Simple modeling allows prediction of steady-state glucose disposal rate from early data in hyperinsulinemic glucose clamps

Pooja Singal,1 Ranganath Muniyappa,2 Robin Chisholm,1 Gail Hall,2 Hui Chen,3 Michael J. Quon,2 and Kieren J. Mather1

1Indiana University School of Medicine, Indianapolis, Indiana; and 2Diabetes Unit, National Center for Complementary and Alternative Medicine, National Institutes of Health, Bethesda, Maryland

Submitted 30 September 2009; accepted in final form 6 November 2009

Insulin resistance is a key contributor to type 2 diabetes mellitus and its cardiovascular complications (9). Insulin resistance also plays an important role in the pathophysiology of many other prominent metabolic diseases, including obesity (17), polycystic ovarian syndrome (10), and nonalcoholic steatohepatitis (4, 7). Therefore, there is a need to quantitatively assess insulin sensitivity in humans that is comparable with that obtained from full-length glucose clamps under steady-state conditions. The additional kinetic parameter provided by this procedure is useful for describing insulin response characteristics. Here, we demonstrate that, using this approach, it is possible to simplify and shorten the glucose clamp procedure while still allowing accurate and precise assessment of insulin sensitivity in vivo that is comparable with that obtained from full-length glucose clamps under steady-state conditions. The additional kinetic parameter provided by this procedure is also meaningful and useful in distinguishing insulin responsiveness across insulin doses and across groups with different insulin sensitivities.

METHODS

Study Information

Glucose clamp data from two institutions were used for these analyses. The approach was developed and assessed in the first data set (the development data set) and then applied without further modification to the second data set to confirm broader applicability (the validation data set). The development data set consisted of euglycemic hyperinsulinemic clamp studies (i.e., experimentally fixing glucose concentrations to 90 mg/dl) performed between 1999 and 2007 under a variety of investigational protocols at the Indiana University (IU) General Clinical Research Center, Indianapolis, IN. All of these protocols were approved by the IU Institutional Review Board. The validation data set consisted of isoglycemic hyperinsulinemic glucose clamp studies (i.e., fixing glucose concentrations to each individual’s spontaneous fasting glucose concentration) conducted at the National Institutes of Health (NIH) Clinical Center, Bethesda, MD, under a variety of investigational protocols between 1999 and 2005. Isoglycemic and euglycemic clamps are comparable for normoglycemic individuals but not for diabetic individuals, and therefore, clamp studies from diabetic subjects were not included. These protocols were approved by the Institutional Review Board of the National Heart, Lung, and Blood Institute. All procedures fol-
adjusted manually by trained, experienced nursing personnel and fellows in response to changes in glucose measurements taken every 5 min, without use of algorithmic or computational methods. Changes in GIRs were made on the basis of prior experience passed on within institutions. The steady-state period was defined as a 30-min period \( \geq 2 \) h after initiation of insulin infusion, during which changes in GIR and blood glucose had a coefficient of variation of \(<5\%\) (5).

The mean glucose disposal rate (GDR; expressed per kg body wt/min) was calculated over each 20-min period within each clamp procedure, with adjustments for shifts in and out of the glucose space over the 20-min interval using an adjustment factor of 0.095 (8). The formula used to calculate this interval GDR is presented below, with the GIR for each of four 5-min intervals averaged. The 20% dextrose solution contains \(~180\) mg/dl glucose:

\[
\text{GDR} = \frac{[(\text{GIR}_1 \times 5 + \text{GIR}_2 \times 5 + \text{GIR}_3 \times 5 + \text{GIR}_4 \times 5) / 20 \times 180 \times \text{weight (kg)}] - (\text{glucose}_4 - \text{glucose}_1) \times 0.095}{5}
\]

The reference GDR for each clamp was calculated as the mean of the last three interval GDRs encompassing the final hour (minutes 180 –240; see Fig. 1). In 16 clamps from the development data set and two clamps from the validation data set, it was necessary to use the three GDRs encompassing minutes 160 –220, because these clamps were terminated prior to the 240-min point. This final hour GDR was prespecified as the reference measurement (refGDR) against which fitting results would be compared.

The 20-min interval GDRs across each individual clamp were used for curve fitting (Prism Plus; GraphPad Software). Using the development data set, an initial exploration of alternate fitting approaches was undertaken, using quadratic and exponential models with one or more fitting parameters. Fitting these data with quadratic models did not produce convergent solutions. The data were well described by monoexponential models. Moreover, single-parameter exponential models were not substantially improved by adding additional parameters. Therefore, we fit data using a simple monoexponential curve \( \text{GDR}_t = f\text{GDR} \left[ 1 - e^{-kt} \right], \) where \( \text{GDR}_t \) is interval glucose disposal rate at time \( t, \) fGDR is final (plateau) GDR, \( k \) is half-time constant, and \( t \) is time in minutes since initiation of the procedure. Data fitting performed with the full data set produced fGDR values designated as fitGDR-full. Data fitting performed using abbreviated data sets incorporating only data from the first 180 or the first 120 min of each clamp study generated fGDR values designated as fitGDR-180 and fitGDR-120, respectively.

We prespecified criteria for adequacy of curve fitting. If the estimated error in fGDR exceeded 100\% of the final estimated value, and/or the \( r^2 \) of the nonlinear fit was <0.25, the fit was rejected as inadequate and excluded from comparisons against the refGDR. Rejected fits were inspected individually to ensure that repairable technical errors were not to blame, and the number of poor fits and reasons for poor fits, if evident, were recorded.

**Data Analysis**

**Accuracy: bias and limits of agreement.** Comparisons of refGDR with fitGDR-full, fitGDR-180, and fitGDR-120 were performed using linear least-squares fitting and Altman-Bland plots (2). The mean difference in the fitGDR vs. refGDR comparison was used as a measure of bias introduced by the fitting procedure. We also calculated the 95\% confidence interval (mean difference \( \pm 1.96 \) SD) for these differences as a numerical index of the precision of the estimate, i.e., limits of agreement for the two measurements (13).

**Precision: calibration modeling.** To derive quantitative measurements of prediction accuracy, we applied a calibration model, \( x_i = \alpha + \beta y_i + e_i, \) where \( x_i \) is refGDR, \( y_i \) is fitGDR-full (or fitGDR-180 or fitGDR-120), and \( e_i \) is the random error for the \( i \)th subject. It was assumed that the random error had Gaussian distribution with mean of 0 and a constant variance. Using this model, cross-validation prediction error (CVPE) and root mean squared error (RMSE) were calcu-

---

**Data Sets**

We analyzed data from glucose clamps performed in lean and obese subjects ranging in age from 18 to 65 yr from two independent laboratories. The development data set contained data from 61 subjects, whereas the validation data set contained data from 122 subjects. Some subjects contributed more than one clamp data set. A total of 310 sets of clamp data were available for analysis (108 in the development data set and 202 in validation data set). Since our analyses only compared fitting procedures with complete clamp data within each individual procedure, each clamp study was treated as an independent study regardless of whether a subject was studied more than once.

**Clamp Procedures**

Clamp studies in both the development and validation data sets were conducted using unprimed constant insulin infusion rates (ranging from 10 to 300 mU·m\(^2\)·min\(^{-1}\) in the development data set and 120 mU·m\(^2\)·min\(^{-1}\) only in the validation data set). Whole blood glucose concentrations were measured at bedside every 5 min with a glucose analyzer (Yellow Springs Instruments, Yellow Springs, OH), using blood sampled from the femoral artery (IU) or an arterialized dorsal hand vein (NIH). Euglycemia (IU) or isoglycemia (NIH) was maintained using a variable infusion of 20\% dextrose. The GIR was

---

**Fig. 1.** Changes in glucose disposal rate (GDR) over time with hyperinsulinemic euglycemic clamps. Mean glucose disposal rate for each 20-min interval, averaged across all subjects for each data set. The solid line represents the mean monoexponential curve fit to these data (using the full data set), and the dashed lines represent the 95\% confidence interval of the fit line. The last 3 data points are those that are used to calculate the reference GDR.
Clinical characteristics of study subjects

| Weight, kg | 82.9 |
| Age, yr | 36 (18–52) | 41 (20–65) |

To compare predictive accuracy of fitGDR-120 and fitGDR-180 with fitGDR-full in terms of CVPE and RMSE, we tested the hypothesis that fitGDR-full had a smaller RMSE/CVPE than fitGDR-120 or fitGDR-180 using a Bootstrap percentile method, with 60,000 replications performed for each comparison (using SAS version 9). For example, for the comparison of CVPE for fitGDR-120 vs. CVPE for fitGDR-full, this method produced a sample of 60,000 differences in CVPE [CVPE (fitGDR-full) – CVPE (fitGDR-120)], and then a P value was estimated as the proportion of the Bootstrap replications greater than zero. P values were calculated in this way for the three pairwise comparisons within each data set.

Kinetics. For each individual clamp data set, the kinetic parameter k was converted to half-time by calculating 0.69/k. Standard descriptive statistics and group comparison techniques were applied to these data.

Analyses by subgroups. The development data set included a range of insulin infusion rates, and therefore, it was well suited to comparing GDR results and half-time results under “low” and “high” insulin action. For this purpose, data from low-dose clamps (10 and 30 mU·m²·min⁻¹) were grouped and compared with data from high-dose clamps (120 and 300 mU·m²·min⁻¹). Both of these groupings included lean and obese subjects; analyses for interactions of phenotypic group with insulin action group were undertaken to evaluate an expected additional effect of phenotype.

The validation data set originated from studies using a single insulin infusion rate (120 mU·m²·min⁻¹) with subjects falling into lean normotensive, lean hypertensive, and obese hypertensive groups (5, 6). Therefore, we were able to undertake comparisons of GDR results and half-time results under these phenotypic groupings.

For subgroup comparisons within data sets, a one-way analysis of variance was applied, with subsequent pairwise testing using the Student-Newman-Keuls procedure, since variances were equal. These analyses were performed using SPSS 16.0 (SPSS, Chicago, IL).

RESULTS

Demographic characteristics of study subjects from the development and validation data sets are presented in Table 1.

Table 1. Clinical characteristics of study subjects

<table>
<thead>
<tr>
<th>Patient Characteristics</th>
<th>Development Data Set</th>
<th>Validation Data Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>61</td>
<td>122</td>
</tr>
<tr>
<td>Age, yr</td>
<td>36 (18–52)</td>
<td>41 (20–65)</td>
</tr>
<tr>
<td>Height, cm</td>
<td>173.2 ± 7.7</td>
<td>170.2 ± 8.8</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>82.9 ± 22.2</td>
<td>87.1 ± 19.8</td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td>27.6 ± 7</td>
<td>30 ± 6.6</td>
</tr>
<tr>
<td>Race (white/nonwhite)</td>
<td>25/36 (41/59%)</td>
<td>81/41 (66/34%)</td>
</tr>
<tr>
<td>Sex (male/female)</td>
<td>39/22 (64/36%)</td>
<td>53/69 (44/56%)</td>
</tr>
<tr>
<td>Patient type (lean/obese)</td>
<td>29/32 (47/53%)</td>
<td>52/70 (43/57%)</td>
</tr>
<tr>
<td>Steady-state glucose, mg/dl</td>
<td>90.0 ± 7.9</td>
<td>82.8 ± 10.7</td>
</tr>
</tbody>
</table>

Results are means ± SD. Where appropriate, the range or percentages are shown.

Since data from lean and obese subjects were included, a wide range of insulin sensitivity was evident in both data sets. Specifically, in the development data set refGDR ranged from 1.2 to 15.6 mg·kg⁻¹·min⁻¹, whereas in the validation data set refGDR ranged from 3.31 to 18.6 mg·kg⁻¹·min⁻¹.

In the development data set with 108 clamp studies, seven studies were excluded due to insufficient data to allow the modeling procedure (i.e., too many missing data points). Applying the prespecified adequacy-of-fit criteria, we found three clamp studies where fitting was complete but the estimated error of fGDR was >100% of its estimated value and two other studies with individual fitting, r² < 0.25 (i.e., these were poor fits by prespecified criteria). Another eight clamp studies were found to have interrupted curve fitting due to significant divergence from the model. Thus, after 20 exclusions, a total of 88 clamp procedures were available for analysis in the development data set (81% of development data set). In the validation data set with 202 clamp studies, 11 studies had individual fits with r² < 0.25 and seven studies had estimated error of fGDR >100% of the estimated value of fGDR. In this data set, fitting was not possible for one study. Thus, after all exclusions, a total of 183 clamp procedures were available for analysis in the validation data set (91% of validation data set).

In the development data set, we found that applying fitting procedures using the full data set produced final GDR estimates that were highly concordant with results obtained using the reference method (Figs. 1 and 2 and Table 2). There was no evidence of heteroscedasticity, with comparable variability in the between-measurement difference (D) across the range of GDR values (Fig. 2). Comparisons of fit-derived estimates with the reference method using abbreviated data sets are presented in Table 2 and Fig. 2. Using 3- or 2-h data resulted in a prediction of final GDR that was also strongly concordant with the reference method using abbreviated data sets are presented in Table 2 and Fig. 2. With the shorter data sets, the limits of agreement worsened (increased RMSE and CVPE; Table 3). The 3-h fit data were not substantially different from those derived using all available data. However, fitting the 2-h data did significantly reduce predictive precision; the increase in CVPE and RMSE was statistically significant compared with values obtained from fitting using data from complete clamp studies (Table 3).

In the validation data set, fitting the full data set again produced GDR values that were strongly concordant with those

<table>
<thead>
<tr>
<th>Table 2. Accuracy of fitting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>refGDR vs.</td>
</tr>
<tr>
<td>fitGDR-full</td>
</tr>
<tr>
<td>fitGDR-180</td>
</tr>
<tr>
<td>fitGDR-120</td>
</tr>
</tbody>
</table>

Comparisons between reference measurement (refGDR) and fitGDR are presented. GDR, glucose disposal rate; r, pearson correlation coefficient; D, mean difference [refGDR – prediction of refGDR from fitGDR]; CI, confidence interval. Negative D values indicate fit values greater than reference values.
derived using the reference method (Table 2 and Fig. 3). Fitting
the 3-h and 2-h data sets produced predicted GDRs that were
concordant with what was seen in fitting the full data set.
However, fitting shorter data sets again resulted in a progres-
sive increase in bias, loss of precision, and statistically larger
prediction errors (Table 3 and Fig. 3).

Altman-Bland plots are presented in Figs. 2 and 3 for the
development and validation data sets, respectively. Two im-
portant pieces of information are evident in these plots. First,
on the y-axis, the offset or bias is evident as the mean
difference (D) in GDRs derived from the two approaches being
compared. Second, the variability in fit-derived GDR estimates
of reference GDR can be evaluated by the dispersion of data
points around the mean D across the range of values observed.
The 95% confidence intervals around the mean values of D
(i.e., limits of agreement) are shown as dashed lines in Figs. 2
and 3. Qualitatively, it is evident that fit-derived GDR values
based on shorter clamp exposures resulted in reduced accuracy
and precision in predicting the final GDR in both development
and validation data sets.

Kinetic parameters are available from the fitting procedure.
In the development data set the half-time values were broadly
spread, consistent with the variety of insulin infusion rates
applied: means $\pm$ SD 51.5 $\pm$ 41.6, range 15.5–299.5 min when
fitting the full data set. In the validation data set using a single
insulin infusion rate, these values were 36.9 $\pm$ 26, range
6.9–176 min, again fitting the full data set.

Both data sets included subjects falling into categories of
white/nonwhite, male/female, and lean/obese. Therefore, we
also compared accuracy across subgroups categorized by race,
sex, and obesity. In the development data set, no significant
differences were detected in D derived from fitGDR-full vs.
refGDR when subgroups for sex (P = 0.26), race (P = 0.97),
or obesity (P = 0.23) were examined. Moreover, the 95%
confidence intervals for D within each of these subgroup
comparisons were overlapping, indicating no significant differ-
ence detected in bias among subgroups (data not shown).
Similar results were obtained with subgroup analysis of the
validation data set (comparisons of D for fitGDR-full vs.
refGDR by sex, P = 0.49; obesity, P = 0.07; race, P = 0.66).

In the development data set, we also compared D across
different insulin infusion rates. The variance differed consid-
erably among the subgroups (Levine statistic 5.5, P = 0.001),
invalidating the assumption of equal variance required for
standard ANOVA. Pairwise testing assuming unequal vari-
ances in subgroups did not reveal specific differences among
insulin infusion rates. Taken together, this argues against
significant differences in D among subgroups with different
insulin infusion rates.

### Table 3. Precision of fitting

<table>
<thead>
<tr>
<th>Development Data Set</th>
<th>Validation Data Set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
</tr>
<tr>
<td>fitGDR-full</td>
<td>1.11</td>
</tr>
<tr>
<td>fitGDR-180</td>
<td>1.27 $\pm$ 0.133</td>
</tr>
<tr>
<td>fitGDR-120</td>
<td>1.56 $\pm$ 0.031</td>
</tr>
</tbody>
</table>

| RMSE, root mean squared error of prediction; CVPE, cross-validation prediction error. Comparisons between fitGDR-full and fitGDR-180 (or fitGDR-120) in terms of CVPE and RMSE. P values are presented for the comparison of CVPE/RMSE for fitGDR-full vs. abbreviated fits. |
In application, these procedures would be used to determine differences between groups of differing underlying insulin sensitivity or changes in insulin sensitivity with an intervention. Therefore, we used the reference method and fitting to compare low and high insulin infusion rates in the development data set and to compare phenotypically differing underlying insulin sensitivity across subject groups in the validation data set (Table 4). We also undertook these comparisons for the half-time parameters derived from the fitting (Table 4). Overall, it is evident that the fit-derived GDR values distinguish low from high insulin responses, and insulin-sensitive from insulin-resistant groups, with statistical power comparable with that of the reference method. The kinetic parameters were also different across these groupings in expected ways when fitting the full data set (longer half-times for lower insulin exposure and for insulin-resistant groups). However, these kinetic parameters were less robust to the use of shorter duration clamp data, although they were still able to distinguish insulin-sensitive from insulin-resistant responses in the validation data set.

**DISCUSSION**

We applied a monoexponential curve to GIR-derived GDRs across the time course of hyperinsulinemic euglycemic clamp studies performed in humans. Fitting the full 4-h data set produced values of glucose disposal rate that were highly concordant with values derived using the reference method, which uses only data in the final hour. Fitting abbreviated data sets, namely 3- and 2-h data, produced predicted GDRs that exhibited nominally but not significantly increased bias (with modest underestimation of the final GDR), but with some loss of precision as indicated by wider ranges of limits of agreement and statistically increased prediction error. This approach was easily applied to an independent second set of data, with

![Fig. 3. Comparisons between refGDR and prediction of refGDR from fitGDR are presented for the validation data set. Linear correlations are shown at top. Altman-Bland plots are shown at bottom.](http://ajpendo.physiology.org/)

**Table 4. Group comparisons using fit data**

<table>
<thead>
<tr>
<th></th>
<th>RefGDR, mg·kg⁻¹·min⁻¹</th>
<th>FitGDR-Full, mg·kg⁻¹·min⁻¹</th>
<th>FitGDR-120, mg·kg⁻¹·min⁻¹</th>
<th>Half-Time-Full, min</th>
<th>Half-Time-120, min</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Development data set</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-dose lean</td>
<td>3.85 ± 1.87*</td>
<td>4.31 ± 2.70*</td>
<td>3.71 ± 2.13*</td>
<td>59.5 ± 41.1*</td>
<td>41.2 ± 21.8</td>
</tr>
<tr>
<td>Low-dose obese</td>
<td>3.79 ± 1.60</td>
<td>4.94 ± 3.25</td>
<td>3.57 ± 1.61</td>
<td>92.7 ± 80.2</td>
<td>55.0 ± 36.0</td>
</tr>
<tr>
<td>High-dose lean</td>
<td>11.43 ± 2.24*</td>
<td>11.62 ± 2.22*</td>
<td>13.19 ± 3.28*</td>
<td>124 ± 10.4</td>
<td>53.9 ± 15.2</td>
</tr>
<tr>
<td>High-dose obese</td>
<td>8.45 ± 2.48*</td>
<td>8.54 ± 2.46</td>
<td>9.99 ± 3.22</td>
<td>42.4 ± 10.4</td>
<td>53.9 ± 15.2</td>
</tr>
<tr>
<td><strong>Validation data set</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lean</td>
<td>10.56 ± 3.62**</td>
<td>10.91 ± 3.66*</td>
<td>11.03 ± 3.69**</td>
<td>27.7 ± 13.4**</td>
<td>28.0 ± 11.8**</td>
</tr>
<tr>
<td>Hypertensive</td>
<td>8.73 ± 2.03*</td>
<td>9.00 ± 2.07</td>
<td>8.94 ± 2.86</td>
<td>38.2 ± 17.2*</td>
<td>35.0 ± 16.4</td>
</tr>
<tr>
<td>Obese</td>
<td>8.20 ± 2.23*</td>
<td>8.95 ± 2.38</td>
<td>8.08 ± 2.52</td>
<td>46.2 ± 25.4*</td>
<td>32.5 ± 13.9</td>
</tr>
</tbody>
</table>

Data presented as means ± SD. *P < 0.001; **P < 0.05 for the main analysis of variance test comparing low-dose vs. high-dose insulin (development data set; see METHODS) or phenotypic groups (validation data set). Statistical significance, P < 0.05 for lean vs. obese within insulin groups in the development data set; a, bstatistical significance, P < 0.05 for pairwise comparisons across groups in the validation data set (different letters indicate statistically different).
similar relationships between the fit and reference estimates of GDR, providing external validation of the approach. In both data sets, the fit procedure produced a kinetic parameter not otherwise available that was sensitive to differing levels of insulin action and differing degrees of insulin sensitivity.

This procedure was readily applied to data sets collected in two separate laboratories, using different personnel and different approaches to the clamp goals. Furthermore, the two data sets included modestly different collections of subjects (see Table 1). Despite all of these differences, the modeling produced similar results in both data sets. This argues that the modeling reflects the underlying biological response to hyperinsulinemic clamping rather than some specific feature of the clamping procedures performed in our laboratories.

**Modeling Vs. Manual GDR**

The current investigations were predicated on widespread observations that insulin-stimulated glucose disposal following institution of a fixed-dose insulin infusion increases exponentially to a plateau (16, 21). The selected monoexponential equation models this behavior well, as evidenced from Fig. 1 (plotting mean data for each data set). In the unselected collection of data we used from two independent laboratories, we found that fitting could not be performed in eight clamps from the development data set and in one clamp from the validation data set. As a general statement, where data could not be fit there were technical problems in the performance of the clamp procedures performed in our laboratories.

**Accuracy and Precision of GDR Modeled from Abbreviated Clamps**

Using the fitting procedure to predict steady-state GDR from abbreviated data sets introduces some increased bias and reduced precision, but even the 2-h data sets allowed accurate prediction of the reference GDR (Tables 2 and 4). The prediction errors for fitGDR estimates derived from 2-h clamp data were significantly increased but in reality only modestly elevated compared with those of fitGDR derived from full clamp data. This appears to be the only consequence of the modeling approach.

**When Would Modeling Abbreviated Clamps Be Useful?**

The potential advantage of modeling abbreviated clamps lies principally in the reduced manpower and resource needs if the clamp procedures could be systematically shortened. As noted above, the accuracy of modeling is adequate, but the real price comes with increased measurement noise.

The main implication of this observation is that this increased variability affects sample size estimation, proportionally increasing the number of subjects needed to demonstrate a given difference between measurement conditions. Overall, we observed a~20% increase in RMSE (Table 3) and a corresponding ~25% increase in SD within the development data set (the increase in SD was smaller in the validation data set, but the SD with refGDR for this population was higher; Table 4).

As a concrete example, the number of subjects needed to demonstrate the observed difference between lean and obese subjects in the high-insulin dose group (GDR difference of ~3.0 mg·kg⁻¹·min⁻¹, group SD ~2.4) would be ~12 subjects in each group if full data sets were used but 17 per group if 2-h clamp studies were used with fitting (using typical parameters of 80% power and α = 0.05). In a study of a treatment that was expected to improve insulin sensitivity, one might expect a 20% improvement in GDR [as seen with PPARγ agonists (18, 22, 26)], i.e., from ~8.3 to 10.0 mg·kg⁻¹·min⁻¹ using 120 mU·m²·min⁻¹ clamps. Within the obese subjects in the validation data set, SD was 2.2 at this
infusion rate when full clamps with reference method were used, and SD was 2.5 for 120-min fit data. Using a repeated-measures design and comparing against placebo, to demonstrate this change (again with 80% power and $\alpha = 0.05$), ~28 subjects per group would be needed if a study using full clamps is designed, and 35 subjects per group would be needed if 2-h clamps with fitting are used.

This assessment of shortened clamps comes in the context of a number of other alternative approaches to assessing insulin resistance. For example, the quantitative insulin sensitivity index and logarithmic transformation of homeostasis model-derived insulin resistance utilize fasting insulin and glucose levels to estimate insulin sensitivity in an overnight-fasted state (14, 19, 27). These indices are known to have good correlation with insulin sensitivity index derived from glucose clamp method in various insulin-resistant states such as type 2 diabetes, polycystic ovarian syndrome, and obesity (3, 5, 18, 20). They have also been shown to have adequate repeatability and good discriminant ratios (18), and sample sizes of 20–30 per group have been shown to be sufficient for demonstrating changes in insulin sensitivity using these indices (24, 26). However, these techniques most directly assess hepatic insulin sensitivity rather than whole body or skeletal muscle insulin sensitivity (8, 10, 18–20). Also, the hyperinsulinemic clamp technique brings other advantages such as the ability to combine with radiolabeled tracers, which can be used to rigorously quantify the contribution of hepatic insulin sensitivity to whole body insulin sensitivity (23, 25) or the sensitivity of adipose tissue to insulin (12, 15). Furthermore, as demonstrated here, modeling of clamp data sets provides an opportunity to derive kinetics of the insulin response. Therefore, the choice between approaches requires carefully assessing which questions are most important in each circumstance as well as the resources and sample size required for each approach.

Limitations

The present analysis did not include clamp studies in diabetic subjects, owing to the structural problem in making this comparison across the two data sets used (isoglycemic vs. eu glycemic clamps). Exploratory analyses in eu glycemic clamp data from diabetic subjects suggests that the modeling procedure will apply equally to this data, but this was not explicitly evaluated because it could not be validated. We did not explore comparative correlations of fasting indices of insulin resistance vs. abbreviated clamps, choosing to focus on the interrelationships of short vs. long clamps only. In our data sets, the fitting procedure could not be used to derive GDR values in ~10% (32/310) of available clamp data. The best handling of such data in a clinical study is not defined here, but an attempt to derive classical GDR values would be tempting. However, it may be necessary to exclude such data rather than apply mixed methods for deriving the main end points.

GRANTS

The data included in the present analyses were generated under funding from a variety of sources, including NIH/National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (DK-42469, M01-RR-00750), the American Diabetes Association (Junior Faculty and Career Development Awards to K. J. Mather), and intramural support through NHLBI and the National Center for Complementary and Alternative Medicine (to M. J. Quon).

K. J. Mather is supported by a Career Development Award from the American Diabetes Association.

DISCLOSURES

The authors report no dualities of interest that impact the material presented in this article.

REFERENCES
