Preprints of the 10th IFAC International Symposium on Dynamics and Control of Process Systems
The International Federation of Automatic Control
December 18-20, 2013. Mumbai, India

Parameter Estimation for Physiologically Based Pharmacokinetics Model Using Bayesian Inference

Dae Shik Kim, * **Jong Hwan Sung,** ** **Jong Min Lee** [1] *

* *School of Chemical and Biological Engineering, Seoul National University, 1, Gwanak-ro, Gwanak-Gu, Seoul, 151-744, Korea*
** *School of Chemical System Engineering and Materials Science, Hongik University, Wausan-ro, Mapo-Gu, Seoul, 121-791, Korea*

Abstract: Physiologically based pharmacokinetics(PBPK) model can predict absorption, degradation, execration and other metabolism in drug delivery system. Thus it can be useful for regulating dose and estimating drug concentration at a particular time during the clinical demonstration. PBPK model is expressed as a set of differential equation with various parameters. Bio-chip experimental data are often noisy and sparse. This makes it difficult to estimate parameters with conventional least squares approaches. The resulting parameters often have a large confidence region. This work presents a Bayesian inference algorithm with an objective function suitable for PBPK model. A Markove Chain Monte Carlo(MCMC) method is employed to estimate the posterior distribution of the parameters. We illustrate the approach with a Tegafur delivery system.

Keywords: Bayesian inference, PBPK model, Drug delivery system, MCMC simulation, Tegafur, Maximum a posteriori method

1. INTRODUCTION

Developing a new drug takes an enormous amount time, money, and effort, mainly because of bottlenecks in the drug discovery process and clinical demonstration. Mathematical models describing drug delivery mechanism in terms of drug concentrations in each organ over the time course can be of significant help in reducing costs and risks. Pharmacokinetics is the study of the course of absorption, distribution, metabolism, and elimination of some substance in a living body and is especially important in the development of drugs(Lindsey et al. (2000)). Not only can we use it for prediction of dynamics of drug delivery; we can also apply it to dose regulation. For these reasons, during development of new drugs, data are collected to construct physiologically based pharmacokinetics (PBPK) models during animal and human trials (Phase I-III) (Gehring et al. (1979)). Experiments for collecting dynamic bio-chip data are expensive and often have poor repeatability. Estimating parameters of a PBPK model with such data set is further complicated by the concentration profiles showing a mixture pattern of declining exponential functions, with the amplitudes and decay times of the different components corresponding to functions of the model parameters (Gelman et al. (1996)). In addition, each individual may have different parameter values depending on their characteristic properties of body. This study presents a Bayesian inference scheme for robust parameter estimation of PBPK model to address the difficulties. In addition, we apply this scheme to Tegafur delivery system. The Bayesian inference scheme is summarized below.

[1] Corresponding author: Jong Min Lee (e-mail: jongmin@snu.ac.kr)

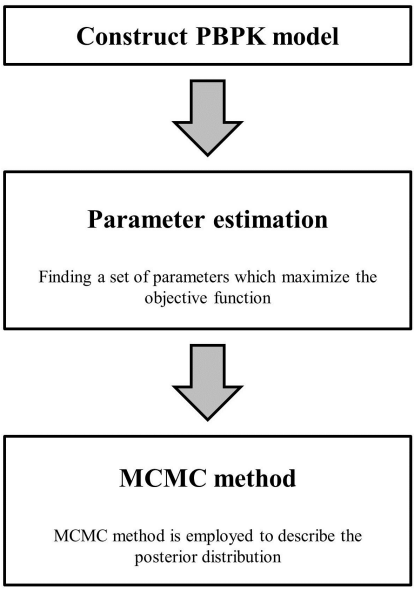

Fig. 1. We construct PBPK model for target system and derive the objective function from Bayes' rule to estimate parameters which maximize the objective function. With this estimation result, we employ MCMC method to estimate the porsterior distribution.

2. PHYSIOLOGICALLY BASED PHARMACOKINETICS MODEL

Physiologically based pharmacokinetics(PBPK) model is the mathematical method to describe concentration of

medicine at each organ. The medicine is dissolved into vessel and transported by blood circulation. At some organs which can perform degradation or clearance, the medicine is transformed into other substance or is excreted from the body. These several metabolisms in each organ can be described as mass balance equations, overall change of medicine concentration in the form of a set of differential equations. The basic mass balance equation of pharmacokinetics model is following(Sung et al. (2009)).

$$V \cdot \frac{dC}{dt} = Q \cdot (C_{in} - \frac{C}{P}) - R_e \tag{1}$$

where C is concentration of medicine, Q is the volumetric flow rate of blood in each organ, C_{in} is the inlet concentration of medicine, P is the tissue/blood partition coefficient of the organ, and R_e is the metabolism rate.

If the drug is injected into a target place, initial concentration should be equal to the dose of the drug. However, in the case of oral administration, The drug dissolves slowly in the internal organ. For drug dissolution model, a simple first order kinetics, referred to as Noyes-Whitney equation, can be used(Costa and Lobo (2001)).

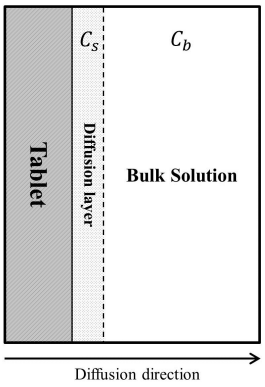

Fig. 2. First order drug dissolution model for drug delivery system. Following the diffusion direction, there exist diffusion layer and bulk solution where concentration of drug is uniformly distributed.

$$\frac{dW}{dt} = \frac{D \cdot A \cdot (C_s - C_b)}{L} \tag{2}$$

Equation (2) is the first order model for drug dissolution where $\frac{dW}{dt}$ is the dissolution rate, C_s is the concentration of the drug in the diffusion layer, C_b is the concentration of the drug in bulk solution, D is the diffusion coefficient, and L is the diffusion layer thickness.

With these two kinds of model equations, we can set up a model for drug delivery system described as the set of differential equations.

3. BAYESIAN INFERENCE

After setting up a PBPK model, we need to estimate unknown parameters with experimental data. However, pharmaceutical experimental data are difficult to obtain and different resulting parameters can be obtained for each test subject. Therefore, a robust parameter estimation technique suitable for a small data set is required. Although the least squares and maximum likelihood estimation method are widely used for parameter estimation, these are not appropriate for PBPK model. Least squares methods only minimize about summation of squared errors. When the number of data is small, it is sensitive to noisy data or outliers. Maximum likelihood estimation may show poor accuracy when the data set is small since it is effective when the data size tends to go infinity(Jang and Gopaluni (2011)). In order to address these difficulties, we propose a Bayesian inference scheme for parameter estimation of PBPK model.

The Bayes' rule is expressed in the following equation.

$$P(\theta|Y) = \frac{P(\theta) \cdot P(Y|\theta)}{\int P(\theta) \cdot P(Y|\theta) d\theta} \tag{3}$$

where θ is the parameter vector to be estimated, Y is the observed data. $P(x|Y)$ is the 'posterior distribution' and $P(\theta)$ is the 'prior distribution' which describes the information of prior knowledge of parameters. $P(Y|\theta)$ is 'likelihood'. The denominator of equation is a normalizing factor. Therefore, we can describe that 'posterior distribution' is proportional to the product of 'prior distribution' and 'likelihood' (Bonate (2006)).

$$Posterior \propto Prior \cdot Likelihood \tag{4}$$

To estimated unknown parameters, we use a 'maximum a posteriori probability(MAP)', which maximizes the posterior probability.

If we have information about a reliable value of each parameter, we can assume that prior distribution follows a Gaussian distribution. Without such prior information, one can also assume that prior distribution follow a uniform distribution between upper and lower limits of parameter.

$$Prior\ distribution = \prod_{i=1}^{n} \frac{1}{(\theta_{max,i} - \theta_{min,i})} \tag{5}$$

where θ_{min} is the set of lower limit of each parameter and θ_{max} is the set of upper limit of each parameter.

If the error between actual concentration and predicted one is assumed to follow the Gaussian distribution, the likelihood term is given by

$$Likelihood = \prod_{i=1}^{m} \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2}} \cdot exp\{-\frac{(Y_i - Y'_i)^2}{2 \cdot \sigma^2}\} \tag{6}$$

$$= \frac{1}{(2 \cdot \pi \cdot \sigma^2)^{\frac{m}{2}}} \cdot exp\{-\sum_{i=1}^{m} \frac{(Y_i - Y'_i)^2\}}{2 \cdot \sigma^2} \tag{7}$$

where σ is the standard deviation of residuals, and m is the number of experimental data. $Y_i - Y'_i$ is residual between observed and predicted values, respectively. The objective function is the product of prior distribution and likelihood. To simplify this objective function, we take negative logarithm. Consequently, the final form of objective function as follows.

$$LnP = \sum_{i=1}^{n} \ln(\theta_{max,i} - \theta_{min,i}) + \frac{m}{2} \cdot \ln(2 \cdot \pi \cdot \sigma^2) + \frac{1}{2 \cdot \sigma^2} \cdot \sum_{i=1}^{m} (Y_i - Y'_i)^2 \quad (8)$$

Our goal is to find the set of parameters that maximize the posterior distribution. Since we take minus logarithm for the objective function, we need to find the set of parameters that minimize the objective function value. Since most of the PBPK model do not have analytic solution and have many parameters to estimate, there can be a large number of optima. To find global minimum point, a 'Particle Swarm Optimization'(PSO), scheme(Schwaab et al. (2008)), is employed. The proposed approach for estimating MAP parameter is summarized as follows:

(1) Determine the number of particles, $p_{i,j}$, where i is the number of iterations and j is the particle index.
(2) Set up the initial position and velocity of $p_{i,j}$ randomly.
(3) Move $p_{i,j}$ with its own velocity and compute objective function value at each position.
(4) Find individual minimum point of each particle, $p_{ind,i,j}$.
(5) Find global minimum point, $p_{glo,i}$, which is the minimum point of individual minimum points.
(6) If the objective function value at global minimum point is greater than the previous global minimum point, let $p_{glo,i} = p_{glo,i-1}$.
(7) Set up a new initial position and velocity considering the individual minimum and global minimum points.
(8) Return to (3) and repeat until no further improvement is achieved

4. MARKOV CHAIN MONTE CARLO METHOD

The result of Bayesian inference is the posterior distribution. Since each patient can have different kinetic parameters of enzyme, It is important to know about the posterior distribution of model parameters. The posterior distribution is useful for setting up the optimal dose of drug for general case to prevent either side effect of overdose or under-dose. However, since the product of prior distribution and likelihood is too complex to calculate due to the integral term of equation (3) numerically, it is difficult to know about exact numerical value of posterior distribution. For this reason, we employ a Markov Chain Monte Carlo(MCMC) method to estimate the posterior distribution of the resulting estimate. Suppose that we can construct a Morkov chain with state space which has a equilibrium distribution. If we run the chain for a long time, simulated values of the chain can be used as a basis for summarizing features of the probability distribution of interest (Smith and Roberts (1993)).

There are various algorithms for MCMC method. This study uses the Metropolis-Hastings algorithm. We suppose the proposal probability density function(p.d.f) is symmetric and closed form. From this proposal p.d.f, we obtain samples and run the algorithm (Chib and Greenberg (1995)). The Metropolis-Hastings algorithm is summarized below.

(1) Draw a new proposal state, x', from the proposal p.d.f.
(2) Calculate $\alpha = min\{1, \frac{P(x'|D)}{P(x_t|D)} \cdot \frac{Q(x_t|x')}{Q(x'|x_t)}\}$ where x_t is the previous state, Q is the proposal p.d.f, and P is the posterior distribution.
(3) When $\alpha \geq 1$, then $x_{t+1} = x'$.
(4) When $\alpha \leq 1$, then we choose $x_{t+1} = x'$ with the probability of α or $x_{t+1} = x_t$ with the probability of $1 - \alpha$.
(5) Return to (1) and repeat until the distribution is converged.

Since $\frac{P(x'|D)}{P(x_t|D)}$ can be calculated only for the proportion of the prior distribution and likelihood, we can calculate numerically posterior distribution without calculating integral term of posterior distribution with Metropolis-Hastings algorithms. From this posterior distribution, we can predict valid range of each parameter of PBPK model.

5. CASE STUDY: TEGAFUR DRUG DELIVERY SYSTEM

Tegafur is widely used in the treatment of a range of cancers, especially of colorectal cancer (Longley and Johnston (2007)). Tegafur is the oral administration drug and transform to 5-fluorouracil by CYP450 enzyme at liver(Sung et al. (2009)), thereby it can perform pharmacological action.

5.1 The PBPK modeling for Tegafur delivery system

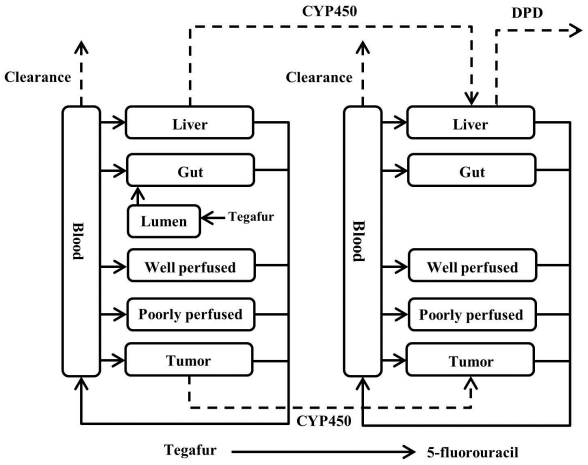

Fig. 3. Tegafur is orally administrated for patients and absorbed into lumen. At liver, Tegafur is transformed to 5-fluorouracil by CYP450 and also 5-fluorouracil degraded by DPD. At blood, both of Tegafur and 5-fluorouracil cleared out from blood. At tumor, the same metabolism is working with that of the liver.

To set up the PBPK model for Tegafur drug delivery system, human body is split into each organ part. The most important parts of the body are liver and tumor where transformation from Tegafur to 5-fluorouracil is occurred. In addition, oral administrated Tegafur dissolves inside of the body and is absorbed at lumen, gut. Drug is delivered by blood and also cleared out at blood. Therefore, The PBPK model can be constructed as in Figure 3. Transformation and degradation are described

Table 1. Parameters to estimate

Parameter	Description
$K_{ml,T}(nmol/min/g$ tissue)	V_{max} for CYP450 enzyme in liver
$V_{ml,T}(nmol/ml)$	Michaelis-Menten constant for CYP450 enzyme in liver
$K_{mt,T}(nmol/min/gtissue)$	V_{max} for CYP450 enzyme in tumor
$V_{mt,T}(nmol/ml)$	Michaelis Menten-constant for CYP450 enzyme in tumor
$K_{ml,FU}(nmol/min/gtissue)$	V_{max} for DPD enzyme in liver
$V_{ml,FU}(nmol/ml)$	Michaelis-Menten constant for DPD enzyme in liver
$K_{mt,FU}(nmol/min/gtissue)$	V_{max} for DPD enzyme in liver
$V_{mt,FU}(nmol/ml)$	Michaelis Menten-constant for DPD enzyme in tumor
$k_{abs}(min^{-1})$	Absorption coefficient of Tegafur
$k(min^{-1})$	Dissolution coefficient of Tegafur
$CL_T(ml/min)$	Clearance rate of Tegafur from plasma
$CL_{FU}(ml/min)$	Clearance rate of 5-flourouracil from plasma

Table 2. Organ volume and blood volumetric flowrate)

Organ	Organ volume(V, ml)	Blood flow rate(Q, ml/min)
Blood(V_b, Q_b)	13.2	76.45
Gut(V_g, Q_g)	7.92	17.1
Liver(V_l, Q_l)	8.8	19
Tumor(V_t, Q_t)	1.0	0.25
Well perfused organs(V_w, Q_w)	8.5	38.9
Poorly perfused organs(V_p, Q_p)	165	18.3

with Michaelis-Menten equation. With Eqs. (1) and (2) the Tegafur delivery system is described as 12 differential equations at each organ.

The notations and determined parameter values are presented in Tables 1, 2, and 3.

Table 3. Tissue/blood partition coefficient

Organ	Tegafur(T)	5-fluorouacil(FU)
Blood(P_b)	0.808	0.794
Gut(P_g)	0.768	0.759
Liver(P_l)	0.895	0.5
Tumor(P_t)	0.336	0.169
Well perfused organs(P_w)	0.834	0.826
Poorly perfused organs(P_p)	0.8	0.795

5.2 The result of Bayesian inference

We used a bio-chip set up an experimental environment similar to the internal body and obtain dynamic drug concentration data. The bio-chip consists of micro organ cells connected by blood vessel (order of micrometers) which copy the real organ.

With PBPK model for Tegafur drug delivery system and experimental data from the bio-chip, we estimate 12 unknown parameters given in Table 1. The bio-chip consists of the organ cells and blood vessel of a rat, and the concentration of Tegafur and 5-fluorouracil at 0.5, 1, 2, 4 hours were measured from gut, liver, tumor cell and blood. The initial Tegafur dose was 15 mg/kg. Since we don't have any prior knowledge of these parameters, uniform distributions were used as the priori distributions, with 32 data points. The estimation result and concentration profiles at each organ are given in Table 4 and Figures 4-11.

5.3 The result of MCMC simulation

To figure out the posterior distribution for the estimation result of Bayesian inference, we conducted MCMC simulation with Metropolis-Hastings algorithm. Iteration of MCMC simulation is 100,000 and , latter 10,000 runs were accepted as the converged posterior distribution. The joint distributions of every two parameters which is very complex distribution are described at figure 12.

Table 4. The estimation result of undetermined parameters

Parameter	Estimated value
$K_{ml,T}$	2.561×10^4 $nmol/min/g$ tissue
$V_{ml,T}$	3.653×10^3 $nmol/ml$
$K_{mt,T}$	2.856×10^4 $nmol/min/gtissue$
$V_{mt,T}$	4.088×10^4 $nmol/ml$
$K_{ml,FU}$	6.319×10 $nmol/min/gtissue$
$V_{ml,FU}$	9.462×10^3 $nmol/ml$
$K_{mt,FU}$	4.687 $nmol/min/gtissue$
$V_{mt,FU}$	2.279×10 $nmol/ml$
k_{abs}	6.768×10 min^{-1}
k	12.8 min^{-1}
CL_T	0.629 ml/min
CL_{FU}	2.203×10 ml/min

6. CONCLUSION

In this study, a Bayesian parameter estimation method for PBPK model by finding maximum point of the objective function is introduced. Despite the large number of unknown parameters, the estimation result is well fitted with the experimental data. Furthermore, it can be helpful for regulating dose for different groups of patients since the estimation result is described as a probability distribution form.

ACKNOWLEDGEMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(grant number 0458-20120012)

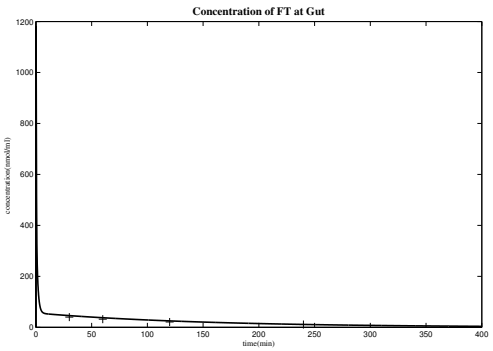

Fig. 4. The Concentration of Tegafur at gut.

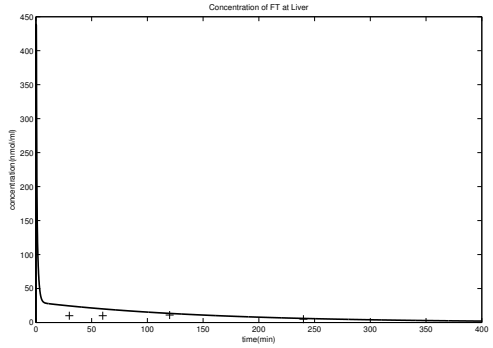

Fig. 5. The Concentration of Tegafur at liver.

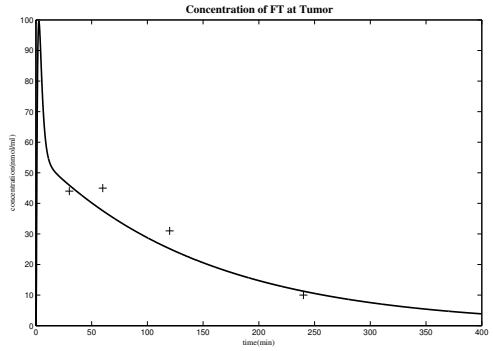

Fig. 6. The Concentration of Tegafur at tumor.

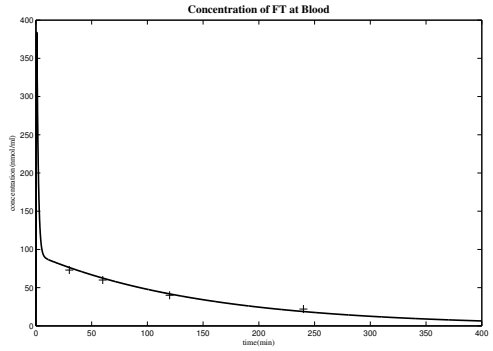

Fig. 7. The Concentration of Tegafur at blood.

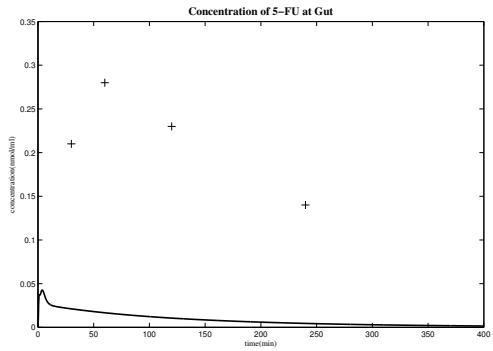

Fig. 8. The Concentration of 5-fluorouracil at gut.

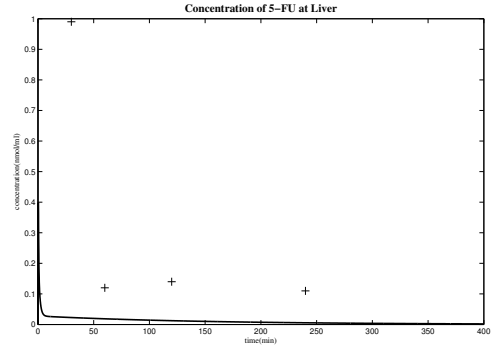

Fig. 9. The Concentration of 5-fluorouracil at liver.

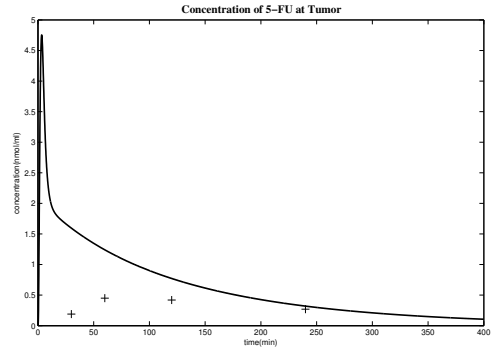

Fig. 10. The Concentration of 5-fluorouracil at tumor.

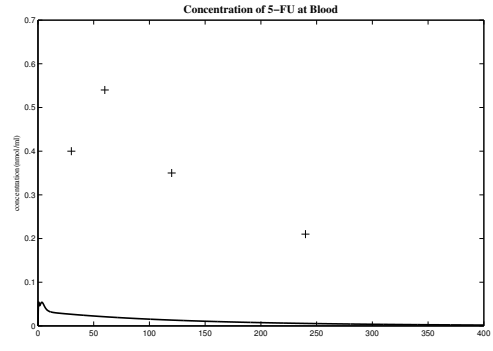

Fig. 11. The Concentration of 5-fluorouracil at blood.

Fig. 12. The joint distribution of the posterior distribution. The probability is getting decrease in order of yellow, red, black and white cite.

REFERENCES

Bonate, P.L. (2006). *Pharmacokinetic-Pharmacodynamic Modeling and Simulation*. Springer Science+Business Media, San Antonio.

Chib, S. and Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The American Statistician*, 49(4), 327–335.

Costa, P. and Lobo, J.M.S. (2001). Modeling and comparison of dissolution profiles. *European Journal of Pharmaceutical Siences*, 13, 123–133.

Gehring, P.J., Watanabe, P.G., and Blau, G.E. (1979). Risk assessment of environmental carcinogens utilizing pharmacokinetic parameters. *Annals of the New York Academy of Sciences*, 329, 137–152.

Gelman, A., Bois, F., and Jiang, J.M. (1996). Physiological pharmacokinetic analysis using population modeling and informative prior distribution. *Journal of the American Statistical Association*, 91, 1400–1412.

Jang, S. and Gopaluni, R. (2011). Parameter estimation in nonlinear chemical and biological processes with unmeasured variables form small data sets. *Chemical Engineering Scinece*, 66, 2774–2787.

Lindsey, J.K., Byrom, W.D., Wang, J., Jarvis, P., and Jones, B. (2000). Generalized nonlineaer models for pharmacokinetic data. *Biometrics*, 56, 81–88.

Longley, D.B. and Johnston, P.G. (2007). *Apoptosis, Cell Signaling, and Human Diseases; Molecular Mechanisms*, volume 1. Humana Press Inc., Totowa.

Schwaab, M., Biscaia, E.C., Monteiro, J.L., and Pinto, J.C. (2008). Non-linear parameter estimation through particle swarm optimization. *Chemical Engineering Science*, 63, 1542–1552.

Smith, A.F.M. and Roberts, G.O. (1993). Bayesian computation via the gibbs sampler and related markove chain monte carlo methods. *Journal of the Royal Statistical Society*, 55(1), 3–23.

Sung, J.H., Dhiman, A., and Shuler, M.L. (2009). A combined pharmacokinetic-pharmacodynamic(pk–pd) model for tumor growth in the rat with uft administration. *Journal of Pharmaceutical science*, 98(5).