Predicting Impaired Glucose Tolerance Using Common Clinical Information

Data from the Third National Health and Nutrition Examination Survey

KARIN M. NELSON, MD, MSHS[1,2]
EDWARD J. BOYKO, MD, MPH[1,2,3]

OBJECTIVE — To develop a score to predict impaired glucose tolerance (IGT) using common clinical data.

RESEARCH DESIGN AND METHODS — We analyzed data from the Third National Health and Nutrition Examination Survey (NHANES III) for 2,746 individuals aged 40–74 years who completed an oral glucose tolerance test. IGT was defined as a 2-h postchallenge glucose ≥140 mg/dl (7.7 mmol/l). We performed bivariate and multivariate analyses to describe the association of IGT with commonly available clinical information. A numerical score to predict IGT was derived from the results of the multivariate logistic regression models.

RESULTS — Fasting glucose levels between 101 and 109 mg/dl (5.6 and 6.0 mmol/l) or between 110 and 125 mg/dl (6.1 and 6.9 mmol/l) were associated with IGT (odds ratio 1.8 and 6.2, respectively; $P < 0.05$). BMI ≥25 kg/m^2, Mexican-American ethnicity, age between 60 and 74 years, hypertension, and triglyceride level ≥150 mg/dl (1.69 mmol/l) were also associated with IGT. The area under the receiver operating characteristic curve for an 8-point scale derived from the multivariate analysis was 0.74 (95% CI 0.72–0.76). Setting a low cut point of 2 on this scale resulted in high sensitivity (86%), whereas a high cut point of 6 yielded high specificity (97%) for the detection of IGT.

CONCLUSIONS — A numerical score based on common clinical data can identify individuals with a low or high likelihood of having IGT.

Diabetes Care 26:2058–2062, 2003

Impaired glucose tolerance (IGT) is an important intermediate stage in the natural history of type 2 diabetes (1) and has been associated with an increased risk for cardiovascular disease and mortality (2–4). Recent clinical trials report that lifestyle modifications or pharmacotherapy among individuals with IGT can reduce their risk of developing diabetes (5,6). The American Diabetes Association recommends screening for IGT or impaired fasting glucose in men and women age ≥45 years, particularly in those who are overweight or obese (7). An oral glucose tolerance test (OGTT) is needed to diagnose IGT, which is defined as a postchallenge glucose value between 140 and 200 mg/dl (7.7 and 11.1 mmol/l) and the absence of a fasting glucose in the diabetic range (8).

In the clinical setting, OGTTs are more costly than fasting or random blood measurements, are burdensome to administer, and are not commonly performed. Using only an elevated fasting plasma glucose to identify individuals at high risk for glucose intolerance is inadequate and will miss a substantial proportion of individuals with IGT (8,9). Some authors advocate screening high-risk populations with OGTTs, regardless of their fasting glucose level, whereas others do not (10–13). The purpose of this study was to examine the association of commonly available clinical characteristics known to be associated with IGT using a nationally representative U.S. dataset and to derive a numerical score based on readily available clinical information to predict presence of IGT.

RESEARCH DESIGN AND METHODS

Study population

The Third National Health and Nutrition Examination Survey (NHANES III) was conducted by the National Center for Health Statistics at 89 U.S. survey locations between 1988 and 1994. The survey used a stratified multistage probability cluster design with oversampling of Mexican Americans, African Americans, and the elderly. The survey consisted of multiple components including a household interview, a physical examination, and laboratory tests. Descriptions of the standardized protocols used for all interview, laboratory, and physical examinations have been previously published (14).

We analyzed data from 2,746 individuals between the ages of 40 and 74 years who had completed a 2-h OGTT. A detailed description of the OGTT methodology has been previously published (8). We excluded individuals who reported a previous diagnosis of diabetes and those who had a fasting plasma glucose ≥126 mg/dl (7.0 mmol/l) (14). For the current study, IGT was defined as a 2-h postchallenge glucose ≥140 mg/dl

From the [1]Primary and Specialty Medical Care Service, VA Puget Sound Health Care System, Seattle, Washington; the [2]Department of Medicine, University of Washington, Seattle, Washington; and the [3]Epidemiologic Research and Information Center, VA Puget Sound Health Care System, Seattle, Washington.

Address correspondence and reprint requests to Karin M. Nelson, MD, MSHS, VA Puget Sound Health Care System, 1660 South Columbian Way, S-111-GIMC, Seattle, WA 98108. E-mail: karin.nelson@med.va.gov.

Received for publication 19 August 2002 and accepted in revised form 17 December 2002.

Abbreviations: IGT, impaired glucose tolerance; NHANES III, Third National Health and Nutrition Examination Survey; OGTT, oral glucose tolerance test; OR, odds ratio; ROC, receiver operating characteristic.

A table elsewhere in this issue shows conventional and Système International (SI) units and conversion factors for many substances.

(7.7 mmol/l). We included 128 individuals with a 2-h postchallenge glucose ≥200 mg/dl (11.1 mmol/l) and a fasting plasma glucose <126 mg/dl (7.0 mmol/l). By definition, these individuals have diabetes and would require further clinical intervention (15). Even though the outcome as we have defined it includes individuals with diabetes diagnosed only by the OGTT, we will refer to the outcome as "IGT" throughout the rest of the article.

Data were collected on other known risk factors for IGT and diabetes (15). Standing height and weight were obtained during the physical examination and were used to calculate BMI. Individuals were considered obese if their BMI was ≥30 kg/m^2 and were considered overweight if their BMI was between 25 and 29.9 kg/m^2 (16). Individuals were considered to have a family history of diabetes if they reported that any first-degree relative had diabetes. Criteria for the metabolic syndrome included abdominal obesity (waist circumference >102 cm for men and >88 cm for women), a triglyceride level ≥150 mg/dl (1.69 mmol/l), and a serum HDL level <40 mg/dl (1.04 mmol/l) for men and <50 mg/dl (1.29 mmol/l) for women (17,18). Individuals were considered to have hypertension if they had a measured blood pressure reading ≥140/90 mmHg or they reported being told by their physician that they had high blood pressure.

Table 1—*Population characteristics and bivariate correlates of IGT in adults aged 40–74 years in NHANES III*

	Total	2-h postchallenge glucose (mg/dl)	
		≥140	<140
n (%)	2,746 (100)	686 (21)	2,060 (79)
Female	1,410 (53)	53 (48–59)	53 (50–55)
Age (years)*			
40–49	945 (40)	25 (20–31)	43 (39–48)
50–59	672 (27)	25 (19–31)	28 (25–31)
60–74	1,129 (33)	50 (44–57)	28 (25–31)
Race/ethnicity†			
White	1,354 (88)	87 (84–87)	88 (86–89)
African American	654 (9)	8 (6–10)	9 (8–11)
Mexican American	618 (3)	5 (4–7)	3 (3–4)
Fasting glucose [mg/dl (mmol/l)]*			
<90 (<4.9)	514 (21)	10 (7–13)	24 (22–27)
90–100 (4.9–5.5)	1,082 (41)	26 (22–31)	45 (42–26)
101–109 (5.6–6.0)	758 (25)	31 (26–37)	24 (21–26)
110–125 (6.1–6.9)	392 (12)	32 (27–38)	7 (5–8)
BMI (kg/m^2)*			
<25	926 (39)	29 (24–35)	41 (38–44)
25–29.9	1,060 (37)	36 (30–42)	38 (35–41)
≥30	758 (24)	35 (31–39)	21 (19–24)
Family history of diabetes	790 (27)	31 (26–37)	26 (24–29)
Hypertension*‡	1,028 (35)	50 (43–56)	31 (28–34)
Triglyceride level ≥150 mg/dl (1.69 mmol/l)*	951 (35)	53 (46–60)	30 (28–33)
Low HDL*§	996 (37)	47 (40–53)	35 (31–39)
Abdominal obesity*‖	1,398 (47)	60 (54–65)	44 (40–47)

Data are *n* (%) or % (95% CI). Percentages are weighted to account for complex survey design. Column totals may vary because of missing data and rounding error. Pearson's χ^2 test: *$P < 0.001$, †$P < 0.05$. ‡≥140/90 mmHg or reported diagnosis of hypertension. §<50 mg/dl (1.29 mmol/l) for women, <40 mg/dl (1.04 mmol/l) for men. ‖Waist circumference: >88 cm for women, >102 cm for men.

Statistical analysis

Data were weighted to account for the unequal probability of selection that resulted from the survey cluster design, nonresponse, and oversampling of certain target populations (19,20). Sampling weights were used to calculate population estimates, and sampling strata and primary sampling units were accounted for to estimate variances and test for significant differences. Statistical analysis was performed using Stata 7.0 software (21,22) to take into account the complex sampling design. All results are presented as unweighted counts (*n*) and weighted percentages and odds ratios (ORs). Pearson's χ^2 test was used to determine bivariate associations between socioeconomic and health characteristics of IGT.

We developed two multivariate models with the presence of IGT as the main dependent variable. The independent variables in the multivariate models included clinical characteristics associated with IGT in bivariate analysis or from previously published literature (8,11). The first model will be referred to as the "clinical model" and includes those data most likely to be available at a routine clinical encounter, including sex, age, race/ethnicity, family history of diabetes, BMI, fasting glucose, and blood pressure. The second model will be referred to as the "full model" and includes these covariates in addition to abdominal obesity and additional laboratory tests (triglyceride and HDL level). All continuous independent variables were entered into models in categorical form.

For the full multivariate model, we assessed goodness of fit with the Hosmer-Lemeshow test (22). We tested interaction terms between age, sex, and race/ethnicity. The risk of IGT predicted for each person by the full multivariate model was used to define the area under the receiver operating characteristic (ROC) curve. To assess the external validity of the full multivariate model, we used an internal bootstrap technique (23) to estimate the area under the ROC curve for the predicted probability of IGT. We also tested the full multivariate model among a subsample of individuals with a BMI ≥24 kg/m^2 and a fasting blood glucose between 95 and 125 mg/dl, entry criteria for two recent diabetes prevention trials (5,6).

We created a simplified 8-point score by replacing significant multivariate model coefficients ($P < 0.05$) with integer values (24) as follows: a value of 1 for a significant multivariate OR between 1.1 and 1.9, 2 for an OR between 2.0 and 2.9, and 3 for an OR of ≥3. An ROC curve was used to describe the diagnostic properties of the 8-point scale.

The ROC curve plots the sensitivity as

Table 2—*Multivariate logistic regression models of IGT in adults aged 40–74 years in NHANES III*

	Clinical model	Full model
Female	1.3 (0.9–1.8)	1.2 (0.8–1.6)
Race/ethnicity		
White	Reference	Reference
African American	0.8 (0.6–1.0)	0.9 (0.7–1.3)
Mexican American	1.7 (1.3–2.2)*	1.7 (1.3–2.3)†
Age (years)		
40–49	Reference	Reference
50–59	1.5 (1.0–2.1)	1.4 (1.0–2.1)
60–74	2.7 (1.9–3.7)*	2.6 (1.8–3.7)*
Fasting glucose [mg/dl (mmol/l)]		
<90 (<4.9)	Reference	Reference
90–100 (4.9–5.5)	0.9 (0.6–1.2)	0.9 (0.6–1.2)
101–109 (5.6–6.0)	1.8 (1.2–2.6)†	1.6 (1.1–2.5)†
110–125 (6.1–6.9)	6.2 (3.5–10.9)*	6.3 (3.5–11.1)*
BMI (kg/m^2)		
<25	Reference	Reference
25–29.9	1.1 (0.8–1.5)	0.8 (0.6–1.8)
≥30	1.6 (1.1–2.2)†	1.1 (0.7–1.2)
Family history of diabetes	1.2 (0.9–1.6)	1.2 (0.8–1.6)
Hypertension‡	1.7 (1.2–2.3)†	1.5 (1.1–2.1)†
Triglyceride ≥150 mg/dl (1.69 mmol/l)	—	1.9 (1.4–2.8)*
Low HDL§	—	1.2 (0.9–1.6)
Abdominal obesity‖	—	1.2 (0.7–1.7)

Data are ORs (95% CI). ORs are adjusted for all listed covariables. OR >1 indicates the increased probability of having an abnormal 2-h OGTT. *$P < 0.001$; †$P < 0.05$. ‡≥140/90 mm Hg or reported diagnosis of hypertension. §<50 mg/dl (1.29 mmol/l) for women, <40 mg/dl (1.04 mmol/l) for men. ‖Waist circumference >88 cm for women, >102 cm for men.

a function of 1 − specificity (the false-positive rate). In this study, cut points of the scale score were examined. The sensitivity refers to the percentage of individuals with IGT whose score was above the cut point. If the sensitivity increases as the cut point increases, with a disproportionately smaller increase in the false-positive rate, the area under the ROC curve will be large. The ROC curve area is smaller if false-positive cases accumulate more rapidly as the cut point increases. The area under the ROC curve represents the probability that a subject chosen at random from a group of people with IGT had a higher point value than an individual without IGT. Likelihood ratios were calculated for each cut point of the score. The likelihood ratio for a positive test result is the ratio of the probability of a positive test result among the truly positive subjects to the probability of the positive test result among truly negative subjects. Similarly, the likelihood ratio for a negative test result is the ratio of the probability of a negative test result among the truly positive subjects to the probability of a negative test result among the truly negative subjects.

Analyses of the area under the ROC curves were used to assess the full multivariate model and the 8-point score. We calculated the SE of the area under the ROC curve and compared areas under the ROC curve (22). Comparisons of the area under the ROC curve were examined for the full model, the use of fasting glucose alone to predict IGT, and the application of the model to a subpopulation eligible for diabetes prevention trials (5,6).

RESULTS — The population characteristics of the sample are displayed in Table 1. The majority of the sample was white, and 53% were women. Over one-third of the participants had high blood pressure, an elevated serum triglyceride level, or a low serum HDL level. The majority of the sample was either overweight or obese, and almost half of the respondents had abdominal obesity. Of the sample, 21% had IGT ($n = 686$). Bivariate correlates of IGT are displayed in Table 1 and include increasing age, Mexican-American ethnicity, and impaired fasting glucose. Characteristics of the metabolic syndrome, including obesity, hypertension, elevated triglycerides, and low serum HDL, were also associated with IGT. An individual's sex or a family history of diabetes was not significantly associated with IGT.

Table 2 presents the results from multivariate logistic regression models predicting IGT (F statistic for clinical and full model, $P < 0.0000$). In the clinical model, using the most commonly available clinical variables, the strongest association with IGT is impaired fasting glucose between 110 and 125 mg/dl (6.1 and 6.9 mmol/l) (OR 6.2, 95% CI 3.5–10.9). Other significant associations include Mexican-American ethnicity, age between 60 and 74 years, obesity, and hypertension. When other characteristics of the metabolic syndrome, including elevated triglycerides, low HDL, and abdominal obesity, are added in the full model, obesity is no longer independently associated with IGT. Interactions between age, sex, and race/ethnicity were not found to be significantly associated with the presence of IGT (data not shown). The area under the ROC curve for the full multivariate model predicting IGT was 0.76 (95% CI 0.72–0.78). The difference in the area under the ROC curve obtained by the bootstrap procedure compared with the original sample was 0.008. The Hosmer-Lemeshow test did not reject the goodness of fit of the full model ($P = 0.11$). The full multivariate model had similar predictive ability for the subsample of individuals with a fasting blood glucose between 95 and 125 mg/dl and a BMI ≥24 kg/m^2 (area under the ROC curve 0.73 [95% CI 0.71–0.75], $P = 0.06$ compared with the entire study sample). When we excluded the 128 individuals with a 2-h OGTT value of ≥200 mg/dl from the analysis, no difference was noted in the predictive ability of the full model (area under the ROC curve 0.76 [95% CI 0.73–0.77], $P = 0.13$ compared with the entire study sample).

Table 3 presents a scoring system based on the results of the multivariate regression analyses. The highest possible score is 8. Because obesity was significant in the first model and not in the full model after the addition of triglyceride level, two scores were created: one using the triglyceride level and one using obesity. Table 4 presents the test characteristics of this

Table 3—*Clinical score*

Clinical characteristic	Points
Fasting glucose (mg/dl)	
100–110	1
110–125	3
Aged between 60 and 74 years	2
Triglyceride ≥150 mg/dl OR obesity (BMI ≥30 kg/m^2)*	1
Mexican-American ethnicity	1
Hypertension	1
Maximum score	8

*Obesity can be used if triglyceride value is not available.

scoring system. The area under the ROC curve is 0.74 (95% CI 0.72–0.76), which is significantly larger than the area under the ROC curve for fasting glucose level alone (0.69 [95% CI 0.67–0.72], $P < 0.002$). The area under the ROC curve was similar when obesity was used instead of triglyceride value to calculate the score (data not shown).

Table 4 highlights two cut points in the score that may be clinically useful. When a low cut point is set (for example, if the test is considered positive with a score ≥2), the sensitivity of the score is high. A total of 86% of individuals *with* IGT had a score ≥2. Forty-eight percent of the sample had a score <2. When the cut point is set high (for example, if the test is considered positive with a score of ≥6), the specificity is high. Of those *without* IGT, 97% had a score <6. The positive likelihood ratio of 6.0 for a score of ≥6 and the negative likelihood ratio of 0.2 for a score <2 both generate a moderate shift in the pretest to posttest probability of IGT (25).

CONCLUSIONS — We present a simple 8-point score based on commonly available clinical data that accurately predicts individuals at low and high risk of having IGT. This scale performs significantly better than fasting glucose level alone in identifying individuals with IGT. When a low cut point of 2 is used to define a positive test, the sensitivity of the test is high (86%). Thus, individuals with a score <2 have a low likelihood of having the disease. Further oral glucose tolerance testing may not be necessary in this low-risk group, which comprises almost half of this nationally representative population of individuals aged 40–74 years.

When a higher cut point of 6 is used to define a positive test, the specificity of the test increases and most individuals (97%) *without* IGT have a score of <6. Thus, individuals with a score ≥6 have a high likelihood of having IGT. Targeting individuals with scores ≥6 for OGTTs may help identify individuals at high risk for IGT and diabetes among individuals with fasting plasma glucose levels in the nondiabetic range. An intermediate score between 2 and 6 will not help clinicians determine the risk of IGT, but lifestyle interventions should be targeted in individuals with low HDL, high triglycerides, and obesity, independent of IGT (18,26). The outcome of the prediction model included 128 subjects with diabetes based on a 2-h glucose of ≥200 mg/dl (11.1 mmol/l) and a fasting glucose in the nondiabetic range. Although diabetes and IGT are distinct categories in terms of glucose tolerance classification, the identification of either may lead to similar clinical interventions.

The strength of this study is the use of nationally representative data (8). Our results add important evidence for screening individuals at high risk of IGT based on commonly available clinical data. Previous authors have not advocated using clinical data to predict IGT (11). In a population-based study conducted in Sweden, a high-risk screening strategy based on obesity and family history did not detect the majority of individuals with IGT (11). However, this study had a lower rate of obesity than in the U.S., did not use triglyceride or HDL levels in analyses, and did not combine multiple risk factors into a numerical score. Limitations of our study include the use of self-reported data of socioeconomic and health characteristics that are subject to recall and other biases. In addition, NHANES III only allows for population estimates among white, African-American, and Mexican-American populations and does not have adequate sample size to make estimates about Asian- or Native-American populations.

Recent clinical trials demonstrated that diabetes can be prevented by behavioral and pharmacological interventions among high-risk populations (5,6). Individuals were included in these clinical trials if they had an abnormal glucose tolerance, defined by a 2-h response to a fixed glucose load. Because of cost and inconvenience, the OGTT is rarely used in clinical practice. We present a simple score based on commonly available clinical data to differentiate individuals at both low and high risk for the presence of IGT.

Table 4—*Score to predict IGT*

Cut point	Proportion with score (%)*	Sensitivity (%)	Specificity (%)	LR$^+$	LR$^-$	Positive predictive value (%)	Possible clinical action
0	27	100	0	1.0		6	↑ No OGTT
1	21	96	22	1.2	0.2	13	
2	19	86	44	1.5	0.3	19	
3	15	71	65	2.0	0.4	29	
4	10	50	82	2.8	0.6	38	
5	4	31	92	3.9	0.8	46	
6	2	18	97	6.0	0.8	59	↓ OGTT
7	2	9	99	9.3	0.9	76	
8	0.02	1	100	7.0	1.0	70	

ROC 0.74 (95% CI 0.72–0.76). LR, likelihood ratio. *Does not add up to 100% because of rounding error.

Acknowledgments— Information in this article has previously appeared in abstract form and was presented at the American Diabetes Association Annual Meeting, New Orleans, Louisiana, 13–17 June 2003.

References

1. Sinha R, Fisch G, Teague B, Tamborlane WV, Banyas B, Allen K, Savoye M, Rieger V, Taksali S, Barbetta G, Sherwin RS, Caprio S: Prevalence of impaired glucose tolerance among children and adolescents with marked obesity. *N Engl J Med* 346: 802–810, 2002
2. Barrett-Connor E, Ferrara A: Isolated postchallenge hyperglycemia and the risk of fatal cardiovascular disease in older women and men: the Rancho Bernardo Study. *Diabetes Care* 21:1236–1239, 1998
3. Shaw JE, Hodge AM, de Courten M, Chitson P, Zimmet PZ: Isolated post-challenge

hyperglycaemia confirmed as a risk factor for mortality. *Diabetologia* 42:1050–1054, 1999
4. Tominaga M, Eguchi H, Manaka H, Igarashi K, Kato T, Sekikawa A: Impaired glucose tolerance is a risk factor for cardiovascular disease, but not impaired fasting glucose: the Funagata Diabetes Study. *Diabetes Care* 22:920–924, 1999
5. Tuomilehto J, Lindstrom J, Eriksson JG, Valle TT, Hamalainen H, Ilanne-Parikka P, Keinanen-Kiukaanniemi S, Laakso M, Louheranta A, Rastas M, Salminen V, Uusitupa M, the Finnish Diabetes Prevention Study Group: Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N Engl J Med* 344:1343–1350, 2001
6. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Nathan DM, the Diabetes Prevention Program Research Group: Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 346:393–403, 2002
7. American Diabetes Association and National Institute of Diabetes, Digestive and Kidney Diseases: The prevention or delay of type 2 diabetes. *Diabetes Care* 25:742–749, 2002
8. Harris MI, Flegal KM, Cowie CC, Eberhardt MS, Goldstein DE, Little RR, Wiedmeyer HM, Byrd-Holt DD: Prevalence of diabetes, impaired fasting glucose, and impaired glucose tolerance in U.S. adults: the Third National Health and Nutrition Examination Survey 1988–1994. *Diabetes Care* 21:518–524, 1998
9. Drzewoski J, Czupryniak L: Concordance between fasting and 2-h post-glucose challenge criteria for the diagnosis of diabetes mellitus and glucose intolerance in high risk individuals. *Diabet Med* 18:29–31, 2001
10. Vaccaro O, Ruffa G, Riccardi G: Is there any use for the oral glucose tolerance test? *Diabetes Care* 23:714–715, 2000
11. Lindahl B, Weinehall L, Asplund K, Hallmans G: Screening for impaired glucose tolerance: results from a population-based study in 21,057 individuals. *Diabetes Care* 22:1988–1992, 1999
12. Unwin N, Shaw J, Zimmet P, Alberti K: Impaired glucose tolerance and impaired fasting glycaemia: the current status on definition and intervention. *Diabet Med* 19:708–723, 2002
13. Hunt KJ, Williams K, Haffner SM, Stern MP: Predicting impaired glucose tolerance (IGT) among individuals with a non-diabetic fasting glucose value: the San Antonio Heart Study (Abstract). *Diabetes* 51 (Suppl. 2):A229, 2002
14. Plan and operation of the Third National Health and Nutrition Examination Survey, 1988–94. *Vital Health Stat* 1:1–62, 1994
15. Expert Committee on the Diagnosis and Classification of Diabetes Mellitus: Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care* 20:1183–1197, 1997
16. Maggio CA, Pi-Sunyer FX: The prevention and treatment of obesity: application to type 2 diabetes. *Diabetes Care* 20:1744–1766, 1997
17. Ford ES, Giles WH, Dietz WH: Prevalence of the metabolic syndrome among US adults: findings from the Third National Health and Nutrition Examination Survey. *JAMA* 287:356–359, 2002
18. Executive Summary of the Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). *JAMA* 285:2486–2497, 2001
19. *National Health and Nutrition Examination Survey III: Weighting and Estimation Methodology*. Rockville, MD, Westat, 1996, p. 1–29
20. Korn EL, Graubard BI: Epidemiologic studies utilizing surveys: accounting for the sampling design. *Am J Public Health* 81:1166–1173, 1991
21. Stata Corporation: *Version 7.0*. College Station, TX, Stata Corporation, 2001
22. Stata Technical Bulletin: *Receiver Operating Characteristic (ROC) Analysis*. College Station, TX, Stata Corporation, 1999 (Stata Technical Bulletin 52)
23. Harrell FE, Lee KL, Mark DB: Multivariate prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors (Review). *Stat Med* 15:361–387, 1996
24. Selby JV, Karter AJ, Ackerson LM, Ferrara A, Liu J: Developing a prediction rule from automated clinical databases to identify high-risk patients in a large population with diabetes. *Diabetes Care* 24: 1547–1555, 2001
25. Jaeschke R, Guyatt GH, Sackett DL: Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA* 271:703–707, 1994
26. National Task Force on the Prevention and Treatment of Obesity: Overweight, obesity, and health risk. *Arch Intern Med* 160:898–904, 2000