

Components of the “Metabolic Syndrome” and Incidence of Type 2 Diabetes

Robert L. Hanson, Giuseppina Imperatore, Peter H. Bennett, and William C. Knowler

The combination of insulin resistance, dyslipidemia, hypertension, and obesity has been described as a “metabolic syndrome” that is a strong determinant of type 2 diabetes. Factor analysis was used to identify components of this syndrome in 1,918 Pima Indians. Prospective analyses were conducted to evaluate associations of identified factors with incidence of diabetes. Factor analysis identified 4 factors that accounted for 79% of the variance in the original 10 variables. Each of these factors reflected a proposed component of the metabolic syndrome: insulinemia, body size, blood pressure, and lipid metabolism. Among 890 originally nondiabetic participants with follow-up data, 144 developed diabetes in a median follow-up of 4.1 years. The insulinemia factor was strongly associated with diabetes incidence (incidence rate ratio [IRR] for a 1-SD difference in factor scores = 1.81, $P < 0.01$). The body size and lipids factors also significantly predicted diabetes (IRR 1.52 and 1.37, respectively, $P < 0.01$ for both), whereas the blood pressure factor did not (IRR 1.11, $P = 0.20$). Identification of four unique factors with different associations with incidence of diabetes suggests that the correlations among these variables reflect distinct metabolic processes, about which substantial information may be lost in the attempt to combine them into a single entity. *Diabetes* 51:3120–3127, 2002

Type 2 diabetes and cardiovascular disease have many risk factors in common, and many of these risk factors are highly correlated with one another (1–3). The relationships among these risk factors may be attributable to a small number of physiological phenomena, perhaps even a single phenomenon. The combination of hypertension, dyslipidemia, insulin resistance, hyperinsulinemia, glucose intolerance, and obesity, particularly central obesity, has been termed the “metabolic syndrome.” It has been proposed that this syndrome is a powerful determinant of diabetes and

cardiovascular disease (3–6). There are few prospective data, however, on the extent to which this syndrome or its constituent components predict incidence of type 2 diabetes.

Factor analysis is a mathematical technique by which a large number of correlated variables can be reduced to fewer “factors” that represent distinct attributes that account for a large proportion of the variance in the original variables. Thus, factor analysis is well suited for identifying components of the metabolic syndrome, and several analyses have been undertaken for this purpose (7–26). Prospective epidemiological studies of factor “scores” from these analyses can further determine relations between components of the metabolic syndrome and incidence of diabetes. In the present study, factor analysis is used to identify components of the metabolic syndrome in Pima Indians, an American Indian population with a high prevalence of type 2 diabetes and obesity (27,28), and relations of these components to incidence of diabetes are examined.

RESEARCH DESIGN AND METHODS

Participants and measurements. The present data come from a longitudinal study of type 2 diabetes that has been conducted among residents of the Gila River Indian Community in central Arizona (27); most of these residents are Pima or Tohono O’odham Indians. Community residents who are ≥ 5 years of age are invited biennially to have an examination, which includes a 75-g oral glucose tolerance test. Plasma glucose and serum insulin concentrations are measured in specimens collected when the participant is fasting and 2 h after an oral glucose load. Insulin concentrations are measured by immunoassay (Concept 4; ICN, Costa Mesa, CA). Diabetes is diagnosed by World Health Organization (WHO) criteria for epidemiological studies (i.e., 2-h plasma glucose concentration ≥ 11.1 mmol/l) or if the diagnosis is made during routine medical care (28,29). Height, weight, and waist circumference were measured with the participant wearing light clothing and no shoes. Blood pressure is measured to the nearest 2 mmHg with the participant lying supine; first and fourth Korotkoff sounds are taken as systolic and diastolic pressures, respectively. Since 1993, fasting serum triglyceride and HDL cholesterol concentrations have also been measured.

The present factor analyses included the following measurements: fasting and 2-h plasma glucose concentrations (G_0 , G_2), fasting and 2-h serum insulin concentrations (I_0 , I_2), systolic and diastolic blood pressure, body weight, waist circumference, serum triglyceride, and HDL cholesterol concentrations. Because many of these variables change dramatically as children mature, analyses were restricted to individuals who were ≥ 20 years of age. For each individual, the first examination at which all relevant measurements were available was taken. Individuals who were currently taking insulin or who were pregnant were excluded. Thus, factor analyses included 1,448 nondiabetic individuals (856 women and 592 men, mean age $[\pm SD]$ 34.6 ± 11.7 years) and 470 diabetic individuals (303 women and 167 men, mean age 43.5 ± 12.8 years). Diabetic and nondiabetic participants were initially analyzed separately, and since similar results were obtained, a combined analysis was conducted. To better represent insulin resistance, particularly when both diabetic and nondiabetic individuals are considered, product ($I_0 \cdot G_0$, $I_2 \cdot G_2$) and ratio (I_0/G_0 , I_2/G_2) indexes were analyzed; these indexes are strongly correlated with measurements of insulin resistance derived from the hyperinsulinemic-euglycemic “clamp” (30,31). However, similar results were ob-

From the Diabetes and Arthritis Epidemiology Section, National Institute of Diabetes and Digestive and Kidney Diseases, Phoenix, Arizona.

Address correspondence to Robert L. Hanson, Diabetes and Arthritis Epidemiology Section, National Institute of Diabetes and Digestive and Kidney Diseases, 1550 E. Indian School Rd., Phoenix, AZ 85014. E-mail: rhanson@phx.niddk.nih.gov.

Received for publication 9 May 2002 and accepted in revised form 17 July 2002.

G.I. is currently affiliated with the Division of Diabetes Translation, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control, Atlanta, Georgia.

G_0 , fasting glucose concentration; G_2 , 2-h plasma glucose concentration; I_0 , fasting insulin concentration; I_2 , 2-h serum insulin concentration; IRR, incidence rate ratio; NCEP, National Cholesterol Education Program; ROC, receiver operating characteristic; WHO, World Health Organization.

tained when I_0 , I_2 , G_0 , and G_2 were analyzed in separate studies of diabetic and nondiabetic individuals. In statistical analyses, the natural logarithm was taken of all variables to reduce skewness. Before factor analyses, all variables were adjusted for age, sex, height, and birth year by linear regression. Analyses were conducted using programs of the SAS Institute (Cary, NC). To examine whether medications that affect blood pressure, lipidemia, or glycemia were influencing results, factor analyses were conducted among the subset of individuals who were not currently taking medicines, and similar results to those presented here were obtained.

Factor analyses. The premise underlying factor analysis is that correlations observed among a set of variables can be explained by a small number of unique unmeasured variables or "factors." Therefore, factor analysis involves two procedures: 1) factor extraction to estimate the number of factors, and 2) factor rotation to determine constituents of each factor in terms of the original variables.

Factor extraction was conducted by the method of principal components. These components are linear combinations of the original variables that are constructed so that each component has a correlation of zero with each of the other components. Each principal component is associated with an "eigenvalue," which represents the variance in the original variables explained by that component (with each original variable standardized to have a variance of 1). The number of principal components that can be constructed is equal to the number of original variables. In factor analysis, the number of factors is customarily determined by retention of only those components that account for more of the total variance than any single original variable (i.e., those components with eigenvalues >1).

Once the number of factors has been established, then factor rotation is conducted to determine the composition of factors that has the most parsimonious interpretation in terms of the original variables. In factor rotation, "factor loadings," which represent correlations of each factor with the original variables, are changed so that these factor loadings are made as close to 0 or 1 as possible (with the constraint that the total amount of variance explained by the factors remains unchanged). A number of methods for factor rotation have been developed; these methods can be distinguished by whether they require the final set of factors to remain uncorrelated with one another (orthogonal methods) or by whether they allow factors to be correlated (oblique methods). In the present analyses, both an orthogonal (varimax rotation) and an oblique (promax rotation) method were used. Because both methods gave similar results and because previous studies of metabolic syndrome variables have used orthogonal methods, results for the varimax method are presented.

In interpretation of factor analysis, the pattern of factor loadings is examined to determine which original variables represent primary constituents of each factor. Conventionally, variables that have a factor loading >0.4 (or less than -0.4) with a particular factor are considered to be its major constituents.

Incidence of diabetes. The examination for which factor analysis was conducted was taken as the baseline examination for analyses of diabetes incidence; (these examinations occurred between February 1993 and May 1998—the time in which all relevant variables were measured). Individuals who were not diabetic at baseline and who had at least one subsequent examination were included. They comprised 890 participants (549 women and 341 men, mean age 33.3 ± 10.3 years). Factor scores were calculated from analysis of all subjects; these scores are linear combinations of the original variables that represent the predicted value of a given factor for an individual. Person-time was calculated from the baseline examination until diagnosis of diabetes or until the last examination before August 2001, whichever came first. Incidence rates were calculated in events per 1,000 person-years for tertile groups for each factor score. Age- and sex-standardized incidence rates (and SEs) were calculated by the direct method as previously described (28) with the 1980 U.S. population as the referent. The proportional hazards model (32) was used to calculate the incidence rate ratio (IRR) and 95% CI for a 1-SD difference in each of the factor scores with adjustment for age and sex. This allows for comparisons of the influence of factors on risk of diabetes by comparison of the magnitude of the IRR. Validity of the proportionality assumption was assessed by use of time-dependent covariates (33) and by comparison of the IRR before and after median time to development of diabetes (3.3 years).

Comparison of the ability of factor scores to predict incident diabetes was determined by analysis of receiver operating characteristic (ROC) curves, in which sensitivity is plotted as a function of 1-specificity. In the present context, area under the ROC curve represents the probability that an individual randomly selected from among those who developed diabetes has a higher factor score than one selected from among those who did not develop the disease (34). Thus, it is a measure of the ability of each factor score to predict diabetes, with values closer to 1 representing stronger prediction. Statistical

significance of the difference between ROC curve areas for pairs of factor scores was tested as described by Hanley and McNeil (34,35). As the purpose of factor analysis is to identify shared components that explain correlations among a set of variables, factor scores are not necessarily optimal predictors of diabetes. For comparative purposes, therefore, a ROC curve was also calculated for an optimally predictive multivariable score calculated as the sum of each of the original variables multiplied by its regression coefficient in a proportional hazards model that simultaneously included all 10 variables.

Incidence of diabetes was also examined according to two recent definitions of metabolic syndrome (36,37). Thus, individuals were defined as having metabolic syndrome by WHO criteria (36) if they had impaired glucose regulation or insulin resistance (fasting plasma glucose ≥ 6.1 mmol/l or 2-h plasma glucose ≥ 7.8 mmol/l or $I_0 \cdot G_0$ in the highest quartile for normoglycemic individuals) and at least two of the following: blood pressure $\geq 140/90$ mmHg; urinary albumin-to-creatinine ratio ≥ 30 mg/g; BMI >30 kg/m² or waist-to-hip ratio >0.90 (men) or >0.85 (women); and serum triglycerides ≥ 1.7 mmol/l or HDL cholesterol <0.9 mmol/l (men) or <1.0 mmol/l (women). Similarly, individuals had metabolic syndrome by National Cholesterol Education Program (NCEP) criteria (37) if they had at least three of the following: blood pressure $\geq 130/85$ mmHg; fasting plasma glucose ≥ 6.1 mmol/l; waist circumference >102 cm (men) or >88 cm (women); triglycerides ≥ 1.7 mmol/l; and HDL cholesterol <1.0 mmol/l (men) or <1.3 mmol/l (women). The present analyses use 1999 WHO criteria, but results are virtually identical if 1998 WHO criteria (38), which have slightly different values for blood pressure and albuminuria, are used.

RESULTS

Correlations among variables are shown in Table 1. For most variables correlations were of modest to moderate magnitude (0.2–0.7). A notable exception was that between body weight and waist circumference, which was consistently >0.9 . In addition, correlations of blood pressure variables with the others tended to be <0.2 . The general pattern of correlations was similar in diabetic and nondiabetic individuals.

Factor analysis in nondiabetic individuals. Among nondiabetic individuals, principal components analysis identified four factors with an eigenvalue >1 ; the largest eigenvalue among the remaining components was 0.64. These four factors, which accounted for 81.2% of the variance in the original 10 variables, were retained for factor rotation. Factor loadings of the original variables with each factor after varimax rotation are shown in Table 2. The first factor was strongly correlated with $I_0 \cdot G_0$, $I_2 \cdot G_2$, I_0/G_0 , and I_2/G_2 as indicated by a factor loading >0.4 with each of these variables. Likewise, the second factor was strongly correlated with body weight and waist circumference, the third factor was strongly correlated with systolic and diastolic blood pressure, and the fourth factor was strongly correlated with serum triglyceride concentration and inversely with HDL cholesterol concentration. Thus, these four factors can be interpreted as representing insulinemia, body size, blood pressure, and lipid metabolism.

Factor analysis in diabetic individuals. Among diabetic participants, there were also four factors that had eigenvalues >1 . The largest eigenvalue among the remaining components was 0.69, and the four factors accounted for 80.7% of the variance among the original variables. Factor loadings after varimax rotation are shown in Table 3. The first factor was correlated with $I_0 \cdot G_0$, $I_2 \cdot G_2$, I_0/G_0 , and I_2/G_2 ; the second factor was correlated with body weight and waist circumference; the third factor was correlated with systolic and diastolic blood pressure, and the fourth factor was correlated with serum triglycerides and $I_0 \cdot G_0$ and inversely correlated with HDL cholesterol. Thus, similar to results in nondiabetic individuals, results in

TABLE 1
Partial correlations (above diagonal) with associated *P* values (below diagonal) between variables used in factor analysis

	WT	WST	HDLC	TG	SBP	DBP	I ₀ · G ₀	I ₂ · G ₂	I ₀ /G ₀	I ₂ /G ₂
Nondiabetic participants (<i>N</i> = 1,448)										
WT		0.93	-0.23	0.15	0.16	0.19	0.52	0.36	0.47	0.32
WST	<0.01		-0.22	0.13	0.17	0.15	0.49	0.35	0.44	0.32
HDLC	<0.01	<0.01		-0.34	0.06	0.02	-0.26	-0.24	-0.23	-0.23
TG	<0.01	<0.01	<0.01		0.14	0.15	0.21	0.21	0.17	0.15
SBP	<0.01	<0.01	0.02	<0.01		0.52	0.12	0.14	0.09	0.07
DBP	<0.01	<0.01	0.47	<0.01	<0.01		0.13	0.12	0.11	0.07
I ₀ · G ₀	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01		0.64	0.98	0.65
I ₂ · G ₂	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01		0.61	0.92
I ₀ /G ₀	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01		0.65
I ₂ /G ₂	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	
Diabetic participants (<i>N</i> = 470)										
WT		0.92	-0.02	-0.18	0.01	0.12	0.31	0.25	0.40	0.28
WST	<0.01		0.01	-0.17	-0.01	0.12	0.31	0.21	0.35	0.22
HDLC	0.70	0.86		-0.44	0.03	0.01	-0.25	-0.11	-0.01	0.09
TG	<0.01	<0.01	<0.01		0.09	0.06	0.09	-0.05	-0.17	-0.20
SBP	0.84	0.83	0.55	0.04		0.59	0.05	-0.02	0.03	-0.02
DBP	0.01	0.01	0.84	0.23	<0.01		0.13	0.02	0.03	-0.04
I ₀ · G ₀	<0.01	<0.01	<0.01	0.06	0.33	<0.01		0.52	0.60	0.14
I ₂ · G ₂	<0.01	<0.01	0.02	0.30	0.64	0.65	<0.01		0.77	0.79
I ₀ /G ₀	<0.01	<0.01	0.79	<0.01	0.50	0.48	<0.01	<0.01		0.80
I ₂ /G ₂	<0.01	<0.01	0.06	<0.01	0.59	0.40	<0.01	<0.01	<0.01	
Nondiabetic and diabetic participants (<i>N</i> = 1,918)										
WT		0.93	-0.21	0.09	0.13	0.19	0.48	0.37	0.44	0.23
WST	<0.01		-0.20	0.09	0.12	0.17	0.47	0.36	0.40	0.20
HDLC	<0.01	<0.01		-0.39	0.04	-0.01	-0.31	-0.25	-0.16	-0.05
TG	<0.01	<0.01	<0.01		0.14	0.15	0.25	0.20	0.05	-0.06
SBP	<0.01	<0.01	0.10	<0.01		0.54	0.13	0.12	0.07	0.01
DBP	<0.01	<0.01	0.74	<0.01	<0.01		0.17	0.13	0.08	-0.01
I ₀ · G ₀	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01		0.65	0.77	0.25
I ₂ · G ₂	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01		0.59	0.66
I ₀ /G ₀	<0.01	<0.01	<0.01	0.03	<0.01	<0.01	<0.01	<0.01		0.67
I ₂ /G ₂	<0.01	<0.01	0.02	0.01	0.79	0.70	<0.01	<0.01	<0.01	

All variables are adjusted for age, sex, birth year, and height. DBP, diastolic blood pressure; HDLC, HDL cholesterol; SBP, systolic blood pressure; TG, triglycerides; WST, waist; WT, body weight.

diabetic participants can be interpreted as representing insulinemia, body size, blood pressure, and lipid metabolism.

Factor analysis in nondiabetic and diabetic individuals. In analyses of both diabetic and nondiabetic individuals combined, there were once again four factors

TABLE 2
Factor loadings for original variables with rotated factors in nondiabetic participants

	Insulinemia	Body size	Blood pressure	Lipidemia
Body weight	0.25	0.93	0.11	0.10
Waist circumference	0.23	0.93	0.09	0.09
HDL cholesterol	-0.15	-0.18	0.18	-0.78
Triglycerides	0.10	-0.01	0.22	0.82
Systolic blood pressure	0.06	0.06	0.86	0.00
Diastolic blood pressure	0.05	0.10	0.85	0.04
Fasting insulin × glucose	0.82	0.37	0.06	0.11
2-h Insulin × glucose	0.89	0.08	0.09	0.13
Fasting insulin/glucose	0.83	0.33	0.03	0.08
2-h Insulin/glucose	0.92	0.05	0.02	0.09
Percentage of variance	31.5	20.4	15.8	13.5

Factor loadings with an absolute value ≥0.4, which indicates that the variable can be considered a major constituent of the factor, are shown in bold. Each factor is named according to the variables which are its major constituents. The percentage of the variance represents the proportion of variance in the original 10 variables accounted for by each factor.

TABLE 3
Factor loadings for original variables with rotated factors in diabetic participants

	Insulinemia	Body size	Blood pressure	Lipidemia
Body weight	0.19	0.95	0.04	-0.05
Waist circumference	0.14	0.95	0.02	-0.05
HDL cholesterol	-0.04	-0.05	0.08	-0.84
Triglycerides	-0.13	-0.19	0.10	0.78
Systolic blood pressure	0.01	-0.05	0.89	0.00
Diastolic blood pressure	0.00	0.12	0.88	0.03
Fasting insulin × glucose	0.52	0.33	0.11	0.45
2-h Insulin × glucose	0.93	0.07	-0.01	0.10
Fasting insulin/glucose	0.91	0.25	0.03	-0.02
2-h Insulin/glucose	0.88	0.05	-0.05	-0.23
Percentage of variance	28.2	20.5	16.1	15.9

Factor loadings with an absolute value ≥0.4, which indicates that the variable can be considered a major constituent of the factor, are shown in bold. Each factor is named according to the variables that are its major constituents. The percentage of the variance represents the proportion of variance in the original 10 variables accounted for by each factor.

TABLE 4
Factor loadings for original variables with rotated factors in nondiabetic and diabetic participants

	Insulinemia	Body size	Blood pressure	Lipidemia
Body weight	0.22	0.94	0.09	0.07
Waist circumference	0.19	0.95	0.08	0.07
HDL cholesterol	-0.11	-0.16	0.14	-0.79
Triglycerides	0.01	-0.03	0.19	0.83
Systolic blood pressure	0.05	0.03	0.87	0.01
Diastolic blood pressure	0.03	0.12	0.86	0.05
Fasting insulin \times glucose	0.63	0.40	0.13	0.36
2-h Insulin \times glucose	0.83	0.16	0.10	0.23
Fasting insulin/glucose	0.85	0.29	0.03	0.05
2-h Insulin/glucose	0.87	0.00	-0.06	-0.16
Percentage of variance	26.6	21.0	16.0	15.5

Factor loadings with an absolute value ≥ 0.4 , which indicates that the variable can be considered a major constituent of the factor, are shown in bold. Each factor is named according to the variables which are its major constituents. The percentage of the variance represents the proportion of variance in the original 10 variables accounted for by each factor.

identified with eigenvalues >1 . The largest eigenvalue among the remaining components was 0.62, and the four factors accounted for 79.1% of the variance in the original variables. Factor loadings after varimax rotation are shown in Table 4. Results were comparable with those obtained in separate analyses of diabetic and nondiabetic individuals. The factors can again be interpreted as being representative of insulinemia, body size, blood pressure, and lipid metabolism.

Incidence of diabetes. Among the 890 individuals who were originally nondiabetic and who had at least one

follow-up examination, 144 (16%) developed diabetes. Median follow-up was 4.1 (range 0.1–8.0) years. Age- and sex-standardized incidence rates in tertile groups for each factor score are shown in Fig. 1. Insulinemia and body size factors were strongly related to diabetes incidence, whereas the lipids factor showed a more modest relationship, and the blood pressure factor was only weakly related to incidence of diabetes. Similar results were obtained when factor scores were analyzed as continuous variables in proportional hazards models. The age- and sex-adjusted IRR for a 1-SD increase in the insulinemia factor score was 1.81 (95% CI 1.48–2.20, $P < 0.01$), whereas the corresponding IRR for the body size factor score was 1.52 (1.30–1.77, $P < 0.01$). The lipids factor score had a more modest, but still statistically significant, association with risk of developing diabetes (IRR 1.37, 95% CI 1.16–1.62, $P < 0.01$), whereas the blood pressure factor score did not significantly predict diabetes (1.11, 0.95–1.31, $P = 0.20$). The risk associated with the lipids factor score was not proportional over the follow-up period, with an IRR of 1.75 (95% CI 1.39–2.21) in the initial period and of 1.10 (0.89–1.38) in the second period.

ROC curves for each of the four factors and for the multivariable score derived from all 10 variables are shown in Fig. 2. In analyses that compared the ability of factor scores to predict diabetes, area under the ROC curve for the insulinemia factor was not significantly greater than that for the body size factor ($P = 0.77$ for comparison of areas) and was not significantly greater than that for the lipids factor ($P = 0.14$). However, the insulinemia factor was a stronger predictor of incident diabetes than the blood pressure factor ($P < 0.01$). Similarly, the body size and lipids factors did not differ

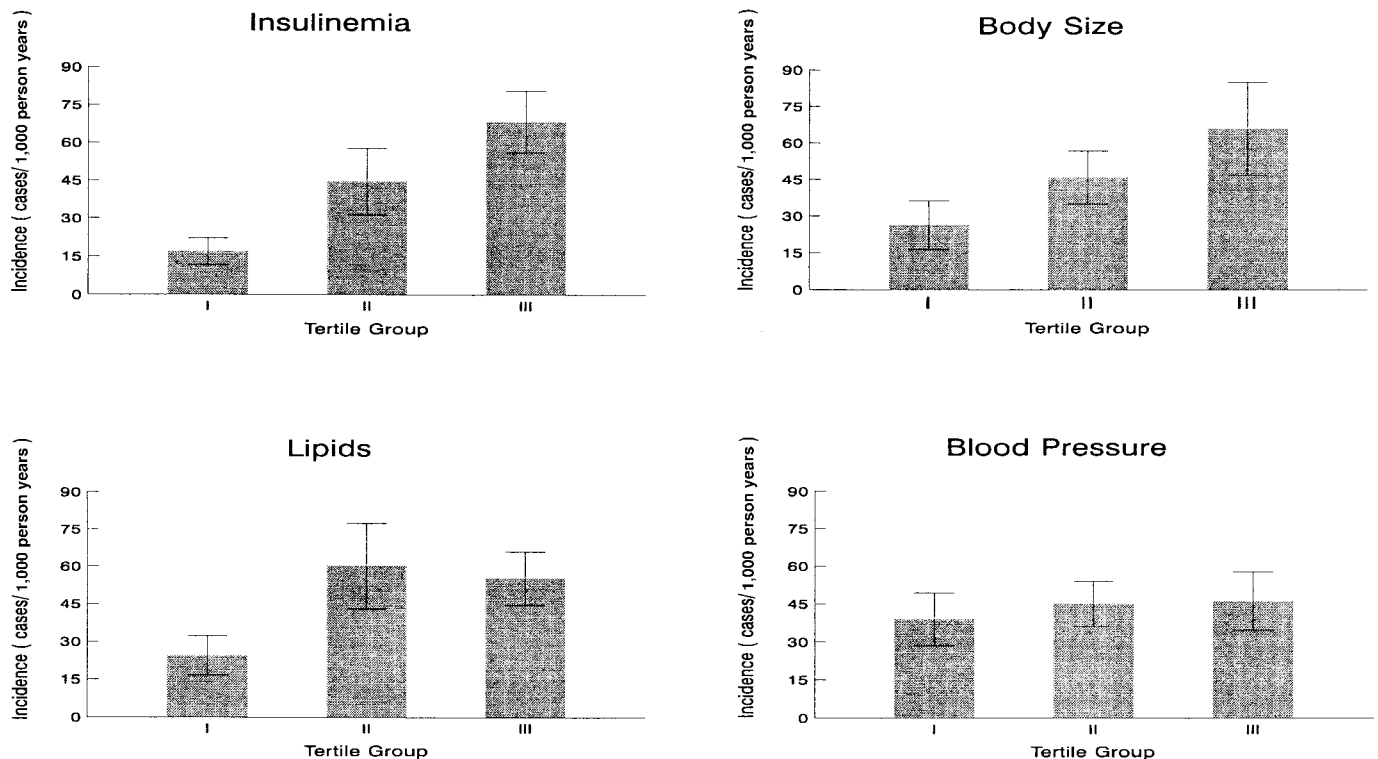


FIG. 1. Age- and sex-standardized incidence rates of diabetes (with 95% CIs) according to tertile groups defined by factor scores for each of the four factors identified in factor analysis.

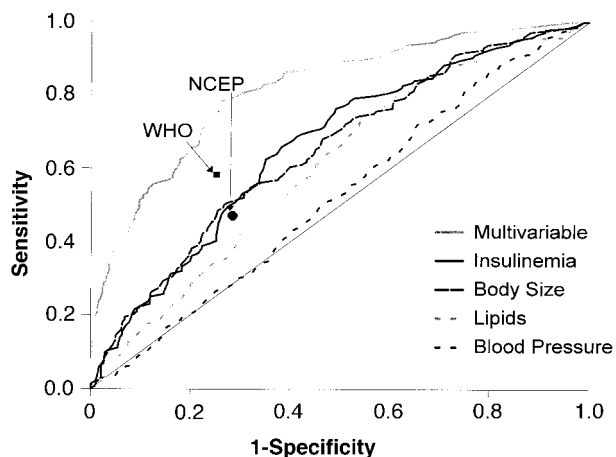


FIG. 2. ROC curves for prediction of development of diabetes. Curves are plotted for each of the four factors identified by factor analysis and for an optimally predictive multivariable score derived from all 10 original variables (see RESEARCH DESIGN AND METHODS). Areas under the ROC curves (\pm SEs) are 0.81 ± 0.022 for the multivariable score, 0.66 ± 0.026 for insulinemia, 0.65 ± 0.027 for body size, 0.60 ± 0.027 for lipids, and 0.52 ± 0.027 for blood pressure. Predictive properties for metabolic syndrome defined by WHO and NCEP criteria are also shown.

significantly in prediction of diabetes ($P = 0.26$), but both ROC curve areas were significantly greater than that for the blood pressure factor ($P < 0.01$ and $P = 0.03$, respectively). Area under the ROC curve for the multivariable score was significantly greater than that for any one of the factors ($P < 0.01$ for each).

At baseline 31% of participants had metabolic syndrome

by WHO criteria and 31% by NCEP criteria. For prediction of diabetes, WHO criteria were more specific and more sensitive than NCEP criteria (Fig. 2). Age- and sex-standardized diabetes incidence rates according to metabolic syndrome by WHO or NCEP criteria are shown in Fig. 3. Those who met WHO criteria had a much higher incidence of diabetes than those who did not (IRR 3.58, 95% CI 2.56–5.00), whereas increased incidence in those who met NCEP criteria compared with those who did not was more modest (2.09, 1.49–2.92). WHO criteria require hyperglycemia or insulin resistance in addition to other metabolic abnormalities, whereas NCEP criteria treat the components more equally. To examine whether this requirement accounts for the stronger prediction of WHO criteria, incidence rates were further stratified according to presence of hyperglycemia or insulin resistance (Fig. 3, right panels). Individuals with hyperglycemia/insulin resistance had a much higher incidence of diabetes than those without it, regardless of whether WHO or NCEP criteria for metabolic syndrome were met. This suggests that the requirement for hyperglycemia or insulin resistance is the major reason for different predictive properties between WHO and NCEP criteria.

DISCUSSION

The observation that many risk factors for diabetes are strongly intercorrelated has led to the hypothesis that they have certain etiological factors in common (3–6). In particular, insulin resistance, obesity, central adiposity, hypertension, hypertriglyceridemia, and hypoalphi-

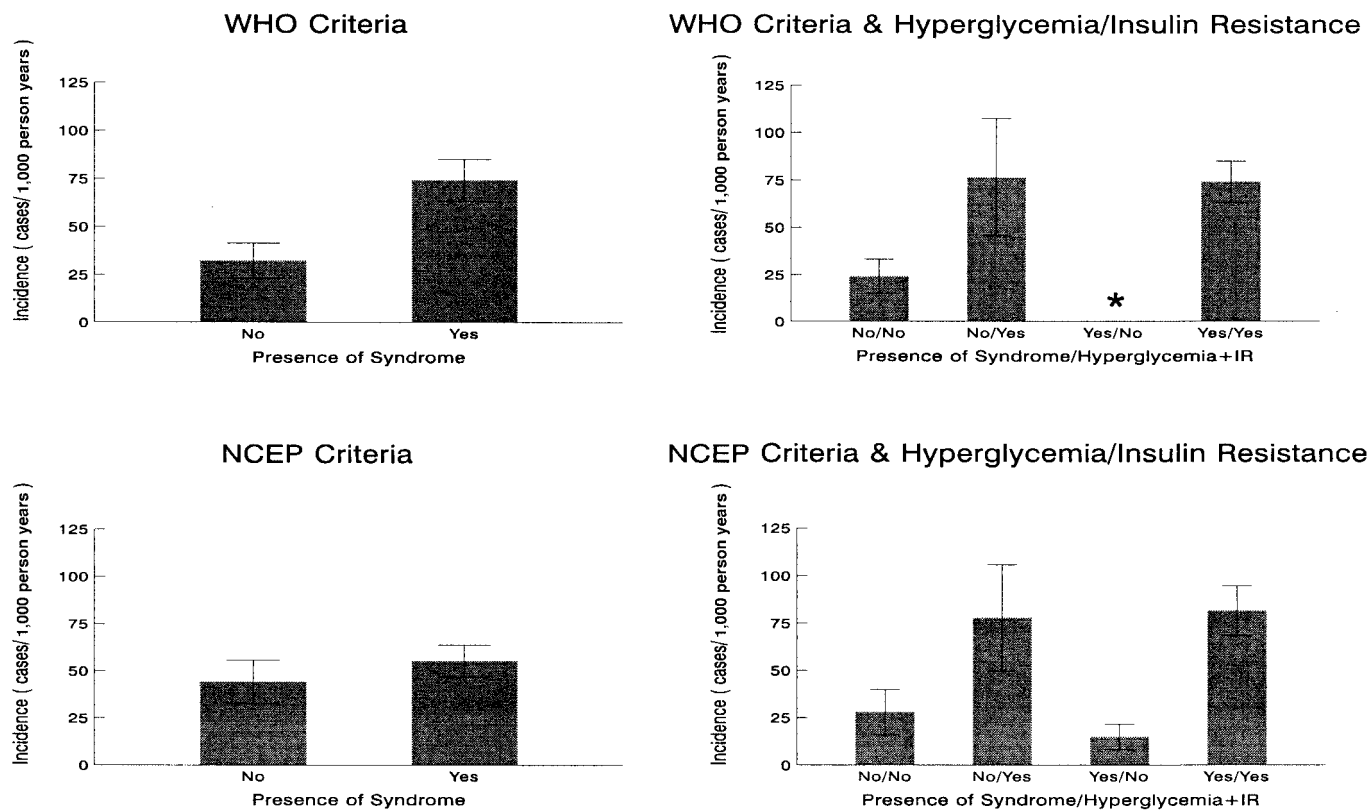


FIG. 3. Age- and sex-standardized incidence rates of diabetes (with 95% CIs) according to presence of metabolic syndrome by WHO or NCEP criteria. The right-hand panels are further stratified according to whether the WHO criterion for hyperglycemia or insulin resistance (IR) is met (see RESEARCH DESIGN AND METHODS). Because the WHO criteria require hyperglycemia or insulin resistance for diagnosis of metabolic syndrome, there are no data for the category indicated by the asterisk.

poproteinemia have all been described as risk factors for diabetes (1,3). Because many of these risk factors cluster together, it has been hypothesized that they may reflect a limited number of etiological metabolic abnormalities or perhaps even a single abnormality. The combination of obesity, insulin resistance, hypertension, and dyslipidemia has, thus, been categorized as a single syndrome, called variously “syndrome X,” “insulin resistance syndrome,” “metabolic syndrome,” or “dysmetabolic syndrome.”

However, the present factor analyses did not identify a single factor underlying the correlation structure in the variables included. Rather, it identified four factors, each of which is interpretable as representative of a proposed component of the metabolic syndrome: insulinemia, body size, blood pressure, and lipid metabolism. This suggests that the 10 variables analyzed are best described as reflecting these 4 independent physiological domains. Factor analysis is designed to reduce a large number of correlated variables to fewer factors, which reflect unique underlying phenomena and which extract much of the variation in the original variables. It is thus well suited to describing and quantifying components of metabolic syndrome. The technique of factor analysis is subject to a number of limitations, however. Most notably, results can be dependent on a number of somewhat arbitrary criteria, i.e., the threshold chosen for selection of number of factors retained, the method of rotation used, and the minimum factor loading chosen to designate a variable as a primary constituent of a factor. Robustness of a particular solution is best evaluated by whether it is consistent across different groups of individuals and different analytic procedures. The present analyses gave very similar results in diabetic and nondiabetic individuals, and these results were consistent using different procedures for factor rotation.

Similar factor analyses have been conducted for a number of other populations (as reviewed by Meigs [39]). Despite the subjective nature of factor analysis and differences as to which variables are included, results of these studies have shown some consistency. Many of them, like the present study, have identified a four-factor solution with factors representing obesity, insulin-glucose metabolism, lipid metabolism, and blood pressure, with some minor variation in composition of individual factors (7–12). Others have identified a three-factor solution that combines some of these entities; commonly, obesity is combined with lipid and/or insulin variables (13–22). A few studies have identified a two-factor solution that further combines variables (23–26). The majority of studies have identified blood pressure as a factor distinct from the others (7–12,14–18,26). Most of the studies that have not identified a separate blood pressure factor (13,20–25) have analyzed only a single blood pressure variable (20–24). Thus, the results of most factor analyses, including the present study, suggest that relationships among the variables typically proposed as constituting the metabolic syndrome are best explained as resulting from multiple physiological processes and that the attempt to reduce these to a single entity will result in a substantial loss of information about these metabolic processes.

Although factor scores are derived to be uncorrelated with one another, this does not necessarily mean that

underlying genetic or environmental determinants of these scores, or of their constituent variables, are uncorrelated. With respect to potential genetic determinants, some family-based studies have suggested substantial genetic correlations between certain variables constituting the metabolic syndrome (40,41), but others have not (42). Among Pimas, genetic linkage analyses of factor scores derived from the present study did not reveal any genomic regions with strong evidence for linkage to multiple factors (43); similar findings were seen in Mexican-Americans (18). This suggests that the major genetic determinants of these factors are separate.

Factor analysis is not designed to identify combinations of variables that are the strongest predictors of diabetes; this information is optimally derived from multivariable prediction models. However, epidemiological analyses that use factor scores can yield useful insights into relationships between various metabolic processes and disease incidence. Although most of the individual variables are risk factors for type 2 diabetes in various populations (1,3), results of analyses of individual risk factors are potentially confounded because of the correlations among these variables. Risk estimates associated with individual variables can be difficult to interpret, when derived from multivariable models that include several correlated variables, because of problems with multicollinearity (although predictive properties of such models tend to be robust [44]). Because factor analysis reduces a large number of correlated variables to fewer uncorrelated factors, it can help to circumvent these problems. In some studies, this approach has been used to examine relationships of metabolic syndrome components with incidence of cardiovascular disease (10,12,23). There are limited data relating metabolic syndrome factors to incidence of type 2 diabetes; one study in Finland identified an insulinemia/obesity/blood pressure/triglycerides factor that was associated with incidence of diabetes (22). The present analyses show that, among the Pimas, hyperinsulinemia and obesity are strong risk factors for type 2 diabetes. The combination of hypertriglyceridemia and low HDL cholesterol is a more modest, but still significant, risk factor, whereas high blood pressure is only weakly, if at all, associated with diabetes incidence. Furthermore, in ROC analyses the insulinemia, body size, and lipids factors were all significantly more strongly predictive of diabetes than the blood pressure factor. These findings suggest that the different physiological processes associated with various components of the metabolic syndrome contain unique information about diabetes risk.

The variable diabetes risk associated with each of these factors is reflected in differences in the extent to which different proposed criteria for metabolic syndrome predict diabetes incidence. WHO criteria were much more strongly associated with diabetes incidence than NCEP criteria. This difference was due to the greater weight given to hyperglycemia/insulin resistance in the WHO criteria than with the NCEP criteria, which essentially give the individual components equal weight. This further illustrates the importance of considering each of the metabolic syndrome components individually, at least with respect to diabetes risk.

The present findings have potential implications for

research and clinical practice. If the abnormalities constituting the metabolic syndrome are the result of largely independent physiological processes, then attempts to study a global syndrome phenotype may be counterproductive. Epidemiological comparisons of metabolic syndrome prevalence across diverse populations, for example, may well reflect very different entities, with substantial heterogeneity in risk for diabetes, or other health outcomes. It may be possible to address these problems by defining the syndrome so as to give each component a different weight, as per the multivariable prediction score in the present analyses. However, such an approach may require a different syndrome definition depending on whether diabetes or cardiovascular risk is the focus of investigation and, as effects of different components may vary among different populations, depending on the population in which it is applied. Furthermore, clinical treatment and prevention strategies that focus solely on the metabolic syndrome may prove suboptimal compared with treatment of the individual components.

ACKNOWLEDGMENTS

This work was presented in part at the 59th Scientific Sessions of the American Diabetes Association, San Diego, CA, June 19–22, 1999.

The authors thank the members of the Gila River Indian Community for participation in these studies and staff of the Diabetes and Arthritis Epidemiology Section for assistance. Additional thanks are due to Dr. David Mott for supervision of insulin assays and to Dr. Jonathan Krakoff for advice.

REFERENCES

1. Hamman RF: Genetic and environmental determinants of non-insulin-dependent diabetes mellitus (NIDDM). *Diabetes Metab Rev* 8:287–338, 1992
2. Castelli WP: Epidemiology of coronary heart disease: the Framingham Study. *Am J Med* 76:4–12, 1984
3. Stern MP: Diabetes and cardiovascular disease: the common soil hypothesis. *Diabetes* 44:369–381, 1995
4. Reaven GM: Role of insulin resistance in human disease. *Diabetes* 37:1595–1607, 1988
5. DeFronzo RA, Ferrannini E: Insulin resistance: a multifaceted syndrome responsible for NIDDM, obesity, hypertension, dyslipidemia, and atherosclerotic cardiovascular disease. *Diabetes Care* 14:173–194, 1991
6. Ferrannini E, Haffner SM, Mitchell BD, Stern MP: Hyperinsulinaemia: the key feature of a cardiovascular and metabolic syndrome. *Diabetologia* 34:416–422, 1991
7. Edwards KL, Burchfiel CM, Sharp DS, Curb JD, Rodriguez BL, Fujimoto WY, LaCroix AZ, Vitello MV, Austin MA: Factors of the insulin resistance syndrome in nondiabetic and diabetic elderly Japanese-American men. *Am J Epidemiol* 147:441–447, 1998
8. Sakkinen PA, Wahl P, Cushman M, Lewis MR, Tracy RP: Clustering of procoagulation, inflammation and fibrinolysis variables with metabolic factors in insulin resistance syndrome. *Am J Epidemiol* 152:897–907, 2000
9. Snehelatha C, Sivasankari S, Satyavani K, Vijay V, Ramachandran A: Insulin resistance alone does not explain the clustering of cardiovascular risk factors in southern India. *Diabet Med* 17:152–157, 2000
10. Kuusisto J, Lempäinen P, Mykkänen L, Laakso M: Insulin resistance syndrome predicts coronary heart disease events in elderly type 2 diabetic men. *Diabetes Care* 24:1629–1633, 2001
11. Hodge AM, Boyko EJ, de Courten M, Zimmet PZ, Chitson P, Tuomilehto J, Alberti KGMM: Leptin and other components of the metabolic syndrome in Mauritius: a factor analysis. *Int J Obes* 25:126–131, 2001
12. Lempäinen P, Mykkänen L, Pyörälä K, Laakso M, Kuusisto J: Insulin resistance syndrome predicts coronary heart disease events in elderly nondiabetic men. *Circulation* 100:123–128, 1999
13. Shmulewitz D, Auerbach SB, Lehner T, Blundell ML, Winick JD, Youngman

- LD, Skilling V, Heath SC, Ott J, Stoffel M, Breslow JL, Friedman JM: Epidemiology and factor analysis of obesity type II diabetes, hypertension and dyslipidemia (syndrome X) on the island of Kosrae, Federated States of Micronesia. *Hum Hered* 51:8–19, 2001
14. Meigs JB, D'Agostino RB Sr, Wilson PWF, Cupples A, Nathan DM, Singer DE: Risk variable clustering in the insulin resistance syndrome: the Framingham Offspring Study. *Diabetes* 46:1594–1600, 1997
15. Chen CH, Lin KC, Tsai ST, Chou P: Different association of hypertension and insulin-related metabolic syndrome between men and women in 8437 nondiabetic Chinese. *Am J Hypertens* 13:846–853, 2000
16. Lindblad U, Langer RD, Wingard DL, Thomas RG, Barrett-Connor EL: Metabolic syndrome and ischemic heart disease in elderly men and women. *Am J Epidemiol* 153:481–489, 2001
17. Gray RS, Fabsitz RR, Cowan LD, Lee ET, Howard BV, Savage PJ: Risk factor clustering in the insulin resistance syndrome: the Strong Heart Study. *Am J Epidemiol* 148:869–878, 1998
18. Arya R, Blangero J, Williams K, Almasy L, Dyer TD, Leach RJ, O'Connell P, Stern MP, Duggirala R: Factors of insulin resistance syndrome-related phenotypes are linked to genetic locations on chromosomes 6 and 7 in nondiabetic Mexican-Americans. *Diabetes* 51:841–847, 2002
19. Edwards KL, Austin MA, Newman B, Mayer E, Krauss RM, Selby JV: Multivariate analysis of the insulin resistance syndrome in women. *Arterioscler Thromb* 14:1940–1945, 1994
20. Leyva F, Godsland IF, Ghatei M, Proudler AJ, Aldis S, Walton C, Bloom S, Stevenson JC: Hyperleptinemia as a component of a metabolic syndrome of cardiovascular risk. *Arterioscler Thromb Vasc Biol* 18:928–933, 1998
21. Anderson PJ, Critchley JAJH, Chan JCN, Cockram CS, Lee ZSK, Thomas GN, Tomlinson B: Factor analysis of the metabolic syndrome: obesity vs insulin resistance as the central abnormality. *Int J Obes* 25:1782–1788, 2001
22. Kekäläinen P, Sarlund H, Pyörälä K, Laakso M: Hyperinsulinemia cluster predicts the development of type 2 diabetes independently of family history of diabetes. *Diabetes Care* 22:86–92, 1999
23. Pyörälä M, Miettinen H, Halonen P, Laakso M, Pyörälä K: Insulin resistance syndrome predicts the risk of coronary heart disease and stroke in healthy middle-aged men: the 22-year follow-up results of the Helsinki Policemen Study. *Arterioscler Thromb Vasc Biol* 20:538–544, 2000
24. Leyva F, Godsland IF, Worthington M, Walton C, Stevenson JC: Factors of the metabolic syndrome: baseline interrelationships in the first follow-up cohort of the HDDRISC Study (HDDRISC-1). *Arterioscler Thromb Vasc Biol* 18:208–214, 1998
25. Donahue RP, Bean JA, DeCarlo R, Goldberg RB, Prineas RJ: Does insulin resistance unite the separate components of the insulin resistance syndrome? *Arterioscler Thromb Vasc Biol* 17:2413–2417, 1997
26. Chen W, Srinivasan SR, Elkasabany A, Berenson GS: Cardiovascular risk factors clustering features of insulin resistance syndrome (syndrome X) in a biracial (black-white) population of children, adolescents and young adults: the Bogalusa Heart Study. *Am J Epidemiol* 150:667–674, 1999
27. Bennett PH, Burch TA, Miller M: Diabetes mellitus in American (Pima) Indians. *Lancet* 2:125–128, 1971
28. Knowler WC, Bennett PH, Hamman RF, Miller M: Diabetes incidence and prevalence in Pima Indians: a 19-fold greater incidence than in Rochester, Minnesota. *Am J Epidemiol* 108:497–505, 1978
29. World Health Organization: *Diabetes Mellitus: Report of a WHO Study Group*. Geneva, World Health Org., 1985 (Tech. Rep. Ser., no. 727)
30. Hanson RL, Pratley RE, Bogardus C, Narayan KMV, Roumain JML, Imperatore G, Fagot-Campagna A, Pettitt DJ, Bennett PH, Knowler WC: Evaluation of simple indices of insulin sensitivity and insulin secretion for use in epidemiologic studies. *Am J Epidemiol* 151:190–198, 2000
31. Anderson RL, Hamman RF, Savage PJ, Saad MF, Laws A, Kades WW, Sands RE, Cefalu W: Exploration of simple insulin sensitivity measures derived from frequently sampled intravenous glucose tolerance tests: the Insulin Resistance Atherosclerosis Study. *Am J Epidemiol* 142:724–732, 1995
32. Cox DR: Regression models and life tables. *J R Stat Soc [B]* 34:187–220, 1972
33. Kalbfleisch JC, Prentice RL: *The Statistical Analysis of Failure Time Data*. New York, Wiley, 1980, p. 119–142
34. Hanley JA, McNeil BJ: The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143:29–36, 1982
35. Hanley JA, McNeil BJ: A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 148:839–843, 1983
36. World Health Organization: *Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications: Report of a WHO Consultation. Part 1: Diagnosis and Classification of Diabetes Mellitus*. Geneva, World Health Org., 1999. Available from http://whqlib.who.int/hq/1999/WHO_NCD_NCS_99.2.pdf. Accessed 10 July 2002

37. 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). *J Am Med Assoc* 285:2486–2497, 2001
38. Alberti KGMM, Zimmet PZ: Definition diagnosis and classification of diabetes mellitus and its complications. Part 1. Diagnosis and classification of diabetes mellitus, provisional report of a WHO consultation. *Diabet Med* 15:539–553, 1998
39. Meigs JB: Invited commentary: insulin resistance syndrome? Syndrome X? Multiple metabolic syndrome? A syndrome at all? Factor analysis reveals patterns in the fabric of correlated metabolic risk factors. *Am J Epidemiol* 152:908–911, 2000
40. Mitchell BD, Kammerer CM, Mahaney MC, Blangero J, Commuzzie AG, Atwood LD, Haffner SM, Stern MP, MacCluer JW: Genetic analysis of the IRS: pleiotropic effects of genes influencing insulin levels on lipoprotein and obesity measures. *Arterioscler Throm Vasc Biol* 16: 281–288, 1996
41. Carmelli D, Cardon LR, Fabsitz RR: Clustering of hypertension, diabetes and obesity in adult male twins: same genes or same environment? *Am J Hum Genet* 55:566–573, 1994
42. Neel JV, Julius S, Weder A, Yamada M, Kardia SLR, Haviland MB: Syndrome X: is it for real? *Genet Epidemiol* 15:19–32, 1998
43. Imperatore G, Knowler WC, Kobes S, Bennett PH, Hanson RL: Genome-wide linkage analysis of factors characterizing the metabolic syndrome in Pima Indians (Abstract) *Diabetes* 48 (Suppl. 1):A182, 1999
44. Cox DR, Snell EJ: The choice of variables in observational studies. *Appl Statist* 23:51–59, 1974