Clinical Chemistry 57:2
326–337 (2011)

Biomarkers in Fasting Serum to Estimate Glucose Tolerance, Insulin Sensitivity, and Insulin Secretion

Allison B. Goldfine,[1*] Robert W. Gerwien,[2] Janice A. Kolberg,[2] Sheila O'Shea,[1] Sarah Hamren,[2] Glenn P. Hein,[2] Xiaomei M. Xu,[2] and Mary Elizabeth Patti[1]

BACKGROUND: Biomarkers for estimating reduced glucose tolerance, insulin sensitivity, or impaired insulin secretion would be clinically useful, since these physiologic measures are important in the pathogenesis of type 2 diabetes mellitus.

METHODS: We conducted a cross-sectional study in which 94 individuals, of whom 84 had 1 or more risk factors and 10 had no known risk factors for diabetes, underwent oral glucose tolerance testing. We measured 34 protein biomarkers associated with diabetes risk in 250-μL fasting serum samples. We applied multiple regression selection techniques to identify the most informative biomarkers and develop multivariate models to estimate glucose tolerance, insulin sensitivity, and insulin secretion. The ability of the glucose tolerance model to discriminate between diabetic individuals and those with impaired or normal glucose tolerance was evaluated by area under the ROC curve (AUC) analysis.

RESULTS: Of the at-risk participants, 25 (30%) were found to have impaired glucose tolerance, and 11 (13%) diabetes. Using molecular counting technology, we assessed multiple biomarkers with high accuracy in small volume samples. Multivariate biomarker models derived from fasting samples correlated strongly with 2-h postload glucose tolerance ($R^2 = 0.45$, $P < 0.0001$), composite insulin sensitivity index ($R^2 = 0.91$, $P < 0.0001$), and insulin secretion ($R^2 = 0.45$, $P < 0.0001$). Additionally, the glucose tolerance model provided strong discrimination between diabetes vs impaired or normal glucose tolerance (AUC 0.89) and between diabetes and impaired glucose tolerance vs normal tolerance (AUC 0.78).

CONCLUSIONS: Biomarkers in fasting blood samples may be useful in estimating glucose tolerance, insulin sensitivity, and insulin secretion.

Impaired glucose tolerance and reduced insulin sensitivity and secretion are established risk factors for type 2 diabetes mellitus (T2DM)[3] *(1–4)*. Although these metabolic disturbances begin before the onset of overt disease, it is difficult to assess these parameters in routine clinical practice. Identification of biomarkers in fasting blood samples that could distinguish individuals at highest risk for developing T2DM would represent a major medical advance and potentially provide novel mechanistic insights into disease pathogenesis. Thus, in this pilot study, we sought to identify biomarkers in fasting blood samples that could estimate glucose tolerance, insulin sensitivity, and insulin secretion, given their importance in diabetes pathophysiology. We note that the cross-sectional study design does not permit evaluating the association of biomarkers to incident diabetes risk. We measured 34 distinct serum protein biomarkers in a small volume of fasting serum from 94 individuals who underwent 75-g oral glucose tolerance testing (OGTT), and we developed models estimating glucose tolerance, insulin sensitivity, and insulin secretion.

Methods

Studies were approved by the Committee on Human Studies of the Joslin Diabetes Center. Participants provided written informed consent. Eighty-four consecutive participants who answered posted advertisements and reported 1 or more risk factors—including body mass index (BMI) >30 kg/m^2, nonwhite ethnicity, pre-

[1] Joslin Diabetes Center, One Joslin Place, Boston, MA; [2] Tethys Bioscience, Emeryville, CA.
* Address correspondence to this author at: Joslin Diabetes Center, One Joslin Place, Boston, MA 02215. Fax 617-713-3403; e-mail allison.goldfine@joslin.harvard.edu.
Received August 31, 2010; accepted November 9, 2010.
Previously published online at DOI: 10.1373/clinchem.2010.156133
[3] Nonstandard abbreviations: T2DM, type 2 diabetes mellitus; OGTT, oral glucose tolerance testing; BMI, body mass index; CISI, composite insulin sensitivity index; HOMA-IR, homeostasis model assessment of insulin resistance; CIR, corrected incremental insulin response; HbA_{1c}, hemoglobin A_{1c}; IL, interleukin; NGT, normal glucose tolerance; IGT, impaired glucose tolerance; PAI-1, plasminogen activator inhibitor 1; ANG, angiogenin; TIMP2, TIMP metallopeptidase inhibitor 2; AHSG, α-2-HS-glycoprotein; IGFBP-1, insulin-like growth factor binding protein 1; GPT, glutamic-pyruvate transaminase; IQR, interquartile range; AUC, area under the ROC curve; TNFRSF1A, tumor necrosis factor receptor superfamily member 1A.

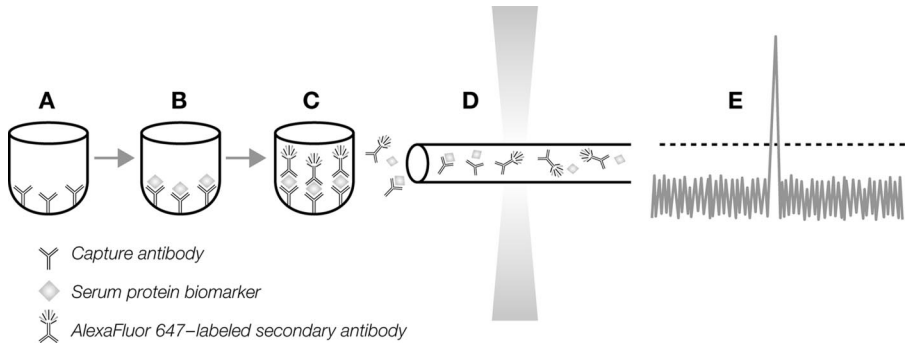

Fig. 1. Schematic of molecular counting technology.

Serum protein biomarkers were quantified with molecular counting technology in a 10-μL volume using a 384-microwell immunoassay format. (A), Attached to the surface of each microwell are capture antibodies specific for an individual biomarker. (B), Upon addition of serum sample, biomarkers are bound to the capture antibodies, followed by the addition and binding of AlexaFluor 647-labeled secondary antibody (C). (D), The fluorescence-labeled antibody complexes are chemically released from each well and pumped through a capillary flow system for detection of laser-induced fluorescence. (E), Photons emitted from AlexaFluor 647–labeled antibody molecules are distinguished from background levels so that each signal represents a molecular counting event.

vious gestational diabetes or offspring >9 lbs (4.1 kg) at birth, parental history of diabetes, history of hypertension, dyslipidemia or ischemic heart disease, or history of "borderline" abnormal glucose (high glucose but not diagnostic for diabetes)—were studied along with 10 persons with no known diabetes risk factors. Additional participant criteria are provided in the Supplemental Materials (which accompany the online version of this article at http://www.clinchem.org/content/vol57/issue2).

Weight, height, and blood pressure were measured (see details in the online Supplemental Materials). Fasting blood samples were obtained for laboratory analysis, and glucose and insulin were measured before and 30, 60, 90, and 120 min after a 75-g glucose load. Participants were classified by glucose tolerance status *(5)*.

We calculated insulin resistance using dynamic composite insulin sensitivity index (CISI) *(6)* and fasting homeostasis model assessment of insulin resistance (HOMA-IR) *(7)*, and insulin secretion using the corrected incremental insulin response (CIR) *(8, 9)*, as described in the online Supplemental Materials.

BIOMARKERS

Using commercial laboratory techniques, we measured widely used clinical laboratory markers—glucose, insulin, hemoglobin A_{1c} (HbA_{1c}), total cholesterol, HDL cholesterol, LDL cholesterol, triglycerides, and fructosamine—in fasting blood samples. Thirty-four additional serum protein biomarkers having a potential role in diabetes development were quantified using sandwich-format immunoassays and molecular counting performed on the ZeptX™ System (Singulex), a predecessor to the Erenna™ System *(10)*. Details of the biomarker assays and detection technology are in the online Supplemental Materials. In brief, serum protein biomarkers are quantified with molecular counting technology in a 10-μL volume using a 384-microwell immunoassay format. Capture antibodies specific for an individual biomarker are attached to the surface of each microwell; serum samples are added, followed by addition and binding of AlexaFluor 647–labeled secondary antibody. Fluorescence-labeled antibody complexes are chemically released from each well and pumped through a capillary flow system for detection of laser-induced fluorescence. A threshold above background is set so that each signal represents a labeled antibody molecule (Fig. 1).

DATA ANALYSIS AND MODEL DEVELOPMENT

We estimated associations among biomarkers on transformed concentrations using Pearson correlation. The correlation matrix was represented as a heat map with marker order decided by nearest-neighbor hierarchical clustering. We estimated univariate Spearman rank correlation between each serum protein biomarker and glucose tolerance (measured as 2-h glucose concentration following oral glucose), insulin sensitivity (calculated as CISI), and insulin secretion (calculated as CIR). These physiologic measures of diabetes pathophysiology provide useful clinical information but require testing at multiple time points. Because our goal is to avoid the need for dynamic testing, we con-

structed models using markers measured in blood samples obtained only in the fasting state and developed multiple linear regression models by searching for the most informative marker subsets. Two nested marker sets were considered: (1) a 46-marker set consisting of 34 serum protein biomarkers and 12 clinical and routine laboratory markers [age, sex, BMI, waist-to-hip ratio, HbA_{1c}, and concentrations of glucose, insulin, fructosamine, cholesterol (total, HDL, and LDL), and triglycerides]; and (2) a 42-marker set that excluded measures of glycemia (glucose, HbA_{1c}, and fructosamine) and insulin from consideration, to better reveal associations of nonglycemic biomarkers with indices of diabetes pathophysiology.

All data were $\log_{10}$ transformed except insulin secretion, which was square root transformed to satisfy distributional assumptions of linear regression. We used 6 marker selection techniques: forward, backward, and stepwise selection based on Akaike and Bayesian information criteria *(11)*. These techniques were executed within 100 bootstrap replicates, and markers selected within each bootstrap sample were tabulated. For each marker, we computed a weighted average "selection-count" based on the number of times it was selected under bootstrap sampling weighted by $1/k$, where k is the number of markers in each resulting model. Markers were selected based on average selection-count exceeding a threshold determined by a permutation test in which outcome was randomly assigned for each model developed, and 100 bootstrap replicates were used to calculate weighted marker count averages of 6 selection techniques. We repeated this permutation process 20 times; the 95th percentile of weighted marker count averages was used as a cutoff to identify markers selected significantly more frequently than random. Selected markers were used to construct multiple linear regression models. Internal validation was estimated using bootstrap model performance (R^2) from the median of 10 000 left-out bootstrap replicates for each model output.

A model incorporating 6 biomarkers [adiponectin, C-reactive protein, ferritin, interleukin (IL)-2Rα, glucose, and insulin] was identified recently as predictive of incident T2DM *(12)*. For comparison in this study, we tested the ability of these markers to estimate glucose tolerance and insulin action and secretion in independent models and compared the performance, based on the coefficient of determination (R^2), to the models we identified.

Results

CLINICAL CHARACTERISTICS

Study participants had 1 or more risk factors for developing diabetes—44, parental history of diabetes; 33, obese (BMI $\geq$30 kg/m^2); 24, history of high cholesterol; 13, history of hypertension; 19, ethnic minority; and 24, history of gestational diabetes or other "borderline" abnormal glucose. Ten participants at low risk for diabetes (no risk factors) were also recruited. On the basis of the glucose tolerance testing in the entire study population, 58 individuals were classified as having normal glucose tolerance (NGT, 2-h glucose $<$7.8 mmol/L). Of the at-risk participants, 25 (30%) had impaired glucose tolerance (IGT, 2-h glucose range $\geq$7.8 to $<$11.1 mmol/L and fasting plasma glucose $<$7.0 mmol/L), and 11 (13%) were identified with newly diagnosed T2DM (2-h glucose $\geq$11.1 mmol/L) (Table 1).

BIOMARKER IMMUNOASSAY ANALYSIS

The mean protein concentrations for the 34 serum protein biomarkers, which ranged over more than 8 orders of magnitude (from 10 ng/L to 4.3 g/L) among all 94 study participants, are shown in Table 2. Univariate Spearman rank correlations between model endpoints and individual biomarkers are also shown in Table 2. Leptin, C-reactive protein, and plasminogen activator inhibitor 1 (PAI-1) have most shared variance with glucose tolerance, insulin sensitivity, and insulin secretion, respectively. A heat map of the univariate correlations among the 34 different markers is provided in Fig. 2. Strong positive associations were noted among total cholesterol and LDL cholesterol, as well as angiogenin (ANG), TIMP metalloproteinase inhibitor 2 (TIMP2), α-2-HS-glycoprotein (AHSG), and apolipoprotein E.

MARKER SELECTION IN MULTIVARIATE MODELS

Performance of the multivariate models in estimating glucose tolerance, insulin sensitivity, and insulin secretion are reported in Table 3. As expected, multivariate models that include glucose and/or insulin have stronger correlation with indices of diabetes pathogenesis than models excluding these markers from consideration.

GLUCOSE TOLERANCE

A subset of 5 markers was associated with glucose tolerance: fasting glucose, leptin, insulin-like growth factor binding protein 1 (IGFBP-1), glutamic pyruvate transaminase (GPT, also known as alanine aminotransferase), and HbA_{1c}. As shown in Fig. 3A, multivariate modeling indicates that 45% of variance between modeled and observed 2-h glucose values is accounted for by these 5 markers alone ($P < 0.0001$), providing a fitted estimate. The bootstrap R^2 value is 0.38 [interquartile range (IQR) 0.16–0.58], providing an estimate of expected performance on an independent data set. By comparison, a model using 6 biomarkers previously shown to assess T2DM risk—adiponectin,

Table 1. Clinical characteristics of study participants according to glucose tolerance at baseline.[a]

	NGT	IGT	T2DM
n	58	25	11
Male, %	27 (46.6)	6 (24)	4 (36.4)
Age, years	41 (29–47)	50 (35–55)	52 (44–54)
BMI, kg/m^2	27.2 (24.7–31.0)	29.7 (25.3–35.0)	33.8 (26.6–38.4)
Waist-to-hip ratio	0.90 (0.84–0.96)	0.86 (0.81–0.94)	0.93 (0.87–0.99)
Systolic blood pressure, mmHg	116 (106–124)	128 (110–134)	130 (123–138)
Diastolic blood pressure, mmHg	75 (68–80)	78 (72–80)	86 (73–90)
Total cholesterol, mmol/L	4.6 (4–5.1)	4.9 (4.3–5.5)	4.3 (3.7–4.6)
LDL cholesterol, mmol/L	3.1 (2.3–3.5)	3.1 (2.4–3.8)	2.7 (2.4–3.5)
HDL cholesterol, mmol/L	1.2 (0.9–1.4)	1.2 (1–1.5)	1 (0.8–1.1)
Triglycerides, mmol/L	1.0 (0.6–1.4)	1.2 (0.7–1.6)	1.3 (0.9–1.9)
HbA_{1c}, %	5.3 (5.1–5.5)	5.4 (5.3–5.6)	6.4 (5.6–6.8)
Urinary albumin, μg/mg creatinine	4.4 (0.6–8.7)	5.6 (3–10.9)	4.9 (3.6–7.6)
Fructosamine, μmol/L	203 (190–218)	214 (203–224)	224 (209–243)
Plasma glucose, mmol/L			
Fasting	5.2 (4.9–5.4)	5.4 (5–5.7)	6.4 (5.8–6.9)
2-h	6.1 (5.3–6.9)	8.8 (8–10)	12.9 (12.6–14)
Serum insulin, pmol/L			
Fasting	56.6 (38.7–79.9)	82.0 (42.4–140.3)	98.6 (65.6–203.8)
2-h	268.8 (193.4–459.6)	713.3 (475–1224.4)	804.9 (476.8–1202.9)
HOMA-IR	1.85 (1.31–2.9)	2.75 (1.63–5.19)	4.48 (2.62–7.67)
CISI	52.9 (31.7–87.4)	50.1 (37.5–63.5)	19.7 (15.0–66.6)
CIR	0.28 (0.17–0.47)	0.24 (0.15–0.32)	0.17 (0.1–0.23)

[a] Data are as n (%) or median (interquartile range). Patients were classified according to glucose tolerance based on the 2006 WHO recommendations *(5)*, where 2-h postload glucose values for NGT are <7.8 mmol/L, for IGT are ≥7.8 and <11.1 mmol/L, and for T2DM are ≥11.1 mmol/L.

C-reactive protein, ferritin, IL-2Rα, glucose, and insulin *(12)*—estimated glucose tolerance with an R^2 value of 0.33 ($P < 0.0001$) and bootstrap R^2 value of 0.21 (IQR 0.14–0.29).

On excluding glycemic markers (glucose, HbA_{1c}, and fructosamine) and insulin from consideration, 5 markers were identified as most informative: leptin, IL-18, GPT, IGFBP-1, and ACE. Multivariate models constructed with these variables yielded an R^2 value of 0.26 ($P < 0.0001$) and bootstrap R^2 value of 0.16 (IQR 0.01–0.39). Although leptin is associated with excess weight, leptin remained an important marker for glucose tolerance even after statistically controlling for BMI ($P = 0.001$, likelihood ratio test).

DISCRIMINATION OF GLUCOSE TOLERANCE STATUS

The best-performing glucose tolerance model (fasting glucose, leptin, IGFBP-1, GPT, and HbA_{1c}) was able to discriminate between individuals based on glucose tolerance status. As shown in Fig. 3B, this model yielded an area under the ROC curve (AUC) of 0.89 for discrimination between those with diabetes vs without diabetes (individuals with IGT and NGT). Further, an AUC of 0.78 was observed for discrimination between those with abnormal glucose tolerance (T2DM or IGT) vs those with NGT, and 0.73 for discrimination between individuals with IGT vs those with NGT.

INSULIN SENSITIVITY

A subset of 5 markers was associated with insulin sensitivity (assessed using the dynamic CISI measure): fasting glucose, insulin, Fas ligand, complement C3, and PAI-1. As shown in Fig. 3C, 91% of variance between predicted and observed CISI values was accounted for by these 5 markers alone ($P < 0.0001$). In addition, a bootstrap R^2 value of 0.90 (IQR 0.83–0.94) indicates that the model could be expected to perform well on an independent data set. By comparison, HOMA-IR, a widely accepted estimate of insulin resistance based on fasting glucose and insulin, explained 88% of the variance of the dynamic measure of insulin sensitivity; thus, addition of Fas ligand, complement

Table 2. Biomarker levels of study participants.[a]

Biomarker	Symbol	Mean μg/L (95% CI)	Spearman correlation (*P*)[b]		
			Glucose tolerance	Insulin secretion	Insulin sensitivity
Angiotensin I converting enzyme (peptidyl-dipeptidase A) 1	ACE	167 (24–1.2×10^3)	0.14 (0.16)	−0.06 (0.54)	0.05 (0.63)
Adiponectin	ADIPOQ	5.0×10^4 (4.2×10^3–6.1×10^5)	−0.11 (0.29)	−0.17 (0.11)	0.29 (0.004)[c]
Advanced glycosylation end product–specific receptor	AGER	0.01 (0.001–0.13)	−0.12 (0.25)	−0.02 (0.86)	0.27 (0.01)[c]
α-2-HS-glycoprotein	AHSG	3.9×10^6 (1.0×10^6–1.5×10^7)	0.09 (0.39)	0.07 (0.48)	−0.16 (0.12)
Angiogenin, ribonuclease, RNase A family, 5	ANG	680 (85–5.4×10^3)	−0.01 (0.90)	0.2 (0.057)	−0.19 (0.07)
Apolipoprotein E	APOE	6.7×10^5 (9.0×10^4–4.9×10^6)	0.09 (0.40)	0.11 (0.31)	−0.20 (0.06)
Complement C3	C3	4.3×10^6 (2.8×10^5–6.6×10^7)	0.10 (0.32)	0.02 (0.86)	−0.09 (0.37)
Chemokine (C-C motif) ligand 2	CCL2	0.09 (0.01–0.88)	0.01 (0.94)	0.13 (0.21)	−0.03 (0.75)
CD 14, soluble	CD14	70 (8.8–562)	0.03 (0.75)	0.01 (0.94)	0.04 (0.71)
C-reactive protein, pentraxin-related	CRP	350 (11–1.1×10^4)	0.32 (0.001)[c]	0.25 (0.01)[c]	−0.55 (<0.0001)[c]
Dipeptidyl peptidase IV	DPP4	3.0 (0.1–115)	−0.16 (0.13)	0.09 (0.39)	0.10 (0.31)
Epidermal growth factor	EGF	0.08 (0.003–2.0)	0.16 (0.12)	0.11 (0.30)	−0.14 (0.19)
Endoglin (Osler-Rendu-Weber syndrome 1)	ENG	6.6 (1.5–29.3)	−0.23 (0.02)[c]	−0.03 (0.81)	0.12 (0.27)
Fas	FAS	1.0 (0.2–6.9)	0.13 (0.22)	0.08 (0.45)	−0.29 (0.005)[c]
Fas ligand	FASLG	0.14 (0.05–0.38)	0.09 (0.38)	0.07 (0.48)	−0.14 (0.18)
Fibrinogen	FGA	5.9×10^4 (1.3×10^4–2.7×10^5)	0.03 (0.80)	0.25 (0.01)[c]	−0.20 (0.06)
Ferritin	FTH1	842 (60–1.2×10^4)	0.11 (0.31)	0 (0.96)	−0.18 (0.07)
Glutamic-pyruvate transaminase	GPT	2.8 (0.6–13.1)	0.14 (0.19)	−0.01 (0.94)	−0.11 (0.30)
Hepatocyte growth factor	HGF	0.19 (0.03–1.34)	0.26 (0.01)[c]	0.22 (0.03)[c]	−0.45 (<0.0001)[c]
Haptoglobin	HP	8.6×10^5 (8.6×10^4–8.5×10^6)	0.19 (0.07)	0.27 (0.008)[c]	−0.33 (0.001)[c]
Heat shock 70-kDa protein 1B	HSPA1B	2.9 (0.4–19.1)	0.05 (0.65)	0.22 (0.05)[c]	−0.12 (0.24)
Insulin-like growth factor binding protein 1	IGFBP1	0.82 (0.07–9.5)	−0.04 (0.71)	−0.28 (0.006)[c]	0.51 (<0.0001)[c]
Insulin-like growth factor binding protein 2	IGFBP2	90.9 (11.2–734.9)	−0.16 (0.13)	−0.11 (0.29)	0.28 (0.006)[c]
Insulin-like growth factor binding protein 3	IGFBP3	17.2 (7.2–41.3)	−0.10 (0.33)	−0.01 (0.96)	0.09 (0.37)
Interleukin-18	IL18	0.22 (0.03–1.5)	0.33 (0.001)[c]	0.1 (0.32)	−0.40 (<0.0001)[c]
Interleukin-2 receptor, α	IL2RA	0.22 (0.05–0.96)	0.18 (0.08)	0.01 (0.9)	−0.29 (0.005)[c]
Leptin (obesity homolog, mouse)	LEP	2.0 (0.1–34)	0.40 (<0.0001)[c]	0.17 (0.10)	−0.52 (<0.0001)[c]
Resistin	RETN	14.8 (4.6–47)	0.17 (0.11)	0.06 (0.59)	−0.18 (0.08)
Plasminogen activator inhibitor-1 (serpin peptidase inhibitor)	PAI-1	132 (22–804)	0.25 (0.01)[c]	0.32 (0.001)[c]	−0.48 (<0.0001)[c]

Continued on page 331

Table 2. Biomarker levels of study participants.[a] *(Continued from page 330)*

Biomarker	Symbol	Mean μg/L (95% CI)	Spearman correlation (*P*)[b]		
			Glucose tolerance	Insulin secretion	Insulin sensitivity
Sex hormone-binding globulin	SHBG	1.1×10^4 (1.1×10^3–1.1×10^5)	−0.07 (0.52)	−0.06 (0.58)	0.19 (0.07)
TIMP metallopeptidase inhibitor 2	TIMP2	206 (26–1628)	0.03 (0.80)	0.13 (0.21)	−0.09 (0.40)
Tumor necrosis factor receptor superfamily member 1A	TNFRSF1A	0.39 (0.06–2.6)	0.14 (0.17)	0.32 (0.0014)[c]	−0.36 (0.0003)[c]
Vascular cell adhesion molecule 1	VCAM1	488 (32–7292)	−0.01 (0.94)	−0.11 (0.31)	0.17 (0.11)
von Willebrand factor	VWF	395 (40–3908)	0.09 (0.40)	−0.14 (0.17)	−0.001 (0.99)

[a] Concentrations for the 34 serum protein biomarkers assessed using molecular counting technology. Mean protein concentrations ranged over 8 orders of magnitude. Univariate Spearman rank correlations between individual protein biomarkers and glucose tolerance, insulin secretion, and insulin sensitivity are provided.
[b] Univariate Spearman rank correlations between model endpoints and individual biomarkers.
[c] Correlations with a significance of $P < 0.05$.

C3, and PAI-1 resulted in a small but significant improvement ($P < 0.0001$, likelihood ratio test) over HOMA-IR in estimating insulin sensitivity by CISI. Further, the model using the 6 biomarkers previously shown to assess incident T2DM risk *(12)* estimated insulin sensitivity with an R^2 value of 0.89 ($P < 0.0001$) and bootstrap R^2 value of 0.87 (IQR 0.85–0.89).

When glycemic markers and insulin were removed from consideration, 5 markers were identified as most informative: IGFBP-1, leptin, PAI-1, GPT, and triglycerides. Multivariate models constructed with these variables yielded an R^2 value of 0.62 ($P < 0.0001$) and bootstrap R^2 value of 0.58 (IQR 0.37–0.73).

INSULIN SECRETION

Five markers were associated with insulin secretion: fasting glucose, insulin, PAI-1, ACE, and IL-2Rα. As shown in Fig. 3D, multivariate modeling indicates that 45% of the variance between predicted and observed square root CIR values is accounted for by these 5 markers alone ($P < 0.0001$), and the bootstrap R^2 value is 0.39 (IQR 0.12–0.61). When glucose, HbA_{1c}, fructosamine, and insulin were removed from consideration, 6 markers were selected: PAI-1, tumor necrosis factor receptor superfamily member 1A (TNFRSF1A), chemokine ligand 2, IL-2 receptor, IGFBP-1, and ACE. Multivariate models constructed with these 6 variables yielded an R^2 value of 0.37 ($P < 0.0001$) and bootstrap R^2 value of 0.29 (IQR 0.07–0.50). By comparison, the model using the 6 biomarkers previously shown to assess incident T2DM risk *(12)* estimated insulin secretion with a larger fitted R^2 value of 0.42 ($P < 0.0001$) but a lower bootstrap R^2 value of 0.32 (IQR 0.24–0.40), suggesting that, within the current population, this latter model is more biased than the 5-marker model defined above.

Discussion

Our data indicate that multivariate models based on multiple biomarkers can be used to estimate glucose tolerance, insulin sensitivity, and insulin secretion. Although our cross-sectional study design cannot identify markers predicting diabetes risk, such composite biomarker profiles may be useful to more readily identify people at high risk of diabetes and improve targeting of preventive strategies. Genetic variants have been shown to provide additive value to clinical risk factors to identify progression to disease *(9, 14)*, and metabolic biomarkers may further improve identification of people with highest risk profiles. The fact that >40% of consecutive clinically at-risk persons in this study had IGT or previously undiagnosed T2DM further emphasizes the need for screening and/or diagnostic tools that extend beyond readily accessible clinical variables. Notably, although the clinical variables age, sex, BMI, and waist-to-hip ratio were evaluated for selection during model development, none of these was selected as among the most informative variables, suggesting that the information was captured in the biomarkers selected by the model development algorithm.

Because the OGTT is less frequently used in clinical practice, individuals with either T2DM or IGT, defined on the basis of high postload glucose values but

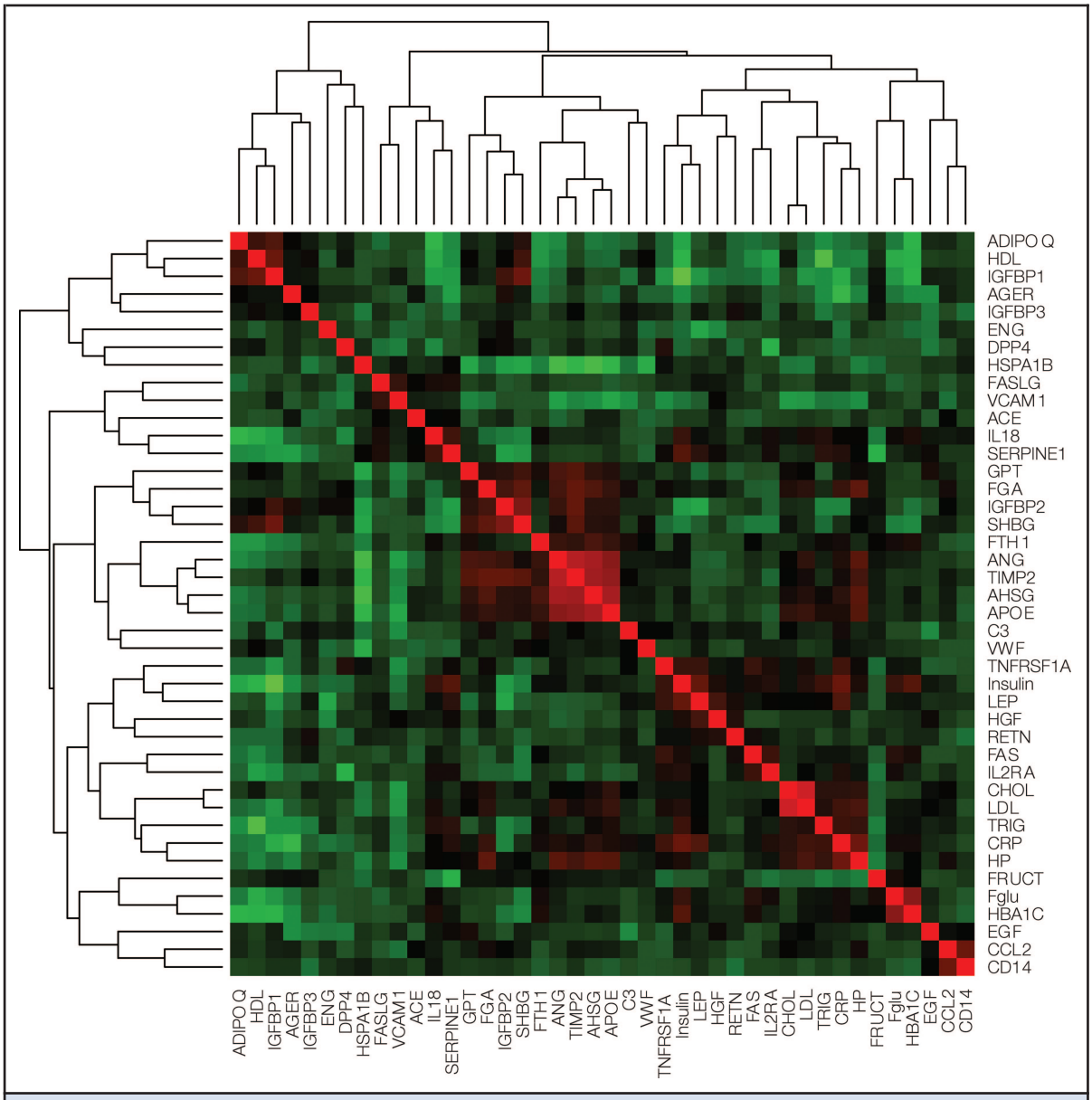

Fig. 2. Heat map of the univariate correlations among the serum protein biomarkers.

Pearson correlation was used to estimate associations among biomarkers on transformed concentrations. For key to abbreviations, see Table 2.

with normal fasting glucose values, may not be diagnosed. This is an important consideration, as postchallenge glucose better predicts increased risk for cardiovascular disease *(13, 14)*. Importantly, the multivariate glucose tolerance model provides strong discrimination between individuals with confirmed newly diagnosed T2DM (based solely on 2-h postload glucose) and those with known risk factors for development of T2DM but without disease (AUC 0.89). The glucose tolerance model also provides relevant discrimination between individuals with IGT and those with NGT (AUC 0.73). These data may therefore provide information to guide selection of patients who warrant more aggressive intervention for risk factors, such as smoking, weight, blood pressure, and lipid management. Recent studies have found combinations of biomarkers—including adiponectin, C-reactive protein, ferritin, IL-2Rα, IL-1 receptor antagonist, apolipopro-

Table 3. Models for the prediction of physiologic indicators of diabetes based on biomarkers measured in fasting blood samples.

Model components[a]	Estimate[b]	Standard error	P[c]	Multiple R^2	Bootstrap R^2
Glucose tolerance (all markers)				0.45	0.38
Glucose	1.025	0.3306	0.0026		
Leptin	0.076	0.0179	<0.0001		
IGFBP-1	0.060	0.0281	0.0355		
GPT	0.082	0.0511	0.1108		
HbA1c	1.162	0.4407	0.0099		
Glucose tolerance (excluding glycemic markers and insulin)				0.26	0.16
Leptin	0.096	0.0241	0.0001		
IL-18	0.246	0.1096	0.0275		
GPT	0.096	0.0600	0.1142		
ACE	0.086	0.0588	0.1448		
IGFBP-1	0.024	0.0319	0.4509		
Insulin sensitivity (all markers)				0.91	0.90
Glucose	−0.646	0.2231	0.0048		
Insulin	−0.852	0.0413	<0.0001		
Fas ligand	−0.011	0.0051	0.0303		
Complement C3	−0.0001	0.00004	0.0073		
PAI-1	−0.122	0.0535	0.0248		
Insulin sensitivity (excluding glycemic markers and insulin)				0.62	0.58
IGFBP-1	0.239	0.0453	<0.0001		
Leptin	−0.171	0.0359	<0.0001		
PAI-1	−0.220	0.0813	0.0083		
GPT	−0.291	0.0856	0.0010		
Triglycerides	−0.223	0.0801	0.0067		
Insulin secretion (all markers)				0.45	0.39
Glucose	−1.782	0.3764	<0.0001		
Insulin	0.390	0.0697	<0.0001		
PAI-1	0.161	0.0694	0.023		
ACE	−0.119	0.0687	0.088		
IL-2Rα	−0.195	0.1014	0.058		
Insulin secretion (excluding glycemic markers and insulin)				0.37	0.29
PAI-1	0.270	0.0685	0.0002		
TNFRSF1A	0.242	0.0854	0.0058		
CCL2	0.124	0.0367	0.0011		
IL-2Rα	−0.341	0.1095	0.0025		
IGFBP-1	−0.096	0.0397	0.0177		
ACE	−0.177	0.0756	0.0218		

[a] Models with all markers considered 34 serum protein biomarkers plus 9 additional markers measured in fasting samples, including glucose, HbA_{1c}, total cholesterol, HDL cholesterol, LDL cholesterol, triglycerides, fructosamine, insulin, and urinary albumin. Models excluding glycemic markers and insulin did not consider glucose, HbA_{1c}, fructosamine, and insulin.

[b] The estimate, β, is the coefficient describing the relationship between the model input and outcome with the associated standard error.

[c] The *P* value tests the hypothesis that $\beta = 0$; the multiple R^2 is the fitted result of the model, and the bootstrap R^2 is intended to estimate its performance on an external data set.

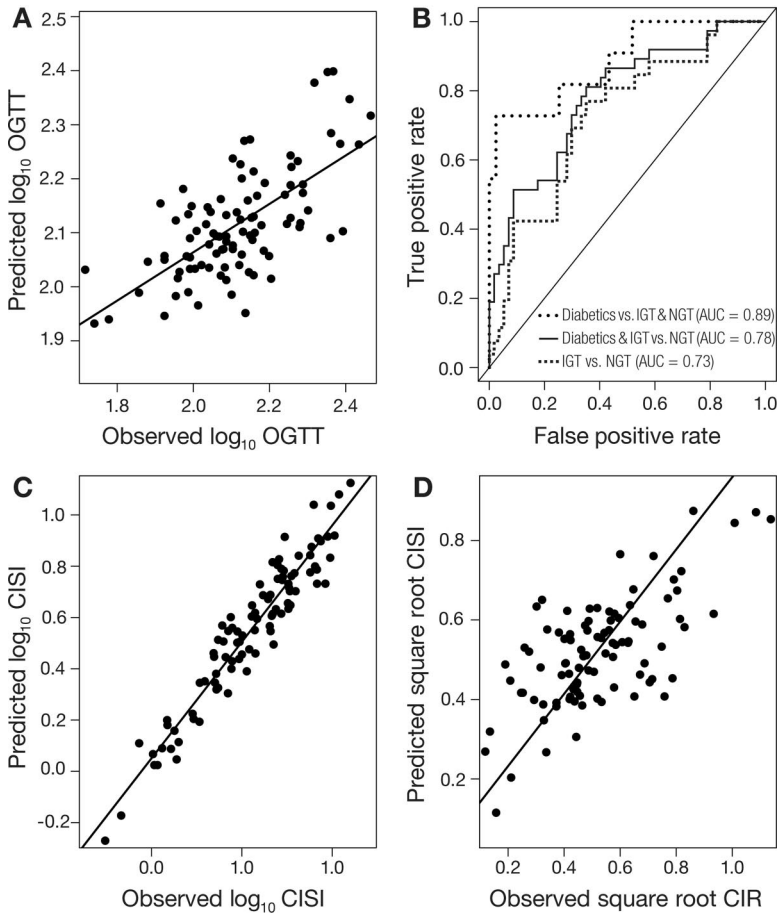

Fig. 3. Performance of the multivariate linear regression models for glucose tolerance, insulin sensitivity, and insulin secretion.

(A), Glucose tolerance model, predicted versus observed $\log_{10}$ OGTT 2-h plasma glucose values, with an R^2 value of 0.45 and a bootstrap R^2 value of 0.38. The components for the glucose tolerance model include fasting glucose, leptin, IGFBP-1, GPT, and HbA_{1c}. (B), ROC analysis for discrimination of T2DM, IGT, and NGT by the glucose tolerance model. (C), Insulin sensitivity model, predicted versus observed $\log_{10}$ CISI values, with an R^2 value of 0.91 and a bootstrap R^2 value of 0.90. The components for the insulin sensitivity model include fasting glucose, insulin, Fas ligand, complement C3, and PAI-1. (D), Insulin secretion model, predicted versus observed square root CIR values, with an R^2 value of 0.45 and a bootstrap R^2 value of 0.39. The components for the insulin secretion model include fasting glucose, insulin, PAI-1, ACE, and IL-2Rα.

tein B, glucose, and insulin—to be of potential value for prediction of incident diabetes *(12, 15)*. The markers we identified to be associated with glucose tolerance differed somewhat. Our study used fasting biomarkers to predict 2-h postload glucose on the same day, whereas the other studies aimed to predict future disease. It is also possible that different biomarkers will predict disease differentially in diverse populations which have dissimilar genetic makeup or environmental exposures.

Biomarkers from fasting specimens also correlated with the dynamic insulin secretion index. Methods to assess β-cell function in clinical studies are limited, and multivariate biomarkers may provide an alternative. Of note, the insulin secretion model is strong under both scenarios, accounting for 45% of the variance (bootstrap estimate 39%) when all markers were considered, and 37.3% of the variance (bootstrap estimate 29%) when glycemic markers and insulin were excluded. This finding suggests that a large component of insulin secretion can be estimated by nonglycemic variables, although confirmation in a unique patient group is needed. This is especially important since the homeostasis model assessment underestimates the

magnitude of β-cell defects demonstrated by the OGTT in IGT and newly diagnosed diabetes *(16)*.

The insulin sensitivity model exhibited the strongest performance, accounting for 91% of the variance between predicted and observed CISI values when all markers were considered. This finding would be anticipated, as fasting insulin and glucose correlate with insulin sensitivity in the widely used HOMA-IR model. However, 62% of the variance in insulin sensitivity could be accounted for by the markers IGFBP-1, leptin, PAI-1, GPT, and triglycerides in models where glycemic markers and insulin were excluded. The strong performance of the insulin sensitivity model is noteworthy, since all measurements are based on fasting samples, yet the full model demonstrates a correlation with CISI values significantly better than that of HOMA-IR alone. In future studies, it will be interesting to compare performance of the insulin sensitivity model with euglycemic-hyperinsulinemic clamp–derived measures of insulin sensitivity, currently the gold standard. Together, these findings underscore the value of multiple select biomarkers over fasting glucose and insulin measurements alone in assessing insulin sensitivity.

A 6-biomarker set was previously identified in the Inter99 cohort to predict incident diabetes: adiponectin, C-reactive protein, ferritin, IL-2Rα, glucose, and insulin *(12)*. On assessing these 6 biomarkers in our study population, we found that each of these biomarkers was involved in glucose tolerance and insulin action in univariate analysis. Although the biomarkers selected in our models for estimating glucose tolerance, insulin secretion, and insulin sensitivity differ from those of the 6-biomarker model previously developed, this could be due to use of different endpoints, differences between study populations, or additional factors related to the pathophysiology of T2DM. Nonetheless, when used to construct linear regression models in our study population, we found that these 6 biomarkers also performed reasonably well in estimating glucose tolerance, insulin sensitivity, and insulin secretion, as demonstrated by bootstrap R^2 values of 0.21, 0.87, and 0.32, respectively, for a 6-biomarker model compared with 0.38, 0.90, and 0.38, respectively, for multivariate models identified solely within our data set. The fact that the 6-biomarker model, selected in a different population for the related but different outcome of development of diabetes, performed as well as it did in estimating glucose tolerance, insulin sensitivity, and insulin secretion in our population is encouraging, in that it suggests that the biomarkers reflect the underlying disease pathology.

The subset of protein biomarkers we identified as most informative are consistent with earlier reports. For example, we found that PAI-1 was important for the prediction of insulin sensitivity and insulin secretion. Earlier studies have shown that serum concentrations of PAI-1 correlate with plasma insulin and insulin resistance and predict the likelihood of developing diabetes *(17)*. Similarly, we identified complement C3 as among the most informative biomarkers for prediction of insulin sensitivity. Complement C3 is strongly associated with insulin resistance and increased risk of T2DM *(18, 19)*. We also identified IL-18 and leptin as important biomarkers for prediction of glucose tolerance. Serum concentrations of IL-18 correlate with fasting glucose *(20, 21)* and with T2DM risk *(22)*. By comparison, beyond the association with excess weight, the relationship between leptin and T2DM risk is less clear. Our results indicate that leptin is an important predictor of glucose tolerance even after statistically controlling for BMI ($P = 0.001$, likelihood ratio test). Although some studies have found high leptin concentrations associated with an increased risk of incident T2DM *(23, 24)*, other studies found either no independent association between leptin concentrations and T2DM risk *(25)* or that high leptin concentrations predicted lower risk *(26)*.

One interpretation of our observations is that biomarkers identified as important for estimating glucose tolerance, insulin sensitivity, and insulin secretion may be causally related to T2DM development. However, statistical association of biomarkers with these parameters does not necessarily imply a causative role in T2DM pathogenesis. Alternatively, biomarkers may change as a result of the underlying disease process or from a common genetic etiology that underlies both biomarker expression and T2DM risk. Additionally, differences in biomarker concentrations could be a result of existing disease rather than a predictor, as several individuals in this cross-sectional cohort had diabetes that was diagnosed during participation. Our study used only biomarkers already linked to diabetes pathophysiology in the published scientific literature and hence does not represent an unbiased survey of all markers that may be important for disease pathogenesis. The criterion we used for selecting markers was whether their weighted counts exceeded random likelihood based on a permutation test, not whether their coefficient was significantly different from zero. The strong correlations we observed between biomarkers (shown in the heat map of Fig. 2) could reflect collinearity, in which case biomarkers could be removed from the models as a statistical artifact due to strong correlation with another marker rather than for more important biological reasons. In addition, although we used bootstrap models to estimate performance in an independent cohort, validation of our findings is important before implementation in clinical practice. Our observations on the relative importance of biomarkers in estimating glucose tolerance, insulin sensitivity, and insulin secretion provide proof-of-concept that prospective studies in large, well-characterized cohorts would be worthwhile.

In conclusion, our findings indicate that measurement of biomarkers in fasting blood samples may be useful for identifying reduced glucose tolerance, insulin sensitivity, and insulin secretion—metabolic disturbances all linked to T2DM risk. Furthermore, it may be possible to use fasting blood samples to identify individuals with IGT who are at risk for not only T2DM but also incident cardiovascular disease. We identified fasting glucose, leptin, IGFBP-1, GPT, and HbA_{1c} as the most informative markers for glucose tolerance; fasting glucose, insulin, Fas ligand, complement C3, and PAI-1 as most informative for insulin sensitivity; and fasting glucose, insulin, PAI-1, ACE, and IL-2Rα as most informative for insulin secretion. It will be important to evaluate performance of these models prospectively in independent cohorts, especially using longitudinal study designs for prediction of T2DM. Ultimately, assessment of biomarkers may complement additional strategies, such as genetic approaches. By identifying individuals with highest levels of insulin resistance, impaired insulin secretion, and/or glucose tolerance, it may be possible to focus lifestyle and drug intervention strategies aimed at preventing or delaying onset of disease.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures or Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:*

Employment or Leadership: R.W. Gerwien, Tethys Bioscience; J. Kolberg, Diabetes Research, Tethys Bioscience; S.J. Hamren, Tethys Bioscience; G.P. Hein, Tethys Bioscience; X.M. Xu, Tethys Bioscience.
Consultant or Advisory Role: A.B. Goldfine, Tethys Scientific Clinical Advisory Board and consultant to Tethys Bioscience.
Stock Ownership: R.W. Gerwien, Tethys Bioscience; J. Kolberg, Tethys Bioscience; S.J. Hamren, Tethys Bioscience; G.P. Hein, Tethys Bioscience; X.M. Xu, Tethys Bioscience.
Honoraria: None declared.
Research Funding: A.B. Goldfine, Tethys Bioscience, NIH grants DK-062948 and DK070648, and American Diabetes Association Career Development Award (06-CD-07); M.E. Patti, NIH grant DK-060837 and the Lilly Foundation. This study was supported by NIH grants DERC P30-DK-36836 and GCRC M01-RR-001032 (Joslin Diabetes Center).
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played a direct role in review and interpretation of data, and preparation and approval of manuscript.

Acknowledgments: The authors gratefully acknowledge the skillful technical assistance of P. Scott Eastman, Emmie Q. Fernandez, Timothy R. Hamilton, and Jillian Meri and thank Linda Wuestehube for editorial assistance and Anthony Sponzilli and Bryan Walton for illustrations.

References

1. Eriksson J, Franssila-Kallunki A, Ekstrand A, Saloranta C, Widen E, Schalin C, Groop L. Early metabolic defects in persons at increased risk for non-insulin-dependent diabetes mellitus. N Engl J Med 1989;321:337–43.
2. O'Rahilly S, Turner RC, Matthews DR. Impaired pulsatile secretion of insulin in relatives of patients with non-insulin-dependent diabetes. N Engl J Med 1988;318:1225–30.
3. Pimenta W, Korytkowski M, Mitrakou A, Jenssen T, Yki-Jarvinen H, Evron W, et al. Pancreatic beta-cell dysfunction as the primary genetic lesion in NIDDM: evidence from studies in normal glucose-tolerant individuals with a first-degree NIDDM relative. JAMA 1995;273:1855–61.
4. Unwin N, Shaw J, Zimmet P, Alberti KG. Impaired glucose tolerance and impaired fasting glycaemia: the current status on definition and intervention. Diabet Med 2002;19:708–23.
5. World Health Organization. Definition and diagnosis of diabetes mellitus and intermediate hyperglycemia: report of a WHO/IDF consultation. Geneva, Switzerland: WHO Press; 2006.
6. Matsuda M, DeFronzo RA. Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. Diabetes Care 1999;22:1462–70.
7. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 1985;28:412–9.
8. Hanson RL, Pratley RE, Bogardus C, Narayan KM, Roumain JM, Imperatore G, et al. Evaluation of simple indices of insulin sensitivity and insulin secretion for use in epidemiologic studies. Am J Epidemiol 2000;151:190–8.
9. Lyssenko V, Jonsson A, Almgren P, Pulizzi N, Isomaa B, Tuomi T, et al. Clinical risk factors, DNA variants, and the development of type 2 diabetes. N Engl J Med 2008;359:2220–32.
10. Todd J, Freese B, Lu A, Held D, Morey J, Livingston R, Goix P. Ultrasensitive flow-based immunoassays using single-molecule counting. Clin Chem 2007;53:1990–5.
11. Venables WN, Ripley BD. Modern applied statistics with S. 4th ed. Springer; 2002.
12. Kolberg JA, Jorgensen T, Gerwien RW, Hamren S, McKenna MP, Moler E, et al. Development of a type 2 diabetes risk model from a panel of serum biomarkers from the Inter99 cohort. Diabetes Care 2009;32:1207–12.
13. Hanefeld M, Fischer S, Julius U, Schulze J, Schwanebeck U, Schmechel H, et al. Risk factors for myocardial infarction and death in newly detected NIDDM: the Diabetes Intervention Study, 11-year follow-up. Diabetologia 1996;39:1577–83.
14. Meigs JB, Shrader P, Sullivan LM, McAteer JB, Fox CS, Dupuis J, et al. Genotype score in addition to common risk factors for prediction of type 2 diabetes. N Engl J Med 2008;359:2208–19.
15. Salomaa V, Havulinna A, Saarela O, Zeller T, Jousilahti P, Jula A, et al. Thirty-one novel biomarkers as predictors for clinically incident diabetes. PLoS One 2010;5:e10100.
16. Festa A, Williams K, Hanley AJ, Haffner SM. Beta-cell dysfunction in subjects with impaired glucose tolerance and early type 2 diabetes: comparison of surrogate markers with first-phase insulin secretion from an intravenous glucose tolerance test. Diabetes 2008;57:1638–44.
17. Festa A, D'Agostino R Jr, Tracy RP, Haffner SM. Elevated levels of acute-phase proteins and plasminogen activator inhibitor-1 predict the development of type 2 diabetes: the insulin resistance atherosclerosis study. Diabetes 2002;51:1131–7.
18. Engstrom G, Hedblad B, Eriksson KF, Janzon L, Lindgarde F. Complement C3 is a risk factor for the development of diabetes: a population-based cohort study. Diabetes 2005;54:570–5.
19. Muscari A, Antonelli S, Bianchi G, Cavrini G, Dapporto S, Ligabue A, et al. Serum C3 is a stronger inflammatory marker of insulin resistance than C-reactive protein, leukocyte count, and erythrocyte sedimentation rate: comparison study in an elderly population. Diabetes Care 2007;30:2362–8.
20. Aso Y, Okumura K, Takebayashi K, Wakabayashi S, Inukai T. Relationships of plasma interleukin-18 concentrations to hyperhomocysteinemia and carotid intimal-media wall thickness in patients with type 2 diabetes. Diabetes

Care 2003;26:2622–7.
21. Esposito K, Nappo F, Giugliano F, Di Palo C, Ciotola M, Barbieri M, et al. Cytokine milieu tends toward inflammation in type 2 diabetes. Diabetes Care 2003;26:1647.
22. Thorand B, Kolb H, Baumert J, Koenig W, Chambless L, Meisinger C, et al. Elevated levels of interleukin-18 predict the development of type 2 diabetes: results from the MONICA/KORA Augsburg Study, 1984–2002. Diabetes 2005;54: 2932–8.
23. McNeely MJ, Boyko EJ, Weigle DS, Shofer JB, Chessler SD, Leonnetti DL, Fujimoto WY. Association between baseline plasma leptin levels and subsequent development of diabetes in Japanese Americans. Diabetes Care 1999;22:65–70.
24. Wannamethee SG, Lowe GD, Rumley A, Cherry L, Whincup PH, Sattar N. Adipokines and risk of type 2 diabetes in older men. Diabetes Care 2007;30:1200–5.
25. Kanaya AM, Wassel Fyr C, Vittinghoff E, Harris TB, Park SW, Goodpaster BH, et al. Adipocytokines and incident diabetes mellitus in older adults: the independent effect of plasminogen activator inhibitor 1. Arch Intern Med 2006;166: 350–6.
26. Schmidt MI, Duncan BB, Vigo A, Pankow JS, Couper D, Ballantyne CM, et al. Leptin and incident type 2 diabetes: risk or protection? Diabetologia 2006;49:2086–96.