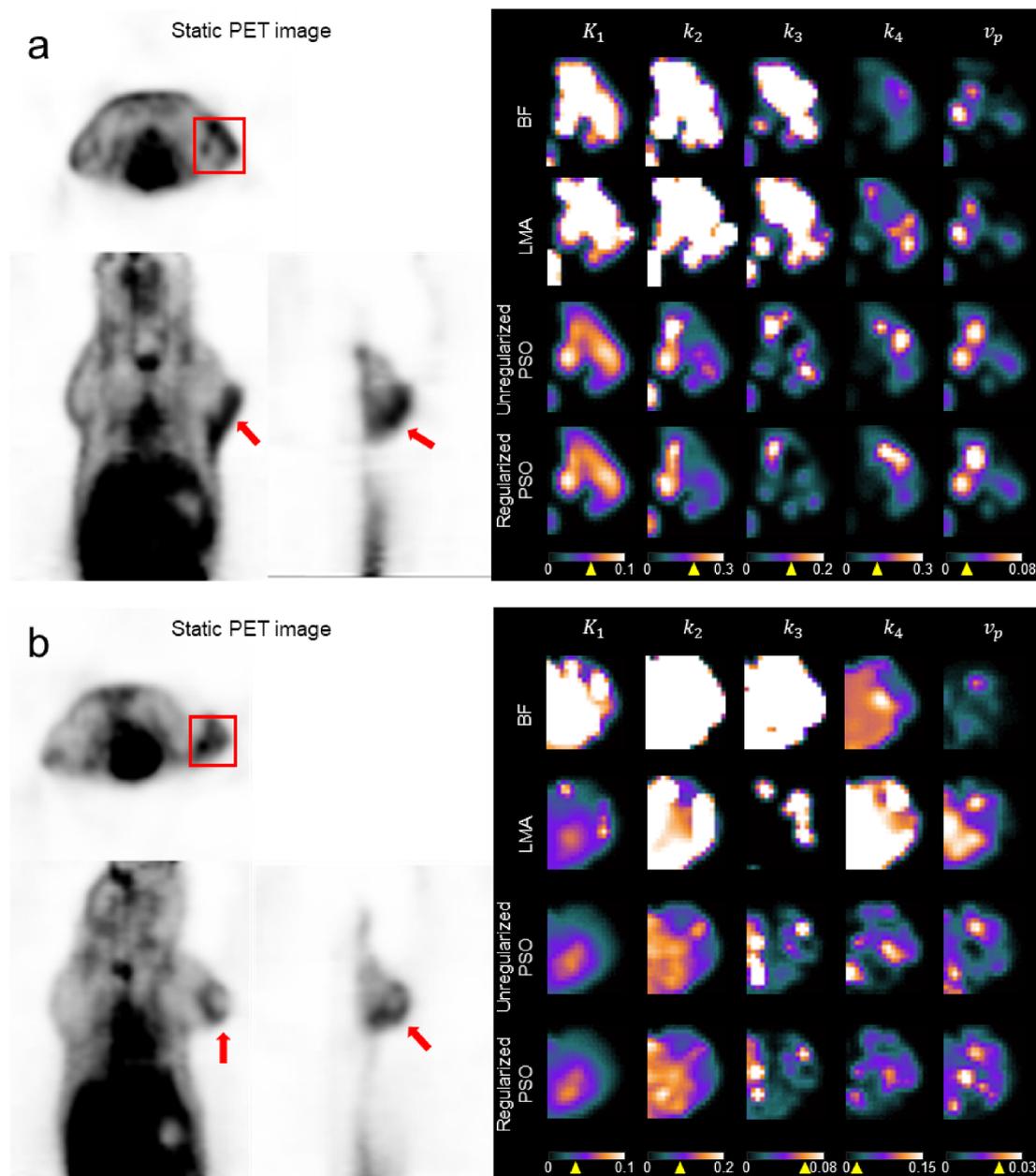


Fig. 9. Static (left) and parametric (right) images of [^{18}F]FDG PET studies performed in BALB/c nude mice. The yellow arrows under the colorbars indicate the parameter values estimated in VOI-based analysis using Regularized PSO. (a) HepG2 tumor model and (b) MDA-MB-231 tumor model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

that the proposed method works well through the wide range of rate constants. On the contrary, the accuracy of LMA depends on the true values of rate constants. The LMA estimation that started with the same initial values for all the pixels hardly reach the global minima in the pixels with true rate constants placed far from the initial values. However, the PSO particles randomly distributed in search space at the start would find the global minima more effectively.

A common regularization strategy for kinetic parameter estimation has been the use of a prior information on the parameters, which was gained from a separate or the same study. There have been several ways of directly exploiting such information to tightly regularize the parameters, like constraining the parameters within physiologically reasonable fixed bounds, or using a specific reliable value for a global parameter, if exists, instead of estimating it [39,40]; however, the latter is more recognized as a model simplification rather than as a regularization framework. More sophisticated and less stringent regularization

approaches have also been developed using a l_2 -norm regularization framework that adaptively exploits the prior information depending on the quality of unregularized estimation [7,26,27,39–50]. Contrary to these conventional approaches, which constrain the parameters of interest using their prior information, the proposed method regularizes only one macro parameter, a combination of rate constants, without using any prior information.

Our infinity-norm regularization framework has a similar formulation (Eq. (8)) with the l_2 -norm regularization framework, in which the weighted least squares cost function is augmented by an l_2 -norm penalty function measuring the estimate's deviation from prior information and a regularization parameter governing the impact of the prior information on the estimation. The statistical performance of the proposed framework depends on a regularization parameter value whereas one of the l_2 -norm regularization relies on not only a selection of the regularization parameter but also a formulation of the penalty function.

For kinetic rate constant estimation, investigators have introduced various formulations of l_2 -norm penalty function with different perspectives on the beneficial prior information, such as a physiological variation of parameters among different subjects [26,27,39–43,47] or local spatial variation within the same subject [45,49]. Although these specification methods have been proved as effective in individual validations, it is based on the premise that appropriate prior information is available; moreover, a comprehensive comparison among those methods is still lacking. To select the regularization parameter, the l_2 -norm-based approaches draw on empirical search or approximation of the theoretical methods developed for linear regression (see [50] for a summary of theoretical methods). Though we also searched for the optimal value through simulation, a future investigation for an approximation of other methods available in the literature will be worthwhile.

There are some limitations of this study. First, the proposed PSO methods does not provide uncertainty in estimated parameters which would be important in VOI-based data analysis. Second, the validity of proposed methods was mainly investigated through simulation studies and limited number of real data was accessed. Third, only the two-tissue compartment model with four unknown parameters was investigated. Incorporation of spillover fraction in the model would be necessary in the further work.

Machine learning is an active research field in medical imaging [51–58]. Deep neural networks trained with large data set show better performance than any other conventional approaches in many problems in medical imaging, such as denoising, segmentation, and registration [55,59–63]. Applying this new approaches to the parameter estimation and direct reconstruction in dynamic PET studies would be

Appendix A

A1. Principle of PSO algorithm

PSO searches for the solution to the optimization problem using a population (or swarm) of candidate solutions called “particles.” Each particle moves around a given search space by iteratively updating the key elements as follows:

- Each particle’s velocity, current position, and local best position visited until current iteration are $v_i(m)$, $x_i(m)$, and $p_i^L(m)$, respectively, for the i th particle at the m th iteration. In addition, the population’s global best position visited until current iteration is $p^G(m)$.
- Simple PSO rules were designed to make the movement of each particle random and toward better positions ($p_i^L(m)$ and $p^G(m)$) with inertia, leading the population to finally converge to the best position.
- The best position of each particle (or the population) means the one yielding the lowest value of the optimization cost function (f) among all positions visited by each particle (or by the population); in addition, $p^G(m)$ is shared by all particles.
- In the kinetic-parameter-modeling problem, the position of a particle corresponds to the set of kinetic parameters to be estimated and, in the particle space, the cost function is defined.

These rules were formulated into mathematical equations (19) through (22). In a nutshell, the PSO finds the optimal solution by comparing the cost functions computed at all positions visited by all particles during the course of the iterations; i.e., the solution is the best among the $N \times M$ candidates that are randomly chosen and evolved through interactions among themselves. Moreover, the PSO makes no assumptions regarding the characteristics of the cost function, such as convexity or differentiability, and requires no initial estimates of the parameters. Based on these properties, we can hypothesize that the random selection of potential candidates (or random movement of particles) through PSO will play a key role in addressing the local minima issues of conventional NLS approaches in kinetic parameter estimations by offering particles a chance to escape from the local minima positions through well-designed randomness.

A2. Proposal of Gamma distribution

In standard PSO, φ_1 and φ_2 are generated from a uniform distribution between zero and 1, where the variance of their distribution is fixed. In addition, w is a constant, and thus has no uncertainty. We wanted to adjust the uncertainty of random distributions through the three positions, $x_i(m-1)$, $p_i^L(m)$, and $p^G(m)$, which are heading to the current position, $x_i(m)$, in Eq. (23). If the cost function values of these three positions are high, particles must move far from the three positions at current iteration (m). Thus, it is favorable for the three weights to have high uncertainty. Two random distributions, φ_1 and φ_2 , must have positive values to force the particles to proceed to the local and global best positions. Accordingly, we adopted a Gamma distribution that is only in the positive range. The initial means and variances for φ_1 and φ_2 were set to have the same values as the original uniform distribution, and were therefore 0.5 and 1/12, respectively. In the same way, the initial mean and variance for w were set to 0.7298 and 1/12. We calculated the logarithmic values of the cost functions because the raw cost function value of each position has too high a deviation. Finally, the weights applied to the initial variances of the Gamma distribution were as follows.

the next step to go, and we warrant the comparison study between the proposed method with those new approaches.

5. Summary and conclusion

This paper presented a new approach based on the infinity-norm regularization for PET compartment modeling, which can constrain the nonlinear kinetic parameters. In solving the infinity-norm regularization that involves the non-smoothness cost function, PSO methods showed much better convergence, owing to the use of multiple initial values, than a proximal gradient algorithm that was susceptible to the initial values. In the PSO implementation, the use of a Gamma distribution to govern random movements improved the convergence rate and stability compared to a uniform distribution. Consequently, the Gamma-based PSO with the infinity-norm regularization outperformed all the methods tested, including conventional LMA, in terms of the statistical properties. In particular, the randomly moving particles in PSO secured the robustness to both the local minima and initial values, which are fatal to the LMA.

Acknowledgments

This work was supported by grants from the National Research Foundation of Korea (NRF) funded by the Korean Ministry of Science, ICT and Future Planning (grant no. NRF-2016R1A2B3014645), and the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant no. HI13C0163020015 and HI14C1135).

$$d_1 = \log(1 + f(\mathbf{p}_1^L(m))),$$

$$d_2 = \log(1 + f(\mathbf{p}^G(m))),$$

$$d_3 = \log(1 + f(\mathbf{x}_i(m - 1))),$$

$$\sigma_1^2 = \frac{1}{12} \times \frac{d_1}{(d_1 + d_2 + d_3)/3},$$

$$\sigma_2^2 = \frac{1}{12} \times \frac{d_2}{(d_1 + d_2 + d_3)/3},$$

$$\sigma_3^2 = \frac{1}{12} \times \frac{d_3}{(d_1 + d_2 + d_3)/3},$$

where σ_1^2 , σ_2^2 , and σ_3^2 are variances for φ_1 , φ_2 , and w , respectively.

A3. Stopping Criteria, Bias, and CV

The proximal gradient and PSO iterate until the following criterion,

$$\frac{\|\theta^k - \theta^{k-1}\|_2^2}{\|\theta^k\|} \leq 10^{-6},$$

is satisfied or until some maximum numbers of iterations are reached. The maximum iterations for the proximal gradient and PSO are 10,000 and 150, respectively.

For 1000 realizations, the bias and CV were calculated using the following equations.

$$\text{Bias}(\%) = \frac{\theta_{\text{mean}} - \theta_{\text{true}}}{\theta_{\text{true}}} \times 100,$$

$$\text{CV}(\%) = \frac{\theta_{\text{std}}}{\theta_{\text{true}}},$$

where θ_{mean} is the mean of the estimated parameters, θ_{std} is the standard deviation, and θ_{true} is the ground truth.

Appendix B. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ejmp.2020.03.013>.

References

- Carson RE. Parameter estimation in positron emission tomography. In: Phelps ME, Mazziotta JC, Schelbert HR, editors. Positron emission tomography and autoradiography: principles and applications for the brain and heart. New York: Raven Press; 1986. p. 347–90.
- Gunn RN, Gunn SR, Cunningham VJ. Positron emission tomography compartmental models. *J Cereb Blood Flow Metab* 2001;21:635–52.
- Kim SJ, Lee JS, Kim YK, Frost J, Wand G, McCaul ME, et al. Multiple linear analysis methods for the quantification of irreversibly binding radiotracers. *J Cereb Blood Flow Metab* 2008;28:1965–77.
- Koeppel RA, Holthoff VA, Frey KA, Kilbourn MR, Kuhl DE. Compartmental analysis of [¹¹C]flumazenil kinetics for the estimation of ligand transport rate and receptor distribution using positron emission tomography. *J Cereb Blood Flow Metab* 1991;11:735–44.
- Lee JS, Lee DS, Ahn JY, Yeo JS, Cheon GJ, Kim SK, et al. Generation of parametric image of regional myocardial blood flow using H(2)(15)O dynamic PET and a linear least-squares method. *J Nucl Med* 2005;46:1687–95.
- Lee J, Lee D. Tracer kinetic analysis for PET and SPECT. *Medical Imaging: CRC Press*; 2013. p. 463–84.
- Seo S, Kim SJ, Kim YK, Lee JY, Jeong JM, Lee DS, et al. Comparative assessment of parametric neuroreceptor mapping approaches based on the simplified reference tissue model using [(1)(1)C]ABP688 PET. *J Cereb Blood Flow Metab* 2015;35:2098–108.
- Turkheimer FE, Hinz R, Cunningham VJ. On the undecidability among kinetic models: from model selection to model averaging. *J Cereb Blood Flow Metab* 2003;23:490–8.
- Watabe H, Ikoma Y, Kimura Y, Naganawa M, Shidahara M. PET kinetic analysis—compartmental model. *Ann Nucl Med* 2006;20:583–8.
- Kanzow C, Yamashita N, Fukushima M. Levenberg–Marquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints. *J Comput Appl Math* 2004;172:375–97.
- Marquardt D. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *J Soc Ind Appl Math* 1963;11:431–41.
- Mankoff DA, Shields AF, Graham MM, Link JM, Eary JF, Krohn KA. Kinetic analysis of 2-[carbon-11]thymidine PET imaging studies: compartmental model and mathematical analysis. *J Nucl Med* 1998;39:1043–55.
- Gunn RN, Lammertsma AA, Hume SP, Cunningham VJ. Parametric Imaging of Ligand-Receptor Binding in PET Using a Simplified Reference Region Model. *Neuroimage* 1997;6:279–87.
- 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;22:1425–39.
- Hong YT, Fryer TD. Kinetic modelling using basis functions derived from two-tissue compartmental models with a plasma input function: general principle and application to [18F]fluorodeoxyglucose positron emission tomography. *Neuroimage* 2010;51:164–72.
- Seo S, Kim SJ, Lee DS, Lee JS. Recent advances in parametric neuroreceptor mapping with dynamic PET: basic concepts and graphical analyses. *Neurosci Bull* 2014;30:733–54.
- Seo S, Kim SJ, Yoo HB, Lee JY, Kim YK, Lee DS, et al. Noninvasive bi-graphical analysis for the quantification of slowly reversible radioligand binding. *Phys Med Biol* 2016;61:6770–90.
- Beck A, Teboulle M. Gradient-based algorithms with applications to signal-recovery problems. In: Palomar DP, Eldar YC, editors. *Convex optimization in signal processing and communications*. Cambridge: Cambridge University Press; 2009. p. 42–88.
- Eberhart R, Kennedy J. A new optimizer using particle swarm theory. *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, 1995. p. 39–43.
- Kennedy J, Eberhart R. Particle swarm optimization. *IEEE International Conference on Neural Networks*, 1995:1995. p. 1942–1948.
- Kennedy J. The particle swarm: social adaptation of knowledge. *Evolutionary Computation*, 1997, *IEEE International Conference on 1997*. p. 303–308.
- Robinson J, Rahmat-Samii Y. Particle swarm optimization in electromagnetics. *IEEE Trans Antennas Propag* 2004;52:397–407.
- Valle Yd, Venayagamoorthy GK, Mohagheghi S, Hernandez JC, Harley RG. Particle swarm optimization: basic concepts, variants and applications in power systems. *IEEE Trans Evol Comput* 2008;12:171–95.
- Huang C-K, Wang W, Tzen K-Y, Lin W-L, Chou C-Y. FDOPA kinetics analysis in PET images for Parkinson's disease diagnosis by use of particle swarm optimization. *2012. 9th IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE; 2012. p. 586–9.
- Ghovvati M, Khayati G, Attar H, Vaziri A. Kinetic parameters estimation of protease production using penalty function method with hybrid genetic algorithm and

- particle swarm optimization. *Biotechnol Biotechnol Equip* 2016;30:404–10.
- [26] Alpert NM, Yuan F. A general method of Bayesian estimation for parametric imaging of the brain. *Neuroimage* 2009;45:1183–9.
- [27] Dean Fang Y-H, El Fakhri G, Becker JA, Alpert NM. Parametric imaging with Bayesian priors: a validation study with ¹¹C-Altropene PET. *Neuroimage* 2012;61:131–8.
- [28] Sitek A, Li Q, El Fakhri G, Alpert NM. Validation of Bayesian analysis of compartmental kinetic models in medical imaging. *Phys Med.* 2016;32:1252–8.
- [29] Kamasak ME, Bouman CA, Morris ED, Sauer K. Direct reconstruction of kinetic parameter images from dynamic PET data. *IEEE Trans Med Imag* 2005;24:636–50.
- [30] Bansal JC, Singh PK, Saraswat M, Verma A, Jadon SS, Abraham A. Inertia weight strategies in particle swarm optimization. *nature and biologically inspired computing (NaBIC), 2011 Third world congress on 2011.* p. 633–640.
- [31] Clerc M, Kennedy J. The particle swarm – explosion, stability, and convergence in a multidimensional complex space. *IEEE Trans Evol Comput* 2002;6:58–73.
- [32] Shi Y, Eberhart R. A modified particle swarm optimizer. *iee international conference on evolutionary computation proceedings. 1998 IEEE; 1998.* p. 69–73.
- [33] Trelea IC. The particle swarm optimization algorithm: convergence analysis and parameter selection. *Inf Process Lett* 2003;85:317–25.
- [34] Krohling RA. Gaussian swarm: a novel particle swarm optimization algorithm. *cybernetics and intelligent systems, 2004 IEEE conference on 2004.* p. 372–376.
- [35] Krohling RA, Coelho LdS. PSO-E: Particle Swarm with Exponential Distribution. 2006 IEEE International Conference on Evolutionary Computation 2006. p. 1428–1433.
- [36] Feng D, Huang S-C, Wang X. Models for computer simulation studies of input functions for tracer kinetic modeling with positron emission tomography. *Int J Biomed Comput* 1993;32:95–110.
- [37] Kim MJ, Lee C-H, Lee Y, Youn H, Kang KW, Kwon J, et al. Glucose-6-phosphatase expression-mediated [18F]FDG efflux in murine inflammation and cancer models. *Mol Imaging Biol* 2019;21:917–25.
- [38] Lammertsma AA. Tracer kinetic modelling. In: Dierckx RAJO, Otte A, de Vries EFJ, van Waarde A, Leenders KL, editors. *PET and SPECT in neurology.* Berlin, Heidelberg: Springer Berlin Heidelberg; 2014. p. 59–73.
- [39] O'Sullivan F, Saha A. Use of ridge regression for improved estimation of kinetic constants from PET data. *IEEE Trans Med Imag* 1999;18:115–25.
- [40] Bertoldo A, Sparacino G, Cobelli C. “Population” approach improves parameter estimation of kinetic models from dynamic PET data. *IEEE Trans Med Imag* 2004;23:297–306.
- [41] Wilson PD, Huang S-C, Hawkins RA. Single-scan bayes estimation of cerebral glucose metabolic rate: comparison with bon-bades single-scan methods using DFG pet scans in stroke. *J Cereb Blood Flow Metab* 1988;8:418–25.
- [42] Wilson PD, Huang SC, Links JM. Improved estimation of local cerebral glucose metabolic rate using bayes regression analysis of PET scan data. *Proceedings of the annual symposium on computer application in medical care.* 1984:128–131.
- [43] Chen B, Sung-Cheng H, Hawkins RA, Phelps ME. An evaluation of Bayesian regression for estimating cerebral oxygen utilization with oxygen-15 and dynamic PET. *IEEE Trans Med Imag* 1988;7:257–63.
- [44] Zhou Y, Huang S-C, Bergsneider M. Linear ridge regression with spatial constraint for generation of parametric images in dynamic positron emission tomography studies. *IEEE Trans Nucl Sci* 2001;48:125–30.
- [45] Zhou Y, Huang S-C, Bergsneider M, Wong DF. Improved parametric image generation using spatial-temporal analysis of dynamic PET studies. *Neuroimage* 2002;15:697–707.
- [46] Zhou Y, Endres CJ, Brašić JR, Huang S-C, Wong DF. Linear regression with spatial constraint to generate parametric images of ligand-receptor dynamic PET studies with a simplified reference tissue model. *Neuroimage* 2003;18:975–89.
- [47] Byrtek M, Sullivan FO, Muzi M, Spence AM. An adaptation of ridge regression for improved estimation of kinetic model parameters from PET studies. *IEEE Trans Nucl Sci* 2005;52:63–8.
- [48] Wang G, Qi J. Spatially penalized methods for linear parametric imaging in dynamic PET. 2006 IEEE nuclear science symposium conference record 2006. p. 1787–1791.
- [49] Kamasak ME. Effects of spatial regularization on kinetic parameter estimation for dynamic PET. *Biomed Signal Process Control* 2014;9:6–13.
- [50] Karakatsanis NA, Lodge MA, Zhou Y, Wahl RL, Rahmim A. Dynamic whole-body PET parametric imaging: II Task-oriented statistical estimation. *Phys Med Biol* 2013;58:7419–45.
- [51] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention. Springer; 2015.* p. 234–41.
- [52] Brebisson Ad, Montana G. Deep Neural Networks for Anatomical Brain Segmentation. *arXiv:150202445.* 2015.
- [53] Ciresan D, Giusti A, Gambardella LM, Schmidhuber J. Deep neural networks segment neuronal membranes in electron microscopy images. *Adv Neural Inf Process Syst* 2012:2843–51.
- [54] Prasoon A, Petersen K, Igel C, Lauze F, Dam E, Nielsen M. Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network. In: Mori K, Sakuma I, Sato Y, Barillot C, Navab N, editors. *Medical image computing and computer-assisted intervention – MICCAI 2013.* Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 246–53.
- [55] Dey D, Chaudhuri S, Munshi S. Obstructive sleep apnoea detection using convolutional neural network based deep learning framework. *Biomed Eng Lett* 2018;8:95–100.
- [56] Mansour RF. Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy. *Biomed Eng Lett* 2018;8:41–57.
- [57] Choi H. Deep learning in nuclear medicine and molecular imaging: current perspectives and future directions. *Nucl Med Mol Imag* 2010;2018(52):109–18.
- [58] Choi JY. Radiomics and deep learning in clinical imaging: what should we do? *Nucl Med Mol Imag* 2010;2018(52):89–90.
- [59] Kang SK, Seo S, Shin SA, Byun MS, Lee DY, Kim YK, et al. Adaptive template generation for amyloid PET using a deep learning approach. *Hum Brain Mapp* 2018;39:3769–78.
- [60] Hwang D, Kang SK, Kim KY, Seo S, Paeng JC, Lee DS, et al. Generation of PET attenuation map for whole-body time-of-flight ¹⁸F-FDG PET/MRI using a deep neural network trained with simultaneously reconstructed activity and attenuation maps. *J Nucl Med* 2019. [jnumed.118.219493](https://doi.org/10.1186/s12949-019-0493-9).
- [61] Hwang D, Kim KY, Kang SK, Seo S, Paeng JC, Lee DS, et al. Improving accuracy of simultaneously reconstructed activity and attenuation maps using deep learning. *J Nucl Med* 2018;59:1624–9.
- [62] Park J, Bae S, Seo S, Park S, Bang J-I, Han JH, et al. Measurement of glomerular filtration rate using quantitative SPECT/CT and deep-learning-based kidney segmentation. *Sci Rep* 2019;9:4223.
- [63] Park J, Hwang D, Kim KY, Kang SK, Kim YK, Lee JS. Computed tomography super-resolution using deep convolutional neural network. *Phys Med Biol* 2018;63:145011.