The disposition index: from individual to population approach

Paolo Denti, Gianna Maria Toffolo, and Claudio Cobelli

Department of Information Engineering, University of Padua, Padua, Italy

Denti P, Toffolo GM, Cobelli C. The disposition index: from individual to population approach. Am J Physiol Endocrinol Metab 303: E576–E586, 2012. First published June 5, 2012; doi:10.1152/ajpendo.00139.2011.—To correctly evaluate the glucose control system, it is crucial to account for both insulin sensitivity and secretion. The disposition index (DI) is the most widely accepted method to do so. The original paradigm (hyperbolic law) consists of the multiplicative product of indices related to insulin sensitivity and secretion, but more recently, an alternative formula has been proposed with the exponent \(\alpha\) (power function law). Traditionally, curve-fitting approaches have been used to evaluate the DI in a population: the algorithmic implementations often introduce some critical issues, such as the assumption that one of the two indices is error free or the effects of the log transformation on the measurement errors. In this work, we review the commonly used approaches and show that they provide biased estimates. Then we propose a novel nonlinear total least square (NLTLS) approach, which does not need to use the approximations built in the previously proposed alternatives, and show its superiority. All of the traditional fit procedures, including NLTLS, account only for uncertainty affecting insulin sensitivity and secretion indices when they are estimated from noisy data. Thus, they fail when part of the observed variability is due to inherent differences in DI values between individuals. To handle this inevitable source of variability, we propose a nonlinear mixed-effects approach that describes the DI using population hyperparameters such as the population typical values and covariance matrix. On simulated data, this novel technique is much more reliable than the curve-fitting approaches, and it proves robust even when no or small population variability is present in the DI values. Applying this new approach to the analysis of real IVGTT data suggests a value of \(\alpha\) significantly smaller than 1, supporting the importance of testing the power function law as an alternative to the simpler hyperbolic law.

It is widely recognized that an accurate and exhaustive evaluation of glucose regulation in a subject must take into account both the ability of insulin to lower glucose level (insulin sensitivity) and the effect of glucose in promoting insulin secretion (\(\beta\)-cell responsivity). These two processes must be interpreted jointly, because it is only the balance of the two that can give an informative picture of the actual status of the system.

The disposition index (DI) is, to this day, one of the most widely used approaches to do so. Although other authors (25) had previously detected a similar relationship, the paternity of the original formulation is attributed to Bergman et al. (9). This paradigm assumes that subjects with similar efficiency of the glucose disposal metabolic system may have different values of insulin sensitivity [SI; e.g., provided by the intravenous glucose tolerance test (IVGTT) glucose minimal model (8)] and \(\beta\)-cell responsivity \(\Phi\), [e.g., provided by the C-peptide minimal model (24)] but share a similar value of their product, DI is thus a mathematical tool to evaluate the efficiency of glucose homeostasis in a given individual (\(\beta\)-cell compensation), which is the ability of the \(\beta\)-cells to upregulate insulin secretion in response to a decrease in insulin sensitivity. Due to the implied inverse proportionality between SI and \(\Phi\), this is widely known as the hyperbolic law. More recently, a more comprehensive paradigm has been tested by Kahn et al. (15) with an exponential parameter \(\alpha\):

\[
\text{DI} = \text{SI}^\alpha \cdot \Phi
\]  

This parameter introduces an additional degree of freedom that is able to accommodate for a shape of the relationship between secretion and sensitivity indices (in the range of observed values) that is different from the hyperbolic law. An estimated value of \(\alpha\) statistically different from one would allow rejecting the null hypothesis that the relationship is hyperbolic. Hereinafter, we will refer to the alternative DI paradigm of Eq. 2 as power function law.

In the original study by Bergman et al. (9), the relationship between minimal model-derived insulin sensitivity and second-phase posthepatic insulin secretion was considered, but ever since then, many different indices (especially for insulin secretion) have been proposed for the calculation of the DI. A large number of studies in the literature (7, 13, 15, 16c, 17, 19, 26–28) employed the DI and addressed the problem of its determination. They also inspected, for different combinations of sensitivity and secretion indices, whether the parameter \(\alpha\) is significantly different from 1, suggesting that the relationship may be more complicated than a rectangular hyperbola. Traditionally, the most widely used approach consists in considering insulin sensitivity and \(\beta\)-cell responsivity as coordinates of a Cartesian plane and then employing a traditional fit approach to find the curve (hyperbola or power function) best explaining the data. However, to avoid the difficulties of a nonlinear two-variable fit, simplifications such as the log transformation of the data or assumptions on the precision of the individual indices have been used in the literature. The first contribution of this work is a novel fit approach based on a nonlinear total least squares (NLTLS) technique that does not employ similar approximations. With the use of simulated data, this new algorithm is tested and compared with the other ones used previously in the literature, and, as shown in results, it provides much more accurate results.

Similarly to all the other curve-fitting approaches, NLTLS implicitly relies on the hypothesis that all the subjects in the population have the same value of DI. Thus, the only variability that is assumed to be present in the data is the estimation uncertainty for the indices of insulin sensitivity and \(\beta\)-cell responsivity in each subject. This is arguably an over-simplification; even if a group of subjects is classified as healthy, it is very unlikely that they will all share exactly the same value of DI. As shown in this work, this assumption is extremely important, and its consequences should not be overlooked.

Address for reprint requests and other correspondence: C. Cobelli, Univ. of Padua, Via Gradenigo 6/B, Padua, Italy 35131 (e-mail: cobelli@dei.unipd.it).
Therefore, the second contribution is a population approach based on a nonlinear mixed-effects model (NLMEM) that accounts separately for between-subject variability (BSV) in the DI values and the residual unexplained variability (RUV) in the data. The BSV quantifies the physiologically driven variability between the subjects in the study, whereas RUV (a term commonly used in the NLMEM jargon) refers to other sources of variability that cannot be ascribed to BSV, i.e., model error and measurement error, the latter being, in this case, the uncertainty affecting the estimates of individual indices of sensitivity and secretion. In the new framework we propose, the population typical values of DI and α are considered as features that characterize the population distribution of the indices of insulin sensitivity and β-cell responsivity. The NLMEM analysis can be carried out with modeling tools such as NONMEM (see Ref. 4 for the NONMEM Users Guide) or SPK (23), which we employed here, but implementations in R (18), Monolix (16b), or WinBUGS (16) are also available.

All of the methods explained in the present work, including the two novel approaches (NLTLs and NLMEM), are compared on simulated data sets, which were generated according to two DI situations, the former with all subjects sharing the same DI value and the latter with variability between subjects in the DI values, but reproducing in both cases the variability observed in real data (2, 3). Finally, results on real data are also presented.

METHODS

Real and Simulated Data Sets

Real data. The real data set derives from a previous study (2, 3) and consists of 204 healthy subjects who underwent an insulin-modified IVGTT with a full sampling schedule (240 min, 21 samples). The individuals belong to two large groups characterized by different ages: 59 young (age 23 ± 3) and 145 elderly subjects (age 69 ± 6). For each of them, identification by the modeling software SAAM II (1) of the IVGTT glucose (8) and C-peptide (24) minimal model provided individual estimates of both the value and precision of insulin sensitivity index (SI) and first- and second-phase secretion indices (Φi and Φς), which were subsequently combined into the total index Φo.

Both indices were approximately log-normally distributed across the patients in the population. Their precision was evaluated by the parameter estimation algorithm implemented in SAAM II by propagating the measurement errors of glucose and C-peptide concentrations. Results on our 204 subjects indicated that precision of SI and Φ estimates are well described by a constant SD and constant coefficient of variation (CV) model, respectively, and the variability of their precision across the population, in both cases, followed a log-normal distribution. The SDs for SI had an average value equivalent to 5% of its population typical value, whereas the CVs for Φ were on average ~8%. The structure and size of the variability detected in the real data were the basis for the generation of the simulated data explained below.

Simulated data. Since the true value of DI in our population is unknown, real data allow only a comparison between the results of the different methods. To assess their real performance, simulated SI-Φ data were generated in keeping with the statistical properties of the real data and satisfying predefined DI laws, namely the hyperbolic (Eq. 1) and power function laws (Eq. 2).

The simulation comprised two steps; first, the “true” values of SI and Φ were generated for each subject, and second, these “true” values were perturbed to mimic noisy experimental data. As regards the first step, data sets were generated either with no variability in DI values, i.e., true values of the indices satisfy either Eq. 1 or 2, or with variability between subjects in the DI values. The general model for the true values of SI, and Φi is a joint population distribution (see APPENDIX for the rationale):

$$\log(\text{SI}_i) \sim N \left( \theta_\text{SI}, \omega^2 \right) \quad \log(\Phi_i) \sim N \left( \theta_\Phi, \omega^2 \right)$$

where θ denotes the mean of the logarithms of the indices SI and Φ and Ω denotes their covariance matrix, with ρ being the coefficient of correlation between the two indices, α the exponent in the DI formula, and ω² a scaling parameter for the variance. The parameter α changes the shape of the DI curve, whereas ω² regulates the size of SI and Φ between-subject variability.

As regards the second step, measurement errors to be added to the true values were generated by first simulating the precisions (constant SD for SI and constant CV for Φ) from a log-normal distribution and then using the selected precision as variability of a Gaussian distribution to generate the actual error (truncated to prevent negative values).

No population variability in the DI values. A first data set was generated with no variability in the DI values. This was done by fixing the parameter ρ to −1, since, as shown in the APPENDIX, a perfect anticorrelation between the two indices (ρ = −1) forces each individual deviation of SI from the mean value to be compensated entirely by an opposite deviation in DI so that all subjects share the same DI value. The value of ω was fixed to −0.5 to reproduce the variability of ~50% for both indices observed in the real data. The average values θSI and θΦ, were tuned to obtain 100 as a symbolic expected value for DI. Three different values of the parameter α were used (0.5, 1, and 2), and for each of them, 100 sets consisting of 1,000 subjects were generated. For each set, the true DI values all superimpose exactly over the same DI curve. When perturbed according to the precision estimated in the real data, their level of scattering was visibly much smaller than in the real data set. A possible explanation could be that the estimated precision of SI and Φ reflects only measurement errors and not other sources of variability, e.g., model error and interday variability. Thus, even if the measurement noise was generated according to the above-mentioned constant CV and SD models, the level of uncertainty used in the simulation was increased by about five times (40% CV for Φ and SD equivalent to 25% of the typical value for SI) to produce simulated data in good agreement with the real ones, which were indeed extremely scattered.

Population variability in the DI values. In a second data set, the assumption that all simulated subjects share the DI value was removed, and some BSV for DI was introduced. SI and Φ were still sampled from the joint distribution in Eq. 3 but with progressively increasing levels of variability in the DI values, which were obtained by changing the value of parameter ρ from its initial value, −1 (−0.8, −0.5, and −0.2 were used). At the same time, the RUV was decreased, maintaining the overall variability at a similar level (with respect to the precision values estimated in the original data set for the secretion and sensitivity indices, the levels of uncertainty were multiplied by 4, 1, and 1/5, respectively). In this way, the focus of the analysis was on the consequences of the different hierarchy underlying the variability rather than its size. The BSV for SI and Φ (dependent only on the values of parameters ω and α) remained the same in these simulations, but due to the alteration of the value of the correlation ρ, the BSV in the DI changed; e.g., for the case α = 1 it was equal to 32 (ρ = −0.8), 30 (ρ = −0.5), and 63% (ρ = −0.2). Since BSV in the DI depends both on ρ and α, as shown in the APPENDIX, the values are lower for larger α-values and vice versa.

Parameter Estimation Methods

Linear fit on log-transformed data. Curve-fitting approaches are based on the interpretation of the individual values of SI and Φ, as coordinates x and y of a Cartesian system. They assume that the
true (error-free) data lie on a curve, on a hyperbola, as in Eq. 1, or on a power function, as in Eq. 2, depending on the value of \( \alpha \). Finding the values of DI and \( \alpha \) characterizing the population consists in identifying the curve that better fits the data, i.e., the curve lying at the minimum distance from the data points. The various methods differ on how they define this distance; many alternatives have been proposed in the literature, mostly employing approximations. Some approaches operate a log transformation to turn the problem into the fit of a straight line. Here, we will assess the most popular approach, which fits a straight line on log-transformed data with a one-variable approach; i.e., it assumes that only the \( y \) variable is affected by noise but uses a coefficient to adjust for the error in the second variable. The standard reference for this method is from Riggs et al. (20), but the first mention we could find is in a publication by Gini (14); thus it will be denoted as Log-Gini in this work. We will also test a linear total least squares approach applied after log transformation of the data (Log-TLS). Details on both methods are included in the APPENDIX.

**NLTLS.** The first novel method proposed here, named NLTLS, extends to the nonlinear case the minimization of the sum of squares of the weighted residuals along both directions, thus considering the presence of errors in both variables. To this purpose, for each data point \((S_i, \Phi_i)\), its projection on the curve, denoted as \((S_i^*, \Phi_i^*)\), is defined as the point lying on the curve and having the shortest distance to the data point, where the distance is calculated by weighting the components along \( x \) and \( y \) with the respective uncertainties (i.e., the estimated precisions \( \sigma_{S_i}^2 \) and \( \sigma_{\Phi_i}^2 \)). Residuals in the TLS fit are equal to these distances; thus the objective function, which is minimized with respect to DI and \( \alpha \), can be written as follows:

\[
F_{\text{NLTLS}} = \sum_i \frac{(S_i - S_i^*(\alpha, \text{DI}))^2}{\sigma_{S_i}^2} + \sum_i \frac{(\Phi_i - \Phi_i^*(\alpha, \text{DI}))^2}{\sigma_{\Phi_i}^2} = \sum_i \frac{\text{res}_{S_i}^2}{\sigma_{S_i}^2} + \sum_i \frac{\text{res}_{\Phi_i}^2}{\sigma_{\Phi_i}^2}.
\] (4)

As can be seen in Eq. 4, uncertainties along the \( x \)- and \( y \)-direction are properly taken into account by weighting the residuals along \( x \) and \( y \), \( \text{res}_x \) and \( \text{res}_y \), with the respective variances derived from the estimation uncertainties of the sensitivity and secretion indices. Efficient algorithms are available for the TLS fit of a straight line, but in the case of a power function we were not able to identify an analytical solution to the calculation of the projection \((S_i^*, \Phi_i^*)\). Thus a numerical minimization procedure was applied, which yields the result at the price of a longer computation time. Within the minimization procedure for the objective function in Eq. 4, a further level of minimization was required to compute the projections \((S_i^*, \Phi_i^*)\) of each data point \((S_i, \Phi_i)\) on the curve. Matlab (16a) was used for this implementation, and code is available upon request.

**NLMEM approach.** All the methods proposed previously, including NLTLS, are based on a curve-fitting approach, and therefore, they
rely implicitly on the existence of a single curve (and thus DI value) that characterizes all subjects in the population. The variability in the data is assumed to be due only to the uncertainty with which the sensitivity and secretion indices are measured. This assumption makes these approaches inconsistent in the case of BSV in the DI values. Our newly proposed approach instead accounts for the presence of BSV in the observed DI values by postulating a joint probabilistic description of SI, and \( \Phi_i \) similar to Eq. 3 and estimates the unknown parameters \( \theta \) and \( \Omega \) by using a population approach, such as NLMEM (5, 10). A model was implemented in NONMEM VI 2.0 and used to estimate the population parameters \( \theta_{SI}, \theta_{\Phi}, \alpha, \) and \( \beta \) of Eq. 3 and consequently combine them as detailed in the APPENDIX to obtain the typical value of the DI. In the estimation process, the estimates of individual indices SI, and \( \Phi_i \) were used as data together with their precision. The NONMEM code for the implementation of this model is available upon request.

Comparison. All methods were assessed on simulated data sets; for each setting of \( \alpha \) and \( \rho \) used in the simulation, the estimates of the 100 repetitions were collected and displayed in boxplots, which show the median value and the size of the quartiles. The median can be used to assess the accuracy of the estimation algorithm, whereas the width of the boxes and whiskers is an indicator of its precision. For the NLTLS approach, we also collected the value of \( \hat{\sigma}^2 \) (the variance a posteriori of the weighted residuals), which was calculated with the following formula:

\[
\hat{\sigma}^2 = \frac{\sum_i \text{res}_{yi}^2 + \sum_i \text{res}_{xi}^2}{2n - p} \tag{5}
\]

where \( n \) is the number of data points (and thus \( 2n \) is the number of residuals), and \( p \) is the number of parameters in the model, in this case 2, DI, and \( \alpha \).

RESULTS

No Population Variability in the DI

Results obtained on the data sets with no BSV in the DI values are summarized in Fig. 1. The estimates provided by the methods using a log transformation are affected by some bias, in particular the DI estimate. Log-Gini tends to underestimate DI (about \(-40\% \) or worse), whereas Log-TLS overestimates it (about \(+40\% \) or more). The precision and accuracy with which \( \alpha \) is estimated are more satisfactory, and Log-TLS performs slightly better (\(-15\% \) for Log-Gini, about \(+10\% \) for Log-TLS). The newly proposed NLTLS and NLMEM methods yield the best performance in terms of both precision and accuracy and for both parameters. The size of the bias for DI is significantly lower compared with the other approaches, at most \(-20\% \) as opposed to at least \(-40\% \). A
similar trend is also evident for the precision of the estimates. In fact, NLTLS and NLMEM results are generally less scattered than the other algorithms. The improvement in the estimate of $\alpha$ is also significant; the bias is at most $\sim 5\%$ instead of $\sim 10\%$ for Log-TLS and $\sim 20\%$ for Log-Gini. The parameter $\alpha$ used in the simulation (0.5, 1, and 2) does not seem to affect the outcome of the comparison.

Population Variability in the DI

Results on data characterized by BSV in the DI values are shown in Figs. 2–4. Figure 2 reports the results for a situation ($\rho = -0.8$) in which the variability in the data is still due mainly to the uncertainty in the estimation of the individual indices rather than the effect of BSV, which is limited. NLMEM provides slightly more precise estimates, but NLTLS and Log-TLS also yield satisfactory results. Interestingly, the performance of the latter two methods seems affected by the value of $\alpha$, whereas this phenomenon is not observed for NLMEM. As the level of BSV in the DI increases, as depicted in Figs. 3 and 4 ($\rho = -0.5$ and $\rho = -0.2$, respectively), the estimates of $\alpha$ and DI provided by the curve-fitting approaches become less and less reliable. This is expected, since when a large fraction of the variability is caused by differences between subjects (BSV) rather than estimation uncertainty of the individual parameters (RUV), the hypothesis underlying the curve-fitting approaches becomes a gross simplification, resulting in a large bias in the estimates of both $\alpha$ and DI.

It is very interesting to notice that the increasing contribution of BSV seems to affect mostly the accuracy of the estimates rather than their precision. The method that seems to be more affected by the introduction of BSV is Log-Gini, whose bias in the estimate of both DI and $\alpha$ is significantly larger than the other methods.

It is also interesting to consider the estimates of $\hat{\sigma}^2$ (the variance a posteriori of the weighted residuals) obtained with the different data sets and reported in Table 1. The value of $\hat{\sigma}^2$ can be used as a marker for goodness of fit. If the a priori knowledge of the level of error in the data is reasonable and the model is correct, a value of $\sim 1$ is expected for $\hat{\sigma}^2$. As can be seen in Table 1, the values are below or $\sim 1$ for the data sets with no or relatively small population variability in the DI ($\rho = -1$ or $-0.8$). Instead, much larger values are obtained when the BSV of the DI is larger and the RUV is smaller ($\rho = -0.5$ or $-0.2$).

The Real Data Set

Finally, both NLTLS and NLMEM were applied to the real data sets. Since NLTLS is not designed to deal with large variability in the population and the fit can become very sensitive to outlier subjects, a jackknife technique (12) was employed to detect and remove such values. On average, the

![Fig. 3. Box plots displaying the parameter estimates (DI in each of the top boxes and $\alpha$ in the bottom boxes) obtained on the simulated data set, including a significant amount of BSV in the DI values, i.e., $\rho = -0.5$. Top to bottom: the 3 panels refer respectively to $\alpha = 0.5$, $\alpha = 1$, and $\alpha = 2$. The dashed vertical line indicates the true parameter value used in the simulation.](http://ajpendo.physiology.org/)
data of one or two subjects were removed from each data set. The analysis was first executed on the young and elderly populations separately and finally on the whole group. In this last case, BSV in the DI is present by hypothesis (since young and elderly patients are expected to have different DI values), so the NLTLS approach was used only as a mean of comparison, whereas the results are somehow expected to be unreliable. Table 2 reports the parameter estimates and their 90% confidence interval provided by a 3,000-sample bootstrap employing the NLTLS and NLMEM algorithms for DI and for SI coupled with $\Phi_1$ and $\Phi_2$. Not surprisingly, some of the results differ significantly between the two methods for both DI and $\Phi_1$. However, the validity of the NLTLS estimates is questionable due to the values of $\hat{\sigma}^2$ detected by NLTLS, which are always much larger than 1, suggesting that the amount of unexplained variability in the data is much greater than indicated by the precision of the individual estimates. This might suggest that either the estimation confidence intervals for the individual indices SI and $\Phi$ are heavily underestimated or the model is inadequate to describe the data, or there is another source of variability that is not properly accounted for (e.g., BSV). This is also evident in Figs. 5, 6, and 7, which provide a graphical representation of the data and the NLMEM fits for the separate groups of young and elderly subjects. The size of the error bars (representing the RUV) does not seem large enough to explain the distance of the individual values from the DI curve. A further confirmation of this comes from the results provided by NLMEM, which detects values of correlation that are consistently different from $\rho = -0.5$, indicating a remarkable amount of scattering of the DI values, particularly when the group of elderly subjects or the whole population with the first-phase index $\Phi_1$ as $\beta$-cell responsivity index is analyzed.

Table 1. The average $\hat{\sigma}^2$ obtained across the 100 runs for the NLTLS method is reported for simulated data sets with increasing levels of population variability.

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$\hat{\sigma}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-1$</td>
<td>0.451</td>
</tr>
<tr>
<td>$-0.8$</td>
<td>0.762</td>
</tr>
<tr>
<td>$-0.5$</td>
<td>11.5</td>
</tr>
<tr>
<td>$-0.2$</td>
<td>401</td>
</tr>
</tbody>
</table>

NLTLS, nonlinear total least squares. Only the results for $\alpha = 1$ are reported.
Table 2. Results obtained on the real data set with NLTLS and NLMEM approach

<table>
<thead>
<tr>
<th></th>
<th>NLTLS</th>
<th></th>
<th></th>
<th>NLMEM</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DI</td>
<td>(\alpha)</td>
<td>(\hat{\sigma}^2)</td>
<td>DI</td>
<td>(\alpha)</td>
<td>BSV (DI)</td>
</tr>
<tr>
<td>SI-(\Phi_n)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>39.7 (27.1, 56.8)</td>
<td>0.51 (0.34, 0.66)</td>
<td>61 (4.2, 8.1)</td>
<td>85.1 (55.6, 146.7)</td>
<td>0.79 (0.62, 1.00)</td>
<td>141% (95, 213%)</td>
</tr>
<tr>
<td>E</td>
<td>31.9 (26.2, 39.6)</td>
<td>0.36 (0.45, 0.68)</td>
<td>8.7 (7.2, 10.6)</td>
<td>33.3 (29.1, 38.7)</td>
<td>0.56 (0.49, 0.64)</td>
<td>92% (78, 110%)</td>
</tr>
<tr>
<td>All</td>
<td>30.6 (25.5, 36.7)</td>
<td>0.49 (0.40, 0.59)</td>
<td>8.9 (7.5, 10.4)</td>
<td>38.6 (33.4, 44.7)</td>
<td>0.57 (0.51, 0.64)</td>
<td>99% (86, 116%)</td>
</tr>
<tr>
<td>SI-(\Phi_c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>937.3 (403.9, 4174.9)</td>
<td>0.85 (0.47, 1.51)</td>
<td>31.0 (23.1, 40.5)</td>
<td>1,133.1 (746.3, 1932.2)</td>
<td>0.81 (0.70, 0.94)</td>
<td>181% (127, 266%)</td>
</tr>
<tr>
<td>E</td>
<td>331.7 (203.6, 618.2)</td>
<td>0.68 (0.43, 1.02)</td>
<td>24.5 (19.5, 30.3)</td>
<td>410.4 (335.2, 507.9)</td>
<td>0.81 (0.71, 0.94)</td>
<td>169% (137, 207%)</td>
</tr>
<tr>
<td>All</td>
<td>361.6 (209.3, 665.6)</td>
<td>0.63 (0.37, 0.92)</td>
<td>34.0 (28.4, 40.0)</td>
<td>519.8 (427.7, 646.9)</td>
<td>0.81 (0.71, 0.94)</td>
<td>174% (147, 207%)</td>
</tr>
<tr>
<td>SI-(\Phi_{\text{res}})</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>44.6 (31.4, 71.0)</td>
<td>0.82 (0.66, 1.04)</td>
<td>149% (102, 224%)</td>
<td>46.9 (4.6, 10.2)</td>
<td>0.63 (0.56, 0.70)</td>
<td>101% (87, 118%)</td>
</tr>
<tr>
<td>E</td>
<td>26.4 (21.0, 34.1)</td>
<td>0.66 (0.54, 0.80)</td>
<td>9.2 (7.7, 11.7)</td>
<td>23.6 (20.4, 27.6)</td>
<td>0.62 (0.54, 0.71)</td>
<td>99% (84, 120%)</td>
</tr>
<tr>
<td>All</td>
<td>24.6 (20.3, 29.8)</td>
<td>0.60 (0.50, 0.70)</td>
<td>10.1 (8.5, 12.1)</td>
<td>25.3 (22.1, 29.3)</td>
<td>0.63 (0.56, 0.70)</td>
<td>101% (87, 118%)</td>
</tr>
</tbody>
</table>

NLMEM, nonlinear mixed-effects model; DI, disposition index; BSV, between-subject variability; \(\Delta OFV\), difference in objective function value; SI, insulin sensitivity; Y, young; E, elderly. The fit was performed both separately on the Y and E subjects and then on the whole population (all). The columns contain the parameter estimates, and the 90% confidence intervals, as provided by 3,000 bootstrap samples, are in parentheses. In addition to DI and \(\alpha\), for NLTLS the value of \(\hat{\sigma}^2\) is also reported. For NLMEM, the corresponding values of BSV in the DI and the correlation \(\rho\) are shown, together with the drop in NONMEM objective function (\(\Delta OFV\)) and the corresponding \(P\) value resulting from the estimation of \(\alpha\), rather than fixing it to 1.

**DISCUSSION**

In this study, the performance of conventional and novel approaches to evaluate the DI law in a population has been assessed on both real and simulated data sets. When no population variability in the DI values is used for the simulations, the novel approaches NLTLS and NLMEM outperform the conventional approaches with nonparametric testing. The likely reason is that although the use of a log transformation also affects the error structure; e.g., a log scale tends to equalize the distance between the data points and the curve. NLMEM also assesses the significance of an additional parameter in a model, which is approximately equivalent to calculating the arithmetic mean after log-transformation of positive numbers. The calculation of the geometric mean is far from the average curve.

In a NLMEM framework, another statistic that is used to assess the drop in objective function value, which is approximately X-square distributed and can therefore be used to obtain a \(P\) value, is the estimate of \(\alpha\). This is based on the fact that a change in the log-transformed plane and thus is not prone to this latter effect, tends to underestimate DI. Log-TLS, on the other hand, assumes the same degree of precision for all of the data points in the log-transformed plane and thus is not prone to this latter effect. Depending on the type of data, the choice of method can substantially change the leverage that each point has in the fit procedure. Moreover, as already mentioned, a constant CV structure also affects the distance between the points and the curve. NLTLS and NLMEM outperform the linear model, which is approximately equivalent to calculating the geometric mean after log-transformation of positive numbers. The calculation of the geometric mean is far from the average curve.

The likely reason is that although the use of a log transformation also affects the error structure; e.g., a log scale tends to equalize the distance between the data points and the curve. NLMEM also assesses the significance of an additional parameter in a model, which is approximately equivalent to calculating the arithmetic mean after log-transformation of positive numbers. The calculation of the geometric mean is far from the average curve.

In a NLMEM framework, another statistic that is used to assess the drop in objective function value, which is approximately X-square distributed and can therefore be used to obtain a \(P\) value, is the estimate of \(\alpha\). This is based on the fact that a change in the log-transformed plane and thus is not prone to this latter effect, tends to underestimate DI. Log-TLS, on the other hand, assumes the same degree of precision for all of the data points in the log-transformed plane and thus is not prone to this latter effect. Depending on the type of data, the choice of method can substantially change the leverage that each point has in the fit procedure. Moreover, as already mentioned, a constant CV structure also affects the distance between the points and the curve. NLTLS and NLMEM outperform the linear model, which is approximately equivalent to calculating the arithmetic mean after log-transformation of positive numbers. The calculation of the geometric mean is far from the average curve.

The likely reason is that although the use of a log transformation also affects the error structure; e.g., a log scale tends to equalize the distance between the data points and the curve. NLMEM also assesses the significance of an additional parameter in a model, which is approximately equivalent to calculating the arithmetic mean after log-transformation of positive numbers. The calculation of the geometric mean is far from the average curve.

In a NLMEM framework, another statistic that is used to assess the drop in objective function value, which is approximately X-square distributed and can therefore be used to obtain a \(P\) value, is the estimate of \(\alpha\). This is based on the fact that a change in the log-transformed plane and thus is not prone to this latter effect, tends to underestimate DI. Log-TLS, on the other hand, assumes the same degree of precision for all of the data points in the log-transformed plane and thus is not prone to this latter effect. Depending on the type of data, the choice of method can substantially change the leverage that each point has in the fit procedure. Moreover, as already mentioned, a constant CV structure also affects the distance between the points and the curve. NLTLS and NLMEM outperform the linear model, which is approximately equivalent to calculating the arithmetic mean after log-transformation of positive numbers. The calculation of the geometric mean is far from the average curve.

In a NLMEM framework, another statistic that is used to assess the drop in objective function value, which is approximately X-square distributed and can therefore be used to obtain a \(P\) value, is the estimate of \(\alpha\). This is based on the fact that a change in the log-transformed plane and thus is not prone to this latter effect, tends to underestimate DI. Log-TLS, on the other hand, assumes the same degree of precision for all of the data points in the log-transformed plane and thus is not prone to this latter effect. Depending on the type of data, the choice of method can substantially change the leverage that each point has in the fit procedure. Moreover, as already mentioned, a constant CV structure also affects the distance between the points and the curve. NLTLS and NLMEM outperform the linear model, which is approximately equivalent to calculating the arithmetic mean after log-transformation of positive numbers. The calculation of the geometric mean is far from the average curve.
performs well, even if it represents an overparameterization of the problem, in the no-population variability case.

The situation is different if population variability is present in the DI values; as already mentioned, the presence of other sources of variability in the data besides the estimation uncertainty of the secretion and sensitivity indices undermines the validity of all of the methods considered here, except NLMEM. And indeed, the performance of all of the curve-fitting approaches degrades significantly when population variability in the DI values comes into play, even if the overall amount of variability in the data is not altered. The underlying reason is that, as the magnitude of population differences (BSV) surpasses the effect of individual uncertainty (RUV), the estimates of precision for SI and \( \Phi \), even if correct, become less and less informative about the actual distance of the data points from the curve characterizing the typical DI of the population. In fact, the precision on the individual parameter estimates (\( \sigma^2_{SI} \) and \( \sigma^2_{\Phi} \)) does not carry any information about how scattered the DI values are in the population. A subject may be characterized by very precise SI and \( \Phi \) estimates, but the resulting DI may be extremely different from the population DI typical value. The curve-fitting approaches are not designed to account for this kind of hierarchical variability, so they fail when different sources of variability come into play, even if no approximation is used, as in the case of NLTLS. A symptom of this is the value of the error \( \hat{\sigma}^2 \); values much larger than 1 are detected when BSV is used to simulate the data. This shows lack of fit and indicates possibly that the a priori level of uncertainty is not large enough to account for all the variability detected a posteriori in the data. As a consequence of poor fit, an extremely large bias is introduced in the parameter estimates, whereas the level of precision is similar. Moreover, the size and direction of the bias becomes strongly dependent on the value of \( \alpha \) used in the simulation of the data set. The different settings of \( \alpha \) introduce asymmetries in the data that lead to systematic over- or underestimation. On the contrary, NLMEM is designed to capture the hierarchical structure of the variability and provides reliable results regardless of the value of \( \alpha \).

Finally, we have some comments on the results obtained using NLTLS and NLMEM on the real data (Table 2).
NLTLS fit approach detected a large value of $\hat{\sigma}^2$, and NLMEM estimates levels of correlation $\rho$ that are different from -1 among the individual indices. Both of these markers are indicators of the presence of considerable variability not explained in terms of precision of the individual estimates, suggesting that NLTLS results should be regarded as unreliable. Therefore, with the purpose of investigating the statistical significance of parameter $\alpha$ and thus the validity of the power function vs. the pure hyperbolic law in our data set, the NLMEM results are used.

The NLMEM results suggest that the typical DI values for the elderly population are lower than for the young one, and the graphical representations in Figs. 5–7 show that the curve for the elderly subjects is closer to the origin than the one for the young ones.

Moreover, although the analysis of other data sets would be needed to confirm the current findings, the estimates of $\alpha$ provided by NLMEM are often lower than 1, supporting the importance of testing the power function law as an alternative to the simple hyperbolic law. A value of $\alpha$ less than 1 produces a curve rotated anticlockwise, with values lower than a rectangular hyperbola for low values of SI, thus suggesting that compensatory mechanisms are less efficient in subjects characterized by low insulin sensitivity.

Conclusions

In this work, we showed the importance of the regression methodology in the analysis of the DI. We pointed out the shortcomings of the currently used methods and proposed the novel NLTLS fit. Although this new approach proved more reliable than its competitors, our results highlighted the inadequacy of all curve-fitting approaches in the case of variability in the DI values across the population. Analysis of real data confirmed that a considerable amount of BSV in the DI values is probably present, even in a population of healthy individuals. Thus we designed and tested a novel population model that was able to efficiently account for BSV. We recommend the use of a population approach in the study of the DI, since neglecting to account for the hierarchical structure of the variability can yield very unreliable results.

In our analysis, a two-step approach was used. First, the individual values of insulin sensitivity and secretion were calculated from the raw blood sampling data, and then the DI analysis on the individual results was performed. A more complete approach would consist in integrating the two steps into a joint population model. Nonlinear mixed-effects modeling has already been successfully applied with the glucose minimal model to estimate SI (11), and similarly it could be used for the estimation for insulin secretion. Thus, one can imagine applying such an approach to jointly estimate the secretion and sensitivity indices directly from glucose and insulin profiles and, at the same time, extracting the information on DI and $\alpha$ from the population covariance matrix.

APPENDIX

Formulation of the DI as a Joint Population Distribution

If SI and $\Phi$ are assumed to be log-normally distributed, their joint population distribution can be written in the following general form:

$$\begin{bmatrix} \log(SI_i) \\ \log(\Phi_i) \end{bmatrix} \sim N\left(\begin{bmatrix} \theta_{SI} \\ \theta_{\Phi} \end{bmatrix}, \begin{bmatrix} \omega_{SI}^2 & \rho \omega_{SI} \omega_{\Phi} \\ \rho \omega_{SI} \omega_{\Phi} & \omega_{\Phi}^2 \end{bmatrix}\right)$$

(7)

where $\theta_{SI}$ and $\theta_{\Phi}$ is their covariance matrix expressed, without loss of generality, in terms of variances ($\omega^2$) and coefficient of correlation ($\rho$). Under the assumptions that both SI and $\Phi$ are log-normally distributed, the distribution of the DI is also log-normal, with mean and variance related to mean values and variances of the sensitivity and secretion indices, as specified in the following equation:

$$\begin{bmatrix} SI_i \Phi_i = DI_i \\ \log(DI_i) \end{bmatrix} \sim N\left(\begin{bmatrix} \theta_{SI} + \theta_{\Phi} \\ \alpha \theta_{SI} + \theta_{\Phi} \end{bmatrix}, \begin{bmatrix} \omega_{SI}^2 + \omega_{\Phi}^2 & \omega_{SI} \omega_{\Phi} + 2\alpha \omega_{SI} \omega_{\Phi} \\ \omega_{SI} \omega_{\Phi} + 2\alpha \omega_{SI} \omega_{\Phi} & \omega_{\Phi}^2 + 2\alpha \omega_{SI} \omega_{\Phi} \end{bmatrix}\right)$$

(8)

From Eq. 8, the conditions under which all subjects share the same DI value, i.e., the variance of DI is equal to zero, are as follows:

$$\omega_{SI}^2 + \omega_{\Phi}^2 + 2\alpha \omega_{SI} \omega_{\Phi} = 0$$

(9)

Apart from the trivial solution $\omega_{SI} = \omega_{\Phi} = 0$ (i.e., sensitivity and secretion indices assume the same value in all subjects), Eq. 8 is satisfied if
that is when \( \log(SI_i) \) and \( \log(\Phi_i) \) are perfectly negatively correlated. Moreover, the ratio between their population variances determines the value of the parameter \( \alpha \).

Therefore, to interpret the SI-\( \Phi \) data with the DI paradigm, we can use the following easy reparameterization

\[
\omega = \frac{\omega_{\Phi}}{\sqrt{\alpha}} = \omega_{SI} \cdot \sqrt{\alpha}
\]

and rewrite, without loss of generality, the joint population distribution as follows

\[
\begin{bmatrix} \log(SI_i) \\ \log(\Phi_i) \end{bmatrix} \sim N \left( \begin{bmatrix} \theta_{SI} \\ \theta_{\Phi} \end{bmatrix}, \begin{bmatrix} \omega^2 & \frac{1}{\alpha} \rho \omega^2 \\ \frac{1}{\alpha} \rho \omega^2 & \frac{1}{\alpha} \omega^2 \end{bmatrix} \right)
\]

The value of the correlation \( \rho \) can then be used to determine the size of BSV in the DI values.

**Log Transformation**

The most widely used curve-fitting approach, introduced by Kahn et al. (15), consists in applying a log transformation to the original data, turning the problem into a linear fit. In fact, the log transformation of both variables, SI and \( \Phi \), reshapes the hyperbola into a straight line, as shown in the following formulae:

\[
DI = SI^\alpha \cdot \Phi \rightarrow \log(DI) = \alpha \cdot \log(SI_i) + \log(\Phi_i)
\]

However, this transformation is nonlinear and does not preserve the Gaussian error structure of the model; in this way, the hypotheses supporting the consistency of the classical fit approach are not satisfied. To maintain at least part of the information on the precision of the individual estimates, a first-order approximation can be used. Loosely speaking, this is equivalent to assuming a Gaussian distribution of the error for the log-transformed estimates, using the values of the CV of the original estimates. However, this simplification is quite poor, especially if the uncertainty is large, since the nonapproximated probability distribution of the log-transformed data is nonsymmetric and skewed to the right. On top of this, one should bear in mind that there is no guarantee that the optimal solution in the log-transformed space preserves its optimality under nonlinear transformations. Therefore, especially if a large variability is present, the two solutions might be very different.

In addition, even if the log transformation simplifies the problem to the fit of a straight line, the error is still present in both variables, and a proper technique must be employed to take this into account. Many approaches are used in literature, but the most widespread is the perpendicular weighted regression presented by Riggs et al. (20), which is not, however, the original source. The oldest reference to the same formula we found is in Gini (14), so we denoted it Log-Gini. This method is essentially an orthogonal fit that uses a correction factor \( \lambda \) (ratio of the variances) to account for the difference of precision along the two directions. Therefore, technically, a fit along just one direction is performed, but the correction factor \( \lambda \) allows one to also consider the second direction. Although a fast linear method is available, and thus no minimization of the objective function is explicitly required, we propose here the formula for the sake of comparison with the other methods:

\[
F_{\text{Log-Gini}} = \sum_i \left( \frac{\lambda \cdot (\log(SI_i) - \log(DI) - \log(\Phi_i^*))^2}{\sigma_{\log(SI_i)}^2} + \frac{\log(\Phi_i) - \log(DI) + \alpha \cdot \log(SI_i^*)^2}{\sigma_{\log(\Phi_i)}^2} \right)
\]

The symbol \( \bot \) denotes the orthogonal projection on the straight line once the axes have been rescaled to accommodate for the heteroscedasticity along the different directions. However, a limitation of Log-Gini is represented by the fact that it assumes the same degree of precision for all the subjects. In this way, all the data points will be considered equally reliable in the fit. Since in our simulations we included individual precisions, a median value was employed to obtain an estimate of \( \lambda \), as shown in the following formula:

\[
\lambda = \text{median} \left( \frac{\sigma_{\log(\Phi_i)}^2}{\sigma_{\log(SI_i)}^2} \right) = \text{median} \left( \frac{CV_{\Phi_i}^2}{CV_{SI_i}^2} \right)
\]

An alternative approach to execute the fit for the straight line after the log transformation has been performed is TLS:

\[
F_{\text{Log-TLS}} = \sum_i \left( \log(SI_i) - \log(\Phi_i^*) \right)^2 + \sum_i \left( \frac{\log(\Phi_i) - \log(DI) + \alpha \cdot \log(SI_i^*)^2}{\sigma_{\log(\Phi_i)}^2} \right)
\]

where \( (\log(SI_i^*), \log(\Phi_i^*)) \) is the prediction in the TLS sense, i.e., the point on the straight line lying at the shortest weighted distance from \( (\log(SI_i), \log(\Phi_i)) \). As mentioned already, in this linear case, the TLS projection can be obtained with a closed-form formula, and therefore, it is computationally very fast. This approach is referred to here as Log-TLS.

**ACKNOWLEDGMENTS**

These results were presented partially at the EASD 2007 Annual Conference in Amsterdam, The Netherlands, and included in the conference proceedings (Denti P, Campioni M, Toffolo GM, Rizza R, Cobelli C. “Is the disposition index law hyperbolic? Importance of the regression methodology.” Abstract 602) and at PAGE 2009 in St. Petersburg, Russia (Denti P, Salinger D, Vicini P, Toffolo GM, Cobelli C. “A NonLinear Mixed-Effects Approach to the Estimation of the Glucose Disposition Index.” PAGE Abstract 1648).

P. Denti thanks Eng. Domenico Dei Tos for performing the analysis on the real data set as part of his Laurea Thesis project, Dr. David Salinger for advice on the NONMEM implementation of the population model, and Dr. Bradley Bell and Paolo Vicini for fruitful discussion.

Present address of P. Denti: Division of Clinical Pharmacology, University of Cape Town, Cape Town, South Africa.

**GRANTS**

P. Denti was partially funded by the scholarship “Borsa Gini”, sponsored by the Gini Foundation, Padua, Italy.

**DISCLOSURES**

No conflicts of interest, financial or otherwise, are declared by the authors.

**AUTHOR CONTRIBUTIONS**

P.D. analyzed the data; P.D., G.M.T., and C.C. interpreted the results of the experiments; P.D. prepared the figures; P.D. drafted the manuscript; P.D., G.M.T., and C.C. edited and revised the manuscript; P.D., G.M.T., and C.C. approved the final version of the manuscript; G.M.T. and C.C. did the conception and design of the research.

**REFERENCES**


