Reproducibility of the Oral Glucose Tolerance Test in Overweight Children

I. M. Libman, E. Barinas-Mitchell, A. Bartucci, R. Robertson, and S. Arslanian

Divisions of Pediatric Endocrinology, Metabolism and Diabetes Mellitus, and Weight Management and Wellness (I.M.L., A.B., S.A.), Children's Hospital of Pittsburgh, Pittsburgh, Pennsylvania 15213; and Department of Epidemiology (E.B.-M.), Graduate School of Public Health, and Center for Exercise and Health-Fitness Research (R.R.), University of Pittsburgh, Pittsburgh, Pennsylvania 15260

Objective: We examined the reproducibility of the oral glucose tolerance test (OGTT) in overweight children and evaluated distinguishing characteristics between those with concordant *vs.* discordant results.

Design: Sixty overweight youth (8–17 yr old) completed two OGTTs (interval between tests 1–25 d). Insulin sensitivity was assessed by the surrogate measures of fasting glucose to insulin ratio, whole-body insulin sensitivity index, and homeostasis model assessment of insulin resistance, and insulin secretion by the insulinogenic index with calculation of the glucose disposition index (GDI).

Results: Of the 10 subjects with impaired glucose tolerance (IGT) during the first OGTT only three (30%) had IGT during the second OGTT. The percent positive agreement between the first and second OGTT was low for both impaired fasting glucose and IGT (22.2 and 27.3%, respectively). Fasting blood glucose had higher reproducibility, compared with the 2-h glucose. Youth with discordant OGTTs, compared with those with concordant results, were more insulin resistant (glucose/insulin 2.7 $\pm$ 1.4 *vs.* 4.1 $\pm$ 1.8, $P = 0.006$, whole-body insulin sensitivity index of 1.3 $\pm$ 0.6 *vs.* 2.2 $\pm$ 1.1, $P = 0.003$, and homeostasis model assessment of insulin resistance 10.6$\pm$ 8.1 *vs.* 5.7 $\pm$ 2.8, $P = 0.001$), had a lower GDI (0.45 $\pm$ 0.58 *vs.* 1.02 $\pm$ 1.0, $P = 0.03$), and had higher low-density lipoprotein cholesterol (117.7 $\pm$ 36.6 *vs.* 89.9 $\pm$ 20.1, $P = 0.0005$) without differences in physical characteristics.

Conclusions: Our results show poor reproducibility of the OGTT in obese youth, in particular for the 2-h plasma glucose. Obese youth who have discordant OGTT results are more insulin resistant with higher risk of developing type 2 diabetes mellitus, as evidenced by a lower GDI. The implications of this remain to be determined in clinical and research settings. (***J Clin Endocrinol Metab* 93: 4231–4237, 2008**)

The prevalence of obesity, in not only adulthood but also childhood, is increasing throughout the world at an unprecedented rate. Parallel to the increase in obesity rates, abnormalities in glucose metabolism, including impaired fasting glucose (IFG), impaired glucose tolerance (IGT), and type 2 diabetes mellitus (T2DM) in youths are on the rise (1).

In the clinical setting and to make a timely diagnosis of T2DM, the American Diabetes Association (ADA) consensus recommends screening of high-risk obese children (2). The recommended screening method is the fasting plasma glucose (FPG) because of its wide availability and ease of performance, compared with the oral glucose tolerance test (OGTT) (2). Recently it has been suggested that the OGTT may be an excellent method for reliably identifying obese children who are at high risk for diabetes (3) in addition to reliably establishing a diagnosis of IGT, in these children, because the intraperson variation is low (4). However, the International Diabetes Federation Consensus Workshop on Type 2 Diabe-

0021-972X/08/$15.00/0

Printed in U.S.A.

doi: 10.1210/jc.2008-0801 Received April 14, 2008. Accepted August 1, 2008.

First Published Online August 19,2008

For editorial see page 4228

Abbreviations: ADA, American Diabetes Association; BMI, body mass index; CVD, cardiovascular disease; FPG, fasting plasma glucose; G, glucose; GDI, glucose disposition index; $HOMA_{IR}$, homeostasis model assessment of insulin resistance; I, insulin; IFG, impaired fasting glucose; IGT, impaired glucose tolerance; LDL, low-density lipoprotein; NGT, normal glucose tolerant; OGTT, oral glucose tolerance test; T2DM, type 2 diabetes mellitus; WBISI, whole-body insulin sensitivity index.

tes in the Young states that the issue of what test to use for the initial screening of T2DM is difficult and recommends further research to determine the role of the OGTT in screening asymptomatic young people (5).

In adults, it is argued that an OGTT may be a better method because it can identify those with IGT, a risk factor for cardiovascular disease and a precursor of T2DM, who would not be picked up otherwise by a fasting glucose level (6, 7, 8). An important requirement for screening asymptomatic individuals is a test with high reproducibility. This has been shown to be one of the major drawbacks of the OGTT in adults (9–12). That is why the ADA requires a second OGTT to confirm the diagnosis of diabetes (13). Very limited conflicting information is available on the reproducibility of the OGTT in obese children. A study in Australian overweight children showed that three of four subjects with IGT, one with IFG, and one with T2DM were found to be normal glucose tolerant (NGT) on repeat testing (14). Another study in four obese children with NGT and six with IGT showed no change in classification when the OGTT was repeated 3 months later (4). These limited studies are hampered by small number of children, the use of medication (14), and a relatively long period (~3 months) between tests.

The purpose of the current investigation was to examine the reproducibility of the OGTT in overweight children and determine whether there are any differences in physical or metabolic characteristics between those with concordant *vs.* discordant test results (for NGT, IFG, or IGT).

Subjects and Methods

Subjects

Overweight [body mass index (BMI) ≥ 85th percentile for age and gender] but otherwise healthy youth, 25 African American, 23 Caucasian, 11 Hispanics, and one biracial (age range 8–17 yr), underwent assessment of glucose tolerance at the Pediatric Clinical and Translational Research Center; previously General Clinical Research Center) at the Children's Hospital of Pittsburgh. Exclusion criteria included prior diagnoses of impaired glucose regulation, diabetes or any chronic illness (such as cystic fibrosis or Cushing's), syndromes associated with obesity, and the use of medications that can influence glucose, lipid metabolism, and blood pressure. Study participants were recruited through several sources: Weight Management and Wellness Center, Pediatric Endocrinology Clinic, Primary Care Center, (all part of Children's Hospital of Pittsburgh), National Youth Sports Program (federally funded project), and newspaper advertisements in the greater Pittsburgh area.

The investigation was approved by the Institutional Review Board of the University of Pittsburgh. Parental informed consent and child assent were obtained.

All participants underwent a physical examination with height and weight measurements, BMI calculation, and assessment of pubertal development by Tanner criteria. Participants underwent two OGTTs (1.75 g/kg, maximum 75 g) with the interval between the two tests (mean ± SD) 8.7 ± 4.9 d, range 1–25 d. Blood samples were obtained at 0, 15, 30, 60, 90, and 120 min for determination of glucose, insulin and C-peptide. IFG and IGT were defined as plasma glucose between 100 and 125 mg/dl at time 0 min and between 140 and 199 mg/dl at 120 min, respectively (13). Insulin sensitivity was calculated using the whole-body insulin sensitivity index (WBISI) as described by Matsuda and DeFronzo (15), which has been shown to represent a good estimate for clamp-derived insulin sensitivity in obese children with normal and IGT (16); homeostasis model assessment of insulin resistance [$HOMA_{IR}$; calculated as fasting insulin (microunits per milliliter) × fasting glucose (mil-

TABLE 1. Subject characteristics at time of first visit

Variables	Visit 1
n	60
Sex (male/female)	25/35
Age (yr)	12.4 ± 2.3
Family history of diabetes (first degree relative) (%)	19 (32)
Family history of diabetes (any relative) (%)	54 (90)
Pubertal, Tanner II–V (%)	47 (78)
Weight (kg)	82.4 ± 27.3
Height (cm)	158.9 ± 12.4
BMI percentile	99.1 [98.1, 99.5]
BMI z score	2.3 ± 0.4
Percent body fat	44.7 ± 5.9
Fat mass (kg)	35.8 ± 12.6
Fat-free mass (kg)	43.8 ± 13.2

Data are n (%), mean ± SD, or median [interquartile range].

TABLE 2. Physiological parameters at time of first and second visits

Variables	Visit 1	Visit 2
n	60	60
Fasting glucose (mg/dl)	90 ± 7	90 ± 6
Fasting glucose (mmol/liter)	5 ± 0.4	5 ± 0.3
Fasting insulin (μU/ml)	26.7 [18.9, 37.7]	27.3 [18.4, 40.0]
Fasting C-peptide (nmol/liter)	2.7 ± 1.0	3.0 ± 1.1
Insulinogenic index at 15 min	3.4 [1.9, 5.9]	3.8 [2.4, 6.8]
Insulinogenic index at 30 min	3.2 [2.0, 5.3]	3.9 [2.2, 6.8]
C-peptide index at 15 min	0.18 ± 0.15	0.13 ± 0.16
C-peptide index at 30 min	0.14 ± 0.10	0.23 ± 0.53
Fasting glucose/insulin	3.2 [2.4, 4.6]	3.3 [2.3, 4.7]
WBISI	1.9 ± 1.0	1.9 ± 1.1
$HOMA_{IR}$	7.1 ± 5.3	6.9 ± 3.4
GDI (with insulinogenic index at 15 min)	0.86 ± 0.94	0.76 ± 0.55
GDI (with insulinogenic index at 30 min)	0.72 ± 0.65	1.21 ± 2.72
2-h glucose (mg/dl)	122 ± 19	121 ± 19
2-h glucose (mmol/liter)	6.8 ± 1.0	6.7 ± 1.0
2-h insulin (μU/ml)	132.5 [69.0, 208.1]	129.9 [85.7, 214.9]
2-h C-peptide (nmol/liter)	8.9 ± 3.2	9.4 ± 4.3
Triglycerides (mg/dl)	93 [68, 122]	
HDL (mg/dl)	42.0 ± 9.7	
LDL (mg/dl)	97.8 ± 28.5	
VLDL (mg/dl)	18.6 [13.6, 24.1]	

Data are mean ± SD or median [interquartile range]. HDL, High-density lipoprotein; VLDL, very low-density lipoprotein.

TABLE 3. Reproducibility of fasting and 2-h plasma glucose

	n	Spearman's correlation coefficient	Intraclass correlation coefficient (95% CI)	Mean absolute difference (mg/dl)
Between OGTTs (within participant)				
Fasting glucose	60	0.73 ($P<0.001$)	0.72 (0.58, 0.82)	4.3
2-h glucose	60	0.37 ($P=0.004$)	0.34 (0.14, 0.57)	16.7

CI, Confidence interval.

limoles per liter)/22.5] (17); and the fasting glucose (milligrams per deciliter) to insulin (microunits per milliliter) ratio, which has been shown to have excellent correlation with the gold standard of the euglycemic clamp (18). Insulin secretion was estimated using the insulinogenic index (the ratio of incremental insulin to glucose during the first 15 and 30 min of the OGTT [Δ insulin (I) to Δ glucose (G) = $I_{15} - I_0/G_{15} - G_0$ and $I_{30} - I_0/G_{30} - G_0$)]. Our group has shown that hyperglycemic clamp and OGTT-derived measures of insulin secretion correlate stronger for the 15-min index than for the 30-min index (19). However, other groups have shown a strong correlation with the 30-min index (4), so both indices were calculated. The glucose disposition index (GDI) was used to adjust insulin secretion for the degree of insulin resistance [insulinogenic index at 30 min to $HOMA_{IR}$ (17) and insulinogenic index at 15 min to $HOMA_{IR}$ (19)]. The same indices were evaluated using C-peptide. Insulin and glucose area under the curve during the OGTT were calculated by the trapezoidal rule. Fasting lipid profile and body composition, by dual-energy x-ray absorptiometry (Lunar, Madison, WI), were assessed at the time of the first visit.

Biochemical measurements

Plasma glucose was measured by a YSI glucose analyzer (Yellow Springs Instrument, Yellow Springs, OH) and insulin by RIA (Millipore, formerly Linco Research Inc., St. Charles, MO), which is 100% specific for human insulin with less than 0.2% cross-reactivity with human proinsulin and no cross-reactivity with C-peptide. The intra- and interassay coefficients of variation for our normal pooled EDTA plasma control are 4.3 and 9.9%, respectively. All samples for the same subject were done on a same assay. C-peptide was measured by double-antibody RIA (Siemens Health Care Diagnostics, formerly Diagnostic Products Corp., Tarrytown, NY), which is 100% specific for C-peptide. The intra- and interassay coefficients of variation for our normal pooled EDTA plasma control are 5.0 and 5.4%, respectively. Plasma lipids were measured in the Nutrition Laboratory of the University of Pittsburgh and certified by the National Heart, Lung, and Blood Institute standardization program.

Statistical analysis

Data are presented as mean ± SD for normally distributed continuous variables, median (interquartile range) for nonnormal continuous variables and n (percentage) for categorical variables. Independent *t* tests were used to compare normally distributed continuous subject characteristics as well as Pearson's χ^2 test to compare proportions. Changes in continuous variables between the two visits were assessed using paired *t* test. Nonparametric analyses were used when appropriate. Statistical analyses were performed using SPSS 15.0 (SPSS, Chicago, IL), unless otherwise indicated. All statistical tests were two tailed, and $P \leq 0.05$ were considered to be statistically significant.

Reproducibility of the OGTT results was assessed in two ways: by evaluating fasting and 2-h glucose levels as continuous variables and by examining the categorical classification of glycemic status at both OGTT visits. Reproducibility was evaluated with Spearman and intraclass correlations and by the absolute value of the difference between replicate values (20, 21). The intraclass correlation coefficient of reliability was calculated as the ratio of the variability of the OGTT measures between participants over the total variation from the various sources of error (22). A repeated-measures ANOVA (SAS software, Procedure VARCOMP, version 9.1; SAS Institute, Cary, NC) was used to partition the total variance into the components originating from between-OGTT visit differences and random error. High values of the intraclass correlation coefficient of reliability indicate greater reliability in that the measurement error is small relative to the between patient variability. Ninety-five percent confidence intervals for the intraclass correlation were calculated using a SAS Macro based on statistical methodology proposed by Shrout and Fleiss (22).

All paired records of fasting and 2-h glucose values were analyzed graphically using Bland-Altman plots according to the recommendations of Bland and Altman (20) for assessing patterns of disagreement between repeated measurements. In the Bland-Altman plots, the difference in paired records is plotted against the mean of the paired records. Ninety-five percent limits of agreement, calculated as within 2 times the SD of the mean difference between paired values, were also plotted.

Percent positive agreement, a reliability index that compensates for the limitations of percent agreement when the prevalence of a condition is low, was calculated for IFG and IGT (23). It is computed by the number of occurrences for which both visits report a positive results, divided by the average number of positives at either visit. The κ statistic, a measure of agreement for categorical measurements, was used as a measure of the extent to which agreement across categories is greater than that expected by chance. A κ value of 0 indicates chance agreement and a value of 1 indicates perfect agreement (24).

Results

The subject physical characteristics at the time of the first OGTT are presented in Table 1 (25 black-Americans, 23 Caucasians, 11 Hispanics, and one biracial). All children had one complete baseline evaluation, but the OGTT was repeated within a period of 1–25 d (8.7 ± 4.9 d, 76% of children within the first 10 d).

Between the two OGTT visits, no therapeutic intervention was implemented nor was there a change in BMI percentile or BMI z scores. Physiological parameters at the time of both visits are presented in Table 2 (except for fasting lipid profile, which was obtained just at the first visit).

Reproducibility of the fasting and 2-h plasma glucoses between the two OGTTs

The Spearman correlation coefficients, intraclass correlation coefficients, and absolute value of the difference for repeat fasting blood glucose and 2-h glucose are reported in Table 3. The fasting blood glucose had higher correlation and agreement between visits than the 2-h glucose (Fig. 1A

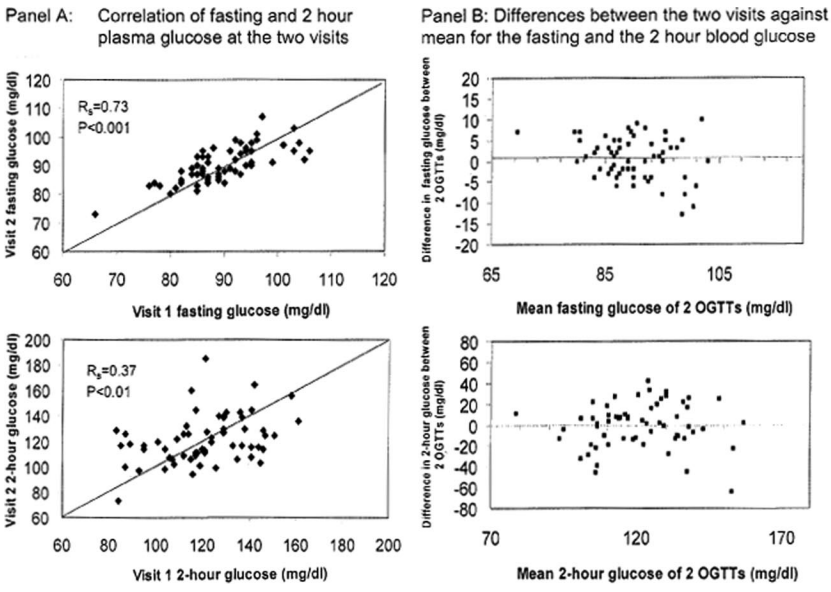

FIG. 1. Reproducibility of the fasting and 2 hour plasma glucoses between the two OGTTs.

TABLE 4. Agreement of IFG between both OGTTs

	Repeat test	
Initial test	**No**	**Yes**
No (n = 54)	52	2
Yes (n = 6)	5	1
Total (n = 60)	57	3

TABLE 6. Concordance of IFG and IGT in first OGTT

	IGT	
	No	**Yes**
IFG		
No (n = 54)	47	7
Yes (n = 6)	3	3
Total (n = 60)	50	10

and Table 3). The mean absolute difference in fasting glucose between visits was 4.3 mg/dl and ranged from 0 to 13 mg/dl. The mean absolute difference in 2-h glucose was 16.7 mg/dl and ranged from 1 to 64 mg/dl. There was no statistically significant systematic difference in fasting or 2-h glucose values between visits. The mean between visit difference (SD) in fasting and 2-h glucose values were 0.8 (5.1) and 0.7 (21.4), respectively. The lack of systematic bias was assessed by paired *t* tests of between-visit values.

Figure 1B depicts the Bland Altman plots for the fasting and 2-h blood glucoses (20). The plots show the differences between the two visits against the mean for fasting blood glucose and 2-h glucose. For the fasting blood glucose, 95% of the repeat fasting glucoses were within 11 mg/dl of the first value. For the 2-h glucoses, 95% of the repeat values were within 43 mg/dl of the first value.

Agreement of IFG and IGT between both OGTTs

Based on both OGTTs, six of 60 (10%) had IFG in the first test and three of 60 (5%) in the second test. Only one child had IFG in both tests. For IFG status, the percent positive agreement between visits was 22.2% and the κ was 0.17 ($P = 0.17$) (Table 4).

Based on both OGTTs, 10 of 60 (17%) had IGT in the first test and 12 of 60 (20%) in the second test. Only three children had IGT in both tests. For IGT status, the percent positive agreement between OGTTs was 27.3%, κ of 0.11 ($P = 0.39$) (Table 5).

Concordance of IFG and IGT in each OGTT

No children were diagnosed with diabetes. During the first OGTT, of the six with IFG, 50% had IGT, and of the 10 with IGT only, 30% had IFG (Table 6). During the second OGTT, of the three with IFG, 33% had IGT, and of the 12 with IGT, only 8% had IFG (Table 7).

Comparison between those concordant for NGT, IFG, or IGT and discordant for both OGTTs

Table 8 shows the subject characteristics by concordance status. There were no differences in physical characteristics between the groups including age, gender, Tanner staging, racial distribution, BMI, or percent body fat. There were no differences in concordance rate by source of referral: 63, 73, 71, and 73% of those referred from the Weight Management and Wellness Center, the Endocrinology Clinic, Primary Care Center, or response to advertisement, respectively, had concordant results ($P = 0.98$).

However, the discordant group was significantly more insulin resistant based on $HOMA_{IR}$, insulin area under the curve (Table 8), fasting glucose to insulin ratio, and WBISI (Fig. 2) and had higher low-density lipoprotein (LDL) and total cholesterol (Table 8). The discordant group had significantly lower insulin secretion relative to insulin sensitivity as measure by the GDI (Fig. 2). These results were consistent when using the data from the first OGTT (Fig. 2) or the average of the two OGTTs. Data were consistent whether the GDI was calculated based on 15- or 30-min insulinogenic index and using C-peptide.

Discussion

Our findings in overweight youths demonstrate that: 1) the yield of abnormalities in glucose metabolism is higher with an OGTT than fasting glucose because only 8–30% of IGT have IFG; 2) there is poor correlation and reproducibility of the 2-h plasma glucose, compared with the fasting value; and 3) those children with discordant OGTT results appear to be more at risk for diabetes, as manifested by more insulin resistance and lower insulin secretion relative to insulin sensitivity, and have worse cardiovascular disease (CVD) profile with higher total and LDL cholesterol than their peers with concordant results.

In the clinical setting and to make a timely diagnosis of T2DM, the ADA recommends screening of high-risk children with fasting plasma glucose. However, it has been argued, more so for adults, that an OGTT may be a better method because it can identify those with IGT or pre-diabetes who are not picked up by a fasting glucose (8). In obese children and adolescents, the prevalence of IGT has been reported to be 25 and 21%, respectively, similar to our study. In these youth with IGT, the preva-

TABLE 5. Agreement of IGT between both OGTTs

	Repeat test	
Initial test	**No**	**Yes**
No (n = 50)	41	9
Yes (n = 10)	7	3
Total (n = 60)	48	12

TABLE 7. Concordance of IFG and IGT in second OGTT

	IGT	
	No	**Yes**
IFG		
No (n = 57)	46	11
Yes (n = 3)	2	1
Total (n = 60)	48	12

TABLE 8. Baseline subject characteristics by concordance status

Variables	Concordant OGTT results	Discordant OGTT results	*P* value
n	43	17	
Days between two OGTTs	8.3 ± 4.8	9.4 ± 5.2	0.43
Race (AA/C/H/Biracial)	18/16/8/1	7/7/3/0	0.85
Sex (male/female)	18/25	7/10	0.59
Age (yr)	12.2 ± 2.3	12.9 ± 2.3	0.28
Pubertal, Tanner II-V (%)	33 (77)	14 (82)	0.66
Family history of diabetes (first degree relative) (%)	12 (28)	7 (41)	0.32
Family history of diabetes (any relative) (%)	40 (93)	14 (82)	0.33
BMI percentile	99.1 [98.0, 99.4]	99.1 [98.2, 99.5]	0.80
BMI z score	2.3 ± 0.3	2.3 ± 0.5	0.61
Percent body fat	44.3 ± 6.4	45.7 ± 3.8	0.44
Fat mass (kg)	34.5 ± 12.3	39.1 ± 12.9	0.19
Fat-free mass (kg)	42.9 ± 12.3	46.3 ± 15.3	0.43
Fasting glucose (mg/dl)	88.9 ± 6.3	91.2 ± 9.9	0.29
Fasting glucose (mmol/liter)	4.9 ± 0.3	5.1 ± 0.5	
Fasting insulin (μU/ml)	22.9 [18.1, 32.7]	35.5 [24.5, 53.5]	0.001
Insulinogenic index at 15 min	4.4 ± 3.6	3.9 ± 2.8	0.65
Insulinogenic index at 30 min	6.5 ± 12.7	3.0 ± 2.8	0.26
$HOMA_{IR}$	5.7 ± 2.8	10.6 ± 8.1	0.001
2-h insulin (μU/ml)	107.7 [61.8, 176.8]	189.4 [125.1, 312.5]	0.008
Area under the glucose curve	25,044.3 ± 3,030.4	29,123.4 ± 3,090.9	0.0005
Area under the insulin curve	28,179.5 ± 15,821.7	45,101.5 ± 26,337.9	0.004
Total cholesterol (mg/dl)	150.5 ± 31.1	179.2 ± 46.4	0.007
Triglycerides (mg/dl)	102.4 ± 55.1	127.4 ± 67.4	0.14
HDL (mg/dl)	42.3 ± 10.9	41.3 ± 5.9	0.73
LDL (mg/dl)	89.9 ± 20.1	117.7 ± 36.6	0.0005
VLDL (mg/dl)	20.4 ± 11.0	25.5 ± 13.5	0.14

Data are presented as n (%), mean ± SD, or median [interquartile range]. AA, African American; C, Caucasian; H, Hispanic; HDL, high-density lipoprotein; VLDL, very low-density lipoprotein.

lence of IFG (based on the former threshold of 111–125 mg/dl) was 0.08% (4).

Data from 710 Italian obese children showed that 4.2% had IGT and 0.4% had IFG, two of whom also had IGT (25). A more recent study in overweight children 4–7 yr old showed that screening with FPG alone would have missed 64% of children with glucose dysregulation, one with diabetes, and six with IGT, reporting a sensitivity of 31% for the FPG and 85% for the OGTT (26). These data and ours suggest that fasting glucose will not identify around 70% of high risk children with abnormalities of glucose metabolism detected by the 2-h glucose value. The fasting and 2-h glucose provide different measures as described by Tuomilehto (8): the first is by definition the lowest glucose level during the day, and the second shows the magnitude of glucose elevation after the glucose load. Even if there is a moderate correlation between these two parameters, they are independent to a certain extent, suggesting that neither of them can be used alone to identify people who have asymptomatic diabetes. In adults, the colinearity between these may be high, as seen in the Pima Indians, but there are some other populations, such as the lean Asians, in which more people have elevated postchallenge than fasting hyperglycemia (7, 8). Despite the limited data, this may also be the case in children.

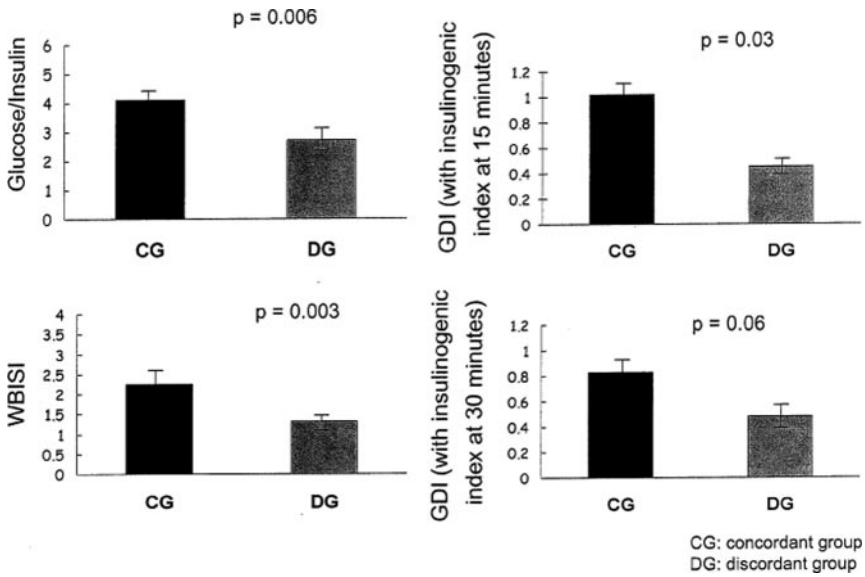

FIG. 2. G/I ratio, WBISI and GDI by concordance status.

For the last 3 to 4 decades, there has been controversy in the adult literature about the use of the OGTT because of its lack of reproducibility (6, 11). That is why the ADA requires a second test to confirm the diagnosis of diabetes. Studies in adults in whom two OGTTs were performed within 2–6 wk showed that the diagnosis of IGT was sus-

tained less than half of the time on the second OGTT (9, 10, 12). The first study demonstrated that of the total IGT subjects during the first OGTT, 48% showed IGT, 39% NGT, and 13% diabetes during the second OGTT (9). Findings were similar in the second study. Of the initial IGT subjects, 44% were IGT, 46% NGT, and 10% had diabetes on repeat testing (10). A more recent study in 64-yr-old women from Sweden showed that almost 50% that were classified as having IGT in the first OGTT had NGT in the second test (within 2 wk) (12).

Limited and conflicting information is available on the reproducibility of the OGTT in obese children. These studies are hampered by small number of children, the use of medication, and a relatively long period (~3 months) between tests. Our results in a larger group of at-risk overweight youth, on no medications, who were tested within 1–25 d, with no change in BMI between the two tests, show poor reproducibility of the OGTT. It is unlikely that this is due to variation in procedural components because all these children were tested in a carefully controlled research setting. Efforts were made to minimize all factors that could affect the results of the testing, same nurses, and physician involved, participants not being aware of results of the first test and no recommendations made in terms of changing their meal plan or activities. The correlation and reproducibility of the 2-h plasma glucose is worse than the fasting glucose. Of the 10 children with IGT during the first OGTT, more than 50% had normal results during the second test, consistent with adult data. Of those with IFG during the first test, half had normal results on the second one. Such observations would raise doubts about making a reliable diagnosis of glucose metabolism abnormalities (IFG and IGT) based on one observation, if they cannot be confirmed.

Despite this quandary, however, our data reveal information that may have important implications. Overweight youth who have discordant results during the two OGTTs show metabolic characteristics that imply higher risk for T2DM. These youth are more insulin resistant and have lower GDI, *i.e.* lower insulin secretion relative to insulin sensitivity than those with concordant results. The Botnia Study demonstrates that a lower GDI is the strongest metabolic predictor for the future development of diabetes in subjects with NGT as well as IFG or IGT (17). Could the discordant OGTT results in a group with lower GDI herald the ultimate progression to permanent abnormalities?

Moreover, is it possible that the β-cell in this group may be more susceptible to acute environmental modulation, nutritionally driven or otherwise, which may result in variable insulin secretion leading to inconsistent results during the OGTT? Also, this group has worse CVD profile with elevated LDL, pointing to a potentially higher risk for CVD and the metabolic syndrome. Do these observations imply that this may be a more at risk group of overweight children who require more aggressive strategies for improvement of insulin resistance, preservation of β-cell function, and control of CVD risk? All these questions are awaiting answers.

In conclusion, from a clinical standpoint, overweight children with an abnormal OGTT should be followed up closely, and repeat OGTT should be entertained, especially if interventions are not successful. From a research standpoint, long-term follow-up of at-risk youth who are found to have IFG or IGT should be conducted to determine the risk of progression to T2DM and future CVD. Finally, it remains to be determined whether the natural history of glucose tolerance is different between youths with discordant *vs.* concordant OGTT results.

Acknowledgments

The authors express their gratitude to the study participants and their parents, the Pediatric Clinical and Translational Research Center nurses, and Resa Stauffer for technical assistance.

Address all correspondence and requests for reprints to: Ingrid M. Libman, M.D., Ph.D., Children's Hospital of Pittsburgh, 3705 Fifth Avenue, 4th A De Soto Wing, Pittsburgh, Pennsylvania 15213. E-mail: ingrid.libman@chp.edu.

This work was supported by Grant M2004-0043 from The Pittsburgh Foundation (to I.M.L.), General Clinical Research Center Grant 5M01 RR00084 and Pediatric Clinical and Translational Research Center Grants UL1 RR024153-01, 2K24-HD-01357 (to S.A.), and K12 DK063704.

Disclosure Statement: The authors have nothing to disclose.

References

1. **Gungor N, Hannon T, Libman I, Bacha F, Arslanian S** 2005 Type 2 diabetes mellitus in youth: the complete picture to date. Pediatr Clin North Am 52: 1579–1609
2. **American Diabetes Association** 2000 Type 2 diabetes in children and adolescents. Diabetes Care 23:381–389
3. **Rocchini A** 2002 Childhood obesity and a diabetes epidemic. N Engl J Med 346:854–855
4. **Sinha R, Fisch G, Teague B, Tamborlane WV, Banyas B, Allen K, Savoye M, Rieger V, Taksali S, Barbetta G, Sherwin RS, Caprio S** 2002 Prevalence of impaired glucose tolerance among children adn adolescents with marked obesity. N Engl J Med 346:802–810
5. **The International Diabetes Federation Consensus Workshop** 2004 Type 2 diabetes in the young: the evolving epidemic. Diabetes Care 27:1794–1811
6. **Barrett-Connor E** 2002 The oral glucose tolerance test—revisited. Eur Heart J 23:1229–1231
7. **Qiao Q, Pyorala K, Pyorala M, Nissinen A, Lindstrom J, Tilvis R, Tuomilehto J** 2002 Two-hour glucose is a better risk predictor for incident coronary heart disease and cardiovascular mortality than fasting glucose. Eur Heart J 23: 1267–1275
8. **Tuomilehto J** 2002 Point: a glucose tolerance test is important for clinical practice. Diabetes Care 25:1880–1882
9. **Mooy JM, Grootenhuis PA, De Vries H, Kostense PJ, Pop-Snijders C, Bouter LM, Heine RJ** 1996 Intra-individual variation of glucose, specific insulin and proinsulin concentrations measured by two oral glucose tolerance tests in a general Caucasian population: the Hoorn Study. Diabetologia 39:298–305
10. **Ko GTC, Cahn JCN, Woo J, Lau E, Yeung BVT, Chow CC, Cockram CS** 1998 The reproducibility and usefulness of the oral glucose tolerance test in screening for diabetes and other cardiovascular risk factors. Ann Clin Biochem 35: 62–67
11. **Davidson MB** 2002 Counterpoint: the oral glucose tolerance test is superfluous. Diabetes Care 25:1883–1885
12. **Brohall G, Behre CJ, Hulthe J, Wikstrand J, Fagerberg B** 2006 Prevalence of diabetes and impaired glucose tolerance in 64-year old Swedish women. Experiences using repeated oral glucose tolerance tests. Diabetes Care 29: 363–367
13. **American Diabetes Association** 2007 Diagnosis and classification of diabetes mellitus. Diabetes Care 30:s42–s47
14. **Conwell LS, Batch JA** 2004 Oral glucose tolerance test in children and adolescents: positives and pitfalls. J Paediatr Child Health 40:620–626
15. **Matsuda F, DeFronzo R** 1999 Insulin sensitivity indices obtained from oral

glucose tolerance testing. Comparison with the euglycemic insulin clamp. Diabetes Care 22:1462–1470
16. **Yeckel C, Weiss R, Dziura J, Taksali SE, Dufour S, Burgert TS, Tamborlane WV, Caprio S** 2004 Validation of insulin sensitivity indices from oral glucose tolerance test parameters in obese children and adolescents. J Clin Endocrinal Metab 89:1096–1101
17. **Lyssenko V, Almgren P, Anevski D, Perfekt R, Lahti K, Nissen M, Isomaa B, Forsen B, Homstrom N, Saloranta C, Taskinen MR, Groop L, Tuomi T; Botnia study group** 2005 Predictors of and longitudinal changes in insulin sensitivity and secretion preceding onset of type 2 diabetes. Diabetes 54: 166–174
18. **Gungor N, Saad R, Janosky J, Arslanian S** 2004 Validation of surrogate estimates of insulin sensitivity and insulin secretion in children and adolescents. J Pediatr 144:47–55
19. **Bacha F, Gungor N, Arslanian S** 2008 Measures of β-cell function during oral glucose tolerance test and liquid mixed meal in youth: how well do they correlate with the gold standard of the hyperglycemic clamp? J Pediatr 152:618–621
20. **Bland JM, Altman DG** 1986 Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1:307–310
21. **Rothwell PM** 2000 Analysis of agreement between measurements of continuous variables: general principles and lessons from studies of imaging of carotid stenosis. J Neurol 415:825–834
22. **Shrout PE, Fleiss JL** 1979 Intraclass correlations uses in assessing rater reliability. Psychol Bull 2:420–428
23. **Szklo MS, Nieto FJ, eds 2007.** Epidemiology beyond the basics. 2nd ed. Chap 8. Sudbury, MA: Jones and Bartlett Publishers
24. **Lachin JM** 2004 The role of measurement reliability in clinical trials. Clin Trials 1:553–566
25. **Invitti C, Guzzaloni G, Gilardini L, Morabito F, Viberti G** 2003 Prevalence and concomitants of glucose intolerance in European obese children and adolescents. Diabetes Care 26:118–124
26. **Ehtisham S, Shaw N, Kirk J, Barrett T** 2004 Development of an assessment tool for screening children for glucose intolerance by oral glucose tolerance test. Diabetes Care 27:280–281