# THE GLUCOSE MINIMAL MODEL: POPULATION VS INDIVIDUAL PARAMETER ESTIMATION

**Alessandra Bertoldo, Paolo Vicini, Claudio Cobelli**

*Department of Information Engineering, University of Padova, Italy*
*Department of Bioengineering, University of Washington, Seattle, WA, USA*

Abstract: Glucose minimal model parameters are commonly estimated by applying weighted nonlinear least squares separately to each subject's data. Because of the model's nonlinearity. the parameter precision of the single-compartment minimal model is not always satisfactory, especially in presence of a reduced sampling schedule. In the current work, the use of population analysis through nonlinear mixed effects models is evaluated and its performance tested against the parameter estimates obtained by the standard individual approach through weighted nonlinear least squares. 

Keywords: Modeling, Parameter estimation, Nonlinear models, Physiological models, Sampling frequency.

## 1. INTRODUCTION

The single-compartment minimal model method (Bergman, *et al*., 1979) is widely used in clinical and epidemiological studies to estimate metabolic indexes of glucose effectiveness ($S_G$) and insulin sensitivity ($S_I$) from an intravenous glucose tolerance test (IVGTT). Use of the minimal model for $S_G$ and $S_I$ determination requires the injection of a glucose bolus at time 0, and subsequently sampling for 3 or 4 hr. As typical in physiological and metabolic modeling, minimal model parameters are commonly estimated by applying weighted nonlinear least squares separately in each subject. After having obtained individual estimates for each subject, the sample mean and the variance of all the model parameter estimates are calculated and assumed to approximate the first- and the second-order moment (expected value and variance) of the subject population distribution. However, due to its complexity, the parameter precision of the single-compartment minimal model is not always satisfactory, especially in presence of a reduced sampling schedule ("data poor" situation). Of note is that a reduced sampling scheme is highly desirable, both for ethical and practical reasons, above all when clinical trials are performed in a large number of subjects: a reduced sampling scheme allows to minimize experimental invasiveness.

To derive accurate and precise individual estimates and, consequently, description of the subject population also in presence of a data poor situation, maximum a posteriori Bayesian estimation has been evaluated (Sparacino, *et al*., 2000). The drawback of this estimation method is that it requires some independent a priori statistical (i.e., mean, variance, covariance) knowledge on the model parameters. This drawback can potentially heavily compromise the parameter estimation process, when a priori information is unavailable or poor quality. However, other estimation approaches focused on ensembles of individuals, like population kinetic analysis through mixed effects models (Beal and Sheiner, 1982), have only recently been applied in this context (De Gaetano, *et al.*, 1996) and have still to be thoroughly evaluated.

Population analysis aims at quantitative assessment of model parameters, taking advantage of the entire collection of measures obtained from a population of individuals. Population analysis directly estimates statistical features of the data set, and finds its natural application in quantification of data poor studies, e.g. when the number of samples available for each individual subject is rather small in comparison with model complexity. It is widely used in the analysis of pharmacokinetic studies. Among all available kinetic data analysis methods, population approaches using nonlinear mixed effects models have become an increasingly important tool, since they not only allow

one to quantify both population and individual parameters, but also to identify the biological sources of between- and within-subject variability.

In this work, we will describe the use of population analysis through nonlinear mixed effects model to identify single-compartment minimal model parameters in a population of subjects composed of healthy and young adults. The performance of the population approach will be tested against the parameter estimates obtained by the standard individual approach, where each subject is analyzed individually by weighted nonlinear least squares. While others have looked at applications of nonlinear mixed effects to minimal modeling in a simulation context (Erichsen, *et al.*, 2004), we have chosen a "data rich" situation for this evaluation. By comparison with the standard estimates, we will evaluate the most common parametric nonlinear mixed effects modeling approaches. By selecting the one performing best, we will have developed an important estimation tool for handling of sparse sampling protocols used in physiological and metabolic modeling.

## 2. MATERIALS AND METHODS

### *2.1 Subjects*

Standard IVGTT [dose 330 mg/kg] studies were performed on 58 nondiabetic young subjects (mean age 23±3 and mean BMI 24.5±2.9 kg/m$^2$) in the Clinical Research Center at the Mayo Clinic, Rochester, MN, USA. Subjects received the glucose bolus at time 0,. Blood samples were collected at -120, -30, -20, -10, 0, 2, 4, 6, 8, 10, 15, 20, 22, 25, 26, 28, 31, 35, 45, 60, 75, 90, 120, 180, and 240 min for measurement of glucose and insulin concentrations.

### *2.2 The Minimal Model*

The classic one-compartment minimal model (Bergman, *et al*., 1979) can be described by:

$$\begin{aligned} \dot{Q}(t) &= -[S_G + X(t)]Q(t) + S_G Q_b \qquad Q(0) = D/V + G_b \\ \dot{X}(t) &= -p_2 X(t) + p_2 S_I [I(t) - I_b] \qquad X(0) = 0 \\ G(t) &= Q(t)/V \end{aligned} \quad (1)$$

where D is the glucose dose, Q(t) (mg/kg) is glucose mass in plasma with $Q_b$ denoting its basal value, G(t) (mg/dl) is plasma glucose concentration, I(t) (μU/ml) is insulin concentration, $G_b$ and $I_b$ are their basal values, and X(t) is insulin action (min$^{-1}$). The model has four uniquely identifiable parameters: $S_G$ (min$^{-1}$), glucose effectiveness, $S_I$ (min$^{-1}$ μU$^{-1}$ ml), insulin sensitivity, $p_2$ (min$^{-1}$), the insulin action parameter, and V (dl/kg), the glucose distribution volume per unit of body mass. The model parameters are estimated by assuming I(t) as a known input (forcing) function.

### *2.3 The Individual Standard Estimation Approach*

We used weighted nonlinear least squares as implemented in SAAM II (Barrett, *et al*., 1998). Assuming that the observed data are statistically related to the individual true parameters $\mathbf{p}_j$ through the measurement equation: $G_j(t_i) = G(\mathbf{p}_j, t_i) + \varepsilon_j(t_i)$; the cost function to be minimized is:

$$\mathrm{WRSS}(\mathbf{p}_j) = \sum_{i=1}^{N} \frac{\left[G_j(t_i) - G(\mathbf{p}_j, t_i)\right]^2}{\sigma_{i,j}^2} \quad (2)$$

where N is the number of glucose samples, $G_j(t_i)$ is the i$^{th}$ time point for the j$^{th}$ of M subjects, $\sigma_{i,j}^2$ is the variance of the measurement error of the i$^{th}$ data point, and $G(\mathbf{p}_j, t_i)$ is the minimal model prediction of glucose concentration. Measurement error was assumed to be additive, uncorrelated, Gaussian, zero mean, and with a standard deviation given by:

$$\sigma_{i,j} = 0.02 \cdot G_j(t_i) \quad (3)$$

After obtaining all the individual estimates, we calculated for each parameter, the sample mean of all the individual parameter estimates and the corresponding sample covariance.

### *2.4 Population Analysis: the Nonlinear Mixed-Effects Model Approach*

Unlike the estimation approach discussed above, a more elaborate statistical model is required by the nonlinear mixed-effects model approach. In particular, the observed data are again supposed to be related to the individual true parameters $\mathbf{p}_j$ thought Eq. 3, but, in addition, it is assumed that the individual parameters $\mathbf{p}_j$ are characterized by some attributes that do not change across the population of M subjects (fixed effects, i.e. values that are common to all subjects) and some others that do (random effects, i.e. values typical of a specific subject). Mathematically, this can be written as:

$$\mathbf{p}_j = d(\boldsymbol{\theta}, \boldsymbol{\eta}_j) \quad (4)$$

where d is a known function that describes the expected value of $\mathbf{p}_j$ as a function of the fixed effects, $\boldsymbol{\theta}$, and the random effects, $\boldsymbol{\eta}_i$. More specifically, the individual parameter can be written as:

$$\mathbf{p}_j = d(\boldsymbol{\theta}, \mathbf{a}_j, \boldsymbol{\eta}_j) \quad (5)$$

with $\mathbf{a}_j$ being known individual specific covariates such as weight, age, body mass index, etc.

Parametric mixed-effects modeling requires to postulate at least some characteristics of the population probability distribution for the random effects (e.g. whether it is Gaussian or lognormal). We assume the random effects to be independent, with: