Performance of Recommended Screening Tests for Undiagnosed Diabetes and Dysglycemia

DEBORAH B. ROLKA, MS[1]
K. M. VENKAT NARAYAN, MD[1]
THEODORE J. THOMPSON, MS[1]
DONA GOLDMAN, BSN, MPH[2]
JOANN LINDENMAYER, DVM, MPH[2]
KATE ALICH, MS, RD[3]
DARCY BACALL, RN, CDE[3]
EVAN M. BENJAMIN, MD[4]
BETTY LAMB, RN, MSN[1]
DENNIS O. STUART, MD[5]
MICHAEL M. ENGELGAU, MD[1]

OBJECTIVE — To evaluate the performance, in settings typical of opportunistic and community screening programs, of screening tests currently recommended by the American Diabetes Association (ADA) for detecting undiagnosed diabetes.

RESEARCH DESIGN AND METHODS — Volunteers aged ≥20 years without previously diagnosed diabetes (n = 1,471) completed a brief questionnaire and underwent recording of postprandial time and measurement of capillary blood glucose (CBG) with a portable sensor. Participants subsequently underwent a 75-g oral glucose tolerance test; fasting serum glucose (FSG) and 2-h postload serum glucose (2-h SG) concentrations were measured. The screening tests we studied included the ADA risk assessment questionnaire, the recommended CBG cut point of 140 mg/dl, and an alternative CBG cut point of 120 mg/dl. Each screening test was evaluated against several diagnostic criteria for diabetes (FSG ≥126 mg/dl, 2-h SG ≥200 mg/dl, or either) and dysglycemia (FSG ≥110 mg/dl, 2-h SG ≥140 mg/dl, or either).

RESULTS — Among all participants, 10.7% had undiagnosed diabetes (FSG ≥126 or 2-h SG ≥200 mg/dl), 52.1% had a positive result on the questionnaire, 9.5% had CBG ≥140 mg/dl, and 18.4% had CBG ≥120 mg/dl. The questionnaire was 72–78% sensitive and 50–51% specific for the three diabetes diagnostic criteria; CBG ≥140 mg/dl was 56–65% sensitive and 95–96% specific, and CBG ≥120 mg/dl was 75–84% sensitive and 86–90% specific. CBG ≥120 mg/dl was 44–62% sensitive and 89–90% specific for dysglycemia.

CONCLUSIONS — Low specificity may limit the usefulness of the ADA questionnaire. Lowering the cut point for a casual CBG test (e.g., to 120 mg/dl) may improve sensitivity and still provide adequate specificity.

Diabetes Care 24:1899–1903, 2001

Screening for undiagnosed diabetes has been favored by some (1–4) but discouraged by others (5,6). A comprehensive review (7) found indirect evidence supporting an opportunistic screening approach (i.e., screening subjects visiting a health care provider for reasons unrelated to diabetes) but noted that currently recommended screening strategies have not been fully evaluated. Understanding the performance of screening strategies will also be important if the interventions of the ongoing Diabetes Prevention Program (8) are found to be effective in reducing the onset of diabetes in subjects with impaired glucose tolerance.

We evaluated the performance, in settings typical of opportunistic and community screening programs, of several screening strategies for type 2 diabetes that are currently recommended by the American Diabetes Association (ADA) (4). The screening tests we evaluated included the ADA risk assessment questionnaire and tests based on casual capillary blood glucose (CBG) measures. The diagnostic criteria for this study were diabetes, impaired fasting glucose (IFG), and impaired glucose tolerance (IGT), as determined by fasting serum glucose (FSG) or 2-h postload serum glucose (2-h SG) concentrations measured as part of a single 75-g oral glucose tolerance test (OGTT).

RESEARCH DESIGN AND METHODS

Between September 1995 and July 1998, 1,471 volunteers aged ≥20 years were recruited by health care systems serving communities in Springfield, MA; Robeson County, NC; and Providence, Pawtucket, and Central Falls, RI. Participants were recruited during routine health center visits and at community health fairs. Informed consent was obtained from all participants, and the study protocol was approved by the institutional review boards at the Centers for Disease Control and Prevention and each of the study sites. Persons who had self-reported previously diagnosed diabetes, had been pregnant or breast-feeding within the previous 3 months, or had been hospitalized within the previous 6 months were not eligible to participate in the study.

Screening tests were administered at recruitment. Eligible participants completed a 14-item questionnaire that included the 7 items needed to score the ADA questionnaire test (Table 1). A portable sensor (Accu Chek Advantage; Roche Diagnostics, Indianapolis, IN) was used to obtain a whole-blood glucose

From the [1]Division of Diabetes Translation, Centers for Disease Control and Prevention, Atlanta, Georgia; the [2]Rhode Island Department of Health, Providence, Rhode Island; the [3]Massachusetts Department of Public Health, Boston, Massachusetts; the [4]Baystate Medical Center, Springfield, Massachusetts; and the [5]Robeson Health Care Corporation, Fairmont, North Carolina.

Address correspondence and reprint requests to Deborah B. Rolka, Mailstop K-10, 4770 Buford Highway NE, Atlanta, GA 30341. E-mail: drolka@cdc.gov.

Received for publication 6 March 2001 and accepted in revised form 26 July 2001.

Abbreviations: 2-h SG, 2-h postload serum glucose; ADA, American Diabetes Association; CBG, capillary blood glucose; FSG, fasting serum glucose; IFG, impaired fasting glucose; IGT, impaired glucose tolerance; OGTT, oral glucose tolerance test; WHO, World Health Organization.

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

Table 1—*Scoring the questionnaire test*

Item	Points
1. Woman who delivered a macrosomic (≥9 lb) infant	1
2. One or more siblings with diabetes	1
3. One or more parents with diabetes	1
4. BMI ≥27 kg/m^2	5
5. Age <65 years and little or no physical activity in most weeks	5
6. Age 45–64 years	5
7. Age ≥65 years	9

Subjects with a total of ≥10 points were considered to have had a positive result of the screening test.

level from a capillary (finger stick) sample from each eligible participant, and time since ingestion of any food or drink except water (postprandial time) was recorded.

Participants were scheduled to return for a 75-g OGTT on a subsequent morning (usually within 7 days) after fasting overnight for ≥10 h. During this visit, fasting and 2-h postload venous blood specimens were collected and FSG and 2-h SG concentrations were analyzed in a clinical laboratory using glucose oxidase methodology.

We computed the sensitivity (i.e., proportion of participants with a positive test, among those who satisfied the criterion) and specificity (i.e., proportion of participants with a negative test, among those who did not satisfy the criterion) of four screening tests for six diagnostic criteria.

To investigate how covariates may effect performance characteristics and the choice of appropriate cut points for the CBG, we fit multiple regression models relating CBG to diabetes (FSG ≥126 mg/dl), age (<45 or ≥45 years), postprandial time (<8 or ≥8 h), sex, and race/ethnicity (Hispanic, non-Hispanic white, or African-American). We also computed the sensitivity and specificity of the four screening tests for FSG ≥126 mg/dl separately by sex and race/ethnicity.

CBG measurements were valid in all but 3 of the 1,471 eligible participants, but postprandial time was not recorded for 44 participants (3.0%). FSG values were not recorded for 380 participants (26%), and 2-h SG values were not recorded for 403 participants (27%). To reduce the potential for bias, we applied the standard statistical technique of multiple imputation (9). Every estimate we report is the arithmetic mean of estimates obtained from 10 imputed data sets. We used the software program NORM (10) to impute missing values and we used SAS software (SAS Institute, Cary, NC) (11) to analyze the data and combine the estimates.

RESULTS — Participants included Hispanics (58%), non-Hispanic whites (19%), African-Americans (12%), Native Americans (4%), and others (7%). The mean age of the participants was 44 years (20–44 years, 43%; 45–64 years, 25%; 65–89 years, 32%), and 70% of the participants were women. A total of 34% of the participants had a parent with diabetes, and 17% had a sibling with diabetes; 67% of the participants reported little or no physical activity in most weeks, and 51% of participants had BMI ≥27 kg/m^2.

A total of 52% of all participants had a positive score (≥10 points) on the ADA questionnaire; 9.5% had CBG ≥140 mg/dl, and 18.4% had CBG ≥120 mg/dl. Fasting and 2-h diagnostic criteria for diabetes, impaired glucose, and normoglycemia resulted in somewhat different classifications of participants (Table 2). We estimated that 157 subjects (10.7%) had undiagnosed diabetes, according to one or both of the two criteria, and that an additional 221 (15.0%) had impaired glucose (IFG or IGT) without satisfying either of the criteria for diabetes.

The ADA questionnaire was moderately sensitive (69–78%) for all diagnostic criteria for diabetes and dysglycemia; however, its specificity did not exceed 54% (Table 3). The cut point of 140 mg/dl for CBG was quite specific (95–97%) for all of the diagnostic criteria but only 56–65% sensitive for diabetes and 28–41% sensitive for dysglycemia.

Empirical receiver operating characteristic curves suggest that a CBG cut point of 120 mg/dl may yield a good balance of sensitivity and specificity (Fig. 1). Indeed, this test was 75–84% sensitive for diabetes, 44–62% sensitive for dysglycemia, and 86–90% specific for all of the diagnostic criteria.

The ADA recommends that, in community screening programs, glycemic testing should be performed only after administration of a risk assessment questionnaire (4). This combination (a positive ADA questionnaire and CBG ≥120 mg/dl) was less sensitive and more specific than either the questionnaire or CBG ≥120 mg/dl alone (Table 3). The ADA also recommends using a capillary blood glucose cut point of 110 mg/dl (instead of 140 mg/dl) for subjects who have fasted for ≥8 h (4). Among study participants who had not eaten for ≥8 h (37% of all participants), CBG ≥110 mg/dl was 82–95% sensitive and 86–89% specific for diabetes and 51–80% sensitive and 89–94% specific for dysglycemia.

The ADA questionnaire was less sensitive (65 vs. 77%) and more specific (56 vs. 47%) for diabetes (FSG ≥126 mg/dl) in men than in women. The CBG tests were more sensitive and less specific among men than in women. CBG ≥140 mg/dl was 81% sensitive and 95% specific in men and 56% sensitive and 96% specific in women. CBG ≥120 mg/dl was 90% sensitive and 86% specific in men and 80% sensitive and 88% specific in women.

We derived estimated receiver operating characteristic curves from a linear regression in which the natural log of CBG was modeled as a function of diabetes (FSG ≥126 mg/dl), age, postprandial time, and sex. We assumed normally

Table 2—*Classification of participants by OGTT results*

	2-h SG			
ISG	<140 mg/dl (normoglycemia)	140–199 mg/dl (IGT)	≥200 mg/dl (diabetes)	Total
<110 mg/dl (normoglycemia)	1,093 (74.3)	124 (8.4)	17 (1.2)	1,234 (83.9)
110–125 mg/dl (IFG)	63 (4.3)	34 (2.3)	15 (1.0)	112 (7.6)
≥126 mg/dl (diabetes)	20 (1.3)	27 (1.9)	78 (5.3)	125 (8.5)
Total	1,176 (79.9)	185 (12.6)	110 (7.5)	1,471 (100)

Data are means (% of total) from 10 imputed data sets. A total of 3% of participants had missing FSG values, and 27% had missing 2-h SG values.

Table 3—*Sensitivity and specificity of four screening tests for six diabetes and dysglycemia criteria*

	ADA Questionnaire		CBG ≥140 mg/dl		CBG ≥120 mg/dl		ADA questionnaire *and* CBG ≥120 mg/dl	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
Diabetes criterion:								
FSG ≥126 mg/dl	72 (69–75)	50 (49–50)	65 (63–68)	96 (95–96)	84 (78–89)	88 (87–88)	63 (58–68)	93 (92–93)
2-h SG ≥200 mg/dl	78 (73–84)	50 (50–51)	62 (55–68)	95 (94–95)	78 (72–84)	86 (86–87)	60 (54–67)	92 (91–93)
FSG ≥126 mg/dl or 2-h SG ≥200 mg/dl	75 (72–79)	51 (50–51)	56 (53–59)	96 (96–96)	75 (70–80)	88 (88–89)	58 (54–62)	94 (93–94)
Dysglycemia criterion:								
FSG ≥110 mg/dl	69 (66–72)	51 (50–52)	41 (39–43)	97 (96–97)	62 (57–66)	90 (89–91)	45 (42–48)	95 (94–95)
2-h SG ≥140 mg/dl	72 (69–75)	53 (52–54)	33 (31–35)	96 (96–97)	48 (45–50)	89 (88–90)	36 (34–39)	94 (94–95)
FSG ≥110 mg/dl or 2-h SG ≥140 mg/dl	69 (67–71)	54 (53–55)	28 (27–29)	97 (97–97)	44 (41–47)	90 (90–91)	32 (30–34)	95 (95–96)

Data are % (95% CI). The 95% CIs account only for the uncertainty due to missing data and are computed as (mean point estimate) ± $[(t_{.975,\,9}) \times (1 + 1/10)^{1/2} \times$ (SD of 10 point estimates)].

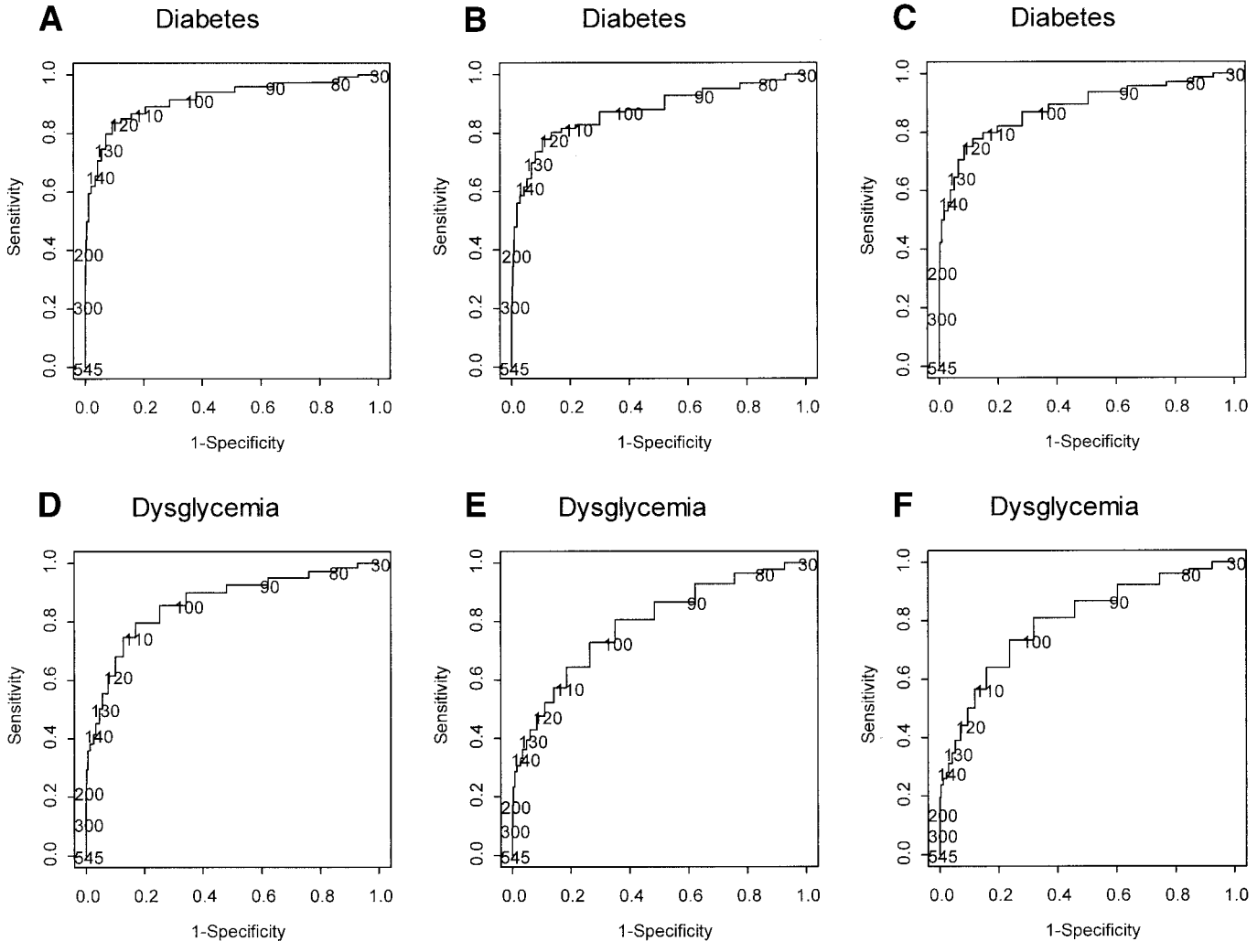

**Figure 1—*Empirical receiver operating characteristic curves. Sensitivity vs. 1-specificity of CBG is plotted over a range of CBG cut points for diabetes* (top row) *and dysglycemia* (bottom row). *Diagnostic criteria are FSG ≥126 mg/dl* (A), *2-h SG ≥200 mg/dl* (B), *FSG ≥126 mg/dl or 2-h SG ≥200 mg/dl* (C), *FSG ≥110 mg/dl* (D), *2-h SG ≥140 mg/dl* (E), *FSG ≥110 mg/dl or 2-h SG ≥140 mg/dl* (F).

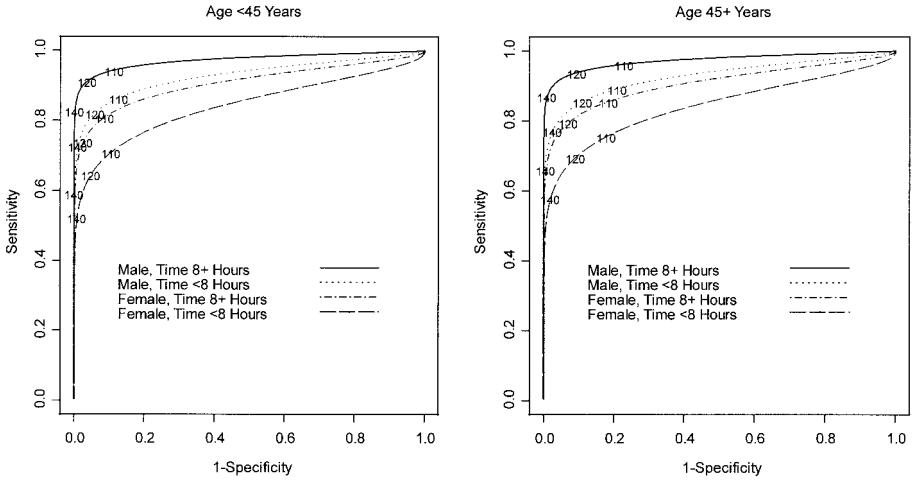

Figure 2—*Estimated receiver operating characteristic curves by age, sex, and postprandial time. Sensitivity and specificity of CBG for the diabetes criterion of FSG ≥126 mg/dl were estimated using the multiple regression model described in the text, in which the natural log of CBG is modeled as a function of diabetes, age, postprandial time, sex, and diabetes × sex.*

distributed errors and heterogeneous variances (varying by diabetes and postprandial time). Cut points for the CBG test that were optimal (maximizing the sum of sensitivity and specificity) tended to be lower for younger subjects and those with longer postprandial times and higher for men. CBG performed somewhat better (larger areas under the curves) for men than for women and for subjects with postprandial time ≥8 h than for those with postprandial time <8 h (Fig. 2).

The sensitivities and specificities of the four screening tests varied little by race or ethnicity, and we did not find substantial racial or ethnic differences in the performance of CBG for diabetes (FSG ≥126 mg/dl) after controlling for age, postprandial time, and sex.

CONCLUSIONS — This is the first comprehensive evaluation of screening tests that use a questionnaire or casual CBG measure to detect undiagnosed diabetes or dysglycemia in patient populations and settings typical of current U.S. screening initiatives. Using several diagnostic criteria for diabetes and dysglycemia, we found that the ADA questionnaire favored sensitivity, whereas CBG ≥140 mg/dl (the recommended cut point) favored specificity.

The ADA questionnaire was developed from the Second U.S. National Health and Nutritional Examination Survey using a binary classification algorithm (12). The ADA questionnaire yielded lower specificity in our study than it did in previous evaluations. In the current study, the questionnaire was 78% sensitive and 50% specific for the World Health Organization (WHO) diabetes criterion (2-h SG ≥200 mg/dl) (13). Sensitivity for this WHO criterion was 79%, and specificity was 65% in the initial evaluation of the ADA questionnaire (12). In an evaluation that was conducted using the Netherlands' Hoorn Study population, sensitivity was 72% and specificity was 56% (14).

CBG screening tests for diabetes have been suggested because they use current self-monitoring technology and require minimal technical skill and laboratory support compared with more laboratory-based tests (e.g., serum glucose or HbA_{1c}). Previous evaluations of CBG screening tests have reported sensitivities of 50–70% at 90% specificity (15,16). In our study, CBG was >70% sensitive for the WHO diabetes criterion (13) at 90% specificity.

The performance of CBG tests may depend on postprandial time and other factors such as age or sex (7,15,17). Consistent with a previous study (15), we found that optimal CBG cut points may be lower for younger subjects and those with longer (≥8 h) postprandial times. In contrast with that study, in which the best performance was observed among those with the shortest postprandial times (15), we found that CBG performed somewhat better in individuals with longer postprandial times than in those with postprandial times <8 h. In our study, we also observed better performance and slightly higher optimal cut points in men than in women.

Diabetes screening tests have been evaluated in homogeneous populations (15,18–22) but rarely in racially heterogeneous populations. We were able to examine the potential effects of race or ethnicity and found that the performance characteristics of the ADA questionnaire and the CBG measure did not vary substantially by race or ethnicity.

Detection of IFG or IGT is not a goal of most current diabetes screening efforts. This may change, however, if the lifestyle and/or medication interventions of the Diabetes Prevention Program (8) are shown to be effective. We included diagnostic criteria for dysglycemia (i.e., diabetes and IFG or IGT) and examined the performance of current diabetes screening tests when applied to these broader diagnostic criteria. Our data suggest that CBG measures do not discriminate dysglycemia from normoglycemia as well as they discriminate subjects with diabetes from those without diabetes.

Our study has some limitations. Because our volunteers and participating clinics were not probability samples, we do not make formal statistical inference beyond the study population. We believe that the participation of subjects from urban and rural areas in three states yielded a study population reflecting the heterogeneity of U.S. populations. However, because it would be inappropriate to use this study population to develop new screening tests and strategies, we focused our evaluation on existing screening tests. Missing data may have biased our estimates for the study population; we attempted to minimize this bias through the use of multiple imputation. Also, clinical diagnosis requires repeat testing, and the diagnostic criteria that we defined are based on a single OGTT. Therefore, our sensitivities and specificities were estimated relative to imperfect criteria.

Our estimates can be used to help project resource needs and expected yields. For example, suppose that a program plans to use a casual CBG test to screen a population of 5,000 individuals for diabetes (FSG ≥126 mg/dl). We estimated that the screening test CBG ≥120 mg/dl is 84% sensitive and 88% specific. If the population prevalence of diabetes is assumed to be 8%, then screening with CBG ≥120 mg/dl can be projected to

yield 8% × 84% × 5,000 = 336 true positives (new cases), 92% × 12% × 5,000 = 552 false positives, and 8% × 16% × 5,000 = 64 false negatives (missed cases). The projected positive predictive value (proportion of actual cases among those who have positive tests) would be 336 ÷ (336 + 552) = 29.8%.

The U.S. Preventive Services Task Force has voiced concern about the lack of a practical screening test that is both sensitive and specific (5). We found that the usefulness of the ADA questionnaire as a screening test may indeed be limited by its low specificity. The casual CBG measure offers better performance and the flexibility to select threshold cut points that balance sensitivity and specificity with the available resources; lowering the cut point (e.g., to 120 mg/dl) may improve sensitivity and still provide adequate specificity.

References

1. World Health Organization Study Group on Prevention of Diabetes Mellitus: *Prevention of Diabetes Mellitus*. Geneva, World Health Organization, 1994 (Tech. Rep. Ser., no. 844)
2. Patterson KR: Population screening for diabetes mellitus. *Diabet Med* 10:77–81, 1993
3. 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
4. American Diabetes Association: Clinical practice recommendations 2000: screening for type 2 diabetes. *Diabetes Care* 23: S20–S23, 2000
5. U.S: Preventive Services Task Force: *Guide to Clinical Preventive Services*. 2nd ed. Alexandria, VA, International Medical Publishing, 1996
6. Canadian Task: Force on the Periodic Health Examination: The periodic health examination. *Can Med Assoc J* 121:1193–1254, 1979
7. Engelgau MM, Aubert RE, Thompson TJ, Herman WH: Screening for NIDDM in nonpregnant adults: a review of principles, screening tests, and recommendations. *Diabetes Care* 18:1606–1618, 1995
8. The Diabetes Prevention Program Research Group: The Diabetes Prevention Program: design and methods for a clinical trial in the prevention of type 2 diabetes. *Diabetes Care* 22:623–634, 1999
9. Schafer JL: *Analysis of Incomplete Multivariate Data*. London, Chapman and Hall, 1997
10. Schafer JL: NORM: multiple imputation of incomplete multivariate data under a normal model, software for Windows 95/98/NT, version 2 [article online], 1999. Available from http://www.stat.psu.edu/~jls/misoftwa.html. Accessed 13 March 2000
11. SAS Institute: *SAS Procedures Guide, Version 6*. 3rd ed. Cary, NC, SAS Institute, 1990
12. Herman WH, Smith PJ, Thompson TJ, Engelgau MM, Aubert RE: A new and simple questionnaire to identify subjects at increased risk for undiagnosed diabetes. *Diabetes Care* 18:382–387, 1995
13. World Health Organization: *WHO Expert Committee on Diabetes Mellitus. Second Report*. Geneva, World Health Organization, 1980 (Tech. Rep. Ser., no. 646)
14. Ruige JB, de Neeling JN, Kostense PJ, Bouter LM, Heine RJ: Performance of an NIDDM screening questionnaire based on symptoms and risk factors. *Diabetes Care* 20:491–496, 1997
15. Engelgau MM, Thompson TJ, Smith PJ, Herman WH, Aubert RE, Gunter EW, Wetterhall SF, Sous ES, Ali MA: Screening for diabetes mellitus in adults: the utility of random capillary blood glucose measurements. *Diabetes Care* 18:463–466, 1995
16. Andersson DK, Lundblad E, Svardsudd K: A model of early diagnosis in type 2 diabetes mellitus in primary health care. *Diabet Med* 10:167–173, 1993
17. Blunt BA, Barrett-Conner E, Windgard D: Evaluation of fasting plasma glucose as a screening test for NIDDM in older adults: Rancho Bernardo Study. *Diabetes Care* 14: 989–993, 1991
18. Hanson RL, Nelson RG, McCance DR, Beart JA, Charles MA, Pettitt DJ, Knowler WC: Comparison of screening tests for non-insulin-dependent diabetes mellitus. *Arch Intern Med* 153:2133–2140, 1993
19. Forrest RD, Jackson CA, Yudkin JS: The glycohaemoglobin assay as a screening test for diabetes mellitus: the Islington Diabetes Survey. *Diabet Med* 4:254–259, 1987
20. Anokute CC: Epidemiologic studies of diabetes mellitus in Saudi Arabia: part I, screening of 3158 males in King Saud University. *J R Soc Health* 110:201–203, 1990
21. Tsuji I, Nakamoto K, Hasegawa T, Hisashige A, Inawashiro H, Fukao A, Hisamichi S: Receiver operating characteristic analysis on fasting plasma glucose, HbA_{1c}, and fructosamine on diabetes screening. *Diabetes Care* 14:1075–1077, 1991
22. Lee CH, Fook-Chong S: Evaluation of fasting plasma glucose as a screening test for diabetes mellitus in Singaporean adults. *Diabet Med* 14:119–122, 1997