Population and Individual Minimal Modeling of the Frequently Sampled Insulin-Modified Intravenous Glucose Tolerance Test

Lars Erichsen, Olorunsola F. Agbaje, Stephen D. Luzio, David R. Owens, and Roman Hovorka

Population approaches are more robust estimators of insulin sensitivity (S_I) and glucose effectiveness (S_G) with the minimal model of glucose kinetics during an intravenous glucose tolerance test (IVGTT). We assessed the performance of 3 population methods, iterative two-stage (ITS), Bayesian hierarchical Markov chain Monte Carlo (MCMC), and NONMEM first-order conditional estimation (FOCE) with interaction (NM), and made a comparison with the standard two-stage method (STS) employing the weighted nonlinear regression analysis. To evaluate accuracy of individual and population estimates, 40 simulated insulin-modified frequently sampled IVGTTs (IM-FSIVGTT) were derived from real IM-FSIVGTTs (0.3 g glucose per kg body weight with 0.02 U/kg insulin at 20 minutes; 30 samples over 180 minutes) performed in 40 healthy Caucasian subjects (male/female, 22/18; age, 46 ± 9 years; body mass index [BMI], 26.7 ± 5.7 kg · m^{-2}; mean ± SD). The population methods assumed a log-normal population distribution of parameters. All methods gave a similar but overestimated population S_G by 9% to 13%. Population S_I was underestimated to a different degree by the methods (STS 6%, ITS 10%, MCMC 13%, and NM 7%). The between-subject variability of S_G was overestimated by STS and underestimated by the population methods (true 33%, STS 40%, ITS 19%, MCMC 24%, NM 24%; coefficient of variation). For S_I, this quantity was well estimated by all methods (true 79%, STS 80%, ITS 82%, MCMC 83%, NM 82%). The results for individual estimates indicate that STS performs better than the population methods when estimating S_I (STS 12%, ITS 16%, MCMC 16%, NM 16%; 1 outlying subject excluded; root mean squared error expressed as percent of mean) but worse for S_G (STS 28%, ITS 21%, MCMC 20%, NM 19%). We conclude that the robust performance of population approaches, preventing parameter estimation failures associated with the nonlinear regression analysis, is not required with IM-FSIVGTT in subjects with normal glucose tolerance. The standard two-stage technique is the preferred method under such circumstances.

THE MINIMAL MODEL of glucose disposal is the reference method to estimate insulin sensitivity (S_I) and glucose effectiveness (S_G) from the intravenous glucose tolerance test (IVGTT).[1,2] The minimal model has been widely used in laboratory and population studies. The individual method, the so-called standard two-stage (STS) approach,[3] which employs the nonlinear regression analysis to analyze data from each subject independently to provide individual parameter estimates, is traditionally used due to its methodological and computational simplicity.

With reduced sampling or in highly insulin-resistant subjects, individual parameters can be estimated with low precision or it might not be possible to obtain estimates at all.[4] Population methods perform better by assuming that all individuals arise from the same homogenous population, enabling information across subjects to be shared. This improves the precision[5] and prevents estimation failures,[6] but may decrease accuracy inducing bias if the assumption about the population distribution is not met. At present it is unclear whether population methods should be preferred over the individual approach in case of frequently sampled data. Several comparisons have been made between the individual approach and population approaches. De Gaetano et al[7] tested the NONMEM mixed effect approach[8,9] and found improved precision of group parameter estimates. Agabaje et al[6] evaluated the Bayesian hierarchical method and documented a more robust performance in subjects with type 2 diabetes. However, a systematic evaluation and comparison is not available.

The present study compares the individual approach with 3 population methods. The performance of STS, iterative two-stage,[5] NONMEM first-order conditional estimation (FOCE) with interaction, and Bayesian hierarchical methods is evaluated on simulated data to facilitate the assessment of accuracy of parameter estimates. Simulations are based on minimal model parameters obtained from experimental data collected in healthy subjects during a frequently sampled insulin-modified IVGTT (IM-FSIVGTT) to represent faithfully the characteristics of the target population.

METHODS

Synthetic Data Sets and Experimental Design

The present study employed synthetic glucose profiles and experimentally measured insulin profiles. This facilitated evaluation of the accuracy of the estimation methods as the "true" values of the minimal model parameters were known.

To achieve plausibility of the synthetic glucose profiles and to represent as faithfully as possible their intersubject variability, the synthetic data were generated from experimental glucose data.

Synthetic data sets were obtained in a 3-step process exploiting experimental data collected during an IM-FSIVGTT (glucose dose 0.3 g/kg body weight given at 0 minutes over 2 minutes, followed by 0.02 U/kg insulin at 20 minutes; Actrapid, Novo Nordisk, Bagsvaerd, Denmark) in 40 healthy Caucasian subjects (male/female, 22/18; age, 46 ± 9 years; body mass index [BMI], 26.7 ± 5.7 kg · m^{-2}; mean ± SD). Blood samples were taken at 0, 1, 2, 3, 4, 5, 6, 7, 8, 10, 12, 14, 16, 19,

From Pharmacometrics, Novo Nordisk A/S, Bagsvaerd, Denmark; Diabetes Modelling Group, Department of Paediatrics, University of Cambridge, Cambridge, UK; and the Diabetes Research Unit, University of Wales College of Medicine, South Glamorgan, UK.

Submitted Janaury 20, 2004; accepted April 19, 2004.

O.F.A. was supported by the Engineering and Physical Sciences Research Council, UK (CASE studentship).

Address reprint requests to Roman Hovorka, PhD, Diabetes Modelling Group, Department of Paediatrics, University of Cambridge, Box 116, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QQ, UK.

0026-0495/04/5310-0042$30.00/0

doi:10.1016/j.metabol.2004.04.011

22, 23, 24, 25, 27, 30, 40, 50, 60, 70, 80, 90, 100, 120, 150, and 180 minutes for the measurement of plasma glucose and insulin.

In the first step, the experimental data were used to estimate individually minimal model parameters using the nonlinear regression analysis. In the second step, these minimal model parameter estimates together with experimental plasma insulin were used to generate new synthetic plasma glucose concentrations. The minimal model Equations (1-2) with the individually estimated model parameters were used in this process employing measured plasma insulin as the driving function to simulate continuous plasma glucose, which was sampled using the sampling schedule as above. In the final third step, a normally distributed random noise with a 7% coefficient of variation (CV) was added to these samples to represent the measurement and model-misspecification error. This level of CV was obtained from the analysis of real data sets. The data were generated using package SAAM II v 1.2 (SAAM Institute, Seattle, WA).

Only synthetic plasma glucose and measured plasma insulin levels were used in subsequent analyses.

Data Analysis

Minimal model of glucose tolerance during IVGTT. The minimal model defines insulin sensitivity (S_I, the ability of insulin to enhance the net glucose disappearance from plasma) and glucose effectiveness (S_G, ability of glucose to promote its own disposal)[1,2] and is described by 2 differential equations:

$$dg/dt = -[p_1 + x(t)]g(t) + p_1 g_b \quad g(0) = g_b + D/V \qquad (1)$$

$$dx/dt = -p_2 x(t) + p_3[i(t) - i_b] \quad x(0) = 0 \qquad (2)$$

where g(t) is the plasma glucose concentration, i(t) is the plasma insulin concentration, x(t) is insulin action above basal, g_b and i_b are end-test glucose and insulin concentrations, D is the amount of exogenous glucose injected at time 0, and p_1, p_2, p_3, and V are model parameters. The parameter p_2 is inversely related to half-time of insulin action, $t_{1/2} = \ln(2)/p_2$. The parameter V represents the distribution volume of glucose.

Insulin sensitivity is defined as the ratio $S_I = p_3/p_2$, and glucose effectiveness as $S_G = p_1$.

Glucose measurements prior to 8 minutes are normally excluded from the parameter estimation to reduce the error made by neglecting mixing by the 1-compartment model.

Estimation Methods

We used 4 estimation methods; a standard individual approach, which employs the weighted nonlinear regression analysis, and 3 population approaches, which estimate, simultaneously, population characteristics (the population mean and variance) and individual parameters. We used a parameterization given by S_G, S_I, p_2, and V on a log-scale to ensure positive estimates. We also assumed that the glucose measurement errors are uncorrelated, normally distributed with a zero mean and a constant CV.

Standard two-stage method (STS)—individual estimation method. This is the simplest approach to estimate individual as well as population means and variances of the minimal model parameters. In the first stage, individual parameters are estimated for each subject by the weighted nonlinear regression analysis. In the second stage population parameters are calculated as the mean and variance of the individually estimated parameters. The package SAAM II v 1.2 was used for individual parameter estimation.

Iterative two-stage method (ITS). The method, proposed by Steimer et al[10] and used by Vicini et al[5] with the minimal model, is based on an empirical Bayes maximum a posteriori probability estimator. It is an expectation maximization (EM) iterative method.[11] In each iteration, a new individual parameter estimate is chosen to balance the distance of model predictions from individual observations, and the distance of the individual parameters from the mean. The STS method was used to generate starting parameter values. Further details are given in the Appendix. The package SAAM II Population Kinetics v 1.0.1 was used to perform the calculations.

First order conditional estimation with interaction (NM). The package NONMEM[12] was used with the FOCE method with interaction to obtain population and individual estimates of the parameters following the work by De Gaetano et al.[7]

This is also an iterative approach. In each iteration, a Laplace-like approximation of the likelihood function is used. This yields an approximate normal likelihood function, which is maximized to obtain new estimates of population parameters. With these new population estimates, the subject-specific parameters are updated by an empirical Bayes maximum posterior probability estimator, similar to that used by ITS.

Bayesian hierarchical analysis (MCMC). A complete Bayesian framework for modeling the time-varying glucose profile during IVGTT and between-subject variability requires a 3-stage hierarchical model as described by Agbaje et al for the minimal model.[6] The first level describes how observations are obtained from individual parameters. The second stage describes how individual parameters are obtained (drawn from) population parameters. The third and final stage describes how the population parameters are drawn from prior distributions. We adopted vague (noninformative) prior distributions representing lack of prior information about parameter values.

The main purpose of the Bayesian inference[13] is to determine the posterior distribution of individual and population parameters. For the minimal model, this cannot be achieved by direct (analytical) computations. Instead, sampling techniques such the Markov chain Monte Carlo (MCMC) method have to be used.[14] These sampling techniques provide a large sample of the posterior distributions normally running into thousands to tens of thousands of samples. The posterior distributions were summarized by the median. We employed for the calculations the public domain WinBUGS program[15] extended by a purpose-made module implementing the numerical solution of Equations (1-2). Details about the Bayesian hierarchical analysis are given in the Appendix.

Statistical Analysis

The Kolmogorov-Smirnov test together with the histogram was employed to test for log-normality of the minimal model parameters estimated from the original experimental data. The population estimates were summarised as geometric mean and between-subject variability expressed as percent of the mean value. The square root of the mean square error (RMSE) was used to assess the accuracy of individual estimates

$$\mathrm{RMSE} = \sqrt{\frac{1}{40}\sum_{i=1}^{40}(p_i - \hat{p}_i)^2}$$

where p_i is the true parameter value and $\hat{p}_i$ is the parameter estimate in subject *i*. RMSE was expressed as percent of the population mean to facilitate comparison among parameters.

RESULTS

Plasma Insulin and Synthetic Glucose Concentrations

Plasma insulin and synthetic glucose concentrations are shown in Fig 1. Distributions of the minimal model parameters estimated from the real IVGTT experiments, and used to generate glucose data, are shown in Fig 2. These do not indicate serious departures from log-normality. The Kolmogorov-Smirnov test of normality was significant for p_2 ($P = .007$) only.

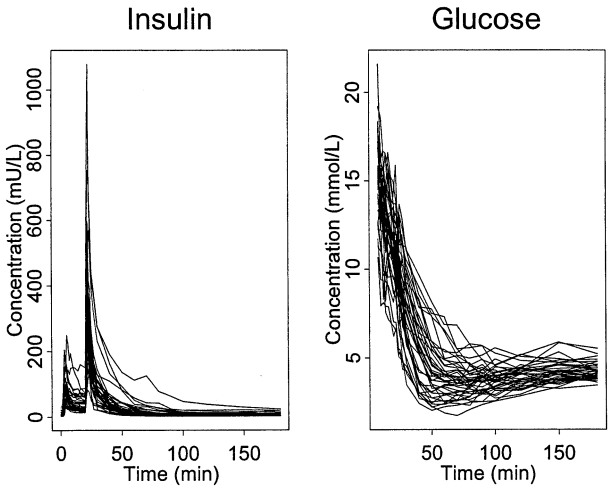

Fig 1. Individual profiles of measured insulin and synthetic plasma glucose values.

This departure from log-normality was ignored in the subsequent analyses.

Comparison of the Four Estimation Methods

The nonlinear regression analysis successfully estimated the minimal model parameters in all subjects; thus, no bias was induced by excluding subjects due to estimation failures.

The population mean and the between-subject variability are shown in Table 1. All methods gave a similar but overestimated population mean of S_G by 9% to 13%. S_I was underestimated to a different degree by the methods (STS 6%, ITS 10%, MCMC 13%, and NM 7%).

The between-subject variability of S_G was overestimated by STS and underestimated by the population methods (true 33%, STS 40%, ITS 19%, MCMC 24%, NM 24%; CV). For S_I, this quantity was well estimated by all methods (true 79%, STS 80%, ITS 82%, MCMC 83%, NM 82%). The between-subject variability of S_I and V is well estimated with all methods.

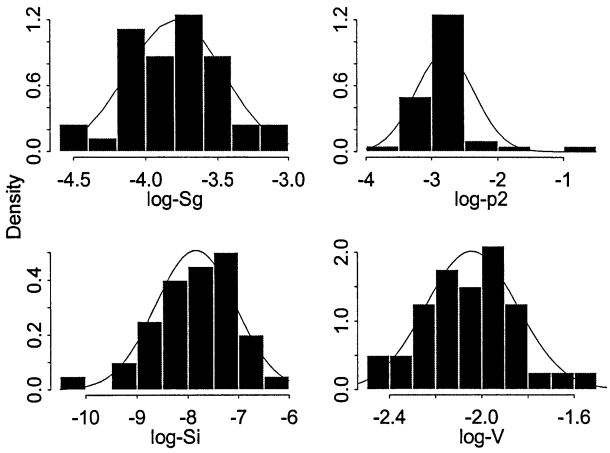

Fig 2. Distributions of log-transformed true minimal model parameters used to generate glucose data, scaled as a density with the normal distribution superimposed.

Table 1. Population Estimates of the Minimal Model Parameters Given as the Geometric Mean and the Between-Subject Variability of the Four Estimation Methods

Method	S_G (min^{-1})	S_I ($10^{-4} \times min^{-1}$ per mU · L^{-1})	p_2 (min^{-1})	V (L · kg^{-1})
True	0.023 (33*)	3.9 (79)	0.061 (45)	0.13 (20)
STS	0.025 (40)	3.7 (80)	0.057 (60)	0.12 (22)
ITS	0.025 (19)	3.5 (82)	0.061 (37)	0.13 (20)
MCMC	0.026 (20)	3.4 (83)	0.062 (38)	0.12 (19)
NM	0.026 (24)	3.6 (82)	0.057 (38)	0.12 (19)

Abbreviations: S_G, glucose effectiveness; S_I, insulin sensitivity; p_2, parameter of the minimal model inversely related to the duration of insulin action; V, volume of glucose distribution; True, true values in the synthetic data set; STS, standard two-stage method; ITS, iterative two-stage method; MCMC, Markov chain Monte Carlo method; NM, NONMEM first-order conditional estimation with interaction.

*Between-subject variability as %CV.

Results related to individual estimates are shown in Table 2 and Figs 3 and 4. STS performed better on S_I when 1 outlier (subject no. 5) was disregarded but worse on S_G and p_2 compared to the population methods. The outlying S_I was well estimated by NONMEM, although this may be a coincidence. The population methods tended to bias individual estimates of S_G towards the mean but the estimates are less scattered than these for STS. The volume of distribution was estimated with the smallest relative RMSE (8% to 9%) and p_2 with the highest RMSE (50% to 63%).

DISCUSSION

The present study demonstrates that in healthy middle-aged subjects with IM-FS IVGTT a drawback of the more sophisticated population approaches outweighs the benefits, and the individual approach is preferable to estimate individual and population insulin sensitivity S_I. This is offset by impaired performance in relation to estimating individual glucose effectiveness S_G.

The drawback of the population methods is that they impose a structure of a log-normal population distribution. On the one hand, this explicit parametric specification has the advantage that in the estimation of subject parameters, information from the remaining subjects facilitates robust performance and reduces the variance of the estimates. However, the imposed structure may induce a bias if the "true" population distribution is misrepresented. A misrepresented distribution, eg, a normal

Table 2. Square Root of the Mean Square Error of Individual Parameter Estimates Expressed as Percentage of the True Mean

Method	S_G (min^{-1})	S_I ($10^{-4} \times min^{-1}$ per mU · L^{-1})	p_2 (min^{-1})	V (L · kg^{-1})
STS	28	12 (17*)	63	9
ITS	21	16 (25)	51	8
MCMC	20	16 (22)	50	9
NM	19	16 (15)	52	8

NOTE. For S_I, results exclude and, in brackets, include an outlying subject no. 5.

*Subject no. 5 included.

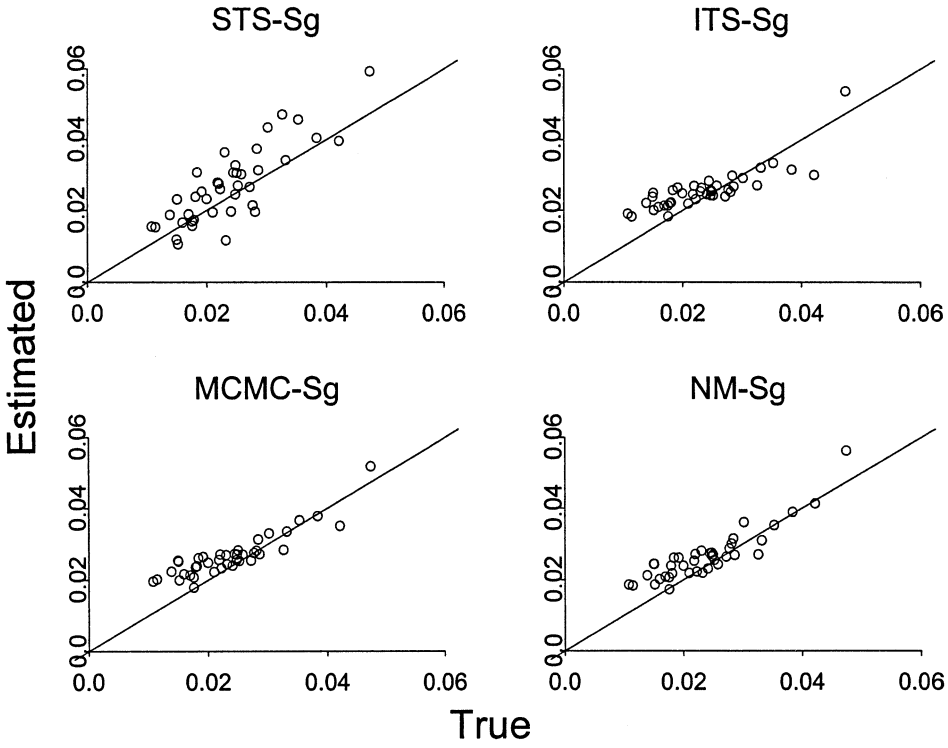

Fig 3. True and estimated individual values of S_G (min^{-1}) with the STS, ITS, MCMC, and NM approaches. The lines indicate identity of true and estimated values.

instead of skewed distribution, may decrease accuracy of individual and population estimates. This is the likely cause for the population methods to under perform slightly compared to the individual approach.

The detailed examination of insulin sensitivity S_I in Fig 2 reveals that its true distribution lies between normal and log-normal. It is skewed left when log-transformed. Although this skewing is not sufficient to fail the normality test a conse-

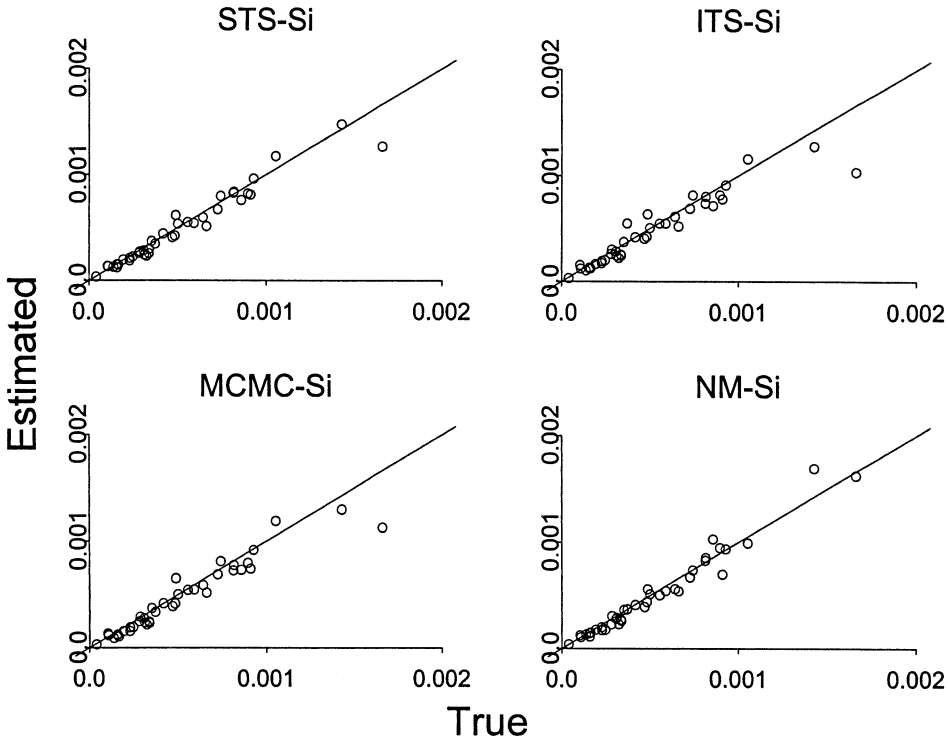

Fig 4. True and estimated individual values of S_I (min^{-1} per $mU \cdot L^{-1}$) with the STS, ITS, MCMC, and NM approaches. The lines indicate identity of true and estimated values.

quence is that population methods underestimate the population mean. This is a well-documented problem associated with approximating a skewed distribution by a normal distribution. The MCMC method appears to be most affected as it implements most rigorously the population structure.

We did not experience any parameter estimation failures with the individual method, which employs the nonlinear regression analysis. The subjects had normal insulin sensitivity, insulin modification provided a strong signal to enhance glucose disposal and to suppress the endogenous glucose production, and frequent sampling allowed separation and identification of the insulin-induced glucose lowering on the glucose excursion curve. This appears to guarantee robust and accurate estimation of insulin sensitivity by the individual approach, matching robustness of population estimators but avoiding the constraining assumption of the parametric population distribution.

It has been shown that the minimal model fails in up to 50% cases in highly insulin-resistant subjects,[16] although with insulin modification and careful data analysis, the failure rate is normally around 10%. Godsland and Walton[17] documented even higher failure rates with a high glucose dose (0.5 g/kg) but without insulin modification. The failure of the individual approach may bias the results since failures often occur in subjects with extreme values for some parameters, eg, at very low values of S_I,[6] especially in combination with reduced sampling schemes. The use of the population methods is then justified as they are more robust and prevent estimation failures. The variance of individual parameter estimates is much larger with sparse data implying that the bias induced by the population methods is a smaller fraction of the total error.[5]

The population means are similar among the 4 methods, with a small overestimation of S_G and an underestimation of S_I. Population methods provide less scattered estimates due to sharing of information across individuals, which also improves the precision of the estimates.[5] In the present study, the population variation is generally overestimated with the individual and underestimated with the population methods. This latter behavior is caused by the individual approach neglecting and the population methods exploiting the resemblance among subjects. The variability among subjects is penalized with the population methods regressing individual subjects towards the mean value.

The greatest advantage of the individual method is that it is simple and fast, whereas NONMEM and MCMC are the most complicated and computationally expensive with running times of about 10 hours on a Pentium III 700 MHz processor for the present data set. A run with full population covariance matrix (results not shown) was performed for the MCMC and NONMEM methods, but this did not improve the estimation of individual parameters.

If normality/log-normality is a reasonable assumption NONMEM seems to be the method of choice to guarantee robust performance. If this assumption is not met, a better performance may be obtained with the MCMC method due to its flexibility in specifying the population distribution. In contrast to the remaining population methods, MCMC can accommodate heavy tailed or skewed distributions such as the t- and gamma-distributions or the population distribution could be specified as a mixture of, eg, normal and log-normal distributions. Further investigations, especially in subjects with type 2 diabetes, should clarify these issues.

In conclusion, while it is important to avoid parameter estimation failures associated with the nonlinear regression analysis, the robust performance of the population approaches is not required with the IM-FSIVGTT in subjects with normal glucose tolerance. The standard two-stage technique is the preferred method under such circumstances.

APPENDIX

All 3 population approaches adopt a hierarchical, 2- or 3-stage structure for the parameter estimation problem. At the first stage, glucose values g_{ij} in subject i at time t_{ij} are obtained as the solution to Equations (1-2):

$$g_{ij} = g(t_{ij}, S_{Gi}, S_{Ii}, p_{2i}, V_i) \cdot (1 + \varepsilon_{ij}) \quad (4)$$

where S_{Gi}, S_{Ii}, p_{2i}, and V_i are parameters of subject i, ε_{ij} is a random term representing the multiplicative measurement error, which is assumed to be normally distributed with a zero mean and an unknown variance σ^2. At the second stage, subjects are assumed to be representatives from the same homogenous population. In particular, we assumed that the individual parameters μ_i are drawn from a multivariate log-normal distribution:

$$\mu_i = [\log(S_{Gi}), \log(S_{Ii}), \log(p_{2i}), \log(V_i)] \sim N(\mu, \Sigma) \quad (5)$$

where μ is an unknown population mean vector, and Σ is an unknown covariance matrix, with all off-diagonal values equal to 0.

For the Bayesian analysis a third stage is required, which specifies prior distributions of population parameters. We assumed:

$$\mu_j \sim \text{Normal}(0, 10^6) \quad j = 1, \ldots, 4 \quad (6)$$

$$\sigma^{-2}, \Sigma_{jj}^{-1} \sim \text{Gamma}(0.001, 0.001) \quad j = 1, \ldots, 4 \quad (7)$$

The particular parameters of the prior distributions were chosen to specify a lack of prior information of parameters although information is contained in the form of the chosen distribution population distribution.

The Bayesian method was implemented with a burn-in of 6,000 samples and using every fourth of 20,000 further samples of the chain. A simple check for convergence was done by adding another 20,000 samples. Since this did not change the population or individual estimates, it was assumed that convergence was obtained.

REFERENCES

1. Bergman RN, Ider YZ, Bowden CR, et al: Quantitative estimation of insulin sensitivity. Am J Physiol 236:E667-E677, 1979

2. Bergman RN, Finegood DT, Ader M: Assessment of insulin sensitivity in vivo. Endocr Rev 6:45-86, 1985

3. Hovorka R, Vicini P: Parameter estimation, in Carson ER, Cobelli C (eds): Modelling Methodology for Physiology and Medicine. San Diego, CA, Academic Press, 2001, pp 107-151

4. Pillonetto G, Sparacino G, Magni P, et al: Minimal model S(I)=0 problem in NIDDM subjects: Nonzero Bayesian estimates with credible confidence intervals. Am J Physiol 282:E564-E573, 2002

5. Vicini P, Cobelli C: The iterative two-stage population approach to IVGTT minimal modeling: Improved precision with reduced sampling. Intravenous glucose tolerance test. Am J Physiol 280:E179-E186, 2001

6. Agbaje OF, Luzio SD, Albarrak AIS, et al: Bayesian hierarchical approach to estimate insulin sensitivity by minimal model. Clin Sci 105:551-560, 2003

7. De Gaetano A, Mingrone G, Castageneto M: NONMEM improves group parameter estimation for the minimal model of glucose kinetics. Am J Physiol 271:E932-E937, 1996

8. Davidian M, Giltinan D: Nonlinear Models for Repeated Measurement Data. New York, NY, Chapman & Hall, 1995

9. Beal SL, Sheiner LB: NONMEM User's Guide. San Francisco, CA, NONMEM Project Group, University of California, 1992

10. Steimer JL, Mallet A, Golmard JL, et al: Alternative approaches to estimation of population pharmacokinetic parameters: Comparison with the nonlinear mixed-effect model. Drug Metab Rev 15:265-292, 1984

11. Schumitzky A: EM algorithms and two stage methods in pharmacokinetic population analysis, in D'Argenio DZ (ed): Advanced Methods of Pharmacokinetic and Pharmacodynamic Systems Analysis. New York, NY, Plenum, 1995, pp 145-160

12. Beal SL, Shreiner LB: NONMEM User's Guide. San Francisco, CA, NONMEM Project Group, University of California, 1992

13. Lee PM: Bayesian Statistics: An Introduction. London, UK, Edward Arnold, 2002

14. Gilks WR, Richardson S, Spiegelhalter DJ: Markov Chain Monte Carlo in Practice. London, UK, Chapman & Hall, 1996

15. Lunn DJ, Thomas A, Best N, et al: WinBUGS—A Bayesian modelling framework: Concepts, structure, and extensibility. Stat Comput 10:325-337, 2000

16. Saad MF, Anderson RL, Laws A, et al: A comparison between the minimal model and the glucose clamp in the assessment of insulin sensitivity across the spectrum of glucose-tolerance. Diabetes 43:1114-1121, 1994

17. Godsland IF, Walton C: Maximizing the success rate of minimal model insulin sensitivity measurement in humans: The importance of basal glucose levels. Clin Sci 101:1-9, 2001