# Bayesian Approach in Pharmacokinetics Models

P.K. ZIEGELMANN
*University of Kent at Canterbury - UK* and
*Federal University of Rio Grande do Sul - Brazil*

P.J. BROWN
*University of Kent at Canterbury - UK*

SUMMARY

Pharmacokinetics studies are carried out to characterize the kinetics of a drug (absorption, distribution, metabolism and excretion). The typical data consists of measurements of the concentration of a drug during a period of time, history dose and covariates. Often, the data comes from a large number of patients but with sparse observations for each one. In this context the Compartment Model Theory has been found a great tool. The Hierarchical framework involving two stages is a natural way to accommodate the two sources of variability (within and among individuals) that are empirically observed in the data. The estimation of the parameters is not straightforward due to the high dimensionality of the model and the nonlinearity in the first stage. In this work a Bayesian analysis is carried out using Metropolis-Hastings algorithm. The novelty is in using both multivariate response and covariate selection.

*Keywords:* PHARMACOKINETICS; BAYESIAN COMPARTMENT MODELS; METROPOLIS HASTINGS; NONLINEAR MODELS.

## 1. INTRODUCTION

Population Pharmacokinetics (PK) is a term widely used to refer to a study of the time course of a drug and its metabolites after its introduction into the human body. The proposal of developing a model to describe the kinetics properties of a drug (absorption, distribution, metabolism and excretion) is to provide a way to predict outcomes for dose regimes that have not been tested. Furthermore, identify potential covariates has extreme clinical importance since physiologic and/or demographic factors may alter significantly the relationship between concentration and time. Following Sheiner, Rosenberg and Melmon (1972), most of the PK analysis published have the same modeling framework, which is based on the recognition of the need to accommodate the two sources of variability empirically observed in the data: within and among individuals. This structure is naturally adjusted in a hierarchical framework involving two stages. In the first stage the within variability is accommodated using a single mathematical form to model the relationship between concentration and time for all individuals. In PK studies the first stage form is usually nonlinear in the parameters and the function comes from the Compartmental Theory (see, for example, Gibaldi and Perrier (1982)). The approach is to represent the body as a system of linear and time-invariant compartments and the result is a model built with parameters that have clinical interpretation. Different values across individuals for the first stage function are allowed through the second stage. It is where the among individuals variability is also explained by incorporating covariates information. A fully parametric Bayesian approach was used by Wakefield and Racine-Poon (1985), Wakefield and Bennett (1996) and Wakefield (1996). Following the Bayesian approach, a third stage, with prior distributions for the population parameters in the stages 1 and 2, is required. The estimation of the parameters is not straightforward due to the high dimensionality of the model and the nonlinearity in the first stage. In this context, MCMC techniques (see Gilks, Richardson and Spiegelhalter (1996)) have been an efficient method. Moreover, PK studies frequently involve more than one response variable. In this paper we work with a multiple response model. The approach used is quite flexible since both the possible correlation among variables is incorporated and there is no obligation of the same number of observations for each variable. This characteristic is extremaly important when we work with data that comes from routine clinical studies, i.e. observations from a large number of individuals but with sparse data for each one.

Section 2 describes the general hierarchical three stage nonlinear model and a summary of estimation techniques. Section 3 presents the changes made in the general model to accommodate multiple response. In section 4, we use a real set of data to work with the methodology presented.

## 2. HIERARCHICAL THREE STAGE NONLINEAR MODEL

The statistical methodology behind PK studies involves what are called Mixed Effects Model for Longitudinal data. The natural framework is hierarchical models with two stages. To follow a Bayesian framework, a third stage, with the prior for the population parameters from the first and second stages, is incorporated.

### 2.1. *Stage 1*

In the first stage, the relationship between concentration and time is modeled for each individual, addressing the within individual variation. Suppose that we have $i = 1, \ldots I$ individuals and $n_i$ measurements of the drug concentration for each individual. Let $\boldsymbol{y_i}$ and $\boldsymbol{\theta_i}$ denote the concentration vector and the PK parameter vector for the individual $i$, respectively, and $\boldsymbol{x_{ij}}$ the dose history for the ith individual at the jth day. Assume that a function $f(\boldsymbol{x}, \boldsymbol{\theta})$ may be specified to model the relationship between $\boldsymbol{y}$ and $\boldsymbol{x}$ (in PK models $f$ is a nonlinear function in the parameters $\boldsymbol{\theta}$). The general model at this stage is given by:

$$\boldsymbol{y_i} = \boldsymbol{f_i}(\boldsymbol{\theta_i}) + \boldsymbol{\epsilon_i},$$

where

$$\boldsymbol{f_i}(\boldsymbol{\theta_i}) = \begin{pmatrix} f(\boldsymbol{x_{i1}}, \boldsymbol{\theta_i}) \\ \vdots \\ f(\boldsymbol{x_{in_i}}, \boldsymbol{\theta_i}) \end{pmatrix},$$

$\boldsymbol{x_{ij}}$ are the dose history: variables such as time, dose, for the individual $i$ at time $j$.
Let $E(\boldsymbol{\epsilon_i}|\boldsymbol{\theta_i}) = \mathbf{0}$ and $Cov(\boldsymbol{\epsilon_i}|\boldsymbol{\theta_i}) = \frac{1}{\tau}\boldsymbol{S_i}(\boldsymbol{\theta_i}, \boldsymbol{x_i}, \xi)$, where $\xi$ is the covariance parameter.

### 2.1. *Stage 2*

In the second stage the among individuals variation is accommodated through the specification of a model for the individuals' parameter vector $\boldsymbol{\theta_i}$. The general model at this stage assume that the variability among $\boldsymbol{\theta_i}$ is due to unexplained variation only. But, physiological characteristics like gender, age, weight and treatment group can be useful to explain these variability. So, to account for these possibilities the second stage is built to model the dependence of $\boldsymbol{\theta_i}$ on both systematic and random components. Suppose all covariates are constant within an individual. In this case the second stage is model by:

$\boldsymbol{\theta_i} = A_i\boldsymbol{\mu} \ + \ \boldsymbol{\delta_i}$,

where $A_i$ is the appropriated design matrix for the individual $i$.
To handle the case where one or more covariates are not constant within an individual some changes are necessary to be done in the first stage. It happens because the change in the covariate values implies the possibility of differences in the individual PK parameter vector along time. In this paper we will use just the former case.

### 2.2. *Stage 3*

In the third stage, prior distribution for the population parameters $\tau$, $\boldsymbol{\mu}$ and $\Sigma$ (and $\xi$ when it is considered unknown) are incorporated. Generally, this prior distribution is assumed to be the product of independent conjugate priors for each of these parameters.

### 2.3. *Inference*

In this paper, the estimation of the parameters of this model will be done using a fully parametric Bayesian approach. The inference is based on the posterior distribution of the parameters, given the data.

The high dimensionality of the model and the nonlinearity in the first stage are characteristics that complicate the inference. To deal with these difficulties Markov Chain Monte Carlo techniques were used and their implementation will be discussed with the example.

## 3. MULTIPLE RESPONSE MODEL

Pharmacokinetics studies usually involve measurements of more than one response and, due to the clinical nature where the data comes from, it is common to have different number of observations for each response. To take into account possible correlation structure between responses without requiring the same number of observations for each one we construct a Combined Hierarchical Model. To simplify the ideas, suppose we have a model with two different responses. The concentration vector $\boldsymbol{y_i}$ for the individual $i$ will be now a vector consisting of the observations of response 1 and the observations of response 2 stacked. The PK individuals parameters vector $\boldsymbol{\theta_i}$ will consist of the respective parameters for the response 1 plus response 2. The correlation between the two responses can be incorporated through the structure of covariance in both the first stage and the second stage. Despite the fact that multiple response can complicate inference, the extension of the model described in the previous section to include multiple responses is straightforward.

## 4. GROWTH HORMONE DATA

A set of data from the "Growth Hormone - Olympic Games 2000" project has been used to work with the methodology described in the previous section.

### 4.1. *Data description*

The data is the result of a double-blind, randomized and placebo controlled study. Fifty-nine healthy volunteers of both genders with ages between 18 and 35 years old and physically active (those who had been training at least twice a week for at least one year) participated in the study. Each one was randomized to receive either Single or Double dose (0.1 or 0.2 IU/Kg/day) of recombinant human Growth Hormone for 28 days followed by a 56 day wash-out period. The data available are the measurements of the concentration of eight different markers (responses) on the days 0 (before starting the treatment), 21 and 28 (during the treatment) and 30, 42 and 84 (during the wash-out period). Information on characteristics like age, height, weight, gender are also available. The preliminary analysis that we will present here does not use data from all the markers. Only the markers IGF1 (Insulin like growth factor I) and PIIIP (Procollagen type III) were chosen to be used. It is important to notice the presence of missing values which reduce the number of observations of some individuals to 5 for each marker. The figure 1 shows IGF1 and PIIIP concentration for the 59 individuals.

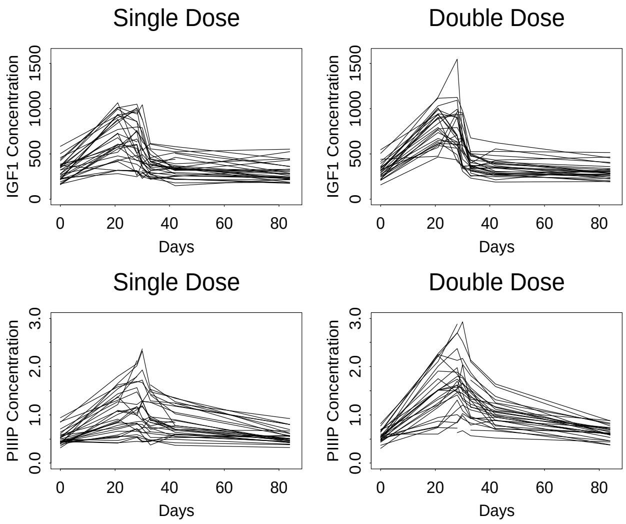


**Figure 1.** *IGF1 and PIIIP concentration-time profiles for all individuals.*

### 4.2. *Objective*

The goal of this analysis is to characterize the kinetics of IGFI and PIIIP, incorporating possible correlation between these two markers and determining those covariates that were good predictors of the PK parameters.

### 4.3. *Compartment Model*

We assume that a one-compartment model with intravenous injection administration and first-order elimination is appropriated for either IGF1 and PIIIP. Also we assume true the assumptions necessary to apply the principle of superposition to work with the multiple dose regime (for all these assumptions see Gibaldi and Perrieri(1982)). Then the predicted concentration (IGF1 or PIIIP) for the individual i at time j (day t after dth dose) is given by:

$$e^{\theta_{i1}} + \frac{Dose_{ij}}{e^{\theta_{i2}}} \times \frac{1 - exp(-d_{ij} \times e^{(\theta_{i3}-\theta_{i2})})}{1 - exp(-e^{(\theta_{i3}-\theta_{i2})})} \times exp(-t_{ij} \times e^{(\theta_{i3}-\theta_{i2})}), \qquad (1)$$

where $Dose_{ij}$ is the dose rate administered, $d_{ij}$ is the number of doses administered until time $j$ and $t_{ij}$ is the number of days after the time $j$ (for example, at $j$ equal day 28, $d_{ij} = 28$ and $t_{ij} = 0$). Also, $\theta_{i1} = \log \alpha_i$, $\theta_{i2} = \log V_i$ and $\theta_{i3} = \log Cl_i$ where $\alpha_i$ represents the baseline, $Cl_i$ is the Clearance and $V_i$ is the Volume of distribution for the individual i.

### 4.4. *Statistical Model*

The nonlinear hierarchical model described in the previous section was used. In this example we have $k = 2$ types of response: IGF1 and PIIIP. Denote $y_1$ the response 1 (IGF1) and $y_2$ the response 2 (PIIIP). Let $\boldsymbol{y_{1i}}$ and $\boldsymbol{y_{2i}}$ denote the $n_{1i}$ dimensional vector of IGF1 and $n_{2i}$ dimensional vector of PIIIP for the individual $i$ $(i = 1, ..., 59)$ respectively, taken at dose history vectors $\boldsymbol{x_{1i}}$ and $\boldsymbol{x_{2i}}$. Then, for example, $y_{1ij}$ is the jth response of type $k = 1$ for the individual $i$. The vector $\boldsymbol{y_i}$ will now represent the single vector of response for the individual $i$ with $\boldsymbol{y_{1i}}$ and $\boldsymbol{y_{2i}}$ stacked. The element $y_{ij}$ is the jth observation of the individual $i$ $(j = 1, ..., n_i)$ with $n_i = n_{1i} + n_{2i}$. Let $\boldsymbol{\theta_i} = [\boldsymbol{\theta_{1i}}^T, \boldsymbol{\theta_{2i}}^T]$ denote the PK parameter vector and $\boldsymbol{\epsilon_i}$ the intra individual covariance parameters vector. Then we have,

**Stage 1**

$$y_{ij} = f_k(x_{ij}, \boldsymbol{\theta_i}) + \epsilon_{ij}, \qquad if\ y_{ij}\ is\ of\ type\ k,\ k = 1, 2\ ,$$

where $\boldsymbol{x_{ij}} = (Dose_{ij}, d_{ij}, t_{ij})$ is the known dose history, $\boldsymbol{\theta_i} = (\theta_{i1}, \theta_{i2}, \theta_{i3})$ is the unknown parameters vector and $f_1$ and $f_2$ have the form given by expression (1),

$$\boldsymbol{\theta_i} = \begin{pmatrix} \theta_{i1} \\ \theta_{i2} \\ \theta_{i3} \\ \theta_{i4} \\ \theta_{i5} \\ \theta_{i6} \end{pmatrix} = \begin{pmatrix} \log \alpha_{1i} \\ \log V_{1i} \\ \log Cl_{1i} \\ \log \alpha_{2i} \\ \log V_{2i} \\ \log Cl_{2i} \end{pmatrix},$$

$$\boldsymbol{\epsilon_i} \sim N(0, \tfrac{1}{\tau}\boldsymbol{S_i}) \ and$$

$$\boldsymbol{S_i}(\boldsymbol{\theta_i}, \boldsymbol{x_i}) = diag[f(\boldsymbol{x_{i1}}, \boldsymbol{\theta_i})^{2\xi}, ..., f(\boldsymbol{x_{in_i}}, \boldsymbol{\theta_i})^{2\xi}]$$

In order to assume constant coefficient of variation for the response variable (IGF1 and PIIIP) along different times we fix $\xi = 1$.

**Stage 2**

Considering a covariate model, the second stage has the form:

$$\boldsymbol{\theta_i} = A_i\mu + \boldsymbol{\delta_i}, \qquad \boldsymbol{\delta_i} \sim N(0, \boldsymbol{\Sigma}).$$

**Stage 3**

We assume that

$$\tau \sim Gamma\left(\tfrac{\delta_0}{2}, \tfrac{\tau_0\delta_0}{2}\right),$$

$$\boldsymbol{\mu} \sim N_q\left(\boldsymbol{\eta}, C\right) \ and$$

$$\boldsymbol{\Sigma}^{-1} \sim Wishart_{6\times 6}\left(\rho, [\rho R]^{-1}\right),$$

with $\delta_0, \tau_0, \boldsymbol{\eta}, C, \rho, R$ known and fixed in order to obtain flat priors.

### 4.5. *Inference*

To sample from the posterior distribution, the Gibbs Sampling technique was used. The parameter vector of interest $(\boldsymbol{\theta}, \boldsymbol{\mu}, \Sigma^{-1}, \tau)$ was divided into the components $\boldsymbol{\theta_1}, ..., \boldsymbol{\theta_I}, \boldsymbol{\mu}, \Sigma^{-1}, \tau$ and each of these was updated one by one. The proposal distribution for updating each component was the respective full conditional distribution and the order used was the same of the parameters vector. The simulation of $\boldsymbol{\mu}, \Sigma^{-1}$ and $\tau$ is straightforward, since the full conditionals are in a closed form. Random Walk Metropolis algorithm was used to simulate each $\boldsymbol{\theta_i}$ since their full conditionals are not in a closed form.

## 6. ANALYSIS

The analysis presented in this section was carried out using a C code written by the author. The proposal distribution chosen to sample $\boldsymbol{\theta_1}$, ..., $\boldsymbol{\theta_I}$ was $N\,(\boldsymbol{\theta_{iB}}, \Omega_i)$, where $\boldsymbol{\theta_{iB}}$ is the current estimate for $\boldsymbol{\theta_i}$ and $\Omega_i$ is the $6 \times 6$ covariance matrix. The upper left $3 \times 3$ submatrix corresponds to a constant $c_1$ times the asymptotic correlation matrix of IGF1. The lower right $3 \times 3$ submatrix corresponds to a constant $c_2$ times the asymptotic correlation matrix of PIIIP. The other entries are all fixed equal to zero. The constants $a_1$ and $a_2$ are known fixed in order to obtain a rate of acceptability between 15% and 50%.

Initially, a preliminary analysis without covariates was carried out. After a 35,000 burn-in iterations, a sample of size 13,000 was collected from 65,000 iterations with spacing of 5 iterations. Geweke and Raftery-Lewis convergence diagnostic was used to monitor the convergence. The figure 2 shows the random effects for each PK-IGF1 parameters plotted versus the covariates. Although it is not showed here, the same plot was considered for PIIIP.

Random Effects versus Covariates - Model without Covariates

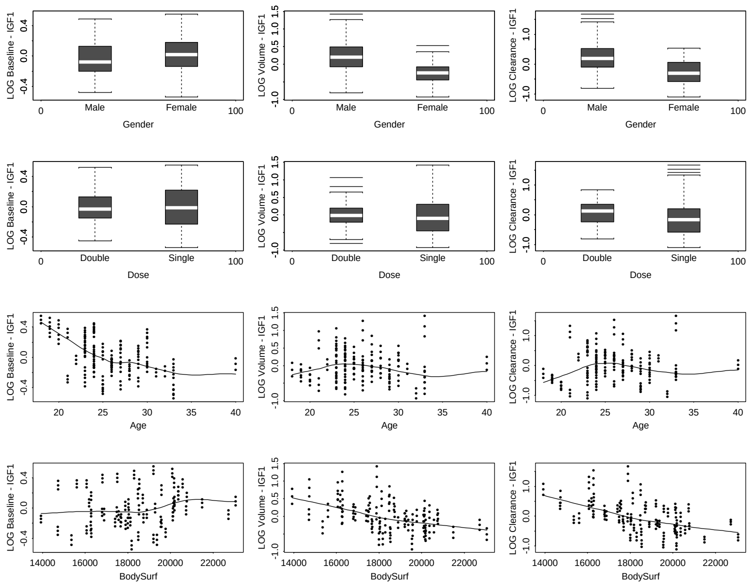


**Figure 2.** *Posterior Quartiles of Log Baseline - IGF1, LOG Clearance - IGF1 and Log Volume - IGF1 Random Effects plotted versus Covariates for all Individuals. Mean values for the Continuous Covariates were used. The Random Effects comes from the Model without Covariates*

In the next step a covariate model with all covariates that seem to have a relationship with the PK individual parameters was proposed. Assuming an additive error structure, the second stage takes the form:

$$\begin{pmatrix} \log \alpha_{1i} \\ \log V_{1i} \\ \log Cl_{1i} \\ \log \alpha_{2i} \\ \log V_{2i} \\ \log Cl_{2i} \end{pmatrix} = \begin{pmatrix} \mu_1 + \mu_2 a_i + \mu_3 \times Age + \epsilon_{i1} \\ \mu_4 + \mu_5 a_i + \mu_6 b_i + \mu_7 \times BodySurface + \epsilon_{i2} \\ \mu_8 + \mu_9 a_i + \mu_{10} b_i + \mu_{11} \times BodySurface + \epsilon_{i4} \\ \mu_{12} + \mu_{13} a_i + \mu_{14} \times Age + \epsilon_{i4} \\ \mu_{15} + \mu_{16} a_i + \mu_{17} b_i + \mu_{18} \times BodySurface + \epsilon_{i5} \\ \mu_{19} + \mu_{20} a_i + \mu_{21} b_i + \mu_{22} \times BodySurface + \epsilon_{i6} \end{pmatrix},$$

where $a_i = \begin{cases} 0 & \text{if } i \text{ is Female} \\ 1 & \text{if } i \text{ is Male} \end{cases}$ and $b_i = \begin{cases} 0 & \text{if } i \text{ received Double Dose} \\ 1 & \text{if } i \text{ received Single Dose} \end{cases}$.

The covariate model was carried out using a single long run with 40,000 burn-in iterations and a further 60,000 iterations. A sample of size 12,000 was collected one out of each 5 iterations. For the Metropolis-Hasting algorithm we fixed $c_1 = 0.02$ and $c_2 = 0.002$. It gives us an acceptance probability approximately equal to 0.3 for all $\boldsymbol{\theta_i}$.
The figure 3 shows the random effects for each PK-IGF1 parameter plotted versus the covariates. Again the PK-PIIIP plots were not showed. On examination of these plots it seems that no relationship between any covariates and PK individual parameters still remains.

Random Effects versus Covariates - Covariate Model.

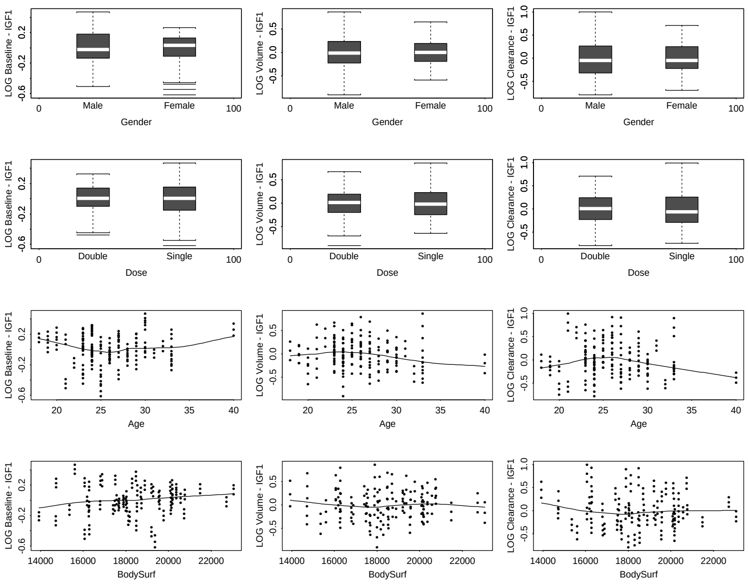


**Figure 3.** *Posterior Quartiles of Log Baseline - IGF1, LOG Clearance - IGF1 and Log Volume - IGF1 Random Effects plotted versus Covariates for all Individuals. Means values for the Continuous Covariates were used. The Random Effects comes from the Covariate Model.*

The figure 4 and figure 5 shows the observed profile and the adjusted curve estimated for some of the 59 individuals.

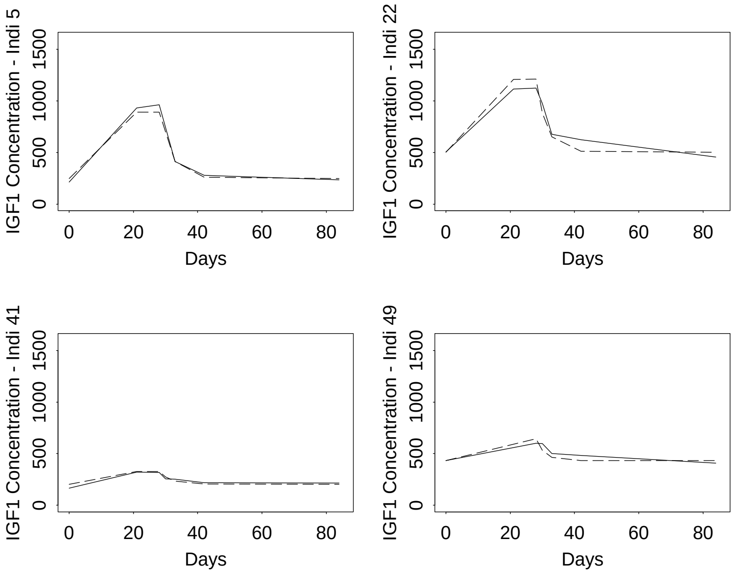


**Figure 4.** *IGF1 concentration-time profiles for four individuals. Observed Profile: solid line. Estimated Profile: dashed line.*

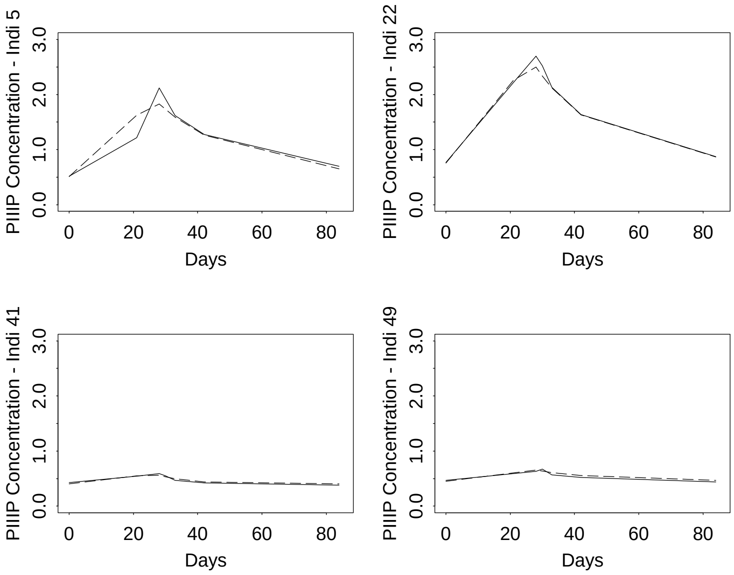


**Figure 5.** *PIIIP concentration-time profiles for four individuals. Observed Profile: solid line. Estimated Profile: dashed line.*

## 7. DISCUSSION

We have been working to develop a statistical model suitable for a class of PK studies. The novelty is in using a Combined Hierarchical Nonlinear Model to accommodate multiple responses. Although the results are quite good (which makes it a promising approach), a number of issues to improve the model are contemplated for future work. One of these involves a formal approach to build a covariate model, where possible interaction among covariates is incorporated. Another feature is related to the accommodation of more than two response variables. Although the extension from the single response model is straightforward its implementation following Bayesian approach has some indesirable implications. The increasing number of parameters associated with the structure of the model can generate convergence difficulties. "Bad" starting points can give poorly mixing chains.

## 8. ACKNOWLEDGMENTS

The work described in this paper is related to the first author's PhD thesis. During her PhD the author has been supported by a grant from CAPES (Brazil).

## REFERENCES

Davidian, M. and Giltinan, D. M. (1995). *Nonlinear Model for Repeated Measurements Data*. London: Chapman and Hall.

Gibaldi, M. and Perrier, D. (1982). *Pharmacokinetics*. New York: Marcel Dekker.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall.

Sheiner, L. B., Rosenberg, B. and Melmon, K. L. (1972). Modeling of Individual Pharmacokinetics for Computer-aided Drug Dosing. *Computers and Biomedical Research.***5**, 441–459.

Wakefield, J. C. (1996). The Bayesian Analysis of Population Pharmacokinetic Models. *J. Amer. Statist. Assoc.* **91**, 62–75.

Wakefield, J. C. and Bennett, J. (1996). The Bayesian Modeling of Covariates for Population Pharmacokinetic. *J. Amer. Statist. Assoc.* **91**, 917–927.

Wakefield, J. C. and Racine-Poon, A. (1985). An application of Bayesian Population Pharmacokinetics/Pharmacodynamic Models to Dose Recommendation. *Statist. in Medicine.***14**, 971–986.