Published in final edited form as:

Phys Med. 2016 October ; 32(10): 1252–1258. doi:10.1016/j.ejmp.2016.09.010.

Validation of Bayesian Analysis of Compartmental Kinetic Models in Medical Imaging

Arkadiusz Sitek¹, Quanzheng Li, Georges El Fakhri, and Nathaniel M Alpert

Massachusetts General Hospital and Harvard Medical School, Radiology Department, 55 Fruit Street, Boston, MA, 02114, USA

Abstract

Introduction—Kinetic compartmental analysis is frequently used to compute physiologically relevant quantitative values from time series of images. In this paper, a new approach based on Bayesian analysis to obtain information about these parameters is presented and validated.

Materials and Methods—The closed-form of the posterior distribution of kinetic parameters is derived with a hierarchical prior to model the standard deviation of normally distributed noise. Markov chain Monte Carlo methods are used for numerical estimation of the posterior distribution. Computer simulations of the kinetics of F18-fluorodeoxyglucose (FDG) are used to demonstrate drawing statistical inferences about kinetic parameters and to validate the theory and implementation. Additionally, point estimates of kinetic parameters and covariance of those estimates are determined using the classical non-linear least squares approach.

Results and discussion—Posteriors obtained using methods proposed in this work are accurate as no significant deviation from the expected shape of the posterior was found (one-sided $P > 0.08$). It is demonstrated that the results obtained by the standard non-linear least-square methods fail to provide accurate estimation of uncertainty for the same data set ($P < 0.0001$).

Conclusions—The results of this work validate new methods for a computer simulations of FDG kinetics. Results show that in situations where the classical approach fails in accurate estimation of uncertainty, Bayesian estimation provides an accurate information about the uncertainties in the parameters. Although a particular example of FDG kinetics was used in the paper, the methods can be extended for different pharmaceuticals and imaging modalities.

Keywords

Bayesian inference; dynamic nuclear imaging; kinetic analysis

¹Currently at Philips Research North America, 2 Canal Park, Cambridge, MA, 02141, USA

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Kinetic compartmental analysis is frequently used to compute physiologically relevant quantitative values from time series of medical images. Traditionally, methods based on nonlinear least squares parameter optimization are used to estimate the values of kinetic parameters and their asymptotic covariance matrix. Although the use of nonlinear least squares optimization is ubiquitous in science, it is not often well suited to estimation of kinetic parameters from clinical data with low signal-to-noise-ratio (SNR) where the classical analysis may erroneously result in variations in parameter values which are outside the known physiological range. Furthermore, the classical approach is rooted in the idea of steady state systems and it is not easily generalized to the analysis of temporal perturbations, such as the effect of amphetamine on dopamine release. These issues are not merely intellectual and conceptual, the limitations have profound effect on the design of experiments and hypothesis test.

Bayesian analysis, among other advantages listed in the discussion, provides the range of probable values. The main difference between the classical approach and the method investigated here is that the result of the Bayesian analysis (BA) is a probability distribution and not a point estimate. This is schematically illustrated in figure 1(A and D). In addition to providing accurate representation of the precision, the Bayesian analysis has an intuitive means of displaying the result, schematically shown in figure 1(C and D) which increases a confidence in the analysis.

The actual shape of the posterior distribution can be visualized, providing a diagnostic tool for determination of which parameters are estimable and which are not. If a parameter is not estimable that will be indicated by a wide 1D posterior (figure 1D). In general, the more information is contained in the prior and the data the narrower the posterior and therefore the method provides a reliable and intuitive tool for visualization of the precision of the estimate represented by the width of posterior distribution.

The classical applications of kinetic compartment modeling is well researched and we briefly describe a general theory in section 2.1. A more complete description is covered in [1]. The Bayesian analysis is a very rich area with many theoretical and computational innovations discovered over the years. Bayesian analysis is applied in many areas of science, engineering, finance, and others [2, 3, 4].

The literature on Bayesian approach to analysis of kinetic models in medical imaging is relatively sparse and recent. In [5] the investigators use estimated posterior for the compartment model-selection task. This approach is useful if the underlying model of the time series is unknown and the most probable model is sought. The method presented in [5] is an approach to find the best model and best parameters corresponding to this model that are consistent with the observed time series. In another application [6] authors investigate Bayesian model of one-compartment data where they model the statistics in the data as well as in the input function. The novelty in this paper, aside from use of Bayesian approach, is the modeling of the noise in the input function. In the overwhelming majority of applications of the compartment modeling in medical imaging the noise in the input function is ignored

because of difficulties in incorporating the noise in the statistical model. In the work presented in this paper, the normality of the data is assumed. In some alternative approaches the Bayesian analysis, the analysis can be performed without the knowledge of the noise model using approximate Bayesian computing (ABC) [7].

In this paper, we focus on how the limitations of the classical approach to modeling stationary systems can be mitigated with Bayesian analysis. We discuss, illustrate and validate a method of Bayesian analysis with simulated data, exemplifying the strengths of the Bayesian approach. In the following sections we detail the implementation of Bayesian methods for two-compartment model and provide validation of the resulting posterior densities. In general, the validation of the posterior is a more difficult task than the validation of the point estimate because the correctness of the entire shape of the distribution needs to be evaluated.

2. Materials and Methods

The mathematical description of the compartmental model and the derivation of the closed form of the posterior distribution from the normal model data and Bayes theorem is given in section 2.1. In section 2.3 we briefly describe the approach for validation of the posterior density distributions which is based on [8]. The numerical Markov chain Monte Carlo (MCMC) methods used to find the approximations of the posterior distributions are given in the appendix 5.1.

Note that we use two different approaches for the validation of posteriors (results of Bayesian methods, section 2.3.1) and for the validation of point estimates (results of classical analysis, section 2.3.2). This is done out of necessity, as the statistical meaning of the results differs for those two approaches.

2.1. Compartment model

We define the instantaneous concentration of tracer in a volume of interest (VOI) as described by a function $C_T(t)$. The tracer molecule may be in different biochemical states (different compartments) in a biological system. Furthermore, we assume that tracer movement between compartments is governed by first-order kinetics. Accordingly, the transfer rates of a tracer molecule from compartment a to a compartment b is linearly proportional to the concentration of the tracer in compartment a and independent of the concentration in compartment b and in any other compartment. The proportionality coefficients between transfer rates and concentrations are referred to as *kinetic parameters*.

In compartment models discussed here, we consider a single mechanism for input of tracer. This will usually be the concentration in the arterial plasma, as it enters the capillary bed. Following the convention frequently used by others we assume the concentration of tracer in the plasma to be known or measurable but we do not require the plasma concentrations to obey compartment definitions. The concentration of the tracer in the plasma is referred to as *input function*. Although in theory the number of compartments can be infinite, the low signal-to-noise of medical imaging data coupled with relatively short measurement periods has limited analyses of kinetic models with more than three tissue compartments.

The two-compartment model pictured in Fig. 2 can be described by a set of ordinary differential equations:

$$\frac{dC_1(t)}{dt} = K_1 C_P(t) + k_4 C_2(t) - [k_2 + k_3] C_1(t) \quad (1)$$

$$\frac{dC_2(t)}{dt} = k_3 C_1(t) - k_4 C_2(t) \quad (2)$$

The total instantaneous concentration in the VOI assuming that all compartments occupy the entire VOI is modeled as

$$C_T(t) = C_1(t) + C_2(t) \quad (3)$$

In this paper we concentrate on PET application and in a typical PET acquisition we do not measure instantaneous concentration. We measure average concentration of the tracer in VOI as y_i during time frames started at t_i and ended at t'_i . Therefore, the model of the measurement during frame i , C_i , using equations 1 through 3 is

$$C_i = \frac{1}{t'_i - t_i} \int_{t_i}^{t'_i} ds C_T(s) \quad (4)$$

2.2. Posterior distribution of kinetic parameters

We assume that measurements y_i are derived from reconstructed PET images by spatial average of voxels contained in a volume of interest (VOI). Index i corresponds to time frames and runs from 1 to I . We further assume that $y_i \sim \mathcal{N}\left(C_i; \frac{C_i \omega^2}{T_i d_i}\right)$ where $\mathcal{N}(\mu; \nu)$ is the normal distribution with mean μ and variance ν . T_i is the i th frame duration and

$d_i = \frac{1}{T_i} \int_{t_i}^{t'_i} dx 2^{x/\tau_{1/2}}$ is the decay correction factor with $\tau_{1/2}$. In the above we introduced parameter ω which is unknown and which will be marginalized. Using the above assumption, we formulate the likelihood $\mathcal{L}(\theta, \omega; \mathbf{y})$ as

$$\mathcal{L}(\theta, \omega; \mathbf{y}) = p(\mathbf{Y} = \mathbf{y} | \theta, \omega) = \prod_{i=1}^I \frac{1}{\sqrt{2\pi\omega^2 C_i(\theta)/(T_i d_i)}} e^{-\frac{(y_i - C_i(\theta))^2}{2\omega^2 C_i(\theta)/(T_i d_i)}} \quad (5)$$

where by θ we indicate the vector of parameters that define the kinetic model specified by k_x s. The symbol ω^2 is used to characterize the absolute scale of the noise variance in the measurement. The symbol $\mathbf{Y}$ used in equation 5 denotes the set of all possible vectors $\mathbf{y}$.

$C_i(\theta)$ is the time-averaged function of θ calculated using equations 1 through 3 and then averaged within the time frame i according to equation 4. We use the notation $\mathbf{Y} = \mathbf{y}$ to denote that $\mathbf{y}$ is measured and known. The set $\{\theta, \omega\}$ constitutes the complete parameter set. We distinguish θ and ω only for clarity of notation. We obtain the posterior of θ and ω using Bayes Theorem

$$p(\theta, \omega | \mathbf{Y} = \mathbf{y}) = \frac{p(\mathbf{Y} = \mathbf{y} | \theta, \omega) p(\theta, \omega)}{p(\mathbf{Y} = \mathbf{y})} \quad (6)$$

In this work we assume a constrained flat prior. By definition $p_F(\theta, \omega) \propto 1$ for $\theta \in \Theta$ and $\omega > 0$ and zero otherwise. Θ is a subset of allowable by the prior parameter values. Other more informative or less informative priors can also be used with our methods if any information about parameters is available before the experiment. In this work we consider the simplest case of the flat prior that indicates a prior belief that any value of parameters is equally probable in allowable range. Considering that $p(\mathbf{Y} = \mathbf{y})$ is constant we obtain the posterior distribution:

$$p_F(\theta, \omega | \mathbf{Y} = \mathbf{y}) \propto \frac{1}{\omega^I} \prod_{i=1}^I \frac{1}{\sqrt{C_i(\omega)/(T_i d_i)}} e^{-\frac{(y_i - C_i(\theta))^2}{2\omega^2 C_i(\theta)/(T_i d_i)}} \text{ for } \theta \in \Theta, \omega > 0 \quad (7)$$

and zero otherwise. The posterior was derived from the improper prior $p_F(\theta, \omega) \propto 1$ and the proportionality sign is carried from this prior. The posterior however is proper (can be normalized to one) but the explicit normalization is not necessary if the posterior is to be estimated using Monte Carlo methods used in this work.

From the posterior expressed by equation 7 different types of Bayesian analysis are derived such as point estimates, confidence sets, hypothesis testing, etc. (for more on this topic see the classical text by Berger [2]). The posterior expressed by equation 7 is multidimensional and as such it is difficult to visualize. Therefore it is processed further in order to create easy to consume information. One way to create a more intuitive representation of the multidimensional posterior is to reduce the dimensionality to 1D or 2D distributions that can be easily visualized and interpreted. A reduction in dimensionality can be achieved by integrating out some of the dimensions, a process called marginalization which is an example of Bayesian analysis. Other Bayesian inferences, such computation of the mean or variance can also be done and also involve integrations over the posterior. Except in the simplest cases, “brute force” integration of the posterior distribution is impractical due to the high dimensionality of the problem. However, an alternative numerical method, Markov Chain Monte Carlo (MCMC) sampling, can be used to obtain adequate approximations. The MCMC method to perform integrations was first used in the 1950's and it has found wide application in science [9]. The MCMC method produces samples from simulated densities (e.g. equation 7) which then can be appropriately averaged to provide the estimate of posterior distribution or desired marginalized distributions or other quantities such means or other expectations.

2.3. Validation

We performed validation of the methods using the Bayesian approach described in section 2.3.1. Using the same data set (creation of the dataset is described in section 2.3.3), we also performed validation of variances obtained by the classical weighted least squares algorithm (section 2.3.2).

2.3.1. Method for posterior validation—The estimates of the posteriors obtained by our method were validated in this work. We closely followed the validation methodology of Cook et al. [8]. In short, we create a computer simulation of L repetitions of an experiment, where an experiment consists of the following steps: (1) Values of parameters θ_0 and ω_0 are drawn from some assumed prior distributions $p(\theta)$ and $p(\omega)$ and L repetitions of this experiment were performed indexed by $l = 1, \dots, L$, (2) Data y_l are generated based on assumed normal model of noise and the kinetic model (equations 1 and 2) for each repetition. (3) For each repetition, R draws from posteriors of θ_k 's and ω using MCMC described in section 5.1 and compute quintiles $q_l(\theta_k)$, $q_l(\omega)$ using

$$q_l(x_0^l) = \frac{1}{R} \sum_{i=1}^R I_{x_0^l > x_i^l} \quad (8)$$

where x_0^l corresponds to true value of either one of elements of θ_0 or value of ω_0 in l th repetition. Value of x_i^l is the result of i th MCMC draw obtained in step 3 for l th repetition. $I_{x_0^l > x_i^l}$ is an indicator function equal to 1 if $x_0^l > x_i^l$ and to 0 otherwise. If the method works properly and the posterior is accurate in describing the probability of the true value of parameters, the $q_l(x_0)$ should be uniformly distributed over $[0, 1]$. After collecting L samples of the quantiles we compute the chi square statistics χ_x^2 with L degrees of freedom for each quantity (elements of θ and ω) using

$$\chi_x^2 = \sum_{l=1}^L \left(\Phi^{-1}(q_l(x)) \right)^2 \quad (9)$$

where Φ represents the standard normal cumulative distribution function. The left-tail P-values for χ_x^2 statistics of each quantity x are computed.

2.3.2. Method for classical variance validation—We also performed validation of the accuracy of covariance matrix estimation using weighted least squares (WLS) using Marquardt-Levenberg nonlinear optimization. The covariance matrix of the solution was estimated using $\widehat{\omega^2} (J^T W J)^{-1}$ where J was Jacobian and W was a diagonal matrix of weights used in the WLS algorithm. In the validation done here an element of W was equal to $w_i = \frac{y_i}{T_i d_i}$. The maximum likelihood estimate $\widehat{\omega^2}$ was obtained and equal to the sum of squares of residuals at the solution divided by the number of degrees of freedom and in our

case equal to $I - 3$ (I is number of time frames). We used the same data as to validate the classical methods as the data used to validate the Bayesian approach. We computed the number of times, n_{obs} , that the true simulated value was inside classical 68% confidence interval. The confidence interval was built assuming normal distribution of the results and normal variance estimated from the diagonal elements of the estimated variance covariance matrix (the typical assumption). The left-tail P-value was computed assuming n_{obs} is a drawn from binomial distribution with probability of success 0.6827. The value of 0.6827 is a fraction of the normal distribution that lie within one standard deviation of the mean.

2.3.3. Irreversible two-tissue compartment model of FDG—We performed validation of the methods described in previous section using a kinetic model of FDG [10, 11]. Equations 1 and 2 were solved using standard numerical methods to obtain the instantaneous compartmental concentrations and then averaged over frame duration, as indicated in 3, yielding a noiseless representation of the tissue data, $\bar{\mathbf{y}}$. Normally distributed noise with variance proportional to $\bar{\mathbf{y}}/(d_i T_i)$ was added. Using the methodology proposed in this paper the estimates of posterior densities of kinetic parameters, ω , as well as the net influx rate of FDG K_i (a macroparameter equal to $K_i = K_1 k_3 / (k_2 + k_3)$) were obtained.

A thousand repetitions ($L = 1000$) of the analysis was performed which resulted in 1000 estimates of the posteriors for each parameters of interest. For the classical analysis described in section 2.3.2 1000 point estimates of parameters and 1000 covariance matrices were obtained. The true values of parameters for every repetition were randomly selected from priors

$p(K_1) = \frac{1}{0.4} I_{K_1 \in [0.1, 0.5]}$, $p(k_2) = \frac{1}{0.4} I_{k_2 \in [0.4, 0.8]}$, $p(k_3) = \frac{1}{0.07} I_{k_3 \in [0.01, 0.08]}$, $p(\omega) = \frac{1}{0.19} I_{\omega \in [0.01, 0.20]}$ where $I_{x \in [x_0, x_1]} = 1$ if $x \in [x_0, x_1]$ and 0 otherwise. We simulated the irreversible model so $k_4 = 0$. The priors were specified based on observed values [13]. The input function was simulated using the model of Feng [12] with value of parameters of the input model are shown in table 1. Ten million MCMC steps were performed for each repetition. Independent samples were selected every 1,000 steps providing a total of $R = 10,000$ samples. Chi square and left-tail P-values for K_1 , k_2 , k_3 , ω , and K_i were computed using the methodology summarized in section 2.3.1 and detailed in [8].

The classical weighted least squares analysis was obtained using Levenberg-Marquardt algorithm [14]. For each repetition 10 starting points were used to ensure the algorithm converged to the same maxima and once the point estimate was established, the variance-covariance matrix was estimated.

3. Results and discussion

Figure 3 presents examples of estimates of posterior densities obtained for two noise levels corresponding to $\omega = 0.0272 \text{ min}^{1/2}$ for low noise and $\omega = 0.1758 \text{ min}^{1/2}$ for high noise level. The true values that generated the data are visualized as well in figure 3 by downward pointing arrows. Scales of x-axes were selected to cover all allowable prior values of parameters. For the high noise, k_2 was difficult to estimate, as indicated by flat posterior density on the entire range allowable by the prior. The horizontal bars in figure 3 represent classical estimates obtained by weighted least squares algorithm. The width of the bars

correspond to standard deviations of those estimates obtained as square roots of diagonal elements of the covariance matrix.

Table 2 presents the χ^2 values and P-values for the posterior validation experiment described in section 2.3.1. No significant deviations of χ^2 statistics were observed, strongly suggesting that our method of posterior estimation performs well and represents an accurate posterior density.

The evaluation of the validity of classical variance estimate (section 2.3.2) is also shown in table 2 in the fourth and fifth column. It is clear from the P-values corresponding to the hypothesis that n_{obs} is drawn from a binomial distribution can be rejected. This implies that either the WLS point estimate is biased or the covariance estimate is inaccurate or some combination of both. The results suggest that the variance is actually underestimated and should be larger. In the 68%-confidence interval that was used, n_{obs} is expected to be around 680 and as seen in table 2 it is statistically significantly smaller. This underestimation is a consistent finding for all parameters.

We have presented a validation of kinetic analysis based on Bayesian statistics. The method is illustrated with applications to simulated data in which the “truth” is known. The final result of Bayesian analysis is the multidimensional posterior distribution (it is not a point estimate as in the case of standard methods based on optimization). The number of dimensions is equal to the number of unknown parameters. In this paper we used methods that estimated this distribution.

The Bayesian method is conceptually different than the classical “fitting” method, but both methods rely on identical likelihood functions (see equation 5). The classical methods used in kinetic modeling provide point estimates as well as the estimation of the covariances associated with the point estimates. The covariance describes the “spread” of point estimates if exactly the same fits are repeated an infinite number of times with different noise realizations added to hypothetical noiseless data. In contrast, in the Bayesian approach we only consider the data at hand without resorting to hypothetical repetition. In the Bayesian method, the results are not represented by “best fit” - values of parameters (point estimate), but instead by a probability distribution (the posterior) which represents beliefs (or a level of uncertainty) about the true values of parameters. In the Bayesian view, parameters are constant but the true values of parameters are uncertain and the formalism of the Bayesian statistical theory describes this uncertainty.

We have shown how the posterior distribution can be used to quantify the uncertainty in our knowledge of the model parameters in well-known situations described by the Sokoloff FDG model. In those examples we marginalized the posterior to study the uncertainty in a single kinetic parameter or macroparameter (figure 3). As explained in methods section, it is often useful to marginalize the posterior to a fewer-dimensional (typically 1 or 2) *marginalized posterior*. Marginalization of some parameters often makes it easier to understand the results of the Bayesian analysis. In the example with FDG (figure 3) the original posterior has four dimensions corresponding to K_1 , k_2 , k_3 and ω . This distribution can be marginalized to 1D or 2D representations that can conveniently be presented to a decision maker. The 1D and

2D marginalized distributions for a given parameter (or macroparameter) directly illustrates the extent of the uncertainty about this parameter (macroparameter) of interest. These methods can be used to study how various processing methods, noise levels, model, constraints (priors), etc., affect the uncertainty of the posterior of parameters or macroparameters, thus providing a method for evaluation of various processing approaches. We believe that this property of the Bayesian methods is their main strength.

We note here that the classical fitting methods provide estimates of covariances which can also provide some insights about reliability; however, as shown in this work these estimates can be inaccurate for non-linear optimization models used in compartmental analyses. This lack of accuracy most likely comes from approximations needed in order to estimate the covariance matrix for the non-linear case. This lack of accuracy (underestimation of the errors) in classical variance estimations in non-linear regression shown in our results is consistent with findings of other investigators ([15, 16]). Although not typically applied in kinetic analysis, the classical confidence intervals can be more accurately estimated using bootstrapping ([17]), but we have not investigated this approach in the current paper.

Another very important strength of the Bayesian approach is the natural and simple extension to Bayesian decision theory [2]. Difficult to put into practice concepts of multiple hypothesis testing can be readily implemented in the Bayesian setting. An example of a three-hypothesis problem can be defined for FDG influx rate as $H_1: K_i \in [0, 0.1]$, $H_2: K_i \in (0.1, 0.5]$, $H_3: K_i \in (0.5, \infty]$. The posterior probabilities of those hypotheses can be easily computed if the marginalized posterior is estimated. These probabilities combined with the loss function would provide a Bayesianly optimal decision in favor of one of the hypothesis H_1 through H_3 . The signal detection which can be regarded as a type of hypothesis testing can be put into practice with ease as well. The signal present/absent corresponds to a simple binary hypothesis. Other types of Bayesian analysis such as odd ratios can also be used.

The method presented here can be extended to more complex analyses. A natural extension is the inclusion of the noise from the input function in the posterior. Currently, we assume that the input function does not have noise or any other uncertainty. Other uncertainty may be coming from the uncertainty in the model for example. We will consider these extensions in the future work.

In the current implementation, we used a flat prior prescribing the initial belief that none of the values of parameters is more/less probable in a set of allowable parameters. We used such a prior, not as a reflection of our actual before-the-experiment beliefs; but rather, more out of convenience, to simplify the discussion and to provide an illustrative implementation. In many cases we will have information that will allow us to specify more informative priors. For example, if we have results of previous experiments the prior beliefs can be specified with more accuracy based on approximate feasible ranges obtained in the past. The incorporation of those “more informative” priors is quite straightforward in the context of the general MCMC theory (section 5.1).

The assumption of normality of the data is an approximation. The actual noise model in real imaging data is typically unknown and depends on a large number of factors including

imaging modality, scanner, reconstruction algorithm, the noise model in the raw data [18], etc. It is our belief that the normal model of the data with a hierarchical prior put on unknown variance as used in this work may be sufficient in practical applications to describe the data variability in nuclear imaging applications. Obviously, this is only a speculation and more research is needed in this area which we plan to undertake in the future. An interesting continuation of this work is to use the approximate Bayesian computation (ABC) [7] in which we do not have to assume the normal model of the data and to compare ABC-results to the results of methods developed in this work.

4. Conclusions

In summary, the Bayesian approach presented in this paper is a new approach that can be used for analysis of stationary compartment models. The result of the Bayesian method is the posterior distribution which provides intuitive graphical representation of the uncertainty about parameters and macroparameters. The validity of the posterior distributions which are the results of the Bayesian analysis was validated for the computer simulation of the two-compartment model of FDG. It is shown that Bayesian estimations of uncertainty are more accurate compared to estimations obtained by the standard least-square approaches.

Acknowledgments

This work was supported by grants from the National Institutes of Health Nos. R21HL106474, R01CA165221, R01HL110241, and S10RR031551.

References

1. Gunn RN, Gunn SR, Turkheimer FE, Aston JAD, Cunningham VJ. Positron emission tomography compartmental models: a basis pursuit strategy for kinetic modeling. *J Cereb Blood Flow Metab.* 2002; 22:1425–1439. [PubMed: 12468888]
2. Berger, JO. Statistical Decision Theory and Bayesian Analysis. Springer Science & Business Media; 2013.
3. Gelman, A., Carlin, JB., Stern, HS., Dunson, DB., Vehtari, A., Rubin, DB. Bayesian Data Analysis. 3. CRC Press; 2013.
4. Robert, C. The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation. Springer Science & Business Media; 2007.
5. Zhou Y, Aston JAD, Johansen AM. Bayesian model comparison for compartmental models with applications in positron emission tomography. *Journal of Applied Statistics.* 2013; 40:993–1016.
6. Malave P, Sitek A. Bayesian Analysis of a One Compartment Kinetic Model Used in Medical Imaging. *J Appl Stat.* 2015; 42:98–113. [PubMed: 25408561]
7. Fan Y, Meikle S, Angelis G, Sitek A. ABC in nuclear imaging. 2016 arXiv:1607.08678.
8. Cook SR, Gelman A, Rubin DB. Validation of software for Bayesian models using posterior quantiles. *Journal of Computational and Graphical Statistics.* 2006; 15:675–692.
9. Metropolis N, Ulam S. The Monte Carlo Method. *Journal of the American Statistical Association.* 1949; 44:335–341. [PubMed: 18139350]
10. Sokoloff L, Reivich M, Kennedy C, Des Rosiers MH, Patlak CS, Pettigrew KD, Sakurada O, Shinohara M. The [¹⁴C]deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure, and normal values in the conscious and anesthetized albino rat. *J Neurochem.* 1977; 28:897–916. [PubMed: 864466]
11. Huang SC, Phelps ME, Hoffman EJ, Sideris K, Selin CJ, Kuhl DE. Noninvasive determination of local cerebral metabolic rate of glucose in man. *Am J Physiol.* 1980; 238:E69–82. [PubMed: 6965568]

12. Feng D, Huang SC, 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. [PubMed: 8449593]
13. Strauss LG, Pan L, Cheng C, Dimitrakopoulou-Strauss A. (18)F-Deoxyglucose (FDG) kinetics evaluated by a non-compartment model based on a linear regression function using a computer based simulation: correlation with the parameters of the two-tissue compartment model. *Am J Nucl Med Mol Imaging.* 2012; 2:448–457. [PubMed: 23145361]
14. Marquardt D. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics.* 1963; 11:431–441.
15. Simonoff JS, Tsai C-L. Jackknife-Based Estimators and Confidence Regions in Nonlinear Regression. *Technometrics.* 1986; 28:103–112.
16. Peddada SD, Haseman JK. Analysis of nonlinear regression models: a cautionary note. *Dose Response.* 2005; 3:342–352.
17. Niedzwiecki D, Simonoff JS. Estimation and inference in pharmacokinetic models: the effectiveness of model reformulation and resampling methods for functions of parameters. *J Pharmacokinet Biopharm.* 1990; 18:361–377. [PubMed: 2231325]
18. Sitek A, Celler AM. Limitations of Poisson statistics in describing radioactive decay. *Physica Medica.* 2015; 31:1105–1107. [PubMed: 26508015]
19. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics.* 1953; 21:1087–1092.
20. Roberts GO, Gelman A, Gilks WR. Weak convergence and optimal scaling of random walk Metropolis algorithms. *Ann Appl Probab.* 1997; 7:110–120.
21. Garthwaite PH, Fan Y, Sisson SA. Adaptive optimal scaling of Metropolis-Hastings algorithms using the Robbins-Monro process. *Communications in Statistics - Theory and Methods.* 2015; 45:5098–5111.

5. Appendix

5.1. Monte Carlo methods

In order to derive inferences about θ and marginalize ω , we used methods similar to those described in our previous work [6] where we used a convolution model of $C_T(t)$ and applied it to a one-tissue compartment model. In this paper, we solve the set of differential equations (equations 1 and 2) and then perform integral (equation 4). We do not assume that ω is known (as in [6]) but marginalize it. We used the Runge-Kutta algorithm to solve the set of equations 1 and 2 and the integral expressed by equation 4 was approximated by taking the average value of $C_T(t)$ at three points uniformly distributed over each time interval i . These three values were then averaged to provide an estimate of the integral in equation 4. We tested the approximation by using five and seven sampling points but no significant differences in inferences were observed. There were slight differences between using one and three sampling points which is why we chose to use three-sample approximation of the integral in equation 4.

In general the MCMC with Metropolis-Hastings sampling consists of two parts. First, new values of parameters are proposed based on current values (selection step) and second, these new values are accepted or not (acceptance step). If new parameters are not accepted the values remain the same as the current values and if they are accepted old values are replaced by new values and the Markov chain proceeds to the next step. In the following we describe in detail the selection and acceptance steps of the algorithm used in this paper. Suppose that

we indicate values of parameters at the s th Markov step as θ^s, ω^s . We propose new values of parameters for the next step, $s+1$, by modifying just one of elements of θ^s or the value of ω^s . This is done by selecting a random integer number $r \in \{1, \dots, K+1\}$. If $r \in \{1, \dots, K\}$ is obtained then the corresponding element of θ is modified; whereas if $r = K+1$, then ω is modified. The modification is done by adding a value of δ_r^s which is a random number from range $[-\delta_r, \delta_r]$ to the current value of parameters. Therefore, δ_r is the *step size* value which in general is different for each parameter. Below, we describe how these step sizes are optimized.

The Θ is a set of vectors θ defined by general inequality constraints $a_k \leq \theta_k \leq b_k$ where a_k and b_k are lower and higher bound for θ_k , this is consider a prior. Although in this paper we use a particular form of prior, other priors can also be implemented in a similar manner. In particular either of these values $\{a_k, b_k\}$ can be equal to some constant (for example $a_k = 0$) or to $\theta_{k'}$ with $k' \neq k$. These constraints are implemented in the algorithm in the selection steps as follows. If the value of $\theta_r^s + \delta_r^s < a_k$ or $\theta_r^s + \delta_r^s > b_k$, then the following selection is made $\theta_r^{s+1} = a_k - (\theta_r^s + \delta_r^s)$ and $\theta_r^{s+1} = 2b_k - (\theta_r^s + \delta_r^s)$. This guarantees the detailed balance [6] of the Markov Chain and ensures that $a_k \leq \theta_r^{s+1} \leq b_k$ if for each k $b_k - a_k$ which we enforce. Similar considerations apply to constraints on parameter ω .

The newly selected set of parameters $\theta^{s+1}, \omega^{s+1}$ is accepted using the Metropolis algorithm [19], that is, it is accepted stochastically with probability

$$\min \left(1, \frac{p_F(\theta^{s+1}, \omega^{s+1} | \mathbf{Y} = \mathbf{y})}{p_F(\theta^s, \omega^s | \mathbf{Y} = \mathbf{y})} \right) \quad (10)$$

where $p_F(\theta^{s+1}, \omega^{s+1} | \mathbf{Y} = \mathbf{y})$ is given by equation 7. A Markov chain as specified above is guaranteed to converge to equilibrium because the condition of detailed balance is fulfilled and while in equilibrium samples from $p_F(\theta^{s+1}, \omega^{s+1} | \mathbf{Y} = \mathbf{y})$ are obtained. Once samples are obtained any integral that involves the posterior can be estimated.

How fast the MC reaches equilibrium, and how efficiently independent samples are obtained once in equilibrium, depends on chain *efficiency*. In this work we optimized the efficiency by using adaptive step sizes δ_k 's and assuming that chain efficiency is closely related to the acceptance rate of Markov steps [20]. The burn-in run (Markov steps used to allow the algorithm to reach equilibrium) was divided into two parts. In the first part the algorithm was run and the ratio α_k of accepted steps was computed for each element of θ and ω every 1000 steps for each parameter. If $\alpha_k < 0.2$ or $\alpha_k > 0.5$ the step size δ_k was decreased or increased by 5%, respectively. Smaller or larger values of step sizes tend to increase or decrease acceptance rate, respectively. Values of 0.2 and 0.5 used in this work were chosen based on theoretical studies of multivariate Gaussians [20]. The value of 5% used to increase/decrease the step size was chosen and the robustness of this parameter was investigated. We found no significant difference in results when varying from 2% to 10% as long as burn-in time was sufficient. In the current implementation we allowed ample time for burn-in's (100,000

steps). The validity of this section was confirmed by the resulting posteriors which were successfully validated. After every adjustment of the step size, we also performed $\alpha_k = \min(\alpha_k, b_k - a_k)$ to ensure that the step size is equal or smaller than allowed size of range for this parameter to avoid numerical instabilities. The reason for this precaution is that if for example $\alpha_k \gg b_k - a_k$ the application of constraints described earlier in this appendix may create serious numerical inefficiencies as the algorithm for step selection would have to be repeated iteratively a large number of times. Some more sophisticated algorithms can be used to optimize the efficiency of the MCMC [21].

The burn-in with the adjustment in step size was followed by the classical burn-in in which the step size was kept constant. Keeping constant step size enforces the detailed balance and assures that the algorithm achieves equilibrium. In all cases investigated, we observed a rapid convergence of the chain to equilibrium and we have not, so far, experienced a failure of the algorithm. When the algorithm reaches equilibrium, 10,000 samples obtained every 1000 steps of Markov chain of the parameters were saved (10^7 total steps). The parameter values were then histogrammed, providing estimates of the posterior distributions. We calculated the mean of the posteriors by averaging of the samples. This mean is the minimum-mean-square-error (MMSE) point estimate of the parameters. For the marginalized 1D posteriors we also compute the interval estimator which corresponds to the shortest interval that contains 95% of samples. We refer to this interval as highest posterior density (HPD) interval.

Highlights

- New Bayesian methods for analysis of compartmental kinetic models
- Comparison of Bayesian and classical methods for estimation of uncertainty
- Bayesian analysis performs better than classical approach

Figure 1.

Standard analysis (A) provides a point estimate of the true value of a parameter and the standard error indicated by the bar. Bayesian approach provides a posterior density distribution estimate (B) indicative of the beliefs that a given parameter value is true. (C) and (D) present examples of an informative narrow posterior and non-informative posterior indicating non-estimable parameter, respectively. Max value is the known upper limit value for a given parameter.

Figure 2.

Two-tissue compartment model. One-tissue compartment model is obtained by setting k_3 and k_4 to zero. $C_P(t)$ and $C_T(t) = C_1(t) + C_2(t)$ are plasma concentrations and the concentration of the tracer in VOI which is the sum of concentrations in tissue compartments 1 and 2.

Figure 3.

Examples of two curves and estimated posterior densities for two repetitions with K_1 , k_2 , k_3 , K_i and ω equal to 0.37, 0.44, 0.044, 0.034, 0.028 and 0.16, 0.42, 0.077, 0.024, 0.18 for low (black color) and high noise (gray color), respectively. The true values are marked with arrows. Horizontal lines correspond to classical weighted least squares estimates with estimated standard deviations

Table 1

Parameters of the input function used in section 2.3.3 from Feng et al.[12].

Parameter	A_1	A_2	A_3	λ_1	λ_2	λ_3
Value	12	1.8	0.45	4	0.5	0.008

Table 2

Summary of the results. Columns 2 and 3 show values of χ^2 for 1000 degrees of freedom for micro- and macro-parameters of the FDG kinetic model and the noise parameter (Expected value of χ^2 is 1000). Columns 4 and 5 show number of observations for which the true value of parameter fell within the 68% confidence interval and P-value corresponding to hypothesis that n_{obs} is binomially distributed with chance of success 0.68 and 1000 trials. N/C=not computed because classical confidence intervals were not computed for these values.

Quantity	Bayesian		Weighted LS	
	χ^2_{obs}	$p(\chi^2 \leq \chi^2_{obs})$	n_{obs}	$p(n_{obs})$
K_1	1064	0.92	589	8.9×10^{-16}
k_2	991	0.43	616	3.1×10^{-18}
k_3	1010	0.60	641	2.8×10^{-3}
$K_i = \frac{K_1 k_3}{k_2 + k_3}$	1056	0.89	N/C	N/C
ω	1008	0.57	N/C	N/C