0013-7227/01/$03.00/0
Printed in U.S.A.

The Journal of Clinical Endocrinology & Metabolism 86(11):5457–5464

Repeatability Characteristics of Simple Indices of Insulin Resistance: Implications for Research Applications

KIEREN J. MATHER, A. EVAY HUNT, HELMUT O. STEINBERG, GIANCARLO PARADISI, GINGER HOOK, ARIE KATZ, MICHAEL J. QUON, AND ALAIN D. BARON

Indiana University School of Medicine (K.J.M., A.E.H., H.O.S., G.P., G.H., A.D.B.), Division of Endocrinology and Metabolism, Indianapolis, Indiana 46250; and Hypertension-Endocrine Branch (A.K., M.J.Q.), National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892

The objectives of this study were to evaluate test characteristics, such as normality of distribution, variation, and repeatability, of simple fasting measures of insulin sensitivity and to use the results to choose among these measures. Duplicate fasting samples of insulin and glucose were collected before 4 h of euglycemic hyperinsulinemic clamping using insulin infusion rates ranging from 40–600 mU/m^2·min. Currently recommended estimates of insulin sensitivity, including the fasting insulin, 40/insulin, the homeostasis model assessment, the logarithmic transformation of the homeostasis model assessment, and the Quantitative Insulin Sensitivity Check Index, were evaluated. The normality of distribution and the variability of the tests (coefficient of variation and discriminant ratio) were compared between the measures and against the "gold standard" hyperinsulinemic clamp. Data from 253 clamp studies in 152 subjects were examined, including 79 repeated studies for repeatability analysis. In subjects ranging from lean to diabetic, the log transformed fasting measures combining insulin and glucose had normal distributions and test characteristics superior to the other simple indices (logarithmic transformation of the homeostasis model assessment coefficient of variation, 0.55; discriminant ratio, 13; Quantitative Insulin Sensitivity Check Index coefficient of variation, 0.05; discriminant ratio, 10) and statistically comparable to euglycemic hyperinsulinemic clamps (coefficient of variation, 0.10; discriminant ratio, 6.4). These favorable characteristics helped explain the superior correlations of these measures with the hyperinsulinemic clamps among insulin-resistant subjects. Furthermore, therapeutic changes in insulin sensitivity were as readily demonstrated with these simple measures as with the hyperinsulinemic clamp. The test characteristics of the logarithmic transformation of the homeostasis model assessment and the Quantitative Insulin Sensitivity Check Index are superior to other simple indices of insulin sensitivity. This helps explain their excellent correlations with formal measures both at baseline and with changes in insulin sensitivity and supports their broader application in clinical research. (*J Clin Endocrinol Metab* 86: 5457–5464, 2001)

THE PAST DECADE has seen a surge of interest in insulin resistance, both as an etiological factor in the pathogenesis of type 2 diabetes mellitus (1) and as a key component of the dysmetabolic cardiovascular syndrome (also known as the insulin resistance syndrome) (2). There is a need for simple, accessible tools for the measurement of insulin sensitivity. Classic steady state hyperinsulinemic clamps (3) arguably represent the current "gold standard." Both steady state and dynamic tests are time intensive for both the subject and the investigator, somewhat invasive, and incur significant costs because they require multiple insulin level measurements. These disadvantages have prevented the general application of these tests in the both clinical and research arenas.

Most large scale epidemiological studies have simply correlated fasting insulin levels with the outcomes of interest. More than 15 yr ago, mathematical modeling of the normal physiological balance of insulin and glucose produced the homeostasis model assessment (HOMA), which provided equations for estimating insulin resistance (HOMA-IR) and β-cell function from simultaneous fasting measures of insulin and glucose levels (4). Although this estimate of insulin resistance has been used in some epidemiological studies (5–11), few clinical studies have relied on this measure alone because of reportedly poor accuracy compared with hyperinsulinemic clamps and a reliance on notoriously variable insulin measurements.

The direct relationships between insulin levels or HOMA-IR and clamp measures of insulin sensitivity are hyperbolic rather than linear (12–15). Recently, a number of groups have suggested that improved correlations of fasting measures with glucose clamps might be obtained with linearizing transformations. Both logarithmic (6, 15–18) and reciprocal transformations (15, 19) have been proposed. We hypothesized that the improved correlations reflected improvements in test characteristics and that examining these characteristics would support a rational choice between the various mathematical transformations. To this end, we evaluated the repeatability, variability, and discriminant power (20) of various calculated indices of insulin sensitivity and compared them to the hyperinsulinemic euglycemic clamp. These findings form the basis for the correlations between the clamps and the fasting indices, so these were compared as well.

Abbreviations: CV, Coefficient(s) of variation; DR, discriminant ratio; EH, euglycemic hyperinsulinemic; GDR, glucose disposal rate; HOMA-IR, homeostasis model assessment of insulin resistance; I40, I120, I300, and I600, insulin doses of 40, 120, 300, and 600 mU/m^2·min; logHOMA-IR, logarithmic transformation of HOMA-IR; logInsulin, logarithm of fasting insulin; QUICKI, Quantitative Insulin-Sensitivity Check Index.

Materials and Methods

Subjects were selected from a database of hyperinsulinemic euglycemic clamp data collected at our institution. Three groups of subjects existed in the database: lean, healthy controls (body mass index ≤ 27); obese nondiabetic (insulin-resistant) subjects (body mass index >27); and subjects with type 2 diabetes mellitus (defined according to American Diabetes Association criteria). Extremes of blood pressure and lipid measures are routinely considered exclusion criteria for our studies, so these subjects differed principally in their degree of insulin resistance. We excluded clamp studies performed before September 1993, at which time a commercial supersensitive insulin assay was adopted for routine use in our core laboratory. The data collected for these studies included duplicate fasting insulin and glucose measurements (separated by 10 min) before the initiation of insulin infusions. Standard measures of weight and height were performed, and percentage body fat was measured by dual energy x-ray absorptiometry or water displacement (for subjects >122 kg, the upper limit for the dual energy x-ray absorptiometry machine). Diabetic subjects receiving oral agents were withdrawn from therapy 2–3 wk before clamp studies. Diabetic patients treated with insulin discontinued long-acting insulin 5 d before and short-acting insulin 24 h before studies.

We chose to divide subjects by clinical characteristics rather than with a post hoc definition of insulin sensitivity based on clamp outcome. Although this results in some overlap of degrees of insulin sensitivity across groups, it provides a better match with the more general prospective approach of classifying subjects on the basis of clinical characteristics, and any statistical effects of this overlap can be accounted for by examining continuous relationships within and across groups.

The database included euglycemic hyperinsulinemic (EH) clamp studies performed at a range of insulin doses, and four levels were included in the present analysis: 40, 120, 300, and 600 $mU/m^2 \cdot min$ (I40, I120, I300, and I600). Except for insulin infusion rate, the design of clamp studies was uniform. Subjects were studied after an overnight (10–14 h) fast. At least 30 min after placement of vascular accesses, an ~20-min baseline period took place during which fasting insulin and glucose samples were taken 10 min apart. An unprimed 4-h insulin infusion was then applied, with a variable infusion of 20% glucose in water to maintain euglycemia. The glucose disposal rate (GDR) was calculated as the mean of the final two 20-min GDRs during the last 40 min of the 4-h clamp. Lean and obese subjects were studied at all four clamp levels, whereas clamp studies in diabetic subjects were available only at the higher two infusion rates. Under these hyperinsulinemic conditions, hepatic glucose output is assumed to be completely suppressed (21), so tracer methods for the measurement of hepatic glucose output were not used.

Insulin determinations were made using a dual-site RIA specific for human insulin with cross-reactivity with proinsulin less than 0.2%. The lower detection limit is 0.56 pmol/liter, and in our laboratory the interassay and intraassay coefficients of variation (CV) are 4.1% and 2.6%, respectively. Because the sensitive insulin assay is known to become increasingly unreliable as measured values approach the lower detection limit, we retested all samples with reported values less than 50 pmol/liter using an ultrasensitive human double antibody insulin assay (detection limit, 0.056 pmol/liter ; intratest and intertest CV, 12.4% and 9.0%), for a repeat correlation analysis.

Our data set included 13 young women with the polycystic ovarian syndrome who had undergone 3 months treatment with oral troglitazone (600 mg daily) and who had undergone I120 clamp studies as described above before and after this treatment. Using these data, the ability of the simple fasting measures to demonstrate the change in insulin sensitivity was compared with the EH clamp.

Test characteristics

Repeat fasting insulin and glucose measures were available for all subjects who had participated in more than one clamp study. The minimum interval for such repeat studies was 4 wk. Where subjects had undergone treatment with identical clamp conditions on more than one occasion, only those studies that had been repeated within 6 months were included. Furthermore, if any significant change in body mass index, drug therapy, or metabolic status (*e.g.* the new development of type 2 diabetes mellitus) had occurred, such repeat studies were excluded from the repeatability analysis.

Four assessments of test characteristics were undertaken. In addition to assessing for a uniform distribution of the data itself, uniform distribution of the error on repeat testing across the measurement range (homoscedasticity) was investigated using Altman-Bland plots. Variability and repeatability were assessed using the CV and a newly proposed measure, the discriminant ratio (DR) (20). For the CV, the standard formula was used:

$$CV = SD/mean.$$

In this case, however, the SD was arrived at using data from repeat tests in a repeated measures ANOVA, taking the square root of the sum of the mean square variations for the repeated test (within-subject error term) and the residual error (22). The two tests results were averaged to arrive at a single value for calculations of group means.

Unlike the CV, the interpretation of the DR is not dependent on the absolute value of the population mean. Also, the DR includes both of the principal sources of systematic error, *i.e.* between-subject and within-subject error. It is easily calculated from a repeated measures ANOVA using the error terms (MS_B, between-subject error term; MS_W, within-subject error term across repeat studies) derived from a standard repeated measures ANOVA table for the repeated tests:

$$DR = \sqrt{[(MS_B - MS_W)/MS_W]}.$$

Because it is a measure of a test's ability to distinguish individuals, data from the entire range of subjects are included in this calculation. Standard 95% confidence intervals can be calculated, and a method of comparing these intervals is provided that allows statistical comparisons of DRs between tests (20). The noncentral F table and χ^2 calculations required for these comparisons were carried out using the UCLA online statistical calculators (http://www.ucla.stat.edu). This measure also provides a means for correcting an observed correlation between tests for the known measurement error of each test (giving an improved estimate of the true underlying correlation between the two tests) using a correction factor incorporating the observed DRs of the two measures being compared (20).

Estimates of insulin sensitivity

Test characteristics were determined for 1) the fasting insulin level itself; 2) 40/insulin (μU/ml); 3) the logarithm of fasting insulin (logInsulin); 4) the HOMA-IR; 5) the logarithm of the HOMA-IR (logHOMA-IR); and 6) the Quantitative Insulin-Sensitivity Check Index (QUICKI) as well as for the clamp-derived GDR.

The HOMA-IR was first put forward in 1985 by Matthews *et al.* (4). The authors recommended using triplicate fasting measures of insulin (reported in μU/ml, at the time measured with single-site antibody assays) and glucose (in mmol/liter). A constant was applied to correct the value to unity in normal subjects, providing the following formula:

$$[insulin \times glucose]/22.5.$$

The most recently proposed derivation using simple fasting measures is the QUICKI (15), which incorporates both inversion and logarithmic transformations. It is calculated as follows (insulin is expressed in mU/ml, and glucose in mg/dl):

$$1/[\log(insulin) + \log(glucose)].$$

The logHOMA-IR and QUICKI are simply related by inversion and otherwise differ only by the normalizing constant applied to the HOMA-IR. The inclusion of the logarithm alone (logInsulin) and the inversion alone (40/insulin) allowed a comparison of the effects of the two mathematical transformations.

Correlations with GDR

Comparisons of the estimates of insulin sensitivity provided by each of the above measures were performed relative to the gold standard measure provided by the GDR. Although it is possible that the GDR is underestimated in the I40 clamps as a result of an unmeasured contribution from hepatic glucose production, this concern is minimized by

the 3- to 4-h duration of the clamp procedure (21). This permits a reliance on the GDR measures at all insulin infusion rates as the standard for comparison. The GDR and fasting measurements were derived from values measured in each individual on the same day. Pearson's correlation coefficients, and the r-to-z estimate of statistical significance, were calculated using StatView 5.0 for Macintosh (SAS Institute, Inc., Chicago, IL). The adjusted correlations, taking into account the measured variability of the tests themselves, were calculated according to the method of Levy *et al.* (20).

Results

Subjects

We evaluated 253 clamp studies in 152 subjects. The characteristics of the subjects are detailed in Table 1. Three of the 11 diabetic subjects were treated with insulin, 1 was treated with a sulfonylurea medication, and the others were newly diagnosed. The obese subjects and type 2 diabetic subjects had lipid levels and blood pressure within the normal range, but they had slightly higher low density cholesterol, lower high density cholesterol, higher triglycerides, and higher blood pressure than the lean subjects. This is consistent with the known characteristics of the insulin resistance syndrome. The racial distribution was 91 (59%) Caucasian, 61 (40%) African American, and 3 (1%) Mexican American. The 48 women studied included 13 with a known diagnosis of the polycystic ovarian syndrome.

Test characteristics were derived from repeated fasting blood samples in 45 lean and 34 obese subjects. Also, 27 lean and 23 obese subjects had repeat measurements of GDR under identical conditions. Correlation analyses included every study in the data set. Except where noted, all results pertain to the standard insulin assay.

Test characteristics

Logarithmic transformations but not inversions normalized the distribution of the data based on fasting insulin levels. The logarithmic transformations also served to normalize the distribution of error across the range of measurements (Fig. 1). The Altman-Bland plot for insulin alone (Fig. 1, *top right*) is clearly heteroscedastic (the variation in the measurement increases across the range of measured values), and although the two-test correlation of insulin appears satisfactory, it is heavily dependent on a minority of data points at the upper end of the range. By contrast, the log-HOMA-IR and QUICKI have much more uniform variability across the range of values and good two-test correlations, with more uniform distributions of measured values. GDR (Fig. 1, *top left*) has these desirable attributes without mathematical transformations.

Variability was assessed using CV and the DR (Table 2). The difficulties of depending on CV alone were apparent in comparing logHOMA-IR and QUICKI, which had markedly different CV despite their simple mathematical relationship. This counterintuitive difference results from the effects of logarithmic and inversion transformations on numbers near one and zero. This makes the CV difficult to interpret, but intuitively the inherent repeatability of these measures must mirror each other. The DRs for these two tests were superior to all of the other simple fasting measures and statistically equivalent to each other and to the clamp-derived GDR (Table 2).

The logInsulin improved the distribution of the results and of the error in the measurement and provided improved CV while maintaining a good DR. However, the DR remained statistically inferior to that of the logHOMA-IR and QUICKI (Table 2).

The alternative approaches of using the insulin concentration alone, the inversion of insulin, or the untransformed HOMA-IR suffered from the heteroscedasticity inherent in the insulin measurement itself. The CV of these values simply reflected the variability of the insulin measurement (Table 2). The DRs of these values were no better than that of insulin alone. The ultrasensitive insulin assay made no significant improvements to any of the above characteristics (data not shown).

In summary, the best combinations of these characteristics (a normal distribution, a favorable distribution of measurement error across the range of values, a low CV, and a high DR) was seen with the GDR derived from EH clamps. The heteroscedasticity of insulin-based indices was improved by both inversion and logarithmic transformations, although only the latter produced a normal distribution of data. The apparent effects of these transformations on the CV are in part an artifact of the mathematics, and these divergent results in fact do not (and logically cannot) represent a significant alteration in this measure of variability. The transformed HOMA and QUICKI values, however, both showed improved DRs equivalent to that of the clamp-derived GDR. This marks these two measures as superior to the other fasting measures of insulin sensitivity.

TABLE 1. Characteristics of the study subjects

	Lean (n = 69)	Obese (n = 72)	Type 2 diabetes mellitus (n = 11)
Gender (male/female)	62/7	38/37	7/4
Age (yr)	38.6 (0.7)	36.3 (0.8)	37.4 (2.5)
Weight (kg)	71.4 (0.8)	99.9 (1.6)	89.5 (7.8)
Body mass index (kg/m^2)	22.8 [17.9–26.9]	33.8 [27.2–55.8]	32.0 [21.7–53.6]
Glucose (mmol/liter)	5.02 (0.03)	5.12 (0.04)	10.43 (0.97)
Insulin (pmol/liter)	52.7 (3.4)	112.0 (9.4)	107.2 (20.4)
Cholesterol (mmol/liter)	3.23 (0.12)	3.83 (0.11)	3.59 (0.28)
LDL cholesterol (mmol/liter)	1.93 (0.11)	2.42 (0.09)	2.08 (0.28)
HDL cholesterol (mmol/liter)	0.93 (0.05)	0.74 (0.03)	0.88 (0.14)
Triglycerides (mmol/liter)	0.99 (0.06)	1.72 (0.16)	1.59 (0.32)
Mean arterial pressure (mm Hg)	97.5 (1.6)	104.9 (2.0)	104.5 (4.4)

Results are presented as mean (SEM) except, where appropriate, the [range] has been presented. LDL, Low density lipoprotein; HDL, high density lipoprotein.

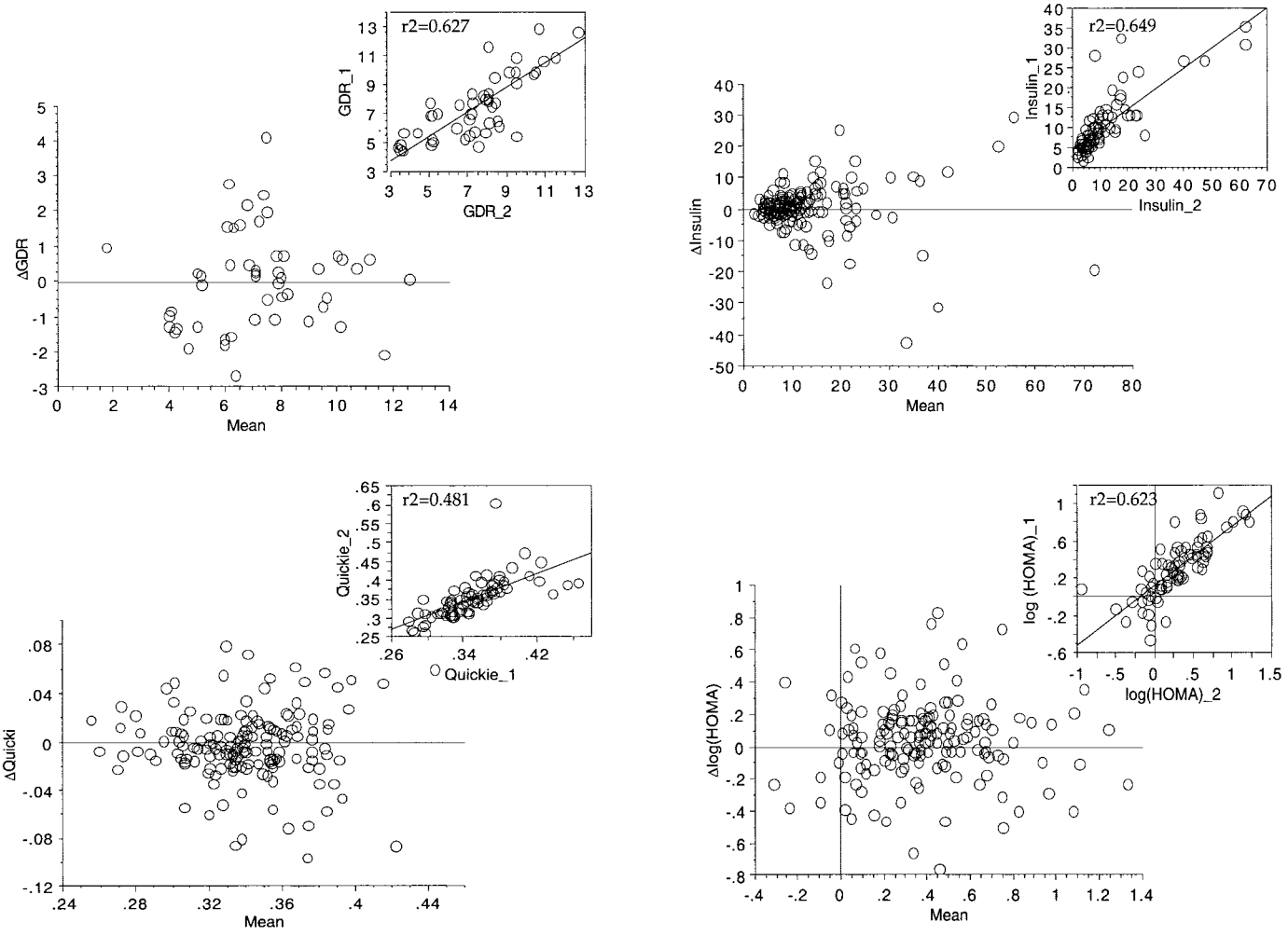

FIG. 1. Altman-Bland plots comparing repeat test characteristics of GDR, fasting serum insulin, QUICKI, and logHOMA-IR. Main plots demonstrate the distribution of variability across the range of values observed. Inset plots reveal inherent correlations between repeat measurements on two separate occasions.

TABLE 2. Measures of repeatability of tests of insulin sensitivity

Test	CV (SD/mean)			DR (95% confidence interval)
	All	Lean	Obese	All
GDR	0.10 (0.80/7.84)	0.09 (0.83/9.05)	0.12 (0.61/5.33)	6.4 (5.0–8.5)
Insulin	0.53 (4.51/8.43)	0.24 (1.39/5.73)	0.59 (8.96/15.18)	3.0 (2.5–3.7)*
40/Insulin	0.52 (3.87/7.44)	0.51 (4.62/9.05)	0.21 (0.81/3.88)	1.8 (1.4–2.3)*
logInsulin	0.17 (0.14/0.82)	0.24 (0.17/0.71)	0.09 (0.10/1.09)	3.5 (2.9–4.3)*
HOMA-IR	0.51 (1.10/2.17)	0.24 (0.32/1.37)	0.58 (2.08/3.54)	3.4 (2.7–4.2)*
logHOMA-IR	0.55 (0.12/0.22)	1.6 (0.13/0.08)	0.23 (0.11/0.45)	13.4 (11.1–16.5)†
QUICKI	0.05 (0.02/0.36)	0.06 (0.02/0.37)	0.03 (0.01/0.33)	10.2 (8.4–12.5)

Results comparing duplicate measurements on two separate occasions. Numbers of subjects included are detailed in *Results*. χ^2 statistic reveals the seven tests to have different DRs.

* $P < 0.05$ compared with logHOMA-IR by pairwise comparison.

† $P < 0.05$ compared with GDR by pairwise comparison.

Correlations to EH clamps

This theoretical grounding informs the assessment of the relationships between these fasting indices and the EH clamp. Simple correlations depend on a normal distribution of the related variables and of their measurement error. These characteristics were achieved by logarithmic transformation, validating correlation analysis with this subset of the fasting measures. The correlations of these measures with GDR, and the corrected correlations after adjustment for the variability of the tests, are presented in Table 3. Untransformed fasting insulin, 40/insulin, and HOMA-IR had only moderately inferior correlation coefficients (data not shown), but they exhibited nonlinear relationships with GDR. Although the observed correlations differed markedly by de-

TABLE 3. Correlations of GDR and simple fasting tests of insulin sensitivity

	I40	I120	I300	I600
Lean subjects	n = 69	n = 12	n = 8	n = 33
logInsulin	−0.41* [0.43]	0.34 [0.36]	0.31 [0.33]	−0.22 [0.24]
logHOMA-IR	−0.38† [−0.40]	0.40 [0.42]	0.37 [0.38]	−0.20 [−0.21]
QUICKI	0.36‡ [0.42]	−0.35 [0.42]	−0.32 [−0.33]	0.20 [0.20]
Obese subjects	n = 28	n = 53	n = 13	n = 17
logInsulin	−0.66* [−0.69]	−0.73* [−0.76]	−0.55§ [−0.59]	−0.51‡ [−0.55]
logHOMA-IR	−0.66* [0.67]	−0.72* [−0.77]	−0.66§ [−0.67]	−0.46‡ [−0.47]
QUICKI	0.65* [0.66]	0.73* [0.88]	0.57§ [0.58]	0.50‡ [0.51]
Type 2 diabetes mellitus subjects			n = 5	n = 6
logInsulin			−0.28† [−0.30]	−0.87§ [−0.94]
logHOMA-IR			−0.57† [−0.58]	−0.93* [−0.95]
QUICKI			0.57† [0.58]	0.94* [0.96]

Data are presented as correlation coefficient [adjusted correlation], with adjustment for the measured test variability as detailed in *Materials and Methods*. Negative values indicate inverse relationships.

* $P < 0.0001$.

† $P < 0.01$.

‡ $P < 0.05$.

§ $P < 0.001$.

gree of obesity (Table 3), they were unchanged overall when examined with the groups divided by age (younger or older than 45 yr: I40, $P = 0.19$; I120, $P = 0.39$; I300, $P = 0.86$; I600, 0.60; age <45 yr, r = −0.466, age >45 yr, r = −0.878, $P = 0.04$) or gender (I40, $P = 0.13$; I120, $P = 0.20$; I300, $P = 0.39$; I600, $P = 0.32$). Similarly, there was no apparent effect of race on the correlation (I40, $P = 0.54$; I120, $P = 0.69$). The best correlations overall were seen in the diabetic subjects when compared with 600 mU/m^2·min insulin infusions, although this includes a small number of subjects. Adjustment of these correlations for the measured variability in the test, to better estimate the true underlying correlation, produced small improvements in the estimated correlation (Table 3).

In contrast to previous reports of such correlations, in lean subjects the correlations of these simple fasting estimates of insulin sensitivity with the measured GDRs under all clamp conditions were comparatively weak (Table 3). The correlations with logHOMA-IR and QUICKI were in the 0.35–0.40 range, achieving statistical significance only under I40 conditions (n = 69). Obese subjects, in contrast, exhibited very good correlations with fasting measures of insulin sensitivity (Table 3). The correlations observed at different clamp levels ranged from 0.46–0.73 (adjusted, 0.47–0.88; Table 3), although where fewer subjects were studied these correlations did not achieve statistical significance. Despite the relatively small number of subjects with type 2 diabetes mellitus included in the data set, excellent correlations of logHOMA-IR and QUICKI with GDR were observed (Table 3), with r values as high as 0.94. These subjects were all obese and had increases in both fasting insulin and glucose levels (Table 1).

The use of the ultrasensitive insulin assay in subjects with low insulin values produced a small leftward shift in the reported insulin values, particularly in those samples in which the originally reported value was less than 30 pmol/liter. The logHOMA-IR and QUICKI derived using these results gave only small improvements in the observed correlations at less than 40 mU (lean subjects: logHOMA-IR, −0.385 to −0.411; QUICKI, 0.359 to 0.389). Under I120 conditions, the correlations were not improved.

Tracking changes in insulin sensitivity

Pharmacological intervention to improve insulin sensitivity using troglitazone (n = 13) produced the changes in GDR presented in Fig. 2. The concurrent changes in fasting insulin and glucose resulted in changes in QUICKI and logHOMA-IR. The GDR (I120 clamps) increased from 5.56 ± 0.64 to 7.29 ± 0.88 mg/kg·min (31% change; $P = 0.005$), whereas the QUICKI increased from 0.308 ± 0.007 to 0.325 ± 0.007 mg/kg·min (23% change; $P < 0.001$) and logHOMA-IR decreased from 0.65 ± 0.07 to 0.50 ± 0.07 mg/kg·min (6% change; $P < 0.001$). As an expected extension of the comparable discriminant powers of these tests, in this small group of patients these changes could be distinguished equally well with the simple measures as with the hyperinsulinemic clamps.

Discussion

We have evaluated the test characteristics of various estimates of insulin sensitivity based on fasting plasma insulin and glucose and compared them with those of the EH clamp. The distribution of values and of measurement error across the range of measured values was assessed, and measures of repeatability, including the CV and the DR, were calculated and compared. The logarithmic transformations achieved the intended normalization of the insulin-dependent data. The measures of variability suggested that the logarithmic transformations (logHOMA-IR and QUICKI) were preferable to the other fasting measures, particularly with regard to the DR.

These characteristics help explain and support the recent favorable comparisons of logHOMA-IR with GDR (13), and we were able to confirm this in our own data set by examining the correlations of the various fasting measures of insulin sensitivity with GDR across a wide range of clamp conditions. The untransformed estimates had the disadvantage of nonlinear relationships in the correlations, in addition to their worse repeatability characteristics. Remarkably good correlations were observed for the logInsulin, logHOMA-IR, and QUICKI with GDR in obese and type 2 diabetic subjects, despite the small number of the latter. In contrast to previous

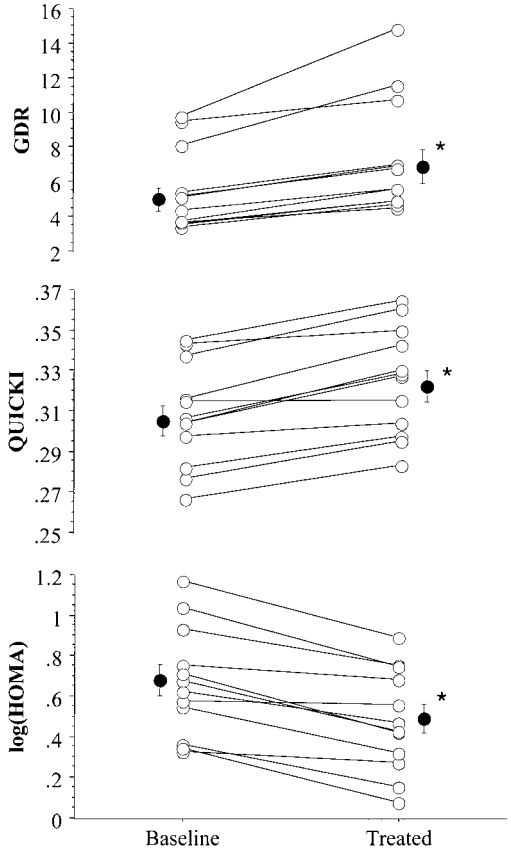

FIG. 2. Changes in GDR (*top*; 120 mU/m^2·min insulin infusion), QUICKI (*middle*), and logHOMA-IR (*bottom*) after 3 months treatment with troglitazone. *, $P < 0.05$ for comparison *vs.* baseline.

reports, we found the correlations observed among lean subjects to be inferior compared with those observed among obese and diabetic subjects. This is in many ways an expected statistical consequence of the greater variability of current insulin assays in the lower end of the range, separate from any physiological explanations. Unfortunately, the newer ultrasensitive assay did little to improve this situation, as might be expected if the variability was attributable to biological rather than test variation.

Finally, the discriminant power and correlations of logHOMA-IR and QUICKI with GDR among insulin-resistant subjects allowed the tracking of changes in insulin after treatment with troglitazone comparably to the EH clamp.

Comparing the simple measures of insulin sensitivity

The logarithmic transformations produced the desired normalization of the insulin-dependent data, including the measurement error, and compressed the range of data to reduce the effect of extreme values. This validates the application of statistical comparisons that assume underlying normal distributions. These effects in turn allowed improved linear modeling of relationships to GDR with the log-transformed measures.

Historically, the repeatability of a test has been expressed as the CV. Although conceptually simple, this measure carries disadvantages; most notably, when the mean value is close to zero, even a test with good precision may have a high CV (20). This was observed in the present report. For lean subjects, a comparatively large CV for logHOMA-IR results from a SD of 0.13, with the numerically small mean of 0.08. In fact, this measure is equally variable across the groups, as is evident in the comparable SDs (Table 2). Furthermore, the QUICKI, which is related to the logHOMA-IR by inversion and a constant term, appeared to have a superior CV, despite the inherently linked variability of these two measures. The DR, a new measure of variability that takes into account both the between-subject and within-subject variations (20), was comparable for the QUICKI and logHOMA-IR. Importantly, the DRs of these two measures and the GDR were statistically comparable. Therefore, the logHOMA-IR and QUICKI are as powerful at discriminating differences in their estimates of insulin sensitivity across the population as the GDR is at discriminating the formally measured insulin sensitivity.

Of course, the utility of these measures in estimating insulin sensitivity *per se* depends on the underlying correlation of the estimate and the formal measure. These correlations were excellent among the obese and diabetic subjects using the logarithmically transformed variables, reflecting the overall improvements in test characteristics. The comparatively poor correlations with GDR of all of these estimates in lean nondiabetic subjects suggests that they are imperfect surrogate measures in insulin-sensitive populations. Better correlations ($r = 0.47$) have been reported in nonobese subjects using QUICKI and the SI_{clamp}, although with this technique the correlations among obese and diabetic subjects were again superior ($r = 0.89$ and 0.7, respectively) (15). This finding is also in contrast to a recent report (13) of equal correlations between logHOMA-IR and GDR in lean and obese as well as diabetic and nondiabetic subsets of subjects. Importantly, the latter data set apparently includes lean diabetic subjects, who would bias toward a correlation among lean subjects. Also, those subjects were studied at insulin concentrations that did not completely suppress hepatic glucose output in all cases; therefore, they provided a measure of hepatic as well as peripheral insulin sensitivity. An analogous balance may have contributed to our finding of a significant relationship in lean subjects only at the lowest insulin infusion rate. Overall, the correlations among lean subjects appear significant but less robust than those among obese and diabetic insulin-resistant subjects.

It is tempting to argue that the logInsulin is a simpler and more accessible correlate of insulin sensitivity than either the QUICKI or the logHOMA-IR, in view of the comparable correlations between these measures among lean and obese subjects. Certainly, under any circumstances in which the glucose was completely normal, this value in effect becomes a constant term added to the equation and therefore contributes little. However, the inclusion of glucose makes the formulas more generalizable to all circumstances with variable glucose levels, including the range of subdiabetic glucose readings seen in some obese subjects. Also, in accounting for this variation, the repeatability characteristics of the test are improved, as is evident in the superior DRs of QUICKI and logHOMA-IR compared with logInsulin. These considerations support

the selection of the QUICKI or logHOMA-IR over measures based on insulin alone when concurrent glucose values are accessible.

The data set included a comparatively small number of diabetic subjects, so conclusions regarding the comparisons of test characteristics and correlations in this group must be weighted accordingly. However, these subjects represent a natural extension of the range of insulin resistance, and the results are consistent with the trend suggested by the larger samples of lean and obese subjects. Also, despite the small numbers, the correlations themselves were the strongest of all the groups, perhaps because the QUICKI and logHOMA-IR account for both the insulin and glucose levels. The calculation of the DR of necessity includes all subjects; therefore, it is strengthened by the inclusion of even this small number of subjects at one extreme of the range.

In summary, the logHOMA-IR and QUICKI were the superior simple measures of insulin sensitivity with regard to test characteristics, and this was reflected in very good correlations across a range of EH clamps among insulin-resistant obese and diabetic subjects.

Limitations of insulin assays

All of these simple fasting estimates of insulin sensitivity are highly dependent on the fasting insulin level. The original description of the HOMA included significant caveats in this regard (4). Although assays have improved, these remain imperfect tests with regard to test-to-test variability (23, 24). Furthermore, the underlying biological variability in insulin levels, arising from the combination of its short serum half-life, the known cyclicity of insulin secretion (25), and the rapid responsiveness to changes in hormonal and metabolic milieu, will remain a source of variation regardless of improvements in insulin assays. The use of a newer, ultrasensitive insulin assay did not improve any of the characteristics of the tests, presumably because it did not improve any of these sources of variability. To account for some of this variability, Matthews *et al.* (4) recommended the mean of three insulin samples, taken over a 15-min period, be used. Our data include two rather than three values, taken 10 min apart, and this produced acceptable results. Cost savings can be obtained by pooling multiple samples before measurement, but the use of at least two samples over a 10- to 15-min period remains prudent. These efforts to reduce the variability of the measurement are necessary to achieve the test characteristics described herein.

Applications

Three main categories of studies are candidates for the use of measures of insulin sensitivity: metabolic studies, intervention studies, and epidemiological studies. With metabolic studies, insulin sensitivity will most likely be a primary end point, and the use of simplified estimates is probably not worthwhile given the availability of expertise in formal EH clamp measurements, which remain the gold standard. Intervention studies targeting insulin resistance or the insulin resistance syndrome are becoming increasingly common, with the advent of PPAR-γ agonists and renewed interest in other interventions such as metformin and changes in diet and exercise. Many such studies will be primarily interested in the changes in insulin sensitivity *per se*, and again the EH clamp should be the preferred method for these studies.

The time and cost of formal clamp testing is prohibitive in large scale epidemiological studies, and to date these studies have largely chosen serum insulin as a surrogate measure of insulin sensitivity. For these studies, the logHOMA-IR or QUICKI would be appropriate surrogate measures of insulin sensitivity, as has been suggested (26). Importantly, large prospective databases already exist with sufficient data to examine associations between these simple indices of insulin resistance and cardiovascular outcomes. However, one important caveat is suggested by our findings: the superior estimate of insulin sensitivity provided among an insulin-resistant subset could systematically affect any apparent correlations with other measures in the population. This will potentially need to be accounted for in statistical analyses of such data sets.

In most interventional studies, changes in insulin sensitivity will be a secondary interest, with the effects on (for example) cardiovascular parameters being the primary end points. In these situations, the logHOMA-IR or QUICKI (using sensitive insulin-specific assays, with at least duplicate samples) would provide a useful, minimally invasive, relatively inexpensive surrogate for the EH clamp. This is further supported by the demonstrated ability of these measures to reveal clinically relevant changes in insulin sensitivity in comparatively small numbers of subjects.

Conclusion

Simple mathematical combinations of fasting insulin and glucose measures (logHOMA-IR or QUICKI) provide estimates of insulin sensitivity with variability and discriminant power comparable to those of EH clamps and superior to measures based on insulin alone. In accounting for both the glucose and insulin levels, these measures are more generalizable to the full range of metabolic conditions associated with insulin resistance. This underlies the excellent correlations of these measures with clamps seen in obese and diabetic insulin-resistant subjects. Furthermore, changes in insulin sensitivity can be demonstrated using these tools in comparatively small groups of subjects. Our results suggest that caution needs to be exercised in applying these estimates to groups including insulin-sensitive subjects. With this caveat, the current report provides the statistical underpinnings for the broader application of these inexpensive, accessible estimates of insulin sensitivity in clinical investigations and large scale epidemiological studies.

Acknowledgments

Received January 16, 2001. Accepted June 13, 2001.

Address all correspondence and requests for reprints to: Kieren J. Mather, M.D., FRCPC, Division of Endocrinology, Indiana University School of Medicine, CL 459, 541 North Clinical Drive, Indianapolis, Indiana 46202. E-mail: kmather@iupui.edu.

This work was supported by NIH Grants DK42469, DK20452, and MO1-RR750-19; and a Veterans Affairs Merit Review Award. K.J.M. was supported by an Alberta Heritage Fund for Medical Research clinical research fellowship.

References

1. **Martin B, Warram J, Krolewski A, Bergman R, Soeldner J, Kahn C** 1992 Role of glucose and insulin resistance in development of type 2 diabetes mellitus: results of a 25-year follow-up study. Lancet 340:925–929
2. **Reaven GM** 1988 Role of insulin resistance in human disease. Diabetes 37: 1595–1607
3. **DeFronzo R, Tobin J, Andres R** 1979 Glucose clamp technique: a method for quantifying insulin secretion and resistance. Am J Physiol 273:E214–E223
4. **Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC** 1985 Homeostasis model assessment: insulin resistance and B cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 28:412–419
5. **UKPDS Study Group** 1988 UK Prospective Diabetes Study V. Characteristics of newly presenting type 2 diabetic patients: estimated insulin sensitivity and islet beta-cell function. Multicentre study. Diabetes Med 5:444–448
6. **Emoto M, Nishizawa Y, Maekawa K, et al.** 1999 Homeostasis model assessment as a clinical index of insulin resistance in type 2 diabetic patients treated with sulfonylureas. Diabetes Care 22:818–822
7. **Kiechl S, Willeit J, Poewe W, et al.** 1996 Insulin sensitivity and regular alcohol consumption: large, prospective, cross sectional population study (Bruneck study). Br Med J 313:1040–1044
8. **Bonora E, Kiechl S, Willeit J, et al.** 1998 Prevalence of insulin resistance in metabolic disorders. The Bruneck Study. Diabetes 47:1643–1649
9. **Haffner S, Miettinen H, Stern M** 1997 The homeostasis model in the San Antonio Heart Study. Diabetes Care 20:1087–1092
10. **Haffner SM, Kennedy E, Gonzalez C, Stern MP, Miettinen H** 1996 A prospective analysis of the HOMA model. The Mexico City Diabetes Study. Diabetes Care 19:1138–1141
11. **Hanson RL, Pratley RE, Bogardus C, et al.** 2000 Evaluation of simple indices of insulin sensitivity and insulin secretion for use in epidemiologic studies. Am J Epidemiol 151:190–198
12. **Bergman R, Finegood D, Ader M** 1985 Assessment of insulin sensitivity in vivo. Endocr Rev 6:45–86
13. **Bonora E, Targher G, Alberiche M, et al.** 2000 Homeostasis model assessment closely mirrors the glucose clamp technique in the assessment of insulin sensitivity. Diabetes Care 23:57–63
14. **Yeni-Komshian H, Carantoni M, Abbasi F, Reaven GM** 2000 Relationship between several surrogate estimates of insulin resistance and quantification of insulin-mediated glucose disposal in 490 healthy nondiabetic volunteers. Diabetes Care 23:171–175
15. **Katz A, Sridhar SN, Albert P, Sumner AE, Quon MJ** 2000 Quantitative Insulin-Sensitivity Check Index: a simple, accurate method for assessing insulin sensitivity in humans. J Clin Endocrinol Metab 85:2402–2410
16. **Cleland S, Petrie J, Morris A, Ueda S, Dorrian C, Connell J** 1996 FIRI: a fair insulin resistance index? (Letter). Lancet 347:770
17. **Hermans MP, Levy JC, Morris RJ, Turner RC** 1999 Comparison of tests of β-cell function across a range of glucose tolerance from normal to diabetes. Diabetes 48:1779–1786
18. **Hermans MP, Levy JC, Morris RJ, Turner RC** 1999 Comparison of insulin sensitivity tests across a range of glucose tolerance from normal to diabetes. Diabetologia 42:678–687
19. **Raynaud E, Perez-Martin A, Brun J-F, Benhaddad A, Mercier J** 1999 Revised concept for the estimation of insulin sensitivity from a single sample. Diabetes Care 22:1003–1004
20. **Levy J, Morris R, Hammersley M, Turner R** 1999 Discrimination, adjusted correlation, and equivalence of imprecise tests: application to glucose tolerance. Am J Physiol 276:E365–E375
21. **Prager R, Wallace P, Olefsky JM** 1986 In vivo kinetics of insulin action on peripheral glucose disposal and hepatic glucose output in normal and obese subjects. J Clin Invest 78:472–481
22. **Celermajer DS, Sorensen KE, Gooch VM, et al.** 1992 Non-invasive detection of endothelial dysfunction in children and adults at risk of atherosclerosis. Lancet 340:1111–1115
23. **Bonora E, Moghetti P, Zancanaro C, et al.** 1989 Estimates of in vivo insulin action in man: comparison of insulin tolerance tests with euglycemic and hyperglycemic glucose clamp studies. J Clin Endocrinol Metab 68:374–378
24. **Rudenski AS, Matthews DR, Levy JC, Turner RC** 1991 Understanding "insulin resistance": both glucose resistance and insulin resistance are required to model human diabetes. Metabolism 40:908–917
25. **Matthews D, Lang D, Burnett M, Turner R** 1983 Control of pulsatile insulin secretion in man. Diabetologia 24:231–237
26. **Bastard JP, Robert JJ, Jardel C, Bruckert E, Grimaldi A, Hainque B** 2001 Is quantitative insulin sensitivity check index a fair insulin sensitivity index in humans? Diabetes Metab 27:69–70