Prediction of Incident Diabetes Mellitus in Middle-aged Adults

The Framingham Offspring Study

Peter W. F. Wilson, MD; James B. Meigs, MD, MPH; Lisa Sullivan, PhD; Caroline S. Fox, MD, MPH; David M. Nathan, MD; Ralph B. D'Agostino, Sr, PhD

Background: Prediction rules for type 2 diabetes mellitus (T2DM) have been developed, but we lack consensus for the most effective approach.

Methods: We estimated the 7-year risk of T2DM in middle-aged participants who had an oral glucose tolerance test at baseline. There were 160 cases of new T2DM, and regression models were used to predict new T2DM, starting with characteristics known to the subject (personal model, ie, age, sex, parental history of diabetes, and body mass index [calculated as the weight in kilograms divided by height in meters squared]), adding simple clinical measurements that included metabolic syndrome traits (simple clinical model), and, finally, assessing complex clinical models that included (1) 2-hour post–oral glucose tolerance test glucose, fasting insulin, and C-reactive protein levels; (2) the Gutt insulin sensitivity index; or (3) the homeostasis model insulin resistance and the homeostasis model insulin resistance β-cell sensitivity indexes. Discrimination was assessed with area under the receiver operating characteristic curves (AROCs).

Results: The personal model variables, except sex, were statistically significant predictors of T2DM (AROC, 0.72). In the simple clinical model, parental history of diabetes and obesity remained significant predictors, along with hypertension, low levels of high-density lipoprotein cholesterol, elevated triglyceride levels, and impaired fasting glucose findings but not a large waist circumference (AROC, 0.85). Complex clinical models showed no further improvement in model discriminations (AROC, 0.850-0.854) and were not superior to the simple clinical model.

Conclusion: Parental diabetes, obesity, and metabolic syndrome traits effectively predict T2DM risk in a middle-aged white population sample and were used to develop a simple T2DM prediction algorithm to estimate risk of new T2DM during a 7-year follow-up interval.

Arch Intern Med. 2007;167:1068-1074

Author Affiliations: Emory Program in Cardiovascular Outcomes Research and Epidemiology, Emory University School of Medicine, Atlanta, Ga (Dr Wilson); General Medicine Division (Dr Meigs) and Diabetes Center and Department of Medicine (Drs Meigs and Nathan), Massachusetts General Hospital and Harvard Medical School, and Department of Mathematics, Boston University (Drs Sullivan and D'Agostino), Boston, Mass; and National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Mass (Dr Fox).

THE OCCURRENCE OF TYPE 2 diabetes mellitus (T2DM) is rising rapidly among middle-aged American adults. It has been estimated that the prevalence of diabetes in the United States increased from 7.3% in 1993 to 7.9% by the year 2000, and greater frequencies are forecast for the future.[1]

Prediction of chronic conditions like T2DM that have a definable onset can help to guide interventions and health policy development. Such a course has been followed for the prediction of coronary heart disease,[2] and similar effects might be obtained with effective prognostication and testing for T2DM.[3] For example, a Diabetes Risk Score Test has been developed that estimates the risk of T2DM on the basis of the birth of a child with macrosomia, parents with diabetes, excess adiposity, self-report of little exercise, and age category.[4] However, the validity of this model has not been fully assessed in diverse populations and in large cohorts followed up for the development of incident T2DM. As another example, a diabetes-predicting model has been developed in high-risk Mexican Americans[5] and further tested in Japanese Americans.[6] These models use a variety of T2DM risk factors to generate a prediction score, including parental history of diabetes and the presence of excess adiposity.[4] Complex algorithms have also been developed that use more than 50 variables to predict the risk of diabetes.[7,8] In addition to age, excess adiposity, and family history, recent research has suggested a large variety of metabolic factors that are potentially involved in the pathophysiology of T2DM.[9,10]

Our investigation predicts the development of diabetes in middle-aged adults during a follow-up interval of 7 years, defining T2DM by fasting and 2-hour post–glucose load criteria at baseline, starting diabetes medication therapy during follow-up, and fasting glucose level at the end of follow-up. We focused on developing a parsimonious prediction model, using a series of perspectives that started simply and considered higher levels of complexity.

METHODS

The population sample included 3140 men and women who attended the fifth clinic examination of the Framingham Offspring Study in the mid-1990s. This population sample is 99% white and non-Hispanic. The baseline examination included information on medication use and self-reported parental history of T2DM, defined as diabetes in one or both natural parents.[11,12] The physical examination included blood pressure measured in the sitting position, height and weight measurements, and waist circumference determined at the umbilicus with the subject standing. Body mass index (BMI) was calculated by dividing the weight in kilograms by the height in meters squared.

Subjects were examined after an overnight fast and had a 2-hour oral glucose tolerance test (OGTT). Persons with a history of diabetes mellitus, who used oral hypoglycemic medications or insulin, or who had a baseline fasting plasma glucose level greater than 126 mg/dL (>7.0 mmol/L) or a baseline post-OGTT plasma glucose level greater than 200 mg/dL (>11.1 mmol/L) were categorized as having diabetes and were not included in this study. An OGTT 2-hour glucose level of 140 to 200 mg/dL (7.8-10.9 mmol/L) defined impaired glucose tolerance. Other laboratory measurements included levels of fasting and 2-hour OGTT insulin (determined using a commercially available assay [DPC Coat-a-Count; Diagnostics Products Corporation, Los Angeles, Calif]), total cholesterol, high-density lipoprotein cholesterol (HDL-C), triglycerides, and C-reactive protein (determined using a commercially available assay [Hemagen Diagnostics Inc, Waltham, Mass]), as previously described.[11,13]

More sophisticated indexes of glucose and insulin control included calculation of the homeostasis model (HOMA) insulin resistance index, the HOMA β-cell index as a measure of reserve pancreatic insulin production, and the Gutt insulin sensitivity index, which includes body weight and OGTT glucose and insulin information and is similar to a glucose disposition index.[14-16] Persons were categorized according to the presence or absence of the metabolic syndrome traits described by the National Cholesterol Education Program Adult Treatment Panel III criteria.[17] Participants with a blood pressure level of 130/85 mm Hg or higher or receiving treatment for hypertension were considered to have elevated blood pressure; those with a fasting glucose level of 100 to 126 mg/dL (5.4-6.9 mmol/L) were considered to have fasting hyperglycemia; a waist circumference greater than 102 cm in men or more than 88 cm in women was considered increased; a fasting triglyceride level of 150 mg/dL or greater (≥1.7 mmol/L) was considered hypertriglyceridemia; and an HDL-C level less than 40 mg/dL (<0.9 mmol/L) in men or less than 50 mg/dL (<1.2 mmol/L) in women was considered low.[18,19]

Participants were followed up from baseline to the sixth (1995-1998) and seventh (1998-2001) Framingham Offspring Study examinations for an average follow-up of 7 years. We used the examination visit date that a new case of diabetes was identified as the date of diagnosis; otherwise follow-up was censored at the last follow-up (examination 6 or 7) for participants remaining nondiabetic. Participants were characterized as developing new diabetes during follow-up if they (1) started receiving oral hypoglycemic agents or insulin or (2) had a fasting glucose level of 126 mg/dL or greater (≥7.0 mmol/L) at 1 of the follow-up Framingham Offspring Study examinations conducted 4 and 7 years after the baseline examination.

Statistical analyses included a series of logistic regression models to predict incident diabetes, using the odds ratio and 95% confidence intervals to estimate relative risk. Alternate analyses using Cox proportional hazards models that accounted for interval censoring gave essentially identical results; only logistic regression results are presented. The rationale for separate models to estimate T2DM risk for diabetes was predicated on evaluation of 3 major levels of health information. The first level, the personal model, was based on information known to an individual without seeking medical advice. The second level, the simple clinical model, was based on personal model variables plus information typically available at a clinic visit with a physician. The third level, a set of complex clinical models, incorporated simple clinical model covariates plus information that is available only with more detailed clinical testing, including data from an OGTT and measurement of insulin levels and inflammatory markers.

Our regression models sequentially included the personal, simple clinical, and more complex clinical models, with evaluation of the discriminatory capability of the models using the C statistic, or the area under the receiver operating characteristic curve (AROC). Between-model comparisons were evaluated by ranking participant risk by decile and performing a χ^2 analysis on the estimates as per Hosmer and Lemeshow.[20] Score sheets to estimate absolute risk for the outcome were derived from the β coefficients of the multivariate logistic regression analysis, as described previously.[2,21]

Table 1. Baseline Characteristics

Factor	Finding*
No. of subjects	3140
Age, mean (SD), y	54.0 (9.8)
Female	53.9
Parental history of diabetes	17.0
BMI, Mean (SD)	27.1 (4.7)
Blood pressure >130/85 mm Hg or hypertension therapy	44.2
HDL-C level <40 mg/dL in men or <50 mg/dL in women	36.9
Triglyceride level ≥150 mg/dL	31.8
Waist circumference >88 cm in women or >102 cm in men	33.6
Fasting plasma glucose level 100-126 mg/dL	27.0
2-Hour OGTT finding 140-200 mg/dL	11.6

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); HDL-C, high-density lipoprotein cholesterol; OGTT, oral glucose tolerance test.

SI conversion factors: To convert HDL-C to millimoles per liter, multiply by 0.0259; glucose to millimoles per liter, multiply by 0.0555; triglycerides to millimoles per liter, multiply by 0.0113.

*Unless otherwise indicated, data are expressed as percentage of subjects.

RESULTS

Among participants attending the baseline examination there were 3.2% with known T2DM, and another 4.7% had T2DM diagnosed by an OGTT result. Those persons were removed from the study group, and the baseline characteristics of the nondiabetic attendees who had an OGTT at baseline are shown in **Table 1**. We included 3140 men and women with a mean age of 54.0 years. Approximately half of the participants were women, the average BMI was 27.1, and impaired glucose tolerance was present in 12.7%. The personal model for diabetes prediction is shown in **Table 2**. Categories of age, sex, parental history of diabetes mellitus, and BMI were considered candidate variables for this model. The age and BMI categories included more than 1 category; being younger than 50 years and having a BMI of less than 25.0 were considered the referent categories. In the multivariate analyses, higher categories of age and BMI and a parental history of diabetes mellitus

were significantly related to development of diabetes during the follow-up interval.

Table 3 shows the results for a multivariate analysis that considered the development of T2DM using 3 similar simple clinical models. These models, using information typically available at a clinic evaluation, differed only in the inclusion of terms for BMI, waist circumference, or both. The variables included the personal information used in the Table 2 analyses as well as an elevated blood pressure, a low HDL-C level, an elevated triglyceride level, an impaired fasting glucose level, and an obesity measure. In this analysis, a significant statistical association with incident diabetes was evident for a parental history of diabetes, elevated blood pressure, low HDL-C level, elevated triglyceride level, and impaired fasting glucose level. The BMI or waist circumference, but not both together, were significant predictors of diabetes development.

Table 2. Multivariate Prediction of T2DM According to Personal Variables

Variable	OR (95% CI)	*P* Value
Age, y		
<50	1 [Reference]	
50-64	1.54 (1.04-2.27)	.03
≥65	1.74 (1.06-2.85)	.03
Male	1.25 (0.89-1.74)	.20
Parental history of diabetes	1.87 (1.28-2.72)	.001
BMI		
<25.0	1 [Reference]	
25.0-29.9	2.35 (1.39-3.96)	.001
≥30.0	6.41 (3.85-10.65)	<.001
Intercept	−4.499	
AROC	0.724	

Abbreviations: AROC, area under the receiver operating characteristic curve; BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); CI, confidence interval; OR, odds ratio; T2DM, type 2 diabetes mellitus.

Each prediction model with BMI alone or waist circumference alone showed that the adiposity measure was statistically related to the development of diabetes during follow-up (Table 3). The prediction model that included 3 BMI categories (<25.0, 25.0-29.9, and ≥30.0) and 2 sex-specific waist categories (normal and increased) did not appreciably increase the ability to discriminate future cases of diabetes, and the increased waist circumference variable was not statistically significant in the model that included the BMI categories. Overall, the AROC for all of these models was approximately 0.85, which indicates an excellent capability to discriminate persons who developed diabetes from those who did not, with virtually no difference in the model's predictive capability according to use of waist circumference or BMI categorical approaches. Results of the simple clinical model using covariates as continuously distributed appear in **Table 4**. Use of predictor variables as continuously distributed yielded better discrimination (AROC, 0.881) than did use of categorical covariates (AROC, 0.852, 0.850, and 0.852, depending on the model).

Results of the complex clinical models are shown in **Table 5**. Each model included the variables in the simple clinical model plus additional factors. The first included impaired glucose tolerance, elevated fasting insulin (≥75th percentile), and C-reactive protein (≥75th percentile) levels; the second included the Gutt insulin sensitivity index (≤25th percentile, in which low values indicate insulin resistance); and the third included the HOMA insulin resistance index (≥75th percentile, in which high values indicate insulin resis-

Table 3. Multivariate Prediction of T2DM According to Simple Clinical Variables

	Simple Clinical Model					
	Obesity by BMI Only		Obesity by Waist Circumference Only		Obesity by BMI and Waist Circumference	
Variable	OR (95% CI)	*P* Value	OR (95% CI)	*P* Value	OR (95% CI)	*P* Value
Age, y						
<50	1 [Reference]		1 [Reference]		1 [Reference]	
50-64	0.98 (0.64-1.50)	.93	0.94 (0.62-1.44)	.79	0.98 (0.64-1.50)	.90
≥65	0.92 (0.54-1.59)	.77	0.83 (0.49-1.43)	.51	0.91 (0.53-1.56)	.70
Male	0.99 (0.70-1.41)	.95	1.09 (0.77-1.54)	.65	1.05 (0.73-1.50)	.80
Parental history of diabetes mellitus	1.76 (1.17-2.64)	.006	1.75 (1.17-2.61)	.007	1.78 (1.19-2.67)	.005
BMI						
<25.0	1 [Reference]		Not included		1 [Reference]	
25.0-29.9	1.35 (0.78-2.34)	.28	Not included		1.21 (0.68-2.14)	.50
≥30.0	2.50 (1.45-4.30)	.001	Not included		1.86 (0.94-3.67)	.07
Blood pressure >130/85 mm Hg or receiving therapy	1.65 (1.10-2.46)	.02	1.73 (1.16-2.59)	.007	1.62 (1.08-2.43)	.02
HDL-C level <40 mg/dL in men or <50 mg/dL in women	2.57 (1.75-3.77)	<.001	2.62 (1.79-3.84)	<.001	2.55 (1.74-3.74)	<.001
Triglyceride level ≥150 mg/dL	1.78 (1.22-2.59)	.003	1.78 (1.23-2.59)	.002	1.75 (1.20-2.56)	.004
Waist circumference >102 cm in men or >88 cm in women	Not included		1.98 (1.37-2.84)	<.001	1.42 (0.88-2.29)	.20
Fasting glucose level 100-126 mg/dL	7.25 (4.89-10.74)	<.001	7.17 (4.86-10.58)	<.001	7.16 (4.83-10.61)	<.001
Intercept	−5.517		−5.434		−5.363	
AROC	0.852		0.850		0.852	

Abbreviations: AROC, area under the receiver operating characteristic curve; BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); CI, confidence interval; HDL-C, high-density lipoprotein cholesterol; OR, odds ratio; T2DM, type 2 diabetes mellitus.

SI conversion factors: To convert HDL-C to millimoles per liter, multiply by 0.0259; glucose to millimoles per liter, multiply by 0.0555; triglycerides to millimoles per liter, multiply by 0.0113.

tance) and the HOMA β-cell index (≤25th percentile, in which low values indicate impaired β-cell function). In each of these models, the relative risks for the individual simple model variables were typically lower than in the simple clinical model, and more sophisticated measures of hyperinsulinemia or insulin resistance were related to the development of diabetes. Inflammation as measured by elevated C-reactive protein levels was not an independent predictor of incident diabetes. The AROCs for the complex clinical models ranged from 0.850 to 0.854 (Table 5). These results were commensurate with the AROC for the simple clinical model shown in Table 3, indicating that the more complex models did not provide additional capability to discriminate persons who developed diabetes from those who did not, even when additional covariates in the complex clinical model were significantly associated with incident cases of diabetes. The **Figure** compares the receiver operating characteristics for the personal, simple clinical, and complex clinical models, showing graphically how the AROC was appreciably less for the personal model analysis and very similar for the other models.

We also investigated the predictive ability of a "best biological model" that included all of the variables in Table 5 and current hormone therapy, current smoking, current weekly alcohol intake, current aspirin or nonsteroidal anti-inflammatory drug use, hemoglobin A_{1C} level, the HOMA insulin resistance index, the Gutt insulin sensitivity index, and the HOMA β-cell index. The AROC for this model was 0.869, and the statistically significant ($P \leq .05$) variables included age of 50 to 65 years, age of 65 years or older, a low HDL-C level, a fasting glucose level of 100 to 126 mg/dL (5.4-6.9 mmol/L), the HOMA insulin resistance index, and the Gutt insulin sensitivity index.

The within-study prediction model validity was assessed using a jackknife procedure. We took 10 random

Table 4. Multivariate Prediction of T2DM According to Continuous Variables

Variable	OR (95% CI)	*P* Value
Age, y	0.99 (0.97-1.01)	.42
Male	0.65 (0.41-1.02)	.06
Parental history of diabetes mellitus, yes/no	1.55 (1.01-2.38)	.04
BMI	1.04 (0.97-1.11)	.24
Systolic blood pressure, mm Hg	1.01 (1.00-1.02)	.11
HDL-C level per mg/dL	0.96 (0.95-0.98)	<.001
Triglyceride level per mg/dL	1.00 (1.00-1.00)	.16
Waist circumference, cm	1.05 (0.97-1.12)	.22
Fasting glucose level, mg/dL*	1.15 (1.12-1.17)	<.001
Intercept	–18.607	. . .
AROC	0.881	. . .

Abbreviations: AROC, area under the receiver operating characteristic curve; BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); CI, confidence interval; HDL-C, high-density lipoprotein cholesterol; OR, odds ratio; T2DM, type 2 diabetes mellitus; ellipses, not applicable.

SI conversion factors: To convert HDL-C to millimoles per liter, multiply by 0.0259; glucose to millimoles per liter, multiply by 0.0555; triglycerides to millimoles per liter, multiply by 0.0113.

*Indicates fasting glucose values of less than 126 mg/dL (<7.0 mmol/L) only.

Table 5. Multivariate Prediction of T2DM According to Complex Clinical Variables

	Complex Clinical Model 1		Complex Clinical Model 2		Complex Clinical Model 3	
Variable	OR	*P* Value	OR	*P* Value	OR	*P* Value
Age, y						
<50	1 [Reference]		1 [Reference]		1 [Reference]	
50-64	1.02	.95	0.99	.96	1.03	.88
≥65	0.83	.53	0.81	.48	0.88	.67
Male	1.25	.26	1.12	.55	1.01	.98
Parental history of diabetes mellitus	1.63	.02	1.73	.01	1.71	.01
BMI						
<25.0	1 [Reference]		1 [Reference]		1 [Reference]	
25.0-29.9	1.08	.80	1.17	.61	1.19	.57
≥30.0	1.32	.45	1.80	.10	1.68	.15
Blood pressure >130/85 mm Hg or receiving therapy	1.53	.05	1.40	.13	1.58	.03
HDL-C level <40 mg/dL in men or <50 mg/dL in women	2.33	<.001	2.18	<.001	2.18	<.001
Triglyceride level ≥150 mg/dL	1.45	.07	1.50	.05	1.57	.03
Waist circumference >88 cm in women or >102 cm in men	1.32	.28	1.25	.38	1.27	.35
Fasting glucose level 100-126 mg/dL	5.37	<.001	5.32	<.001	5.09	<.001
2-Hour OGTT finding 140-200 mg/dL	2.87	<.001	NI		NI	
Fasting insulin level >75th percentile	1.23	0.33	NI		NI	
C-reactive protein level >75th percentile	1.43	0.07	NI		NI	
Log Gutt insulin sensitivity index <25th percentile	NI		2.28	<.001	NI	
Log HOMA insulin resistance index >75th percentile	NI		NI		2.05	.001
HOMA β-cell index <25th percentile	NI		NI		1.88	.002
Intercept	–5.506		–5.427		–5.620	
AROC	0.854		0.850		0.851	

Abbreviations: AROC, area under the receiver operating characteristic curve; BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); HDL-C, high-density lipoprotein cholesterol; HOMA, homeostasis model; NI, not included; OGTT, oral glucose tolerance test; OR, odds ratio; T2DM, type 2 diabetes mellitus.

SI conversion factors: To convert HDL-C to millimoles per liter, multiply by 0.0259; glucose to millimoles per liter, multiply by 0.0555; triglycerides to millimoles per liter, multiply by 0.0113.

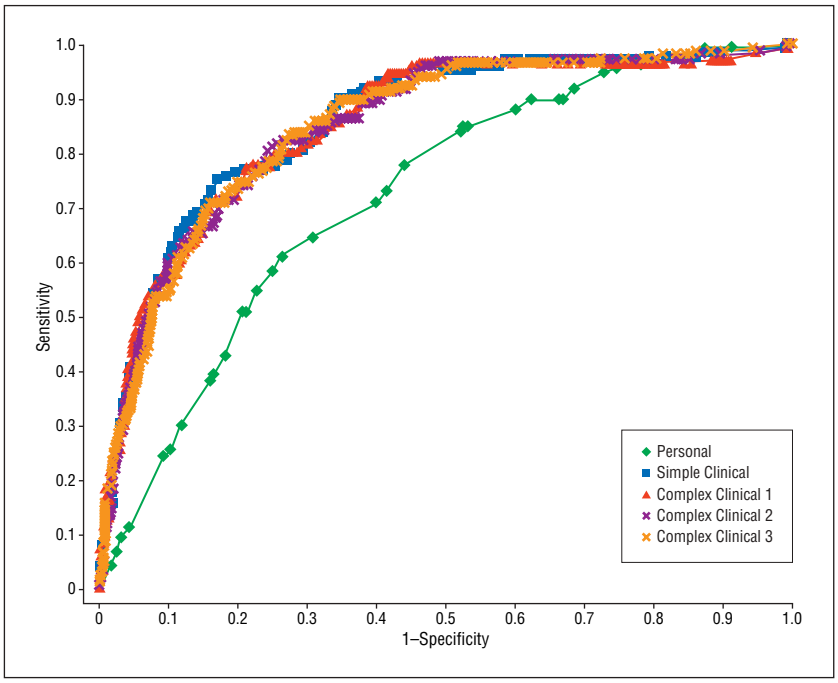

Figure. The receiver operating characteristic curves for the personal, simple clinical, and 3 complex clinical models.

Table 6. Algorithm to Estimate Risk for T2DM Using Simple Clinical Model*

Items	Item Points	Item Point Total	8-Year Risk of T2DM, %
Fasting glucose level 100-126 mg/dL, yes/no	10	≤10	≤3
BMI 25.0-29.9, yes/no	2	11	4
BMI ≥30.0, yes/no	5	12	4
HDL-C level <40 mg/dL in men or <50 mg/dL in women, yes/no	5	13	5
Parental history of diabetes mellitus, yes/no	3	14	6
Triglyceride level ≥150 mg/dL, yes/no	3	15	7
Blood pressure ≥130/85 mm Hg or receiving treatment, yes/no	2	16	9
Item Point Total		17	11
		18	13
		19	15
		20	18
		21	21
		22	25
		23	29
		24	33
		≥25	>35

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); HDL-C, high-density lipoprotein cholesterol; T2DM, type 2 diabetes mellitus.

SI conversion factors: To convert HDL-C to millimoles per liter, multiply by 0.0259; glucose to millimoles per liter, multiply by 0.0555; triglycerides to millimoles per liter, multiply by 0.0113.

*The unshaded portion of the table shows the possible item point totals and their corresponding 8-year risk of T2DM. After calculating the item point total, check the last 2 columns to determine the risk of T2DM.

samples of 90.0% of the participants to test the ability of that sample to discriminate future diabetes cases.[21] The AROCs for these 10 iterations ranged from 0.73 to 0.91, indicating a high reliability of discrimination for the model in repeated random-sample subsets. We developed a point score system to estimate diabetes risk using the intercept and the β coefficients of the simple clinical model that used BMI as the adiposity measure. This approach allows manual estimation of the 8-year risk of developing diabetes, as shown in **Table 6**. An impaired fasting glucose finding (10 points), a BMI of 30.0 or greater (5 points), and a low HDL-C level (5 points) had the greatest effects in the point scores, and successively smaller effects were evident for a positive parental history (2 points), a triglyceride level greater than 150 mg/dL (>1.7 mmol/L) (2 points), a BMI of 25.0 to 29.9 (2 points), and elevated blood pressure (2 points). The age categories and sex were not related to development of diabetes, and no age or sex variables were included in the point calculations. By using the point score, we determined that 63.8% of the sample had a less than 3% risk, 20.7% had a 3% to 10% risk, and 15.6% had a greater than 10% risk of incident diabetes during an 8-year interval. The AROC for the point score prediction was 0.850.

COMMENT

The principal finding of this study is that some information beyond personal awareness of diabetes risk factors is important to determine risk of T2DM, but complex models are not needed. We started with a personal model that included information generally available to persons before a clinic visit with a physician. The simple clinical model that included common metabolic traits efficiently identified subjects at elevated risk for T2DM diabetes, suggesting that clinical screening adds value beyond screening using personal knowledge alone. Consideration of 3 complex clinical models showed negligible improvement in assessment of T2DM risk over and above the simple clinical model.

Others who have reported using questionnaire data in cross-sectional studies to identify persons with undetected T2DM or to increase the yield of glucose testing have found that greater age, higher BMI, and ethnicity were especially important predictors.[22,23] Where investigated, hypertension history, physical activity, and parental history of diabetes have been shown to be predictive of an abnormal OGTT result.[24]

To test the utility of questionnaire data, investigators in Cambridge, England, undertook an external validation study and found that age, sex, BMI, use of corticosteroids and antihypertensives, smoking, and parental history of diabetes mellitus were predictive elements of prevalent T2DM with an AROC of 0.80.[25] Other longitudinal studies have used the Cambridge risk score prognostically to track individuals for deterioration in glycemic status, shown by a hemoglobin A_{1c} level of greater than 7.0%, and have shown good success in predicting such deterioration in glycemia with an AROC of 0.74 in a British cohort.[26]

Others have examined these factors as determinants of T2DM in cross-sectional and longitudinal studies, including the San Antonio Heart Study,[27] Insulin Resistance Atherosclerosis Study,[28,29] Rancho Bernardo,[30] and Munster[31] cohorts. An analysis of the Atherosclerosis Risk in Communities data, which included more than 7900 adults aged 45 to 64 years, showed a high degree of model discrimination of future T2DM cases during follow-up when the National Cholesterol Education Program metabolic syndrome variable count of 0 to 5 was used in the analysis (AROC, 0.78).[32] The authors concluded that rules based on the metabolic syndrome are reasonable alternatives for estimating risk for T2DM,[32] similar to a recent report from the Framingham experience[33] in which the National Cholesterol Education Program metabolic syndrome trait count was highly related to a greater risk for developing T2DM.

The traits that constitute the metabolic syndrome are especially important in the determination of risk for T2DM.[33] The simple clinical models presented in Table 3 show that each of the metabolic syndrome traits is highly associated with the development of T2DM, and T2DM risk varies considerably across the variables. This supports a predictive approach wherein each variable should be used individually.

Because the simple clinical approach represented an easy and effective approach to estimate risk for the development of incident diabetes, we transformed the simple clinical model that used BMI into a point score that can be used in the office setting. As has been the case for prediction of coronary heart disease,[2,21] the availability of a simple clinical tool to estimate disease risk should improve the prediction of events and enhance prevention strategies.

Others have undertaken to predict or identify risk of diabetes mellitus with a variety of approaches. The American Diabetes Association prediction algorithm is based on the experience of the second National Health and Nutrition Examination Survey.[4] The American Diabetes Association model used a decision tree, and a point score was developed to estimate risk. The authors reported an AROC of 0.78 with this approach.[4] The key variables in that formulation were the birth of a child with macrosomia, obesity, sedentary lifestyle, and a parental history of diabetes mellitus. Their approach used self-reported personal information that identified individuals cross-sectionally and did not predict incident diabetes over time.

San Antonio researchers have developed a diabetes prediction rule that included simple clinical variables.[5] Their model predicted the development of T2DM during a 7.5-year interval, and the key prediction variables were age, sex, Mexican American ethnicity, fasting plasma glucose level, systolic blood pressure, HDL-C level, BMI, and parental history of diabetes. In South Texans of Hispanic descent, the absolute risk for T2DM is much greater than in white subjects from suburban Boston or in Europe; the discriminatory capacity of their approach was high (AROC, 0.843-0.845), and the metabolic syndrome variables included were important diabetes predictors. Just as in the present Framingham analyses, more sophisticated measures, such as postchallenge plasma glucose level, did not add to the discriminatory capacity of more simple models. The utility of the San Antonio model has been tested in a German cross-sectional cohort and in a Japanese American prospective cohort.[6] In the latter setting, the authors reported that the multivariate clinical model was better than the fasting glucose level for predicting development of T2DM after 5 or 6 years, but not after 10 years; the clinical model's predictive capability was similar to the predictive capability of the fasting or 2-hour glucose level in older Japanese Americans.

Investigators from Finland developed a diabetes risk score and predicted T2DM during 5 years of follow-up in a middle-aged population sample that identified cases by initiation of diabetes medications.[34] They found that age, BMI, waist circumference, history of blood pressure therapy, high blood glucose level, physical activity, and dietary components were predictive of events. This approach, with separate identification and weighting of metabolic factors, most closely parallels the results we obtained with the simple clinical model, but they did not use a formal OGTT at the beginning of their study.

In summary, we found that complex models are not needed to predict T2DM and that information from a typical clinic visit adds to T2DM prediction beyond personal awareness of diabetes risk factors. The simple clinical model we developed should be tested in other population samples to validate our approach, as has been done for prediction of coronary heart disease events.[21,35,36]

Accepted for Publication: February 1, 2007.
Correspondence: Peter W. F. Wilson, MD, Emory Program in Cardiovascular Outcomes Research and Epidemi-

ology, Emory University School of Medicine, 1256 Briarcliff Rd, Suite 1N, Atlanta, GA 30306 (pwwilso@emory.edu).
Author Contributions: *Study concept and design:* Wilson, Meigs, and Nathan. *Acquisition of data:* Wilson, Meigs, and Nathan. *Analysis and interpretation of data:* Wilson, Meigs, Sullivan, Fox, and D'Agostino. *Drafting of the manuscript:* Wilson, Meigs, and Sullivan. *Critical revision of the manuscript for important intellectual content:* Wilson, Meigs, Sullivan, Fox, Nathan, and D'Agostino. *Statistical analysis:* Sullivan and D'Agostino. *Obtained funding:* Meigs and Nathan. *Administrative, technical, and material support:* Wilson, Fox, Nathan, and D'Agostino. *Study supervision:* Wilson, Nathan, and D'Agostino.
Financial Disclosure: Dr Wilson is a consultant to adjudicate outcomes for a Lilly clinical trial and has research grants from GlaxoSmithKline and Sanofi-Aventis. Dr Meigs has research grants from GlaxoSmithKline, Wyeth, and Sanofi-Aventis and serves on safety or advisory boards for GlaxoSmithKline, Merck, and Lilly.
Funding/Support: This study was supported by a grant from the Centers for Disease Control and Prevention; a cooperative agreement with the American Association of Medical Colleges; a Career Development Award from the American Diabetes Association (Dr Meigs); the Visiting Scientist Program, which is supported by Astra USA, Hoechst Marion Roussel, and Sevier Canada, Inc (Dr D'Agostino); the Charlton Family Trust (Dr Nathan); and contract N01-HC-25195 from the National Heart, Lung, and Blood Institute's Framingham Heart Study.

REFERENCES

1. Mokdad AH, Ford ES, Bowman BA, et al. Diabetes trends in the U.S.: 1990-1998. *Diabetes Care*. 2000;23:1278-1283.
2. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998; 97:1837-1847.
3. Smith SC Jr, Jackson R, Pearson TA, et al. Principles for national and regional guidelines on cardiovascular disease prevention: a scientific statement from the World Heart and Stroke Forum. *Circulation*. 2004;109:3112-3121.
4. Herman WH, Smith PJ, Thompson TJ, Engelgau MM, Aubert RE. A new and simple questionnaire to identify people at increased risk for undiagnosed diabetes. *Diabetes Care*. 1995;18:382-387.
5. Stern MP, Williams K, Haffner SM. Identification of persons at high risk for type 2 diabetes mellitus: do we need the oral glucose tolerance test? *Ann Intern Med*. 2002;136:575-581.
6. McNeely MJ, Boyko EJ, Leonetti DL, Kahn SE, Fujimoto WY. Comparison of a clinical model, the oral glucose tolerance test, and fasting glucose for prediction of type 2 diabetes risk in Japanese Americans. *Diabetes Care*. 2003;26:758-763.
7. Eddy DM, Schlessinger L. Archimedes: a trial-validated model of diabetes. *Diabetes Care*. 2003;26:3093-3101.
8. Eddy DM, Schlessinger L. Validation of the Archimedes diabetes model. *Diabetes Care*. 2003;26:3102-3110.
9. Ginsberg HN, Stalenhoef AF. The metabolic syndrome: targeting dyslipidaemia to reduce coronary risk. *J Cardiovasc Risk*. 2003;10:121-128.
10. Eckel RH, Grundy SM, Zimmet PZ. The metabolic syndrome. *Lancet*. 2005;365: 1415-1428.
11. Rutter MK, Meigs JB, Sullivan LM, D'Agostino RB Sr, Wilson PW. C-reactive protein, the metabolic syndrome, and prediction of cardiovascular events in the Framingham Offspring Study. *Circulation*. 2004;110:380-385.
12. Murabito JM, Nam BH, D'Agostino RB Sr, Lloyd-Jones DM, O'Donnell CJ, Wilson PW. Accuracy of offspring reports of parental cardiovascular disease history: the Framingham Offspring Study. *Ann Intern Med*. 2004;140:434-440.
13. Meigs JB, Mittleman MA, Nathan DM, et al. Hyperinsulinemia, hyperglycemia, and impaired hemostasis: the Framingham Offspring Study. *JAMA*. 2000;283: 221-228.
14. Gutt M, Davis CL, Spitzer SB, et al. Validation of the insulin sensitivity index ($ISI_{0,120}$): comparison with other measures. *Diabetes Res Clin Pract*. 2000;47:177-184.
15. Hanley AJ, Wagenknecht LE, D'Agostino RB Jr, Zinman B, Haffner SM. Identification of subjects with insulin resistance and β-cell dysfunction using alternative definitions of the metabolic syndrome. *Diabetes*. 2003;52:2740-2747.
16. Hanley AJ, Williams K, Stern MP, Haffner SM. Homeostasis model assessment of insulin resistance in relation to the incidence of cardiovascular disease: the San Antonio Heart Study. *Diabetes Care*. 2002;25:1177-1184.
17. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. 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*. 2001; 285:2486-2497.
18. Kahn R, Buse J, Ferrannini E, Stern M; American Diabetes Association; European Association for the Study of Diabetes. The metabolic syndrome: time for a critical appraisal: joint statement from the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care*. 2005;28:2289-2304.
19. Genuth S, Alberti KG, Bennett P, et al. Follow-up report on the diagnosis of diabetes mellitus. *Diabetes Care*. 2003;26:3160-3167.
20. Hosmer DW, Lemeshow S. *The Multiple Logistic Regression Model: Applied Logistic Regression*. New York, NY: John Wiley & Sons Inc; 1989:25-37.
21. D'Agostino RB Sr, Grundy S, Sullivan LM, Wilson P. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. *JAMA*. 2001;286:180-187.
22. Diabetes Prevention Program Research Group. Strategies to identify adults at high risk for type 2 diabetes: the Diabetes Prevention Program. *Diabetes Care*. 2005;28:138-144.
23. Baan CA, Ruige JB, Stolk RP, et al. Performance of a predictive model to identify undiagnosed diabetes in a health care setting. *Diabetes Care*. 1999;22:213-219.
24. Glümer C, Carstensen B, Sandbaek A, Lauritzen T, Jorgensen T, Borch-Johnsen K. A Danish diabetes risk score for targeted screening: the Inter99 study. *Diabetes Care*. 2004;27:727-733.
25. Griffin SJ, Little PS, Hales CN, Kinmonth AL, Wareham NJ. Diabetes risk score: towards earlier detection of type 2 diabetes in general practice. *Diabetes Metab Res Rev*. 2000;16:164-171.
26. Park PJ, Griffin SJ, Sargeant L, Wareham NJ. The performance of a risk score in predicting undiagnosed hyperglycemia. *Diabetes Care*. 2002;25:984-988.
27. Lorenzo C, Okoloise M, Williams K, Stern MP, Haffner SM. The metabolic syndrome as predictor of type 2 diabetes: the San Antonio Heart Study. *Diabetes Care*. 2003;26:3153-3159.
28. D'Agostino RB Jr, Hamman RF, Karter AJ, Mykkanen L, Wagenknecht LE, Haffner SM; Insulin Resistance Atherosclerosis Study Investigators. Cardiovascular disease risk factors predict the development of type 2 diabetes: the Insulin Resistance Atherosclerosis Study. *Diabetes Care*. 2004;27:2234-2240.
29. Hanley AJ, Karter AJ, Williams K, et al. Prediction of type 2 diabetes mellitus with alternative definitions of the metabolic syndrome: the Insulin Resistance Atherosclerosis Study. *Circulation*. 2005;112:3713-3721.
30. Kanaya AM, Wassel Fyr CL, de Rekeneire N, et al. Predicting the development of diabetes in older adults: the derivation and validation of a prediction rule. *Diabetes Care*. 2005;28:404-408.
31. von Eckardstein A, Schulte H, Assmann G. Risk for diabetes mellitus in middle-aged Caucasian male participants of the PROCAM study: implications for the definition of impaired fasting glucose by the American Diabetes Association: Prospective Cardiovascular Munster. *J Clin Endocrinol Metab*. 2000;85:3101-3108.
32. Schmidt MI, Duncan BB, Bang H, et al; Atherosclerosis Risk in Communities Investigators. Identifying individuals at high risk for diabetes: the Atherosclerosis Risk in Communities Study. *Diabetes Care*. 2005;28:2013-2018.
33. Wilson PW, D'Agostino RB, Parise H, Sullivan L, Meigs JB. Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. *Circulation*. 2005;112:3066-3072.
34. Lindström J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. *Diabetes Care*. 2003;26:725-731.
35. Liu J, Hong Y, D'Agostino RB Sr, et al. Predictive value for the Chinese population of the Framingham CHD risk assessment tool compared with the Chinese Multi-Provincial Cohort Study. *JAMA*. 2004;291:2591-2599.
36. Ferrario M, Chiodini P, Chambless LE, et al; CUORE Project Research Group. Prediction of coronary events in a low incidence population: assessing accuracy of the CUORE Cohort Study prediction equation. *Int J Epidemiol*. 2005;34:413-421.