Population approaches to estimate minimal model indexes of insulin sensitivity and glucose effectiveness using full and reduced sampling schedules

Kevin M. Krudys,1 Steven E. Kahn,2 and Paolo Vicini1

1Resource Facility for Population Kinetics, Department of Bioengineering, University of Washington, and 2Division of Metabolism, Endocrinology and Nutrition, Department of Medicine, Veterans Affairs Puget Sound Health Care System, University of Washington, Seattle, Washington

Submitted 28 July 2005; accepted in final form 30 March 2006

Krudys, Kevin M., Steven E. Kahn, and Paolo Vicini. Population approaches to estimate minimal model indexes of insulin sensitivity and glucose effectiveness using full and reduced sampling schedules. Am J Physiol Endocrinol Metab 291: E716–E723, 2006; doi:10.1152/ajpendo.00346.2005.—The intravenous glucose tolerance test (IVGTT) interpreted with the minimal model provides individual indexes of insulin sensitivity (SI) and glucose effectiveness (SG). In population studies, the traditional approach, the standard two-stage (STS) method, fails to account for uncertainty in individual estimates, resulting in an overestimation of between-subject variability. Furthermore, in the presence of reduced sampling and/or insulin resistance, individual estimates may be unobtainable, biasing population information. Therefore, we investigated the use of two population approaches, the iterative two-stage (ITS) method and nonlinear mixed-effects modeling (NM), in a population (n = 235) of insulin-sensitive and insulin-resistant subjects under full (FSS, 33 samples) and reduced [RSS(240-min), 13 samples and RSS(180-min), 12 samples] IVGTT sampling schedules. All three population methods gave similar results with the FSS. Using RSS(240), the three methods gave similar results for SI, but SG population means were overestimated. With RSS(180), SI and SG population means were higher for all three methods compared with their FSS counterparts. NM estimated similar between-subject variability (19% SG, 53% SI) with RSS(180), whereas ITS showed regression to the mean for SG (0.01% SG, 56% SI) and STS provided larger population variability in SI (29% SG, 91% SI). NM provided individual estimates for all subjects, whereas the two-stage methods failed in 16–18% of the subjects using RSS(180) and 6–14% using RSS(240). We conclude that population approaches, specifically NM, are useful in studies with a sparsely sampled IVGTT (~12 samples) of short duration (~3 h) and when individual parameter estimates in all subjects are desired.

NONMEM: two-stage; parameter estimation; minimal model; insulin sensitivity

THE INTRAVENOUS GLUCOSE TOLERANCE TEST (IVGTT) interpreted with the minimal model of glucose kinetics has found wide application in the field of intermediary metabolism, primarily due to its ability to provide indexes of insulin sensitivity (SI) and glucose effectiveness (SG). Many years after the inception of the minimal model and its subsequent validation studies, the IVGTT has made the leap from small scale studies to large clinical and epidemiological studies aimed at unearthing insulin sensitivity’s genetic basis and insulin resistance’s association with components of the metabolic syndrome (16, 18). In large studies, the traditional approach to obtaining population information on SI and SG is the standard two-stage (STS) method, which constructs the sample mean and covariance from the individual estimates. One problem, however, is that the numerical algorithm used to estimate model parameters may fail in some individuals, so that individual parameter estimates are not available for all subjects in the population. To recover optimal parameter estimates in an individual, the numerical algorithm tries to find the best prediction of the glucose data, based on the insulin time course and the model parameters, by minimizing some objective function. The most common reason for model failure occurs when this objective function does not have a well-defined minimum. Model failure may also occur when parameters are estimated, but their reliability (precision) cannot be determined. Although computationally trivial, STS can introduce bias when individual parameter estimates cannot be precisely estimated, as is often the case in subjects with insulin resistance (12, 13). Furthermore, STS ignores uncertainty in the individual estimates, leading to an estimate of between-subject variability that is upwardly biased (4, 8). These concerns have led some to investigate the use of more sophisticated population analysis techniques, which exploit the fact that subjects belong to a population of individuals that share certain quantitative traits.

De Gaetano et al. (9) were the first to use a population approach, nonlinear mixed-effects modeling (NM), to estimate minimal model parameters in a small group of 20 healthy subjects, demonstrating reduced variability in group parameter estimates relative to the STS approach. Individual estimates, however, were not reported. Vicini and Cobelli (17) used an iterative population approach, the iterative two-stage (ITS) method, which relies on the concept of maximum a posteriori probability empirical Bayes estimation. In a group of 16 healthy subjects, they reported improved precision of individual estimates of SI and SG for the ITS method under both full and reduced sampling schedules of the IVGTT. Most recently, Agbaje et al. (1) investigated the use of a Bayesian hierarchical analysis in 63 subjects with newly developed type 2 diabetes. The Bayesian approach provided individual estimates of SI and SG for all subjects, whereas the individual nonlinear regression technique used in the STS method failed to obtain individual estimates in 6% of the population. Thus there are at least three different population approaches (NM, ITS, and Bayesian) that
have shown applicability in relatively small, selected populations.

Recent simulation studies, however, suggest that population analysis methods, which are usually computationally demanding, may not be a significant improvement to the STS method for subjects with normal glucose tolerance and well-determined insulin sensitivity (10). This approach, however, was somewhat limited because it involved simulation and assessments only in insulin-sensitive subjects. Therefore, we examined two population approaches (ITS and NM), compared with STS, in a situation that may be less than ideal, IVGTTs in a relatively large population ($n = 235$) of subjects spanning a wide range of sensitivity to insulin. Furthermore, we evaluated the two population approaches under both full (FSS) and reduced (RSS) sampling schedules of the IVGTT, the latter of which may typically be thought to be less desirable from a parameter estimation standpoint but may be attractive in the interests of experimental design and reduced cost.

SUBJECTS AND METHODS

Database. The database consisted of 235 IVGTTs (100 men, 135 women; age $52 \pm 10$ yr; body mass index $26.4 \pm 4.4$ kg/m$^2$; means $\pm$ SD) performed in individuals recruited by advertisement to participate in a study to examine the relationship between insulin sensitivity and plasma lipoproteins after egg consumption. The study was approved by the Human Subjects Review Committee at the University of Washington. Details regarding study design, protocol, and measurements can be found elsewhere (11).

Briefly, for the IVGTT, baseline samples were taken at $-15$, $-10$, and $-1$ min before glucose (11.4 g/m$^2$) was injected intravenously over 1 min beginning at time 0. Tolbutamide (125 mg/m$^2$) was administered intravenously over 30 s at time 20 min. During the IVGTT, blood samples were drawn at 2, 3, 4, 5, 6, 8, 10, 12, 14, 16, 19, 22, 23, 24, 25, 27, 30, 35, 40, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 220, and 240 min. These samples comprise the FSS. We considered a 180-min reduced sampling schedule, RSS(180), based on a previous suggestion (14) that included 12 samples taken at 0, 2, 4, 8, 19, 22, 30, 40, 50, 70, 90, and 180 min. Because a 4-h test allows for glycemic levels to return closer to baseline and has been recommended (12) for a previous suggestion (14) that included 12 samples taken at 0, 2, 4, 8, 19, 22, 30, 40, 50, 70, 90, and 180 min. Because a 4-h test allows for glycemic levels to return closer to baseline and has been previously used in a similar context (17), we also considered a second reduced sampling schedule, RSS(240), which is equivalent to the RSS(180) schedule, but with an additional data point included at 240 min.

Minimal model. The minimal model of glucose disappearance describes the kinetics of glucose after a bolus injection. Glucose disappearance is enhanced by glucose independent of insulin as well as by glucose dependent on insulin acting from a compartment remote from plasma. Remote insulin levels increase as a result of rising plasma insulin levels and are cleared through a first-order process. The minimal model of glucose disappearance (5) is given by the equations

$$\frac{dG(t)}{dt} = -[S_G + X(t)]G(t) + S_GG_b \quad G(0) = \frac{D}{V} + G_h$$

$$\frac{dX(t)}{dt} = -p_2[X(t) - S[I(t) - I_b]] \quad X(0) = 0$$

where $G(t)$ (mg/dl) is plasma glucose concentration, $I(t)$ ($\mu$U/ml) is plasma insulin concentration, and $G_b$ and $I_b$ are their basal values. $X(t)$ (min$^{-1}$) is insulin action and $D$ is glucose dose. The model equations yield four uniquely identifiable parameters: $V$ (dl/kg), the volume of distribution of glucose; $p_2$ (min$^{-1}$), the insulin action parameter; $S_G$ (min$^{-1}$), glucose effectiveness, reflecting the ability of glucose per se to stimulate glucose disposal and inhibit glucose production; and $S_I$ (min$^{-1}$-$\mu$U$^{-1}$-ml), insulin sensitivity, representing the ability of insulin to enhance the glucose simulation of glucose disposal and glucose inhibition of glucose production. For the purposes of this study, we report results only on the two most clinically relevant parameters, $S_I$ and $S_G$.

Parameter estimation. For all parameter estimation techniques, the error associated with glucose concentrations was assumed to be independent, Gaussian, with a zero mean and a constant coefficient of variation proportional to the model predicted values. Glucose data before 8 min were not used for parameter estimation.

Standard two-stage. In the first stage of the STS approach, each individual’s parameter values were identified with a weighted nonlinear least squares estimator implemented in the commercially available software SAAM II (2). If, for any individual, the SAAM II estimation routine failed to converge on a solution, parameter estimates for that subject were deemed unobtainable. In the second stage, population information was inferred as the sample mean and sample variance of the individual estimates. If an individual’s parameter estimates were unobtainable in the first stage of the analysis, they were subsequently excluded from the second stage.

Iterative two-stage. The ITS approach was first proposed by Steiner et al. (15) to estimate population pharmacokinetic parameters. In the first step of an ITS analysis, population information is obtained from the STS approach. In the following step, individual parameter estimation is performed again, this time incorporating the population information from the previous step as informative Bayesian priors on the model parameters. The parameter estimates from the second step are often referred to as empirical Bayes estimates and are used to construct population priors in a subsequent step. This recursive procedure is repeated until convergence is reached. We used the software SAAM II Population Kinetics for this approach and determined convergence to be reached when estimates of the population mean, population variance, and individual parameter estimates differ by <1%. The ITS approach, like STS, requires parameter estimation in each individual. Therefore, subjects in which individual estimates were unobtainable were also excluded from the ITS analysis.

Nonlinear mixed-effects modeling. NM models assume a hierarchical structure that specifies parameters that vary (random effects) and do not vary (fixed effects) in the population. In the $i$th individual, the model takes the general form

$$y_i = f_i(\beta_i) + \varepsilon_i$$

where $y_i$ are the measurements, $f_i(\beta_i)$ are the model functions (i.e., minimal model), $\beta_i$ is the parameter vector of the $i$th subject, and $\varepsilon_i$ is the measurement error. The vector $\beta_i = \beta + (\eta_i)$ consists of the fixed effects (i.e., mean $S_G$, mean $S_I$) vector $\beta$, and the subject specific random effect $\eta_i$, representing the $i$th individual’s deviation from the mean.

Population models were estimated using the NONMEM software (4) with first order conditional estimation (FOCE) with interaction. The FOCE method provides population parameters and individual estimates simultaneously. To ensure positive parameter estimates, we assumed log-normal population distributions for all minimal model parameters.

Statistical analysis. Because of the lognormal distribution of parameter estimates, population statistics are presented as geometric means and percent coefficients of variation (%CVs). Agreement between individual parameter estimates was tested with the use of Lin’s concordance correlation coefficient (11a). Lin’s concordance correlation coefficient is a reproducibility measure used to evaluate the agreement between two readings by determining the extent to which observations deviate from the line of identity. Relative error was calculated as $100\% \times (\text{Full} - \text{Reduced[Population]})/\text{Full}$, where Full refers to the estimates obtained from individual nonlinear regression and Reduced[Population] represents the individual estimates obtained from reduced sampling schedule and/or a population approach. Statistical analysis was assessed by one-sample Student’s t-test.
RESULTS

Individual plasma glucose and insulin data are shown in Fig. 1. Fasting levels of glucose and insulin in the population were 97 ± 8 mg/dl and 10 ± 7 μU/ml (means ± SD), respectively. A total of 78 of 235 subjects had impaired fasting glucose (IFG) levels, defined as baseline glucose levels between 100 and 125 mg/dl. Reference estimates of SI and SG obtained from individual linear regression with the FSS are plotted in Fig. 2 to show the distribution of these parameters in the study population.

FSS. Estimates of population means and between-subject variability for SI and SG using the three methods are presented in Table 1. All methods estimated similar population means, but the ITS and NM methods provided similarly lower estimates of between-subject variability. The ITS method resulted in lower individual %CVs for SG 13% (5–18%) median (5th–95th percentile) and SI 7% (3–19%) than the STS [SG, 19% (6–60%); SI, 10% (3–27%)]. NM provides estimates of the percent relative standard error (%RSE) in the estimation of the population mean, which for the case of FSS were 1.88% for SG and 3.81% for SI.

To assess individual estimates of SI and SG obtained by NM and ITS, we assumed estimates obtained from nonlinear regression in each individual separately (the first stage of a two-stage analysis) are the “gold-standard” or reference estimates. Parameter estimates were unobtainable in five subjects (2%) by the two-stage methods and were therefore left out of the analyses. Model failure occurred in four subjects with IFG (5%) and one subject with normal fasting glucose (0.6%). The NM approach provided estimates in all subjects. Figure 3 shows the performance of the population techniques compared with the reference estimates obtained in the STS. The three SI estimates falling the farthest from the line of unity compared with NM and the ITS methods were those in which the STS provided its most unreliable estimates, with individual %CVs in these subjects ranging from 92–450% (Fig. 3). These unsatisfactory estimates were therefore excluded from %RE and correlation calculations, as well as one outlier that had an
unphysiological estimate that was $>50 \times 10^{-4}$ min$^{-1}$·μU$^{-1}$·ml. In the case of $S_G$, the least amount of agreement can be seen at both low and high values, reflecting regression to the mean in the population approaches. Nevertheless, individual estimates obtained from ITS and NM showed strong correlation and minimal bias compared with the STS reference estimates (Table 2).

### Table 1. Population estimates of $S_I$ and $S_G$

<table>
<thead>
<tr>
<th>Method</th>
<th>Number of Subjects Included</th>
<th>$S_I$, $10^4$ min$^{-1}$·μU$^{-1}$·ml</th>
<th>$S_G$, min$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Full sampling schedule</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard 2-stage</td>
<td>230 (98%)</td>
<td>5.20 (64)</td>
<td>0.0177 (31)</td>
</tr>
<tr>
<td>Iterative 2-stage</td>
<td>230 (98%)</td>
<td>5.05 (55)</td>
<td>0.0181 (22)</td>
</tr>
<tr>
<td>Nonlinear mixed-effects modeling</td>
<td>235 (100%)</td>
<td>5.17 (55)</td>
<td>0.0184 (24)</td>
</tr>
<tr>
<td><strong>Reduced sampling schedule (240 min)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard 2-stage</td>
<td>221 (94%)</td>
<td>5.16 (67)</td>
<td>0.0189 (29)</td>
</tr>
<tr>
<td>Iterative 2-stage</td>
<td>202 (86%)</td>
<td>5.06 (58)</td>
<td>0.0192 (12)</td>
</tr>
<tr>
<td>Nonlinear mixed-effects modeling</td>
<td>235 (100%)</td>
<td>5.18 (56)</td>
<td>0.0201 (19)</td>
</tr>
<tr>
<td><strong>Reduced sampling schedule (180 min)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard 2-stage</td>
<td>197 (84%)</td>
<td>6.12 (92)</td>
<td>0.0194 (29)</td>
</tr>
<tr>
<td>Iterative 2-stage</td>
<td>193 (82%)</td>
<td>5.39 (57)</td>
<td>0.0201 (12)</td>
</tr>
<tr>
<td>Nonlinear mixed-effects modeling</td>
<td>235 (100%)</td>
<td>5.45 (53)</td>
<td>0.0199 (19)</td>
</tr>
</tbody>
</table>

Parameter estimates are given as their geometric means. Numbers in parentheses are estimates of between-subject variability given as percent coefficient of variation (%CV). The number of subjects included in each method varies according to the exclusion of those subjects with estimation failures in the 2-stage methods. $S_I$, insulin sensitivity; $S_G$, glucose effectiveness.

For $S_G$, all methods estimated virtually identical population means for $S_I$ relative to the FSS (Table 1). Similarly, estimates of between-subject variability of $S_I$ were robust compared with the FSS. Of particular note is that $RSS(240)$ does not induce a significant overestimation of between-subject variability (67% CV) vs. the FSS (64%). For the estimation of $S_G$, all methods overestimated the population mean relative to the FSS and showed similar extents of deviation (7% STS, 6% ITS, 9% NM). STS and NM estimated similar levels of between-subject variability for $S_G$, but the ITS method showed significant shrinkage to the mean (1.2% CV).

Individual estimates obtained from the $RSS(240)$ were compared with the reference estimates from individual linear regression and the FSS (Fig. 4). Model failures occurred in 6% of subjects for the STS and ITS analyses, respectively. For the STS method, parameters were unobtainable in 6% of both IFG and normal fasting glucose subjects. For the ITS, estimates could not be obtained in 13% of IFG subjects and 15% of normal fasting glucose subjects. NM provided estimates of $S_I$ and $S_G$ for the entire population. For $S_I$, all three methods showed strong agreement to the reference estimates (Fig. 4 and Table 2), although relative error for the STS estimates was significantly different from 0 ($P < 0.05$). For $S_G$, agreement was weaker for all three methods (Fig. 4 and Table 2), with relative errors all being significantly different from 0 ($P < 0.05$). The weakest correlation was for the ITS method ($r_c = 0.83$), owing to the considerable regression to the mean observed for these estimates. The STS method exhibited the strongest correlation ($r_c = 0.0001$), being negligible as the estimates are essentially those from the first stage of the standard 2-stage (STS) analysis. The line of unity is plotted on all graphs for visual comparison.
the FSS (Table 1). The largest deviation of the RSS(180) estimate of SI from its FSS counterpart was seen for the STS method (18%), whereas ITS (7%) and NM (5%) were more robust. All methods showed similar deviations in estimation of SG relative to FSS (10% STS, 11% ITS, and 8% NM). STS estimated a much higher degree of between-subject variability (92% CV) in SI relative to the other two population approaches. Conversely, ITS showed a dramatic reduction in its estimate of between-subject variability of SG (1.2% CV), compared with any other method. This was also reflected in the estimated individual %CVs of SG, which all converged to 0.01% CV. The same effect was not seen in individual %CVs of SI, however, where the %CV was 9% (4–25%). As expected, individual %CVs from the STS method were larger than those provided by ITS [SG, 27% (6–136%); SI, 18% (6–125%)]. Precision of population estimates (%RE) obtained from NM (SG, 1.66%; SI, 3.96%) were similar to those obtained by the FSS.

Individual estimates were again compared with the reference estimates obtained from individual nonlinear regression using the FSS (Fig. 5). Parameter estimates were unobtainable in 16–18% of the subjects by the two-stage analysis methods. Model failure occurred in 31% of IFG subjects and 9% of subjects with normal fasting glucose for the STS method. Similarly, for the ITS, model failure occurred in 31% of IFG subjects and 12% of subjects with normal fasting glucose. Again, NM provided estimates of SG and SI in all individuals. The individual estimates obtained from STS exhibited the most bias and the worst agreement to the reference estimates of SI (Fig. 5 and Table 2). In addition to providing estimates in all individuals, NM also showed the best agreement to the reference values (SI, 0.88). Individual estimates of SG showed much weaker agreement than SI for all three parameter estimation methods (Table 2). The RSS(180) clearly affects ITS estimation the most, as individual estimates of SG show complete shrinkage to the mean (Fig. 5).

Table 2. Comparisons of individual estimates of SI and SG

<table>
<thead>
<tr>
<th>Method</th>
<th>SI</th>
<th>SG</th>
</tr>
</thead>
<tbody>
<tr>
<td>%RE, mean ± SD</td>
<td>SI</td>
<td>SG</td>
</tr>
<tr>
<td>Full sampling schedule</td>
<td>0.95</td>
<td>0.72 ± 6.12</td>
</tr>
<tr>
<td>Reduced sampling schedule</td>
<td>0.98</td>
<td>0.26 ± 7.56</td>
</tr>
<tr>
<td>Standard 2-stage</td>
<td>0.95</td>
<td>-2.8 ± 18*</td>
</tr>
<tr>
<td>Iterative 2-stage</td>
<td>0.95</td>
<td>-2.5 ± 22</td>
</tr>
<tr>
<td>Nonlinear mixed-effects</td>
<td>0.91</td>
<td>-2.2 ± 17</td>
</tr>
<tr>
<td>Reduced sampling schedule</td>
<td>0.09</td>
<td>-5.9 ± 249*</td>
</tr>
<tr>
<td>Standard 2-stage</td>
<td>0.83</td>
<td>-11 ± 38*</td>
</tr>
<tr>
<td>Nonlinear mixed-effects</td>
<td>0.88</td>
<td>-9 ± 29*</td>
</tr>
</tbody>
</table>

%RE, relative percent error. Individual parameter estimates are related to estimates obtained with individual nonlinear regression (first stage of the 2-stage analysis). Coefficients listed in the table are Lin’s concordance coefficient (p<). %RE is calculated as 100% [Full-Reduced(Population)]/Full, where Full refers to estimates from individual nonlinear regression and Reduced(Population) refers to estimates obtained from a reduced sampling schedule and/or a population approach. *Statistically different from 0 (P < 0.05).

Fig. 4. Individual estimates of SI and SG using the reduced sampling schedule, RSS(240). Estimates obtained from the 3 methods (STS, ITS, NM) are compared with the “gold-standard” estimates obtained from the first stage of the STS, using full sampling schedules. The line of unity is plotted on all graphs for visual comparison.
a high degree of regression to the mean for $S_G$ but compares similarly to the STS method in its agreement with the reference estimates. Relative error estimates were significantly different from 0 ($P < 0.05$) for all estimation methods, suggesting a consistent negative bias when a RSS(180) is considered.

**DISCUSSION**

Our results agree with previous simulation studies (10) that suggest there are no major advantages in using more sophisticated population analysis methods in relatively healthy subjects undergoing a frequently sampled IVGTT. The NM method, however, did provide estimates of $S_I$ and $S_G$ in all individuals. Furthermore, both NM and ITS used population information to improve individual estimates that were poorly estimated with individual regression. Overall, individual estimates showed strong agreement.

Reduced sampling schedules have been suggested (7, 14, 17) as modifications to the IVGTT that would enable its economical use in larger population studies. The STS method is more likely to introduce bias in population mean estimates and inflate between-subject variability under this circumstance, because individual estimates tend to be more uncertain. Indeed, this was the case in the present study in which the estimate of between-subject variability of $S_I$ increased considerably when using the RSS(180). Surprisingly, when an additional data point was added at 240 min, this phenomenon did not markedly manifest itself. In addition, parameter identification failed in 16–18% of subjects with the RSS(180), resulting in an upward bias in the population mean, because those subjects most resistant to insulin are most prone to estimation failure. Also, individual values of $S_I$ estimated with the highest %CV also tended to be the largest in magnitude. The RSS(240) had a lower failure rate (6–14%) for the two-stage methods, which did not appear to have a major impact on estimates of population mean. The lower failure rate with the addition of data at 240 min can most likely be explained by the fact that glucose and insulin levels are closer to baseline at 240 min than at 180 min, especially in those subjects with insulin resistance, providing a more complete picture of glucose disappearance.

More sophisticated methods of population parameter estimation have been routinely applied in pharmacokinetic studies, in which large populations are often sampled sparsely. One of these techniques, the ITS, was proposed for use in metabolic studies as an approach to improve parameter precisions, especially in cases where individual %CVs are expected to be poor, such as with a reduced sampling schedule (17). In a relatively small population of 16 subjects, the ITS method indeed was able to provide reliable estimates of $S_G$ and $S_I$ with reasonable precisions, suggesting its usefulness in larger population studies. Here, we examined the ITS in a much larger population of 235 subjects and found that individual estimates of $S_G$ collapsed to the population mean for both RSS(180) and RSS(240), rendering it essentially useless in studies that aim to discern small differences in glucose effectiveness across a large population of subjects. In a simulation study, Erichsen et al. (10) also observed a bias toward the mean for $S_G$ using ITS, although with only 40 subjects, not to the extent observed in the present study. It is clear from Figs. 4 and 5 that NM also causes shrinkage around the mean for $S_G$ estimates, although not to the same extent as ITS.
Regression to the mean in population approaches is not a new concept and so comes as no surprise. After all, population approaches by their design use population information to refine individual estimates. What is noteworthy, however, is the observation that the estimation of $S_G$ is affected by this feature much more than $S_I$. We explain this phenomenon as the combination of two characteristics inherent to $S_G$ estimation: poor precision (high %CV) and low between-subject variability. From the STS analysis with a FSS, we found the median %CV of $S_G$ to be larger than that of $S_I$ (19 vs. 10%), in keeping with previous findings (17). As expected, in the case of RSS(180), precisions for $S_G$ and $S_I$ weakened to median values of 27 and 18%, respectively. Previous studies have also suggested the negative effect of a reduced sampling schedule is felt more by $S_G$ than $S_I$. In addition to the poor precision of $S_G$ estimates relative to those of $S_I$, between-subject variability of $S_G$ was roughly one-half that of $S_I$ (Table 1) using the FSS. When RSS(180) was applied, this ratio fell to about one-third. It is likely that the presence of both of these factors contributed to the greater degree of regression to the mean for $S_G$ estimates as opposed to $S_I$ estimates. This assertion is further supported by the results of the two minimal model parameters that were not reported in the text, $V$ and $p_2$. The parameter $V$ showed the least variability in the population but was also very well estimated (low %CV). Conversely, $p_2$ estimates varied the most in the population and were estimated with the least precision. Neither one of these parameters, however, showed nearly the same degree of regression to the mean as $S_G$.

Another important consideration when using population methods is the assumption of a normal or log-normal distribution of the parameters in the population, which may introduce bias if it misrepresents the “true” underlying population distribution. In the case of reduced sampling schedules, especially the RSS(180), this is less of an issue because the variance of the individual parameter estimates is much larger than the bias that may be introduced by a misspecified distribution.

Because an important feature of the NM approach in a reduced sampling schedule setting is its ability to recover parameter estimates when two-stage methods cannot, it is important that these estimates be accurate. It is possible that NM will place these estimates at the population mean, because individual information is imprecise. We therefore compared NM estimates of $S_I$ in the 38 subjects in whom model identification failed for the RSS(180) STS method to the reference estimates obtained from individual nonlinear regression with the FSS. Although agreement was not as strong in this subset as in the entire population ($p_c = 0.56$ vs. $p_c = 0.88$), the estimates were not significantly different (paired $t$-test). Average relative error was $4 \pm 18\%$ and was not significantly different from zero.

Another important consideration is the type of subject for which parameter estimation is more likely to fail. With a full sampling schedule and the traditional STS approach, 4 of 78 subjects with IFG failed to provide parameter estimates, whereas for those with normal fasting glucose model failure occurred in only 1 of 157 subjects. Although with RSS(240) a similar rate of model failure occurred in IFG and normal subjects ($\sim 6\%$), we found a drastic difference with the RSS(180) where individuals with IFG were more than three times more likely to produce model estimation failures than normal subjects (31 vs. 9%). Thus, in populations that may include subjects with even a limited extent of impaired glucose metabolism, we conclude that the traditional approach to parameter estimation with a 180-min reduced sampling schedule is not an appealing option.

Others have used simulation to assess population approaches in a different setting (10). One advantage to simulation is that the true values of population mean and variance can be known and set in advance. Although the true value of population variance in experimental studies, such as ours, is not known a priori, it is clear that when using the STS with RSS(180) population variance for $S_I$ estimation is inflated compared with the FSS. For the case of individual estimates, the parameter estimates we obtain from individual nonlinear regression (first stage of the two-stage analysis) are in a sense the true estimates, because they are the ones used in clinical studies. Therefore, we think that our analysis of experimental data offers a complementary, practical approach to simulation studies.

In conclusion, we show that population approaches offer advantages over the traditional method in the estimation of $S_I$ and $S_G$ from large populations with reduced sampling schedules of the IVGTT. Specifically, the nonlinear mixed-effects approach was able to provide parameter estimates in all individuals, whereas the traditional two-stage methods failed in 6% of the population with a 240-min reduced sampling schedule and 16% with a 180-min schedule. Thus a population approach like NM is more important for the shorter, 180-min sampling schedule, because individual and population estimates of $S_I$ obtained with the NM method showed the best agreement with the reference estimates obtained from the FSS. Furthermore, NM provided individual estimates in subjects with IFG, whereas the traditional approach failed in almost a third of these subjects. With a 240-min schedule, the choice of population approach was not as critical in our study cohort. One major caveat with using population approaches with an RSS is regression to the mean of parameter estimates, particularly $S_G$ and especially with the ITS method. The RSS itself also appeared to induce a consistent negative bias in $S_I$ and $S_G$ estimation, regardless of the parameter estimation technique used. Further research is warranted to develop a truly “optimal” reduced sampling schedule to ensure the accuracy of parameter estimates. Nevertheless, the implementation of population approaches in conjunction with reduced sampling schedules represents a sensible approach to estimating metabolic indexes in large populations in a more cost-efficient manner.

ACKNOWLEDGMENTS

We thank Robert H. Knopp and Barbara M. Retzlaff for making the data available to us. We also thank the staff involved in the study who were responsible for generating the data and the subjects who participated.

GRANTS

We acknowledge the National Institutes of Health (NIH) Clinical Nutrition Research Unit at the University of Washington, DK-35816, the Diabetes Endocrinology Research Center, DK-17047, the American Egg Board, the Resource Facility for Population Kinetics, P41 EB-001975, and the Medical Research Service of the Department of Veterans Affairs for support of this study. K. M. Krudys was supported by NIH Grant T32 DE-07023. S. E. Kahn is the recipient of an American Diabetes Association Distinguished Clinical Scholars Award.
REFERENCES


