34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Identification of the Glucose Minimal Model by Stochastic Nonlinear-mixed Effects Methods

Anna Largajolli, Alessandra Bertoldo, *Member, IEEE*, Claudio Cobelli, *Fellow, IEEE*

Abstract— **The nonlinear mixed effects models (NLMEM) are widespread modeling techniques in PKPD analysis and epidemiological studies because they can produce a description of not only the individual but also of the population features. Moreover, they are able to deal with individual data sparseness by borrowing the lack of information from the entire population. In this way, the NLMEM do not fail where instead other techniques, such as the traditional individual weighted least squares (WLS), sometimes do. The NLME approach relies on the maximization of a likelihood function that due to model parametric nonlinearity not always has an explicit solution. Various techniques have been proposed to solve this problem including the first order (FO) and the first order conditional (FOCE) estimation methods that approximate the likelihood function through a linearization; the expectation maximization algorithm (EM) that maximize the exact likelihood; the Bayesian estimation method where a third stage of variability, the distribution of the population parameters, is taken into account [1]. Recently, new estimation methods that rely on the EM algorithm have been implemented in the last release of the population software NONMEM [2]. These methods are: the iterative two stage (ITS), Monte Carlo importance sampling EM (IMP), Monte Carlo importance sampling EM assisted by Mode a Posteriori estimation (IMPMAP) and the Stochastic Approximation EM (SAEM). Moreover, another new method is available, the Markov Chain Monte Carlo Bayesian Analysis (BAYES), next to the more known FO and FOCE. With this article we want to complete the Denti et al [3] simulation study by evaluating the newest population methods applied on the IVGTT glucose minimal model.**

I. Introduction

The IVGTT glucose minimal model (MM) is a well known tool to study the glucose-insulin system in different pathophysiological states after an intravenous glucose perturbation. Indeed, its parameter S_I, the insulin sensitivity, that is the overall effect of insulin to stimulate glucose uptake and inhibit glucose production, represents an important metabolic index in clinical and epidemiological trials. By now the model has been identified using both individual and population approaches. At the beginning the weighted least square (WLS) single subject technique was used and applied on the data of each individual. The WLS, though, in typical epidemiological conditions such as sparse and noisy data per individual does not produce satisfactory estimations. In order to improve this aspect, the population approaches were then introduced to identify the MM by exploiting the not used information spread on the subject collection. Vicini et al [4] identified the MM by the iterative two stage (ITS), a population technique. This method is made up of two steps: first each subject's data is separately fitted and then the population parameter estimates are obtained. This procedure is repeated until convergence. Afterwards, Agbaje et al [5] used a different population approach to identify the same model: the Bayesian hierarchical method [6]. This method adds a third stage of knowledge to the individual and population step that is the prior distribution of the population parameters. More recently [3], the MM was quantified using both iterative methods, like the Global two stage (GTS) and ITS and other techniques like the first order (FO) and the first order conditional (FOCE). These last two methods are approximated solutions of the nonlinear mixed effects models (NLMEM) approach that aims to characterize the individual and population description by maximizing a likelihood function. Due to nonlinear parametric dependencies it is almost impossible to have an explicit solution of this optimization problem. FO and FOCE methods fix this by approximating the likelihood through a linearization. This work is the natural follow up of Denti et al [3] where different population estimation methods, implemented in the software SPK [7], were tested in a simulation study. In particular the study aims are mainly two. The first is to complete the population analysis done so far in the MM with the latest methods implemented in the software NONMEM. Three new optimization techniques have never been applied so far to quantify the MM parameters: the Monte Carlo importance sampling EM (IMP) [8], the Monte Carlo importance sampling EM assisted by Mode a Posteriori estimation (IMPMAP) [2] and the Stochastic Approximation EM (SAEM) [9]. These three, together with the ITS method, are implemented in NONMEM by exploiting the characteristic two steps of the Expectation-Maximization algorithm (EM). In the first step (the expectation step) the expectation of the log-likelihood given current estimates of the population parameters is calculated and, in the second step (the maximization step), new population parameters that maximize the expectation are computed. This procedure is repeated until there are no visible changes in the objective function. Note that the four different implementations of the EM algorithm are different approximations of the expectation step for which no analytical solution is available. The second aim of this work is to test the robustness of the different methods in a data poor context by comparing their performances on two randomly generated datasets obtained by removing respectively 50% and 75% of the original samples respectively. All our analysis is carried out using the software NONMEM VII [2].

A. Largajolli, A. Bertoldo and C. Cobelli* are with the Department of Information Engineering, University of Padova, Via G. Gradenigo 6/B, 35131 Padova, Italy (* corresponding author, phone: +39 049 827 7661; fax: +39 049 827 7826; e-mail: cobelli@dei.unipd.it).

II. Material and Methods

A. Synthetic Data

As already mentioned this work is the Denti et al [3] natural follow up but carried out with different software and estimation methods. The dataset used is the same. The dataset (dataset A1) consists of 58 simulated insulin modified IVGTT profiles (dose 330 mg/kg glucose at time 0, 0.02 units/kg of insulin at time 20). This dataset was obtained through two steps. Firstly, the MM was identified using the individual estimation method WLS, implemented in the software SAAMII [10], in 58 nondiabetic young subjects (mean age 23±3 and mean BMI 24.5±2.9 kg/m^2) that underwent an IVGTT in the Clinical Research Center at the Mayo Clinic, Rochester, MN. 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. Secondly, the profiles were simulated using the individual estimates obtained at the first step and a measurement noise equal to the 2% of the simulated profile was added. In order to exploit the potentials of the population technique and to test the robustness of the estimates we evaluate the different methods performance in a data poor context. In particular the simulated dataset was reduced by randomly removing the original samples. The first time 50% of the original samples were randomly removed (dataset A2), while the second time instead 75% of the original samples were removed (dataset A3). In this way the typical condition of the epidemiological studies, that is few and noisy data per individual, was recreated. Note that only the glucose data were reduced whereas the insulin data that acts as a forcing function in the model was not. This choice was made because the work aim is to test different estimation methods and not really simulate a real experiment data sparseness situation.

B. Glucose Minimal Model

The IVGTT MM is described by:

$$\begin{cases} \dot{Q}(t) = -(S_G + X(t)) \cdot Q(t) + S_G \cdot Q_b & Q(0) = G_b \cdot V + D \\ \dot{X}(t) = -p_2 \cdot X(t) + p_2 \cdot S_I \cdot (I(t) - I_b) & X(0) = 0 \end{cases} \quad (1)$$

where Q is the glucose mass in plasma (mg/kg) and Q_b its basal value, G_b is the basal glucose concentration in plasma (mg/dL), I is insulin plasma concentration (pmol/L) and I_b its basal value, X is insulin action (min^{-1}). The uniquely identifiable parameters are: glucose effectiveness SG (min^{-1}), insulin sensitivity SI (min^{-1} pmol^{-1} L), insulin action parameter p2 (min^{-1}) and volume V (dL/kg). The model is not designed to take into account the first 8 minutes of glucose so the corresponding measurements were excluded from the modeling analysis.

C. Population Assumptions

In the population analysis done using the NLME approach data are described by the model:

$$y_{ij} = f(\mathbf{p}_i, x_{ij}) + \varepsilon_{ij} \qquad 1 \le i \le n, \qquad 1 \le j \le m_i \quad (2)$$

where y_{ij} is the *jth* observation of the *ith* subject at some known time instant x_{ij}. Here, n is the number of individuals and m_i is the number of observations of the individual i. P_i is the vector of model parameters for the ith individual. The model parameters across the population are assumed to be lognormal distributed. In particular they can be described by:

$$p_{ki} = \theta_k e^{(\eta_{ki})} \quad (3)$$

$$\boldsymbol{\eta}_i \sim N(0, \boldsymbol{\Omega}) \quad (4)$$

where p_{ki} is the *kth* model parameter of the *ith* subject, θ_k is the typical value of the *kth* parameter common to the entire population and η_{ki} is the random effect of the *kth* model parameter of the *ith* subject. $\boldsymbol{\eta}_i$ is assumed to be independently distributed with zero mean and Gaussian with $\boldsymbol{\Omega}$ being a positive definite covariance matrix (4). The $\boldsymbol{\Omega}$ values define the Between-Subject Variability (BSV). The omega set up matrix was chosen coherently with Denti et al [11] including just the correlations term between the S_I-P_2 and S_G-V. The variability due to measurement and model errors, known as the residual unknown variability (RUV), instead is described by ε_{ij} which is assumed to be independently distributed with zero mean and Gaussian with standard deviation described by σ (proportional error variance) being an additional parameter to estimate:

$$\varepsilon_{ij} \sim N(0, (\sigma y_{ij})^2) \quad (5)$$

D. Nonlinear Mixed Effects Methods

The reference estimates (REF) were obtained by the WLS approach implemented in SAAM II [10] applied on the original data. Then the other methods were applied on the simulated dataset. At first we investigated the standard two-stage (STS) performance which is another individual WLS that we implemented in NONMEM. Then we applied the population approach NLME that provides different estimation methods due to computational non feasibility of the exact solution of the likelihood maximization. Firstly, we applied the FOCE algorithm that is a linearization of the likelihood function. Then, we used the EM algorithm based estimation methods. These methods are the ITS, the IMP, the IMPMAP and the SAEM. Finally we applied the BAYES method that adds the third stage of variability due to the population parameters. The priors that were given to the population estimates were vague as in Agbaje et al [5], representing the lack of information about parameter distributions.

E. Analysis of Results

In order to assess the different methods performance both the individual and the population estimates were evaluated and compared to REF estimates. As far as the population results are concerned, we evaluated the percentages of discrepancy between the estimated (fixed effects and square root of the BSVs) and the true values. The true values are the geometrical mean and standard deviation of the individual estimates REF. As far as the individual results are concerned, we evaluated the goodness of the individual estimates assessed by the square Root of the Mean Square Error (RMSE):

$$RMSE = \sqrt{\sum_{i=1}^{N} \frac{(p_i - \hat{p}_i)^2}{n}} \quad (6)$$

where p_i is the true parameter value (REF) for subject i, $\hat{p}_i$ its estimate, and n is the number of subjects involved in the analysis. For readability purposes, these values were indicated as percentage of the true population mean of each parameter.

III. RESULTS

As already said, we exploited the population approaches to the full by analyzing both the population and the individual results in the simulated dataset (A1) and in its two reduced versions (A2-A3). Before proceeding it, is important to make a remark. All the methods were successful and no subject was excluded from the analysis. However some methods are more sensitive to initial estimates. In order to make them run smoothly we used a software feature that allows minimizing in cascade different methods; i.e. starting from the most stable whose final estimates are given as initial estimates to the subsequent less stable method.

TABLE I. POPULATION PARAMETERS

	POPULATION PARAMETERS DATASET A1			
dataset A1	ΔSG	ΔVOL	ΔSI	ΔP2
STS	-2 (74)	0 (173)	0 (41)	0 (47)
ITS	4 (-25)	0 (-1)	-2 (4)	1 (-8)
FOCE	4 (-28)	0 (-2)	-2 (4)	1 (-10)
SAEM	0 (-30)	1 (1)	-11 (6)	-7 (-9)
BAYES	2 (-17)	-1 (21)	0 (8)	2 (-2)
IMP	3 (-25)	-1 (0)	-11 (7)	-7 (-7)
IMPMAP	5 (-26)	-1 (-1)	-13 (10)	-8 (-5)

	POPULATION PARAMETERS DATASET A2			
dataset A2	ΔSG	ΔVOL	ΔSI	ΔP2
STS	1 (93)	-1 (184)	1 (46)	-13 (65)
ITS	10 (-27)	-1 (0)	-1 (2)	-9 (-7)
FOCE	10 (-29)	-1 (-2)	-1 (3)	-8 (-10)
SAEM	7 (-32)	0 (-2)	-9 (6)	-17 (-5)
BAYES	8 (-17)	-1 (21)	-1 (8)	-6 (-1)
IMP	7 (-28)	-1 (-1)	-6 (4)	-12 (-7)
IMPMAP	7 (-29)	-1 (-2)	-10 (9)	-14 (-4)

	POPULATION PARAMETERS DATASET A3			
dataset A3	ΔSG	ΔVOL	ΔSI	ΔP2
STS	-11 (122)	-3 (257)	-13 (94)	-4 (107)
ITS	13 (-40)	-1 (-18)	-1 (-10)	-3 (-31)
FOCE	13 (-53)	-1 (-26)	-1 (-9)	-4 (-24)
SAEM	13 (-42)	-1 (-20)	-4 (-6)	-3 (-39)
BAYES	10 (-44)	-1 (-18)	-1 (-4)	-1 (-19)
IMP	12 (-53)	-1 (-25)	-5 (-7)	-6 (-20)
IMPMAP	15 (-53)	-2 (-25)	-6 (-6)	-9 (-19)

The distance of the estimated values for both the fixed effects and the square root of the BSV (in brackets) from the true values are reported as percentage differences normalized to the true values for the 3 datasets.

These methods were SAEM, IMP, IMPMAP and BAYES and each one was preceded in the minimization by ITS which is a more stable and at the same time fast technique in producing reliable estimates.

A. Population Results

For the population estimates in dataset A1, in general, all methods provide results coherent with the ones that were used for the simulation. However, not all the methods behave the same. Looking at Table I, we can see that the best performing methods are ITS, FOCE and BAYES where all the fixed effects estimates discrepancy percentage do not exceed the 4% modulus and the discrepancies percentage of the BSV square root (values in brackets) do not exceed 28% modulus. The parameter that is worst estimated in these three methods in both the fixed effects and in the BSV is SG. Whereas SAEM, IMP and IMPMAP tend to underestimate SI (fixed effects percentage discrepancy values from -11% using SAEM or IMP to the -13% using the IMPMAP) and p2 (fixed effects percentage discrepancy values from -7% using SAEM or IMP to the -8% using IMPMAP). The parameter whose mean is estimated more precisely using all the estimation methods is V, whereas SG, SI and p2 are affected by a slightly larger error. The population approach, anyway, works better than the individual approach STS. Looking at Table I, one can see that STS presents the largest BSV. The overestimation of the variance of the population was expected as it is already well known in literature [12]. The population approach improvement due to the information borrowed across the population is expected to grow with the paucity of data. Looking now at the reduced dataset A2 and the furtherly reduced dataset A3, we can see that the discrepancy of the estimates, as expected, becomes larger with the samples reduction. This is true apart from some cases where instead there is the opposite effect. In particular, looking at table I we can see that if we consider the parameter SI estimated with the SAEM method and we move from dataset A1 to dataset A3, the parameter fixed effect seems to be estimated better with less samples. This effect is a typical feature of the population approach, especially in a poor data context. In fact when there is not enough individual information (i.e. few samples per individual), a condition that is merely tolerated by the individual approach, a sort of constraint is generated between the individual estimates that tends to bring them together towards the population mean. This phenomenon is known in literature as shrinkage [13]. Also in these two reduced dataset, in general the parameter whose mean is estimated more precisely is V, whereas SG, SI and p2 are affected by a slightly larger error. The individual approach worsens its performance moving from dataset A2 to A3 as one can see clearly in Table I from the increase of discrepancy percentage of the BSV square root.

B. Individual Results

As far as the individual results are concerned, all the different population estimation methods perform comparable apart from the STS. In Table II the RMSE percentage of the individual estimates are presented. Looking at dataset A1, all the methods estimate well V and SI, whereas SG and p2 have

a RMSE percentage larger than 13%. Also in this case, analyzing the individual results, the population approaches behave better than the individual approach represented by STS. Moreover, moving the attention to the reduced datasets, the same trend that was previously observed in the population results, is present here: RMSE percentage increases as expected with the lack of samples. V and SI, as in A1 dataset individual results, are the parameters that are estimated more precisely. Regarding the comparison between the individual and the population approach, the population technique improvement due to the borrowed information across the subjects is larger in the two reduced datasets. In particular, the difference between the %RMSE values of the STS method and the other corresponding values of the different population techniques increases moving from dataset A2 to dataset A3. In other words, the population approach features can be better appreciated in severely reduced datasets.

C. *Residual Unknown Variability*

In the population approach the residual unknown variability (RUV) represents the variability due to model error and measurement error. RUV is estimated with the parameter σ that was not fixed to 2% of the data (the error structure that was used to individually generate the data) but was left free to be optimized by the algorithm. In Table III the estimated σ in the dataset A1 in the different population methods are compared with the true value (2%). The measurement error was on average well estimated by all the methods apart from STS. BAYES, IMP and ITS seem to slightly underperform as we can see from Table III.

TABLE II. INDIVIDUAL PARAMETERS

	INDIVIDUAL PARAMETERS DATASET A1			
dataset A1	SG	VOL	SI	P2
STS	16.94	2.88	4.56	21.23
ITS	13.54	2.84	4.17	15.65
FOCE	13.53	2.85	4.22	15.60
SAEM	13.49	2.86	4.24	15.58
BAYES	13.73	2.78	4.04	16.21
IMP	13.56	2.84	4.24	15.65
IMPMAP	13.59	2.86	4.32	15.63

	INDIVIDUAL PARAMETERS DATASET A2			
dataset A2	SG	VOL	SI	P2
STS	33.39	7.73	11.44	40.99
ITS	16.05	3.77	6.14	21.32
FOCE	16.00	3.79	6.18	21.36
SAEM	15.60	3.74	6.34	21.18
BAYES	15.48	3.91	6.10	19.69
IMP	15.55	3.74	6.13	20.74
IMPMAP	15.33	3.69	6.10	20.43

	INDIVIDUAL PARAMETERS DATASET A3			
dataset A3	SG	VOL	SI	P2
STS	59.69	15.91	25.20	100.15
ITS	23.75	8.23	11.32	26.09
FOCE	23.30	8.42	11.26	24.67
SAEM	24.40	8.39	11.30	28.40
BAYES	23.06	8.07	11.84	23.11
IMP	23.07	8.37	11.20	24.03
IMPMAP	23.91	8.44	11.37	25.04

Square root of the mean square error (RMSE) of the individual parameter estimates expressed as percentage of the true population mean for the three datasets.

TABLE III. RESIDUAL UNKNOWN VARIABILITY

	RESIDUAL UNKNOWN VARIABILITY
dataset A1	estimated CV
TRUE	2.000%
STS	1.686%
ITS	1.967%
FOCE	1.970%
SAEM	1.970%
BAYES	1.957%
IMP	1.967%
IMPMAP	1.970%

Estimated CV with the different methods for dataset A1.

IV. CONCLUSIONS

We have confirmed that the population approach behaves better than the individual approach and that this trend is more evident with the samples reduction in the dataset. Not all the population estimation methods perform equally as well. We suggest to use ITS and FOCE since they are stable to initial estimates and at the same time they produce reliable estimates. BAYES is less stable but it produces comparable and maybe improvable estimates if more informative population priors are given. Finally, from this analysis we do not recommend to use SAEM, IMP and IMPMAP as they perform slightly worse and are less stable.

REFERENCES

[1] R. J. Bauer , S. Guzy , and C. Ng, "A Survey of Population Analysis Methods and Software for Complex PKPD Models with Examples", *APPS J*, vol. 9, no. 1, pp. E60-E83, 2007.

[2] S. L. Beal, L. B. Sheiner, and A. J. Boeckmann, "NONMEM Users Guides", Icon Development Solutions, Maryland, USA, 1989-2012.

[3] P. Denti, A. Bertoldo, P. Vicini, and C. Cobelli, "Nonlinear Mixed Effects to Improve Glucose Minimal Model Parameter Estimation: A simulation Study in Intensive and Sparse Sampling", *IEEE Trans Bio Eng,* vol. 56, no. 9, pp. 2156-66, 2009.

[4] P. Vicini, and C. Cobelli, "The iterative two-stage population approach to IVGTT minimal modeling: improved precision with reduced sampling. Intravenous glucose tolerance test", *Am J Physiol. Endocrinol. Metab.,* vol. 280, no. 1, pp. E179-86, Jan 2001.

[5] O. F. Agbaje, S. D. Luzio, A. I. Albarrak *et al.*, "Bayesian hierarchical approach to estimate insulin sensitivity by minimal model", *Clin Sci (Lond),* vol. 105, no. 5, pp. 551-60, Nov, 2003.

[6] DJ. Lunn, N. Best, A. Thomas, J. Wakefield, D. Spiegelhalter, "Bayesian analysis of population PK/PD models: general concepts and software", J PKPD,vol. 29, pp. 271-307, 2002.

[7] The RFPK team - University of Washington. "System for Population Kinetics (SPK)," http://spk.rfpk.washington.edu.

[8] R. Bauer, S. Guzy, Monte Carlo expectation maximization (MC-PEM) method for analyzing population PK/PD data. In: D'Argenzio DZ, ed. Advanced Methods of PK/PD System Analysis, Kluwer Academic Publishers,vol. 3,pp.135-163, 2004.

[9] B. Deylon, M. Laveille, E. Moulines, "Convergence of a stochastic approximation version of the EM algorithm" Ann Stat, vol. 27, no. 1, pp. 94-128, 1999.

[10] P. H. Barrett, B. M. Bell, C. Cobelli *et al.*, "SAAM II: Simulation, Analysis, and Modeling Software for tracer and pharmacokinetic studies" *Metabolism,* vol. 47, no. 4, pp. 484-92, Apr 1998.

[11] P. Denti, A. Bertoldo, P. Vicini, and C. Cobelli, "IVGTT glucose minimal model covariate selection by NLME approach", *Am J Physiol Endocrinol Metab.,*vol. 298,no. 5,pp. E950–E960, May 2010.

[12] M. Davidian, and D. M. Giltinan, *Nonlinear models for repeated measurement data*, Boca Raton, Fla.: Chapman & Hall/CRC, 1998.

[13] R. M. Savic, and M. O. Karlsson, "Importance of Shrinkage in Empirical Bayes Estimates for Diagnostics: Problems and Solutions", *AAPS J*, vol. 11, no. 3, pp. 558-569, September 2009.