Effectiveness of diabetes mellitus screening recommendations

Florence J. Dallo and Susan C. Weller*

Sociomedical Sciences Division, Department of Preventive Medicine and Community Health, University of Texas Medical Branch, Galveston, TX 77555-1153

Communicated by A. Kimball Romney, University of California, Irvine, CA, June 20, 2003 (received for review May 6, 2003)

Screening guidelines proposed by the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus have been endorsed by several medical societies. However, one-third of cases are undiagnosed, and complications at the time of diagnosis indicate that disease may have been present for several years before diagnosis. This study evaluates the effectiveness of the guidelines for detecting new cases of diabetes mellitus. By using a cross-sectional, representative sample of the United States (National Health and Nutritional Examination Survey, NHANES III), the guidelines are tested on adults, 20 years and older without a prior diagnosis of diabetes. Individuals are classified as nondiabetics ($n = 6,241$) or as having undiagnosed diabetes ($n = 274$) based on their blood glucose. Screening when one risk factor is present, as stated in the guidelines, has a true-positive rate of 100% and would require that 83% of the population be tested. Screening when two risk factors are present is more efficient, with a comparable true-positive rate (98%), but requires that only 59% of the population be tested. A notable finding is the earlier age of onset among minorities, which may be associated with other health disparities. Because diabetes occurs at younger ages in minorities, screening whites who are ≥40 and minorities ≥30 years of age has a high true-positive rate (95%) and also reduces testing (60%). The screening guidelines would be effective, if followed, and would essentially eliminate undiagnosed cases of diabetes.

The Expert Committee on the Diagnosis and Classification of Diabetes Mellitus (1) has proposed guidelines for diabetes mellitus testing in asymptomatic, undiagnosed individuals. Although screening is not recommended in the general adult population, it is recommended for individuals with one or more risk factors (Table 1). Screening is recommended at 3-year intervals beginning at age 45 years, with earlier and/or more frequent screening for individuals with other risk factors (1). Because these guidelines are generally considered the standard for diabetes screening (several professional societies have adopted the guidelines), we evaluate their ability to detect new cases of diabetes in a U.S. national sample. The recommendations proposed by the Expert Committee are based on a review of scientific evidence, but their sensitivity for detection of new cases of diabetes has not been tested.

Diabetes is a national health problem in the Unites States. It is associated with micro- and macrovascular disease resulting in significant morbidity and mortality. Diabetes is a class of metabolic disorders characterized by hyperglycemia. The two most prevalent types are distinguished by a lack of insulin (type 1) or insufficient insulin (type 2) to metabolize glucose (1). Among adults, diabetes is among the leading causes of death (2, 3) and is the number one cause of new cases of legal blindness (20- to 74-year-olds) (4), end-stage renal disease (5), and lower-extremity amputations (6). Direct costs for medical services and indirect costs because of lost productivity due to diabetes approached 132 billion dollars in 2002 (7). The prevalence of diabetes (physician-diagnosed and undiagnosed) among adults (≥20 years) is 7.8% (8) and has risen during the past decade. For those 40–74 years of age, the prevalence of diabetes has increased from 8.9% (1976–1980) to 12.3% (1988–1994) (8). The increased prevalence is most likely due to increased incidence and not decreased mortality (9) or improved case detection. Because intervention trials (10–14) have demonstrated that improved glycemia can prevent or delay complications, an early diagnosis and appropriate management of diabetes should be a priority.

Table 1. Criteria for testing for diabetes in asymptomatic, undiagnosed individuals

Testing should be considered for those 45 years of age and older and repeated every 3 years, if results are normal.
Testing should be considered younger or more frequently for those who:
(*i*) are African-, Hispanic-, Native-, Asian-, or Pacific Island-American
(*ii*) are overweight (BMI, kg/m^2 ≥ 25);
(*iii*) have had gestational diabetes or delivered a baby weighing >9 pounds;
(*iv*) have a positive family history of diabetes (parents or siblings);
(*v*) have hypertension (blood pressure ≥140/90 mmHg);
(*vi*) have low HDL cholesterol (≤35 mg/dl) and/or high triglyceride level (≥250 mg/dl);
(*vii*) have had impaired glucose (110 ≤ FPG < 126 mg/dl or 140 ≤ OGTT < 200 mg/dl).

Adapted from table 6 of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus (1). BMI, body mass index; HDL, high-density lipoprotein; FPG, fasting plasma glucose; OGTT, oral glucose tolerance test.

Diagnostic and screening guidelines have been developed by an expert panel (1). Diagnosis is made when hyperglycemia is present, as measured by a FPG test (FPG ≥ 126 mg/dl) (1). Screening is recommended for those 45 years of age and older; with repeated testing every 3 years if results are normal. Screening also should be considered at younger ages or at more frequent intervals for those who have diabetes risk factors (Table 1) (1).

The guidelines have been adopted by the American Academy of Family Physicians (15), the American Association of Clinical Endocrinologists (www.aace.com), and the International Diabetes Center (16). The latter two groups added risk factors (acanthosis nigricans and cardiovascular disease), and the endocrinologists recommended screening at 30 years, if other risk factors are present. The U.S. Preventive Services Task Force (17), the Canadian Task Force (18), and the American College of Physicians (19) have more ambiguous recommendations and suggest screening those who are obese, older, or the member of a racial/ethnic minority, or who have a family history of diabetes.

Unfortunately, one third of diabetes cases (probably type 2) are undiagnosed (8). Also, the presence of microvascular disease (specifically, retinopathy) at the time of diagnosis suggests that disease may be present for 2–10 years before diagnosis (20). The question is whether the guidelines are sufficient to detect most cases of diabetes. The use of risk factors to identify individuals

Abbreviations: FPG, fasting plasma glucose; NHANES, National Health and Nutrition Examination Survey; BMI, body mass index; OR, odds ratio; C.I., confidence interval; HDL, high-density lipoprotein.

*To whom correspondence should be addressed. E-mail: sweller@utmb.edu.

 www.pnas.org/cgi/doi/10.1073/pnas.1733839100

with diabetes has had mixed results. Studies comparing the presence of risk factors and blood glucose values have found true-positive rates of 62–85% (21–29). These studies, however, relied on a variety of variables measured in clinical samples. Samples without variation in ethnicity (all white) or samples with a truncated age range (21–23, 25, 26, 28) are not appropriate for generalizing to the U.S. population. Alternatively, it is possible that the guidelines are sufficient, but that they are not being followed. In the 1989 National Health Interview Survey, only 38.6% of those without a history of diabetes but with three diabetes risk factors reported being screened and only 57.1% of those with four or more risk factors reported being screened (30). Minorities have disproportionately higher rates of complications and mortality because of diabetes, and it is not known if these disparities are due to a delay in diagnosis (31).

This study examines the performance of screening guidelines for identifying undiagnosed diabetes in a national sample. We focus on only the risk factors included in the guidelines and test their ability to detect new or undiagnosed cases of diabetes in the U.S. population. The National Health and Nutrition Examination Survey (NHANES III) (32, 33) provides information that parallels the screening problem, namely, the detection of undiagnosed cases of diabetes in the general population. The NHANES III includes data on seven of the risk factors in Table 1. In addition, the study includes FPG results for a sample of individuals without a prior diagnosis of diabetes. The relative importance of risk factors in identifying new cases of diabetes can be obtained by comparing those with undiagnosed diabetes (FPG $\geq$ 126) with those without diabetes (FPG $<$ 126). More importantly, their efficacy in detecting new cases of diabetes can be estimated by calculating the true-positive rate for the screening guidelines.

Methods

Data Source and Sample. The NHANES III is based on a stratified, multistage probability sample of noninstitutionalized U.S. civilians (1988–1994), with an oversampling of non-Hispanic blacks and Mexican Americans. The complex survey design included sample weights for population estimates, because each person does not have the same probability of selection. Only subjects identified as non-Hispanic white, non-Hispanic black, and Mexican American were eligible for the study. Included individuals who did not specifically self-identify as one of these groups were categorized as "other" (32, 33). In this article, these groups are referred to as white, African American, Mexican American, and other.

The sample for this study included nonpregnant adults ($\geq$20 years of age) who had a physical examination, morning venipuncture, and an interview (n = 7,703). Of those, 574 had a prior diagnosis of diabetes (excluding gestational diabetes) and were excluded. For those without a prior diagnosis of diabetes, only those who had fasted between 8 and 24 h and had a FPG measurement were included (n = 6,515).[†] Those who had a FPG $\geq$ 126 mg/dl were categorized as undiagnosed diabetics (n = 274), and those who had a FPG $<$ 126 mg/dl were categorized as nondiabetics (n = 6,241).

Measures and Analysis. The main outcome variable, diabetes status, was defined with two categories of subjects: those with undiagnosed diabetes and those without diabetes. The two groups were compared in terms of their risk factors: age ($\geq$45 years); ethnicity; positive family history (parent or sibling with diabetes); overweight/obesity (measured by BMI $\geq$ 25 and BMI $\geq$ 30 kg/m^2) (34); history of gestational diabetes; a prior diagnosis of hypertension or a current hypertensive reading; and their lipid profile (HDLs and triglycerides). Information was not available for history of macrosomic infant (>9 pounds) or dysglycemia (e.g., 110 $<$ FPG $\leq$ 126 mg/dl). Thus, the first seven risk factors in Table 1 were considered for analysis. Descriptive statistics (percentages and odds ratios) were obtained for the groups for all risk factors. A simple, main effects logistic regression (35) model was used to estimate the relative strength of the risk factors in distinguishing those with and without diabetes. Odds ratios [and 95% Taylor series confidence intervals (C.I.)] for risk factors were obtained from logistic regression models. All risk factors were dichotomized to parallel the screening guidelines and to facilitate comparisons of the relative magnitude of the odds ratios across variables. All analyses include weighting and design factors with SAS (36) and SUDAAN (37). The purpose of weighting the sample data is to allow estimates of statistics that would have been obtained if the entire sampling frame (the United States) had been surveyed. The true-positive (hit rate) and false-positive (false alarm) rates were estimated for the presence or absence of different combinations of risk factors for detecting diabetes, and the proportion of the population with each combination of risk factors.

Results

Relative Importance of Risk Factors for Detecting Undiagnosed Diabetes. Half of the sample was female (51.80%), 23.24% had less than a high school education, and 11.82% had no health insurance (Table 2). Those who were 45 years of age and older constituted 42.35% of the total sample, 83.97% of those with undiagnosed diabetes, and 41.11% of those without diabetes. Those who were 45 years and older were 7.50 (95% C.I. = 4.37, 12.89) times more likely to have undiagnosed diabetes than were those who were younger than 45 years. Compared with persons without diabetes, those with diabetes were older, had a first-degree relative with diabetes, were overweight or obese, had hypertension, and had a poor lipid profile (Table 2). The effect of most variables was reduced when all were considered simultaneously. Each variable had <2% missing data with ≈5% missing for the combined logistic regression analysis (n = 6,214; undiagnosed diabetes n = 261 and n = 5,953 in nondiabetics). In model 1 (Table 3), all the risk factors except gestational diabetes were used. In model 2, risk factors and gestational diabetes were tested for women only.

In the general population, the strongest risk factor is older age. The bivariate odds ratio (OR) (Table 2) indicates that those $\geq$45 years of age are at 7.50 times greater risk for having undiagnosed diabetes. When other risk factors are controlled for (Table 3, model 1), those $\geq$45 years of age are at 5.84 (95% C.I. = 3.41, 10.02 in model 1) times greater risk of having undiagnosed diabetes than those <45 years of age. The odds of having diabetes increase 1.05 times for each year increase in age (OR = 1.05, 95% C.I. = 1.04, 1.06 with age as a continuous variable replacing age $\geq$45 years in model 1, not shown). Diabetes, however, occurs at an earlier age in the minority ethnic groups. Among those with undiagnosed diabetes, the youngest individuals were 34 years of age among whites, 27 years among African Americans, and 24 years among Mexican Americans. The distribution of undiagnosed diabetes cases shows that 11.51% of white, 28.09% of African American, and 52.84% of Mexican-American individuals are <45 years of age (Table 4). The 45-year cutpoint for whites (where ≈12% of undiagnosed cases are younger) is comparable with 34 years of age for African Americans and 31 years for Mexican Americans.

The effect of ethnicity appears to be due to the younger age distribution of undiagnosed cases among minorities. The simple, bivariate OR (Table 2) indicates that the minority ethnic groups are at ≈1.4 times greater risk than whites for having undiagnosed

[†]The sample size (n = 6,515) differs slightly from that of Harris *et al*. (8) (n = 6,587), because we used the recommended fasting time of at least 8 h (instead of 9) and we omitted women who were pregnant at the time of the study.

Table 2. Distribution of factors overall and within each study group

Demographic and risk factors	Distribution, % Total	Undiagnosed diabetes* (n = 274)†	No diabetes‡ (n = 6,241)†	Odds ratio (95% C.I.)
Gender (female)	51.80	47.05	51.94	0.82 (0.54, 1.25)
Education (<12 yr)	23.25	38.81	22.78	**2.15 (1.44, 3.21)**
No health insurance	11.82	9.29	11.90	0.76 (0.45, 1.27)
Age (≥45 yr)	42.35	83.97	41.11	**7.50 (4.37, 12.89)**
Ethnicity				
Non-Hispanic white	77.11	70.67	77.30	1.00
African American	10.07	12.97	9.98	**1.42 (1.04, 1.94)**
Mexican American	5.02	6.27	4.98	1.38 (1.00, 1.89)
Other	7.81	10.08	7.74	1.42 (0.80, 2.52)
First degree relative	23.08	43.82	22.46	**2.69 (1.83, 3.95)**
Obesity				
Overweight (BMI ≥ 25 kg/m^2)	54.07	85.85	53.12	**5.36 (3.51, 8.18)**
Obesity (BMI ≥ 30 kg/m^2)	21.36	56.84	20.30	**5.17 (3.72, 7.18)**
History of gestational diabetes	0.36	1.75	0.32	**5.56 (1.48, 20.86)**
Women only	0.69	3.75	0.61	**6.31 (1.69, 23.49)**
Hypertension				
Previous diagnosis	22.10	54.79	21.12	**4.53 (2.96, 6.93)**
Current hypertension reading	16.93	41.98	16.18	**3.75 (2.59, 5.42)**
Previous diagnosis or current hypertension	29.71	66.90	28.61	**5.05 (3.27, 7.78)**
Lipid profile				
Total cholesterol (≥200 mg/dl)	50.27	65.90	49.81	**1.95 (1.30, 2.93)**
HDL cholesterol (≤35 mg/dl)	12.59	31.48	12.02	**3.36 (2.20, 5.14)**
Triglycerides (≥250 mg/dl)	8.26	32.69	7.52	**5.97 (3.57, 9.99)**
HDL (≤35 mg/dl) or triglycerides (≥250 mg/dl)	17.42	48.93	16.47	**4.86 (3.38, 6.98)**

Statistically significant ($P < 0.05$) odds ratios are in bold.
*Percent of risk factor in undiagnosed diabetes group = true-positive rate.
†Individual variables have <2% missing data.
‡Percent of risk factor in undiagnosed diabetes group = false-positive rate.

diabetes. In the multivariate model (Table 3, model 1), Mexican Americans and African Americans are twice as likely as whites to have undiagnosed diabetes. However, when the age variable is recoded to capture a comparable proportion of cases from each ethnic group (≥45 for whites, ≥35 for African Americans, ≥32 for Mexican Americans, and ≥45 for other, model not shown), the effects of ethnicity are no longer significant (African Americans OR = 1.28, 95% C.I. = 0.91, 1.81; and Mexican Americans OR = 1.15, 95% C.I. = 0.79, 1.69), and the remaining variables are relatively unaffected, with the exception of an increase in the effect of age (OR = 6.64, 95% C.I. = 3.36, 13.11).

Table 3. Odds ratios for risk factors from logistic regression models

Risk factor	Odds ratios (95% C.I.) Model 1 (n = 6,214)	Model 2 (n = 3,265)
Age (≥45 yr)	**5.84 (3.41, 10.02)**	**7.79 (3.88, 15.66)**
Ethnicity		
African American	**1.97 (1.33, 2.92)**	**2.70 (1.51, 4.82)**
Mexican American	**2.10 (1.37, 3.22)**	1.88 (0.93, 3.81)
Other	**1.83 (1.05, 3.20)**	2.29 (0.83, 6.32)
First degree relative	**2.29 (1.56, 3.35)**	**1.86 (1.09, 3.17)**
BMI (≥25 kg/m^2)	**2.83 (1.77, 4.52)**	**4.24 (2.01, 8.97)**
Hypertension	**2.44 (1.56, 3.79)**	1.96 (0.99, 3.87)
HDL (≤35 mg/dl) or triglycerides (≥250 mg/dl)	**4.11 (2.83, 5.98)**	**6.36 (3.38, 11.94)**
Gestational diabetes		**15.35 (3.79, 62.10)**

Statistically significant ($P < 0.05$) odds ratios are in bold.

A poor lipid profile (HDL ≤ 35 mg/dl or triglycerides ≥ 250 mg/dl) was the second strongest factor predicting undiagnosed diabetes. The bivariate odds ratios indicated that risk was increased approximately three times with a low HDL level, whereas high triglycerides increased risk almost six times. In the multivariate analysis, those with a low HDL or high triglyceride level were four times more likely (OR = 4.11, 95% C.I. = 2.83, 5.98) to have diabetes than were those with high HDLs and low triglycerides.

Obesity was the next strongest risk factor. The bivariate ORs (Table 2) indicated that overweight or obese individuals were at five times higher risk of diabetes. When all risk factors were considered (Table 3, model 1), those who were overweight (BMI ≥ 25 kg/m^2) were 2.83 (95% C.I. = 1.77, 4.52) times more likely to have undiagnosed diabetes than those with BMI < 25 kg/m^2. When obesity was specified, those who were overweight but not obese (25 ≤ BMI < 30 kg/m^2) were 1.70 (95% C.I. = 1.01, 2.86) times more likely to have undiagnosed diabetes, and

Table 4. Age distribution within each ethnic category for new cases of diabetes

Ethnic group	Distribution, % <30 yr	30–39 yr	40–44 yr	45–64 yr	≥65 yr
White (n = 112)	0.00	6.50	5.01	48.52	39.97
African American (n = 74)	0.92	21.08	6.09	48.76	23.15
Mexican American (n = 75)	4.38	23.08	25.37	37.96	9.20

Table 5. Performance of decision rules for identifying cases of diabetes

TPR	PV+	TNR	PV−	% of population	Screening rule
100.00	3.46	17.5	100.00	83.00	1. One risk factor
99.67	4.08	30.65	99.97	70.22	2. One risk factor in those ≥30 yr
98.30	4.79	42.20	99.88	58.97	3. Any two risk factors
97.97	5.26	47.88	99.88	53.44	4. Any two risk factors in those ≥30 yr
83.97	5.75	58.89	99.19	42.35	5. Anyone ≥45 yr
99.61	3.69	22.22	99.95	78.42	6. Anyone ≥30 yr
88.60	5.50	54.47	99.38	46.78	7. White ≥45 *or* black ≥35 *or* Mexican American ≥32 (*or* others ≥45) yr
95.01	4.61	41.22	99.64	59.84	8. Whites ≥40 *or* non-White ≥30 yr
85.85	4.61	46.88	99.11	54.07	9. BMI ≥ 25 kg/m^2
100.00	3.76	23.55	100.0	77.14	10. Whites ≥40 *or* non-white ≥30 yr) *or* BMI ≥25 kg/m^2
48.93	8.19	83.53	98.20	17.42	11. HDL ≤35 or triglycerides ≥250 mg/dl
98.45	4.34	34.90	99.87	66.07	12. Whites ≥40 *or* non-white ≥30 yr) *or* (HDL ≤35 or triglycerides ≥250 mg/dl)
1.75	13.99	99.68	97.16	0.36	13. Gestational diabetes
96.09	4.61	40.99	99.72	60.08	14. (Whites ≥40 *or* non-white ≥30 yr) *or* gestational diabetes

TPR, true-positive rate or sensitivity; PV+, predictive value positive; TNR, true-negative rate or specificity; PV−, predictive value negative.

those who were obese (BMI ≥ 30 kg/m^2) were 4.51 (95% C.I. = 2.66, 7.62) times more likely to have undiagnosed diabetes than those who were not overweight or obese (BMI < 25 kg/m^2; three categories of BMI in model 1, not shown). For each unit increase in BMI, risk increased 1.09 times (95% C.I. = 1.06, 1.12; BMI as a continuous variable in model 1 instead of dichotomous, not shown).

Hypertension and a family history of diabetes each approximately doubled the risk of diabetes (Table 3, model 1). In contrast, women who had had gestational diabetes were at very high risk of diabetes mellitus. In the multivariate model, the effect of gestational diabetes was stronger (OR = 15.35, 95% C.I. = 3.79, 62.10 in model 2) than for age, lipids, or obesity. When gestational diabetes was tested for the entire population, it continued to have a stronger effect than other variables (OR = 11.71, 95% C.I. = 2.84, 48.38 in model 1 with men and women, not shown).

Accuracy of Risk Factors for Detecting New Cases of Diabetes. Performance of the risk factors for detecting diabetes in the general population can be estimated by evaluating decision rules to select individuals for screening. Performance can be described with the proportion of diabetes cases that would be detected (hit rate or true-positive rate) and the proportion of the population that would need to be screened (e.g., the proportion of the population with those risk factors). A high true-positive rate (with a corresponding low false-negative rate) is an important attribute of a decision rule identifying high-risk individuals for screening for diabetes, because false-positive errors (resulting in unnecessary FPG tests) are less costly than are false-negative errors (missing cases of diabetes). The goal of using risk factors to identify individuals for screening is to rule out disease, rather than a diagnostic role to identify when disease is present. Thus, a high true-positive rate (sensitivity) and high negative predictive value are desirable. Predictive value is the Bayesian posterior probability of disease, given the presence or absence of risk factors and the prevalence of undiagnosed diabetes (38). Negative predictive value is the likelihood that someone does *not* have diabetes, given that they do *not* have any risk factors. Positive predictive value is the likelihood that someone has diabetes, given that they have risk factors.

The Expert Committee guidelines (1) recommend screening those who are 45 years of age or older and those with other risk factors at a younger age. In this study, the screening recommendations were operationalized in the following ways. The first strategy was equivalent to screening in the presence of at least one risk factor (including age ≥45 years as a risk factor), regardless of age. The true-positive rate or sensitivity for this strategy is 100.00% (Table 5, rule 1). Negative predictive value is also 100.00%. Because 83.00% of the adult population has one or more risk factors, 83.00% would be identified for testing. If screening at a "younger age," when risk factors are present, were operationalized instead as screening only those 30 years of age and older with one other risk factor (rule 2), 70.22% of the population would be tested with 99.67% sensitivity. Screening, at any age, when two or more risk factors are present (rule 3) has sensitivity of 98.30% and would require testing 58.97% of the population. Limiting screening (in rule 3) to those 30 years of age and older with any two or more risk factors reduces testing to 53.44% of the population (rule 4).

Simpler rules, such as universal screening for everyone 45 years of age and older (rule 5), results in a sensitivity of 83.97% and would identify 42.35% of the population for testing. (Note that the sensitivity of each variable, taken singly, appears in the second numeric column of Table 2.) However, this rule has sensitivity of 88.49% for whites, 71.91% for African Americans, and 47.16% for Mexican Americans because of the earlier age of onset among minorities (Table 4). If ethnicity were incorporated into rule 5 by using the age distribution of undiagnosed cases, so that whites 45 years of age and older, African Americans 35 and older, Mexican Americans 32 and older, and others 45 and older (rule 7) were identified for screening; sensitivity would be equalized for each ethnic group at ≈88%. A slight simplification of rule 7 would be to screen whites at 40 years and non-whites at 30 (rule 8). This simple rule has sensitivity of 95.01% (with sensitivity ≥93% within each ethnic group) and 59.84% of the population would be identified for testing.

Overweight, poor lipid profile (low HDLs or high triglycerides), and gestational diabetes were also identified as strong predictors of undiagnosed diabetes. Screening for overweight (rule 9) has sensitivity similar to that of age (rule 5), but not age and ethnicity (rule 8). Screening for older age within ethnic groups or overweight (rule 10) is equivalent to screening younger people if they were overweight. This rule has sensitivity of 100.0% and would identify 77.14% of the population for testing

and performs as well as all risk factors together (rule 1). A poor lipid profile (rule 11) has a high true-negative rate (83.53%) and a high negative predictive value (98.20%). Combination of age and ethnicity with lipid profile would detect almost all cases (98.45%) and identify 66.07% for FPG testing. Gestational diabetes has a high true-negative rate (specificity) and although few people had gestational diabetes, those that have had it were very likely to have diabetes. Screening based on two or three of the risk factors, such as age and ethnicity (rule 8); age, ethnicity, and overweight (rule 10); age, ethnicity, and lipids (rule 12); or age, ethnicity, and gestational diabetes (rule 14) perform as well as the entire set of risk factors (rule 1).

Conclusions

Risk factors included in the screening guidelines have a strong association with diabetes. Having hypertension or a positive family history of diabetes can double the risk of having diabetes. Age, obesity, a poor lipid profile, and gestational diabetes, however, more than double the risk of diabetes. Previous screening guidelines identified obesity (BMI $\geq$ 27 kg/m^2) (39) as a risk factor, but it is clear that risk increases as BMI increases and that being "overweight" is a significant risk factor without the presence of obesity. If clinician screening practices followed recommended guidelines and individuals with one or more risk factors (Table 1) were screened, then the approximately one-third of diabetes cases that are undiagnosed would be essentially eliminated.

The strongest risk factors performed as well as all risk factors together in detecting new cases of diabetes. Age is the risk factor most strongly associated with the detection of undiagnosed cases of diabetes, but it does not have a simple cut-point appropriate for all ethnic groups. Clinical screening strategies focused on older adults ($\geq$45 years of age) will miss minorities with diabetes. The best performance occurs when white patients are screened beginning at age 40 years and Mexican Americans and African Americans are screened beginning at age 30 years. This simple rule would require testing of only 59.84% of the population and would detect most cases (95%) of diabetes. The 45-year cutoff in the screening guidelines should be reconsidered, given the risk to younger minorities.

Identification of risk factors can help guide who should be screened and may reduce unnecessary testing. An efficient strategy would rule out disease with a high probability, accurately detect new cases, and minimize unnecessary testing. Negative predictive value estimates the probability of classifying individuals as disease-free and, in this case, its small numerical range (97–100%) indicates that ruling out disease is possible with most strategies. The difficulty lies in detecting the small proportion of the general population who have undiagnosed diabetes.

By focusing on those at highest risk for diabetes, the screening guidelines can accurately detect new cases and reduce unnecessary testing in the general population. If everyone were screened (100% of the adult population), all undiagnosed diabetes cases would be detected (sensitivity = 100%). If any one of the risk factors is used to indicate who should be screened, sensitivity is also 100%, but only 83% of the population would need to be screened. The presence of two or more risk factors sacrifices a small amount of sensitivity, doubles the specificity, and further reduces necessary testing from 83% to 59%. This additional 24% reduction in testing could have a large impact on the cost of testing in the United States.

Using risk factors to identify individuals with diabetes has had mixed results in previous studies primarily because of limited samples. Nonrepresentative, clinical samples or samples with little variation in ethnicity or age are not appropriate for generalizing to the U.S. population. An advantage of using the NHANES data is that the sample allows for generalization of results to the U.S. population. Previously, Herman *et al.* (24) attempted to identify individuals for screening by using the NHANES II data. The Herman *et al.* study, however, did not test the screening guidelines, and they used an *ad hoc* (non-maximum likelihood) modeling technique. They developed four rules combining two or more risk factors (age, obesity, family history, and sedentary lifestyle) and correctly identified 79% of "newly diagnosed" diabetics. In the NHANES III data, the current recommended guidelines and simpler decision rules have detection rates >95%.

The NHANES data set offers a wonderful opportunity to validate the screening guidelines on a representative U.S. sample that is sufficiently large to contain adequate numbers of undiagnosed diabetics. Conclusions based on the NHANES data, however, are limited by the set of variables included in the survey and the accuracy of those variables. One limitation is the reliance on a single FPG measurement to diagnose diabetes. Although FPG has replaced the oral glucose tolerance test as the diagnostic standard, a clinical diagnosis is confirmed by an additional test (FPG or oral glucose tolerance test) on a subsequent day (1). A single FPG would have more classification errors (higher false-negative rate) than a two-step, two-test method. Even screening criteria estimated to detect 100% of diabetes cases would miss some cases because of the false-negative rate of the diagnostic test. Another limitation is that information is unavailable for history of macrosomic infant and dysglycemia. Sensitivity of the available risk factors, however, is so high that inclusion of additional variables could not appreciably improve sensitivity, but might increase the number of people identified for screening. Even without a complete set of risk factors, almost all cases of diabetes are detected.

The Expert Committee (1) guidelines would detect almost all cases of diabetes in the general population, if followed. Although the guidelines have been adopted by several professional societies, the prevalence of undetected cases of diabetes in the general population (8) suggests that guidelines are not being followed. Efficacious guidelines exist, but implementation of the guidelines remains the greatest challenge.

We thank Alvah R. Cass, M.D., for comments throughout the development of this paper. This project was supported in part by National Science Foundation Grant BCS-0108232 (to S.C.W.).

1. The Expert Committee on the Diagnosis and Classification of Diabetes Mellitus (2002) *Diabetes Care* **25,** S5–S20.
2. Gu, K., Cowie, C. C. & Harris, M. I. (1998) *Diabetes Care* **21,** 1138–1145.
3. Arias, E. & Smith, B. L. (2003) *Deaths: Preliminary Data for 2001*, National Vital Statistics Report, March 14, 2003 (National Center for Health Statistics, Hyattsville, MD), pp. 1–45.
4. Klein, R. K. B. (1995) in *Diabetes in America*, eds. Harris, M. I., Cowie, C. C., Stern, M. P., Boyko, E. J., Reiber, G. E. & Bennett, P. H. (Natl. Inst. of Health, Bethesda), pp. 293–338.
5. Nelson, R. G., Knowler, W. C., Pettitt, D. J. & Bennett, P. H. (1995) in *Diabetes in America*, eds. Harris, M. I., Cowie, C. C., Stern, M. P., Boyko, E. J., Reiber, G. E. & Bennett, P. H. (Natl. Inst. of Health, Bethesda), pp. 349–400.
6. Reiber, G. E., Boyko, E. J. & Smith, D. G. (1995) in *Diabetes in America*, eds. Harris, M. I., Cowie, C. C., Stern, M. P., Boyko, E. J., Reiber, G. E. & Bennett, P. H. (Natl. Inst. of Health, Bethesda), pp. 409–428.
7. Hogan, P., Dall, T. & Nikolov, P. (2003) *Diabetes Care* **26,** 917–932.
8. Harris, M. I., Flegal, K. M., Cowie, C. C., Eberhardt, M. S., Goldstein, D. E., Little, R. R., Wiedmeyer, H. M. & Byrd-Holt, D. D. (1998) *Diabetes Care* **21,** 518–524.
9. Burke, J. P., Williams, K., Gaskill, S. P., Hazuda, H. P., Haffner, S. M. & Stern, M. P. (1999) *Arch. Intern. Med.* **159,** 1450–1456.
10. The Diabetes Control and Complications Trial Research Group (1993) *N. Engl. J. Med.* **329,** 977–986.
11. U.K. Prospective Diabetes Study Group (1998) *Lancet* **352,** 837–853.
12. U.K. Prospective Diabetes Study Group (1998) *Lancet* **352,** 854–865.
13. Reichard, P., Nilsson, B. Y. & Rosenqvist, U. (1993) *N. Engl. J. Med.* **329,** 304–309.
14. Wang, P. H., Lau, J. & Chalmers, T. C. (1993) *Lancet* **341,** 1306–1309.
15. Mayfield, J. (1998) *Am. Fam. Physician* **58,** 1355–1370.

16. Mazze, R. S., Strock, E. S., Simonson, G. D., Bergenstal, R. M. & Etzwiler, D. D. (2002) *Stated Diabetes Management: A Systematic Approach* (International Diabetes Center, Minneapolis).
17. Reid, M. C., Sox, H. C., Comi, R. & Atkings, D. (1996) in *Guide to Clinical Preventive Services: Report of the U.S. Preventive Services Task Force* (Williams & Wilkins, Baltimore), pp. 193–208.
18. Beaulieu, M. D. (1994) *Screening for Diabetes Mellitus in the Non-pregnant Adult* (Canadian Task Force on the Periodic Health Examination, Health Canada, Ottawa), pp. 602–609.
19. Singer, D. E., Samet, J. H., Coley, C. M. & Nathan, D. M. (1991) in *Common Screening Tests*, ed. Eddy, D. M. (Am. College of Physicians, Philadelphia), pp. 154–178.
20. Harris, M. I., Klein, R., Welborn, T. A. & Knuiman, M. W. (1992) *Diabetes Care* **15,** 815–819.
21. Azzopardi, J., Fenech, F. F., Junoussov, Z., Mazovetsky, A. & Olchanski, V. (1995) *Diabetic Med.* **12,** 271–276.
22. Baan, C. A., Ruige, J. B., Stolk, R. P., Witteman, J. C., Dekker, J. M., Heine, R. J. & Feskens, E. J. (1999) *Diabetes Care* **22,** 213–219.
23. Griffin, S. J., Little, P. S., Hales, C. N., Kinmonth, A. L. & Wareham, N. J. (2000) *Diabetes Metab. Res. Rev.* **16,** 164–171.
24. Herman, W. H., Smith, P. J., Thompson, T. J., Engelgau, M. M. & Aubert, R. E. (1995) *Diabetes Care* **18,** 382–387.
25. Lindstrom, J. & Tuomilehto, J. (2003) *Diabetes Care* **26,** 725–731.
26. Park, P. J., Griffin, S. J., Sargeant, L. & Wareham, N. J. (2002) *Diabetes Care* **25,** 984–988.
27. Rolka, D. B., Narayan, K. M., Thompson, T. J., Goldman, D., Lindenmayer, J., Alich, K., Bacall, D., Benjamin, E. M., Lamb, B., Stuart, D. O., *et al.* (2001) *Diabetes Care* **24,** 1899–1903.
28. Ruige, J. B., de Neeling, J. N., Kostense, P. J., Bouter, L. M. & Heine, R. J. (1997) *Diabetes Care* **20,** 491–496.
29. Tabaei, B. P. & Herman, W. H. (2002) *Diabetes Care* **25,** 1999–2003.
30. Cowie, C. C., Harris, M. I. & Eberhardt, M. S. (1994) *Diabetes Care* **17,** 1158–1163.
31. Carter, J. S., Pugh, J. A. & Monterrosa, A. (1996) *Ann. Intern. Med.* **125,** 221–232.
32. U.S. Department of Health and Human Services (1996) *Third National Health and Nutrition Examination Survey, 1988–1994*, National Center for Health Statistics, Centers for Disease Control and Prevention (Centers for Disease Control and Prevention, Hyattsville, MD), NHANES III Laboratory Data File (CD-ROM), Public Use Data File Documentation No. 76200.
33. U.S. Department of Health and Human Services (1996) *Third National Health and Nutrition Examination Survey, 1988–1994*, National Center for Health Statistics, Centers for Disease Control and Prevention (Centers for Disease Control and Prevention, Hyattsville, MD), NHANES III Examination Data File (CD-ROM), Public Use Data File Documentation No. 76200.
34. National Heart, Lung, and Blood Institute (2000) *The Practical Guide: Identification, Evaluation, and Treatment of Overweight and Obesity in Adults* (Natl. Inst. of Health, Bethesda), NIH Publication No. 00-4084, pp. 1–44.
35. Hosmer, D. W. & Lemeshow, S. (2000) *Applied Logistic Regression* (Wiley, New York).
36. SAS Institute (2000) SAS *Procedures Guide* (SAS Inst., Cary, NC), Version 8.1.
37. Shah, B. V., Barnwell, B. G. & Bieler, G. S. (2000) SUDAAN *User's Manual* (Research Triangle Institute, Research Triangle Park, NC), Version 7.5.6.
38. Weinstein, M. C., Fineberg, H. V., Elstein, A. S., Frazier, H. S., Neuhauser, D., Neutra, R. R. & McNeill, B. J. (1980) *Clinical Decision Analysis* (Saunders, Philadelphia).
39. Expert Committee on the Diagnosis and Classification of Diabetes Mellitus (2000) *Diabetes Care* **23,** Suppl. 1, S4–S19.