Novel Risk Factors and the Prediction of Type 2 Diabetes in the Atherosclerosis Risk in Communities (ARIC) Study

L.A. Raynor, phd[1]
James S. Pankow, phd[2]
Bruce B. Duncan, md, phd[3,4]
Maria I. Schmidt, md, phd[3,4]
Ron C. Hoogeveen, phd[5]
Mark A. Pereira, phd[2]
J. Hunter Young, md[6,7]
Christie M. Ballantyne, md[5]

OBJECTIVE—The objective of this study was to determine potential added value of novel risk factors in predicting the development of type 2 diabetes beyond that provided by standard clinical risk factors.

RESEARCH DESIGN AND METHODS—The Atherosclerosis Risk in Communities (ARIC) Study is a population-based prospective cohort study in four U.S. communities. Novel risk factors were either measured in the full cohort or in a case-control sample nested within the cohort. We started with a basic prediction model, previously validated in ARIC, and evaluated 35 novel risk factors by adding them independently to the basic model. The area under the curve (AUC), net reclassification index (NRI), and integrated discrimination index (IDI) were calculated to determine if each of the novel risk factors improved risk prediction.

RESULTS—There were 1,457 incident cases of diabetes with a mean of >7.6 years of follow-up among 12,277 participants at risk. None of the novel risk factors significantly improved the AUC. Forced expiratory volume in 1 s was the only novel risk factor that resulted in a significant NRI (0.54%; 95% CI: 0.33–0.86%). Adiponectin, leptin, γ-glutamyl transferase, ferritin, intercellular adhesion molecule 1, complement C3, white blood cell count, albumin, activated partial thromboplastin time, factor VIII, magnesium, hip circumference, heart rate, and a genetic risk score each significantly improved the IDI, but net changes were small.

CONCLUSIONS—Evaluation of a large panel of novel risk factors for type 2 diabetes indicated only small improvements in risk prediction, which are unlikely to meaningfully alter clinical risk reclassification or discrimination strategies.

Diabetes Care 36:70–76, 2013

A number of risk prediction tools for type 2 diabetes have been developed that could be used for opportunistic screening in clinical practice; however, at this time, there is no widely accepted risk prediction score that has been developed and validated in routine clinical practice (1,2). Developing a tool that successfully identifies those at high risk of type 2 diabetes is important because the disease is largely preventable through lifestyle and/or pharmacologic interventions (3). Therefore, the successful identification of at-risk individuals, via risk prediction models, would create greater opportunities for clinicians to intervene to prevent or delay the onset of type 2 diabetes and the complications associated with this disease (4).

From the [1]Department of Pediatrics, Division of Academic General Pediatrics, University of Minnesota, Minneapolis, Minnesota; the [2]Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota; the [3]Graduate Studies Program in Epidemiology, School of Medicine, Federal University of Rio Grande du Sol, Porto Alegre, Brazil; the [4]Department of Epidemiology, School of Public Health, University of North Carolina, Chapel Hill, North Carolina; the [5]Department of Medicine, Baylor College of Medicine, Houston, Texas; the [6]Department of Medicine, Johns Hopkins University, Baltimore, Maryland; and the [7]Department of Epidemiology, Johns Hopkins University, Baltimore, Maryland.
Corresponding author: L.A. Raynor, rayno007@umn.edu.
Received 29 March 2012 and accepted 1 July 2012.
DOI: 10.2337/dc12-0609
This article contains Supplementary Data online at http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc12-0609/-/DC1.

Within the last decade, a number of potential new risk factors for type 2 diabetes have been identified that are related to chronic inflammation, metabolic abnormalities, endothelial dysfunction, oxidative stress, and a prothrombotic state. Many of these factors have been found to be independently associated with type 2 diabetes in prospective cohort studies, including the Atherosclerosis Risk in Communities (ARIC) Study (5–21). Likewise, a number of common gene variants have been identified that are associated with type 2 diabetes in both candidate gene and genome-wide association studies. Because there is a possibility that one or more of these novel risk factors could serve in a tool for predicting type 2 diabetes, allowing clinicians to intervene and prevent the onset of disease, it is important to identify those risk factors that may refine and improve tools for risk prediction. Therefore, the purpose of this analysis is to identify novel risk factors that could improve type 2 diabetes risk prediction.

RESEARCH DESIGN AND METHODS

—The ARIC Study began in 1987–1989 and recruited a population-based cohort from four U.S. communities including: Forsyth County, NC; Jackson, MS; the northwest suburbs of Minneapolis, MN; and Washington County, MD (22). Participants received an extensive examination, including medical, social, and demographic data. The baseline examinations (visit 1) were conducted between 1987 and 1989; visit 2 was held between 1990 and 1992; visit 3 between 1993 and 1995; and visit 4 was conducted between 1996 and 1998. Of participants still alive at the time of follow-up visits, response rates for visits 2, 3, and 4 were 93, 86, and 81%, respectively.

Case-control study

For some analyses, we used data from a case-cohort study design previously used to examine the role of inflammation in the development of type 2 diabetes in ARIC (8). Prior to sampling, the following

individuals were excluded: 2,018 with prevalent diabetes, 95 members of minority ethnic groups with small numbers, 853 individuals who did not return to any follow-up visit, 26 with no valid diabetes determination at follow-up, 7 with restrictions on stored plasma use, 12 with missing baseline anthropometric measurements, and 2,506 in previous ARIC case-control and case-cohort studies involving cardiovascular disease for whom stored plasma was either previously exhausted or held in reserve. After exclusions, the sampling frame consisted of 10,275 individuals.

Case subjects were defined as participants who met any of the following criteria for type 2 diabetes at one or more follow-up visits: *1*) self-reported use of hypoglycemic medications; *2*) casual serum glucose of ≥200 mg/dL; *3*) fasting (>8 h) serum glucose of ≥126 mg/dL; or *4*) self-reported physician diagnosis of type 2 diabetes. There were 1,155 incident cases identified among the participants in the sampling frame. Due to budget constraints, all eligible type 2 diabetes cases were not selected for the case-cohort design. Instead, a stratified random sample of cases was selected with oversampling of African Americans. A subcohort was selected from all eligible cohort members to serve as the comparison group. Because risk-prediction software could not readily accommodate sample weights necessary for case-cohort analysis, we excluded incident case subjects who were independently selected for the cohort random sample but not the case sample (N = 23), resulting in a final sample size of 529 case and 543 control subjects.

Phenotypic measurements

Prevalent diabetes at baseline was defined as a nonfasting glucose ≥200 mg/dL, a fasting glucose ≥126 mg/dL, self-reported diagnosis of diabetes by a physician, or the current use of medications. Parental history of diabetes was defined as a report of diabetes in either parent. Subjects were asked to fast for 12 h prior to the clinical examination. Anthropometric measurements were taken with participants dressed in scrub suits without shoes. Technicians measured waist girth at the umbilical level. Blood pressure was measured three times with the subject in the sitting position after 5 min of rest using a random-zero sphygmomanometer, and the last two measurements were averaged.

After informed consent, blood was drawn from the antecubital vein of seated participants. Serum glucose was measured using a hexokinase/glucose-6-phosphate dehydrogenase method. Triglycerides were measured using an enzymatic method and HDL cholesterol (HDL-C) was measured enzymatically after dextran sulfate-Mg^{2+} precipitation of other lipoproteins.

Details regarding the measurement of novel risk factors are found in Supplementary Table 1.

Genotyping

Genotyping, quality control, and imputation procedures for the ARIC genome-wide association study have previously been described (23). The genetic risk score was created by adding together the number of genotyped or imputed risk alleles of 30 genes or regions, thus assuming an additive model of inheritance. The selection of genetic variants was based on a recent large-scale association analysis of European Americans that combined genome-wide association data from multiple studies to identify genetic variants associated with type 2 diabetes (24). The risk alleles modeled were those used in Voight et al. (24), which indexed alleles to the forward strand of National Center for Biotechnology Information Build 36. Because most of the variants were discovered and validated in Caucasian populations, the genetic risk score was created only for Caucasian study participants.

Statistical methods

Total cohort. Baseline characteristics of the study population were examined by incident type 2 diabetes status and shown as means ± SD or N (%) and compared by t or χ^2 tests. For prediction analyses, we started with a simple or basic prediction model, previously validated in ARIC (25), that includes age, parental history of diabetes, race/ethnicity, fasting glucose, fasting triglycerides, systolic blood

Table 1—*Baseline characteristics (mean or percentage) of the total ARIC cohort by incident type 2 diabetes status*

	Type 2 diabetes (N = 1,457)	No type 2 diabetes (N = 10,820)	P value
Basic risk factors			
Age (years)	54.1	53.9	0.30
Parental history of diabetes (%)	33.9	21.3	<0.0001
Race (African American, %)	32.3	20.1	<0.0001
Systolic blood pressure (mmHg)	125.5	118.8	<0.0001
Waist circumference (cm)	104.6	94.5	<0.0001
Height (cm)	169.1	168.4	0.006
Fasting triglycerides (mg/dL)	155.1	120.5	<0.0001
HDL-C (mg/dL)	45.8	53.4	<0.0001
Fasting glucose (mg/dL)	108.1	97.3	<0.0001
Novel risk factors			
WBC count (1,000/mm^3)	6.4	6.0	<0.0001
Fibrinogen (mg/dL)	310.0	296.7	<0.0001
Albumin (g/dL)	3.9	3.8	0.22
vWF (%)	120.3	113.5	<0.0001
aPTT (s)	29.0	29.3	<0.0001
Factor VIII (%)	135.1	126.1	<0.0001
Magnesium (mg/dL)	1.63	1.65	<0.0001
FEV_1 (L)	2.8	2.9	<0.0001
FVC (L)	3.7	3.9	<0.0001
Hematocrit (%)	42.4	41.5	<0.0001
Heart rate (bpm)	67.0	65.4	<0.0001
Low-frequency-power heart rate variability (ms)	23.9	23.4	0.34
Leg length (cm)	80.1	79.6	0.0005
Hip circumference (cm)	109.0	103.3	<0.0001
Blood viscosity (centipoise)	6.0	5.9	0.001
Genetic risk score (number of risk alleles)*	29.1	29.8	<0.0001

*Genetic risk score was available only in Caucasian participants.

Table 2—*Measures of risk prediction for the total cohort*

	N	C statistic	95% CI*	Difference	NRI	95% CI*	IDI	95% CI*
Basic**	12,277	0.8411	0.8316–0.8457					
WBC count	12,277	0.8430	0.8429–0.8520	0.0019	0.0015	−0.0036 to 0.0142	0.0043	0.0019–0.0068
Fibrinogen	12,232	0.8415	0.8396–0.8471	0.0004	−0.0011	−0.0004 to 0.0048	0.0003	−0.0001 to 0.0011
Albumin	12,277	0.8409	0.8318–0.8456	−0.0002	−0.0011	−0.0037 to 0.0122	0.0002	0.0001–0.0013
vWF	12,235	0.8415	0.8364–0.8450	0.0004	0.0022	−0.0029 to 0.0035	0.0000	−0.0000 to 0.0004
aPTT	12,226	0.8415	0.8413–0.8485	0.0004	0.0022	−0.0016 to 0.0082	0.0001	0.0000–0.0011
Factor VIII	12,229	0.8418	0.8387–0.8483	0.0007	0.0016	−0.0036 to 0.0109	0.0009	0.0008–0.0025
Magnesium	12,277	0.8409	0.8404–0.8474	−0.0002	−0.0010	−0.0050 to 0.0011	0.0002	0.0000–0.0007
FEV_1	11,095	0.8426	0.8419–0.8559	0.0015	0.0054	0.0033−0.0086	0.0001	−0.0006 to 0.0018
FVC	11,095	0.8427	0.8349–0.8451	0.0016	0.0042	−0.0146 to 0.0081	0.0002	−0.0005 to 0.0008
Hematocrit	12,088	0.8414	0.8283–0.8560	0.0003	0.0033	−0.0060 to 0.0039	0.0000	−0.0000 to 0.0010
Heart rate	11,702	0.8427	0.8347–0.8490	0.0016	0.0015	−0.0001 to 0.0080	0.0006	0.0005–0.0014
Low-frequency-power heart rate variability	11,702	0.8424	0.8389–0.8450	0.0013	−0.0017	−0.0025 to 0.0018	−0.0001	−0.0000 to 0.0001
Leg length	12,274	0.8412	0.8400–0.8509	0.0001	0.0025	−0.0010 to 0.0024	0.0001	−0.0001 to 0.0002
Hip circumference	12,277	0.8422	0.8332–0.8514	0.0011	−0.0013	−0.0059 to 0.0072	0.0019	0.0009–0.0027
Blood viscosity	12,088	0.8413	0.8380–0.8449	0.0002	0.0034	−0.0015 to 0.0108	−0.0000	−0.0000 to 0.0001
Genetic risk score***	8,067	0.8496	0.8385–0.8604	0.0012	−0.0032	−0.0130 to 0.0108	0.0018	0.0011–0.0024

The basic model was rerun for each covariate, and the difference between the basic and expanded models was calculated accordingly due to changing sample sizes. *The 95% CIs were bootstrapped. **The basic model includes age, parental history of diabetes, race/ethnicity, fasting glucose, fasting triglycerides, systolic blood pressure, HDL-C, height, and waist circumference. ***Modeled only in Caucasians.

pressure, HDL-C, height, and waist circumference, all measured at visit 1. The expanded model for the full cohort considered the following measures obtained at visit 1 and reported to be associated with incident type 2 diabetes in previous ARIC publications:

- White blood cell (WBC) count
- Fibrinogen
- Albumin
- von Willebrand factor (vWF) antigen
- Activated partial thromboplastin time (aPTT)
- Factor VIII coagulant activity
- Serum magnesium
- Forced vital capacity (FVC)
- Forced expiratory volume in 1 s (FEV_1)
- Total blood viscosity
- Hematocrit level
- Leg length
- Hip circumference
- Heart rate
- Low frequency power heart rate variability
- Genetic risk score that includes variants from the following 30 genes or regions: *NOTCH2* (rs10923931), *THADA* (rs7578597), *BCL11A* (rs243021), *PPARG* (rs1801282), *ADAMTS9* (rs6795735), *IGF2BP2* (rs1470579), *WFS1* (rs10010131), *ZBED3* (rs4457053), *CDKAL1* (rs7754840), *JAZF1* (rs849134), *KLF14* (rs972283), *TP53INP1* (rs896854), *SLC30A8* (rs13266634), *CHCHD9*(rs13292136), *CDKN2A/B* (rs10811661), *CDC123/CAMKID* (rs12779790), *HHEX/IDE* (rs1111875), *TCF7L2* (rs7903146), *KCNQ1* (rs231362), *KCNJ11* (rs5215), *CENTD2* (rs1552224), *HMGA2* (rs1531343), *TSPAN8/LGR5* (rs7961581), *HNF1A* (rs7957197), *ZFAND6* (rs11634397), *PRC1* (rs8042680), *FTO* (rs9939609), *HNF1B* (rs75210), *MTNRB1* (rs1387153), and *IRS1* (rs7578326)

We used Cox proportional hazards regression models, with incident type 2 diabetes as the outcome, to calculate the C statistic for each individual risk factor. We defined incident type 2 diabetes as described above. The date of type 2 diabetes incidence was estimated by linear interpolation using glucose values at the ascertaining visit and the previous one (8). We constructed models by adding each novel risk factor one at a time to the basic risk prediction model. C statistics were compared between the baseline model and the model with the novel risk factor, and if the variable produced an incremental change of at least 0.005, it was included in the final model (25).

We used the macro derived by Chambless et al. (26) to calculate the area under the curve (AUC), net reclassification index (NRI), and integrated discrimination index (IDI) for our risk-prediction models. The AUC is calculated via a nonparametric method, which produces the AUC at time *t* in a setting of risk prediction from survival analysis and takes censoring into consideration (26). All analyses were conducted using SAS version 9.2 (SAS Institute, Cary, NC).

We excluded 2,018 individuals who had prevalent type 2 diabetes at baseline, 95 individuals from underrepresented minority groups, 314 individuals with missing information on the risk factors included in the basic risk model, 267 individuals who did not fast for at least 8 h, and 821 individuals who had no follow-up time data to ascertain type 2 diabetes, thus leaving 12,277 individuals for the analysis. Because the genetic risk score was created only for Caucasian study participants, there were 8,067 individuals available for the analysis of the addition of a genetic risk score to the basic model.

Case-control study

We started with the same aforementioned basic risk-prediction model (25). The expanded model for the case-control subsample considered the following novel measures obtained from visit 1 blood samples:

- Total adiponectin
- High-molecular-weight adiponectin
- Leptin
- γ-Glutamyl transferase (GGT)
- Alanine aminotransferase

- Fetuin A
- Ferritin
- C-reactive protein
- Intercellular adhesion molecule 1 (ICAM-1)
- Orosomucoid
- Sialic acid
- Interleukin-6
- Interleukin-18
- Complement C3
- Asymmetric dimethylarginine (ADMA)
- Retinol binding protein 4 (RBP-4)
- Free fatty acids
- Oxidized LDL
- Lactate

For the case-control analysis, we used logistic regression models, with incident type 2 diabetes as the outcome, to calculate the C statistics for each individual risk factor. We constructed models by adding each novel risk factor independently to the basic risk-prediction model, as we did in the total cohort, and looked for an incremental change of at least 0.005. We used the nriidi macro created by Sundstrom et al. (27) to calculate the NRI, IDI, and associated *P* values.

RESULTS

Total cohort

There were 1,457 (11.9%) incident cases of type 2 diabetes with a mean of >7.6 years of follow-up. Unadjusted baseline characteristics of the type 2 diabetes and non–type 2 diabetes groups for the total cohort are summarized in Table 1. Individuals with incident type 2 diabetes were more likely to have a parental history of diabetes, be African American versus white, have a higher systolic blood pressure, higher mean waist circumference, greater mean height, higher levels of fasting triglycerides and glucose, and lower levels of HDL-C. In terms of novel risk factors, individuals with incident type 2 diabetes had statistically significantly different levels of all novel risk factors except albumin and low-frequency-power heart rate variability when compared with those without incident type 2 diabetes.

Multivariate-adjusted hazard rate ratios for the variables included in the basic risk model are shown in Supplementary Table 2. Supplementary Table 3 shows the correlations between novel and basic risk factors. All of the novel risk factors except for the genetic risk score were significantly correlated with at least three of the six basic risk factors, and the majority of novel risk factors are significantly correlated with five or more basic risk factors. However, it is important to note that not all of these correlations were strong. By contrast, the genetic risk score was only correlated with baseline glucose levels.

The basic model had an AUC of 0.8411 (95% CI: 0.8316–0.8457) (Table 2). There were no novel risk factors that improved the AUC of the basic model by an increment of at least 0.005. All of the novel risk factors, with the exception of FEV_1, did not have statistically significant NRIs. The NRI when adding FEV_1 to the basic model was 0.54% (95% CI: 0.33–0.86%). Finally, the addition of WBC count, albumin, aPTT, factor VIII, magnesium, heart rate, hip circumference, or the genetic risk score statistically significantly increased the IDI.

Nested case-control sample

Baseline characteristics of incident type 2 diabetes case and control subjects in the nested case-control subsample are summarized in Table 3. All novel risk factors were statistically significantly different between case and control subjects. Supplementary Table 4 shows that all of the novel risk factors were significantly correlated with at least two of the basic risk factors, and the majority of novel risk factors were significantly correlated with three or more basic risk factors. As with the total cohort, not all of these risk factors were strongly correlated.

The basic model, which included the aforementioned risk factors from the Schmidt et al. (25) analysis, had an AUC of 0.8607 (95% CI: 0.8386–0.8828) (Table 4). None of the novel risk factors improved the C statistic by at least 0.005. In terms of model fit, the addition of ICAM-1 had the greatest improvement as it came closest to achieving the 0.005 increment of change in the AUC. None of the novel risk factors exhibited a statistically

Table 3—*Baseline characteristics (mean or percentage) of case-control sample by type 2 diabetes status*

	Type 2 diabetes (*N* = 522)	No type 2 diabetes (*N* = 526)	*P* value
Basic risk factors			
Age (years)	53.1	52.3	0.03
Parental history of diabetes (%)	35.3	22.2	<0.0001
Race (African American, %)	47.9	41.1	0.03
Systolic blood pressure (mmHg)	124.8	119.2	<0.0001
Waist girth (cm)	105.0	94.0	<0.0001
Height (cm)	167.8	167.3	0.39
Triglycerides (mg/dL)	146.1	108.6	<0.0001
HDL-C (mg/dL)	47.9	56.3	<0.0001
Glucose (mg/dL)	108.1	96.7	<0.0001
Novel risk factors			
Adiponectin (μg/mL)	6.9	9.3	<0.0001
High-molecular-weight adiponectin (g/mL)	1.9	3.0	<0.0001
Leptin (ng/mL)	26.1	19.4	<0.0001
GGT (U/L)	21.0	15.9	0.002
Alanine aminotransferase (U/L)	14.7	12.6	0.01
Fetuin A (g/mL)	501.5	486.5	0.009
Ferritin (ng/mL)	202.9	144.4	<0.0001
C-reactive protein (mg/mL)	3.8	2.7	<0.0001
Lactate (mg/dL)	9.4	8.3	<0.0001
Oxidized LDL (U/L)	42.7	38.9	<0.0001
ICAM-1 (ng/mL)	277.3	255.8	<0.0001
Orosomucoid (mg/dL)	94.6	86.6	<0.0001
Sialic acid (mg/dL)	100.8	93.5	<0.0001
Interleukin-6 (pg/mL)	3.3	2.7	0.003
Interleukin-18 (pg/mL)	254.0	234.4	0.01
Complement C3 (mg/dL)	173.2	151.8	<0.0001
RBP-4 (μg/mL)	31.2	29.3	0.0002
Nonesterified free fatty acids (g/L)	0.9	0.8	0.0009
ADMA (μmol/L)	0.25	0.24	0.04

Table 4—*Measures of risk prediction and model fit for the case-control sample*

	N	C statistic	95% CI	Difference*	AIC	NRI (SE)	*P* value	IDI (SE)	*P* value
Basic**	1,048	0.8607	0.8386–0.8828		979				
Adiponectin	1,048	0.8624	0.8404–0.8844	0.0017	996	0.0133 (0.0083)	0.11	0.0034 (0.0017)	0.04
High-molecular-weight adiponectin	1,043	0.8604	0.8382–0.8827	0.0002	997	0.0019 (0.0051)	0.71	0.0002 (0.0005)	0.70
Leptin	1,048	0.8617	0.8396–0.8838	0.0010	976	0.0076 (0.0060)	0.21	0.0031 (0.0015)	0.04
GGT	1,046	0.8613	0.8392–0.8834	0.0004	973	0.0000 (0.0054)	1.00	0.0038 (0.0019)	0.04
Alanine aminotransferase	1,046	0.8607	0.8385–0.8828	−0.0002	976	−0.0019 (0.0042)	0.66	0.0015 (0.0013)	0.25
Fetuin A	1,038	0.8614	0.8393–0.8836	0.0005	966	−0.0019 (0.0051)	0.70	0.0022 (0.0014)	0.10
Ferritin	1,017	0.8644	0.8423–0.8865	0.0020	937	0.0019 (0.0065)	0.77	0.0051 (0.0023)	0.03
C-reactive protein	1,048	0.8610	0.8389–0.8831	0.0003	978	0.0057 (0.0042)	0.18	0.0006 (0.0008)	0.50
Lactate	1,030	0.8613	0.8391–0.8836	0.0001	960	0.0000 (0.0027)	1.00	0.0001 (0.0020)	0.70
Oxidized LDL	1,048	0.8608	0.8386–0.8829	0.0001	979	0.0038 (0.0026)	0.16	0.0002 (0.0041)	0.56
ICAM-1	1,024	0.8644	0.8423–0.8865	0.0028	947	0.0078 (0.0067)	0.24	0.0062 (0.0023)	0.01
Orosomucoid	1,048	0.8613	0.8392–0.8833	0.0006	978	0.0076 (0.0047)	0.10	0.0006 (0.0008)	0.40
Sialic acid	1,048	0.8608	0.8387–0.8830	0.0001	979	−0.00001 (0.0038)	0.99	0.0004 (0.0006)	0.43
Interleukin-6	1,048	0.8611	0.8390–0.8832	0.0004	978	0.0076 (0.0038)	0.05	0.0006 (0.0007)	0.41
Interleukin-18	1,037	0.8621	0.8399–0.8842	0.0000	965	0.0000 (0.0000)	—	0.0000 (0.0001)	0.69
Complement C3	1,043	0.8623	0.8402–0.8844	0.002	970	0.0037 (0.0076)	0.62	0.0044 (0.0019)	0.02
RBP-4	1,038	0.8612	0.8390–0.8833	0.0003	968	−0.0058 (0.0051)	0.26	0.0030 (0.0006)	0.63
Nonesterified free fatty acids	1,047	0.8614	0.8393–0.8835	0.0010	975	0.0019 (0.0069)	0.78	0.0027 (0.0016)	0.09
ADMA	1,033	0.8585	0.8361–0.8810	0.0000	971	0.0000 (0.0000)	—	−0.0000 (0.0002)	0.97

*The basic model was rerun for each covariate, and the difference between the basic and expanded models was calculated accordingly due to changing sample sizes.
**The basic model includes age, parental history of diabetes, race/ethnicity, fasting glucose, fasting triglycerides, systolic blood pressure, HDL-C, height, and waist circumference.

significant NRI. ADMA, interleukin-18, GGT, and lactate exhibited no movement between risk categories, and therefore, an NRI calculation was not made. Adiponectin, leptin, GGT, ferritin, ICAM-1, and complement C3 had statistically significantly improved IDIs.

CONCLUSIONS—None of the novel risk factors significantly improved the AUC in the total cohort or nested case-control sample. However, FEV_1 did significantly, albeit modestly, improve the NRI in the total cohort. None of the risk factors statistically significantly improved the NRI in the case-control study sample; however, the addition of adiponectin, leptin, GGT, ferritin, ICAM-1, and complement C3 did statistically significantly but moderately improve the IDI. Likewise, in the total cohort, the novel risk factors WBC count, albumin, aPTT, factor VIII, magnesium, heart rate, hip circumference, and the genetic risk score exhibited significant but modest improvements in IDI. These results suggest that of these novel risk factors, only FEV_1 may be helpful for type 2 diabetes risk stratification in the ARIC cohort study. Several novel risk factors did modestly improve the IDI, which indicates that the difference in average predicted probabilities between individuals with and without type 2 diabetes significantly increased when these risk factors were added to the basic model; however, critics argue that it is unclear whether a significant IDI indicates that the novel risk factor in the model is clinically useful (28,29).

Despite the fact that many of the novel risk factors are independent risk factors for type 2 diabetes in the total cohort, none of these risk factors appeared to provide additional value to type 2 diabetes risk prediction. Previous studies that have incorporated one or more novel risk factors into a risk prediction model have been limited, and although these analyses may have found increased AUCs with the inclusion of novel risk factors, they are also often single studies in very specific populations (4,30,31). Our own study failed to replicate the contributions of WBC count, heart rate, or alanine aminotransferase to the improvement in the AUC, as found in the aforementioned studies (4,30,31). It is important to note that although novel risk factors may be associated with type 2 diabetes, it does not mean they will contribute to risk prediction, as these are separate issues of etiology and prediction (32). All of the novel risk factors modeled in the total cohort and case-control analyses were significantly associated with type 2 diabetes in ARIC; however, none of them significantly contributed to improved risk prediction when C statistics were calculated with and without the novel risk factors.

It is difficult to improve upon existing risk factors for type 2 diabetes. Specifically, when a single measurement of obesity or glycemia is included in a risk model, the AUCs already range from 0.66–0.77. When obesity and glycemia measures are combined with readily available clinical variables, such as those included in the basic model, the AUC increases greatly, making it difficult to improve the risk prediction (32). Furthermore, the correlation between novel risk factors and traditional risk factors must also be considered, as correlated risk factors provide less independent information about type 2 diabetes risk. We found this to be true in our own analysis, as many statistically significant correlations existed between traditional risk factors and novel risk factors in both the total cohort and the case-control analysis.

Recent advances in the identification of a number of genetic variants associated with type 2 diabetes have generated interest in the clinical utility of combining the loci associated with type 2 diabetes into a genetic risk score, which could be used for risk prediction. Thus far, the use of genetic risk scores in type 2 diabetes risk prediction models prior to this analysis has been limited, often involved a smaller number of genetic variants, and yielded varied results (33).

Our own analysis did not find a statistically significant contribution to the AUC or NRI with the addition of a genetic risk score; however, it did moderately improve the IDI. The incorporation of a genetic risk score into future type 2 diabetes risk prediction models could be more useful, once an ideal set of variants is identified, as genes are not prone to the biological variability or measurement error that often accompanies other risk factors. Further, the genotype does not change over one's lifetime, and this offers opportunities for earlier screening and identification of individuals at risk (34). In fact, de Miguel-Yanes et al. (35) found that the incorporation of a genetic score into a risk model was actually more beneficial in younger subjects. Identifying individuals at risk earlier in the disease process will allow for interventions that can either reverse the course of the disease or control its accompanying risk factors such as dyslipidemia and hypertension.

Limitations to this study include the absence of an oral glucose tolerance test or hemoglobin A_{1c} test results to classify type 2 diabetes and the use of a single baseline value for the novel risk factors, which does not capture the variation in levels over time for risk factors. Further, not all novel risk factors are included in this analysis. We chose to only include biomarkers that had not previously been included in risk prediction analyses in ARIC and biomarkers that were measured and not self-reported.

Another limitation is the inclusion of only 30 SNPs in the genetic risk score, which account for only a small fraction of the heritability of type 2 diabetes (36). Finally, there were 35 novel risk factors evaluated, resulting in multiple testing that may yield false positives. A strength of this analysis was the availability of a large, population-based cohort of white and African American men and women with follow-up data. Further, there were standardized data collection methods for both predictors and type 2 diabetes outcomes.

In conclusion, our modeling indicates that no novel risk factor contributed significantly to risk prediction, as measured by the AUC. There was a modest improvement in risk classification with the addition of FEV_1 and a small improvement in the IDI with the addition of WBC count, aPTT, albumin, factor VIII, magnesium, heart rate, hip circumference, and the genetic risk score in the total cohort and adiponectin, leptin, GGT, ferritin, ICAM-1, and complement C3 in the case-control sample. However, these improvements are small and unlikely to motivate refinement of clinical risk reclassification or discrimination strategies. Further study by prospective, population-based cohort studies is needed to confirm the generalizability of these findings.

Acknowledgments—The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN-268201100005C, HHSN-268201100006C, HHSN-268201100007C, HHSN-268201100008C, HHSN-268201100009C, HHSN-268201100010C, HHSN-268201100011C, and HHSN-268201100012C), grants R01-HL-087641, R01-HL-59367, and R01-HL-086694; National Human Genome Research Institute contract U01-HG-004402; and National Institutes of Health (NIH) contract HHSN-268200625226C, with the ARIC carotid MRI examination funded by U01-HL-075572-01. Infrastructure was partly supported by grant UL1-RR-025005, a component of the NIH and NIH Roadmap for Medical Research.

No potential conflicts of interest relevant to this article were reported.

L.A.R. analyzed data and wrote the manuscript. J.S.P. contributed to data analysis and was the primary editor. B.B.D., M.I.S., R.C.H., M.A.P., J.H.Y., and C.M.B. contributed to data collection and reviewed and edited the manuscript. L.A.R. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Parts of this work were presented in poster form at the 45th Annual Society for Epidemiologic Research Meeting, Minneapolis, Minnesota, 27–30 June 2012.

The authors thank the staff and participants of the ARIC Study for important contributions.

References

1. Schwarz PE, Li J, Lindstrom J, Tuomilehto J. Tools for predicting the risk of type 2 diabetes in daily practice. Horm Metab Res 2009;41:86–97
2. Hippisley-Cox J, Coupland C, Robson J, Sheikh A, Brindle P. Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore. BMJ 2009;338:b880
3. Gillies CL, Abrams KR, Lambert PC, et al. Pharmacological and lifestyle interventions to prevent or delay type 2 diabetes in people with impaired glucose tolerance: systematic review and meta-analysis. BMJ 2007;334:299
4. Bindraban NR, van Valkengoed IG, Mairuhu G, et al. Prevalence of diabetes mellitus and the performance of a risk score among Hindustani Surinamese, African Surinamese and ethnic Dutch: a cross-sectional population-based study. BMC Public Health 2008;8:271
5. Ma J, Folsom AR, Melnick SL, et al. Associations of serum and dietary magnesium with cardiovascular disease, hypertension, diabetes, insulin, and carotid arterial wall thickness: the ARIC Study. Atherosclerosis Risk in Communities Study. J Clin Epidemiol 1995;48:927–940
6. Carnethon MR, Jacobs DR Jr, Sidney S, Liu K; CARDIA study. Influence of autonomic nervous system dysfunction on the development of type 2 diabetes: the CARDIA study. Diabetes Care 2003;26:3035–3041
7. Wang L, Folsom AR, Zheng ZJ, Pankow JS, Eckfeldt JH; ARIC Study Investigators. Plasma fatty acid composition and incidence of diabetes in middle-aged adults: the Atherosclerosis Risk in Communities (ARIC) Study. Am J Clin Nutr 2003;78:91–98
8. Duncan BB, Schmidt MI, Pankow JS, et al.; Atherosclerosis Risk in Communities Study. Low-grade systemic inflammation and the development of type 2 diabetes: the Atherosclerosis Risk in Communities Study. Diabetes 2003;52:1799–1805
9. Duncan BB, Schmidt MI, Pankow JS, et al. Adiponectin and the development of type 2 diabetes: the Atherosclerosis Risk in Communities Study. Diabetes 2004;53:2473–2478
10. Pankow JS, Duncan BB, Schmidt MI, et al.; Atherosclerosis Risk in Communities Study. Fasting plasma free fatty acids and risk of type 2 diabetes: the Atherosclerosis Risk in Communities Study. Diabetes Care 2004;27:77–82
11. Yeh HC, Punjabi NM, Wang NY, Pankow JS, Duncan BB, Brancati FL. Vital capacity as a predictor of incident type 2 diabetes: the Atherosclerosis Risk in Communities Study. Diabetes Care 2005;28:1472–1479
12. Hoogeveen RC, Ballantyne CM, Bang H, et al. Circulating oxidised low-density lipoprotein and intercellular adhesion molecule-1 and risk of type 2 diabetes mellitus: the Atherosclerosis Risk in Communities Study. Diabetologia 2007;50:36–42

13. Schmidt MI, Duncan BB, Vigo A, et al.; ARIC Investigators. Leptin and incident type 2 diabetes: risk or protection? Diabetologia 2006;49:2086–2096
14. Carnethon MR, Prineas RJ, Temprosa M, Zhang ZM, Uwaifo G, Molitch ME; Diabetes Prevention Program Research Group. The association among autonomic nervous system function, incident diabetes, and intervention arm in the Diabetes Prevention Program. Diabetes Care 2006;29:914–919
15. Reich LM, Heiss G, Boland LL, Hirsch AT, Wu K, Folsom AR. Ankle-brachial index and hemostatic markers in the Atherosclerosis Risk in Communities (ARIC) Study cohort. Vasc Med 2007;12:267–273
16. Jehn ML, Guallar E, Clark JM, et al. A prospective study of plasma ferritin level and incident diabetes: the Atherosclerosis Risk in Communities (ARIC) Study. Am J Epidemiol 2007;165:1047–1054
17. Tamariz LJ, Young JH, Pankow JS, et al. Blood viscosity and hematocrit as risk factors for type 2 diabetes mellitus: the Atherosclerosis Risk in Communities (ARIC) Study. Am J Epidemiol 2008;168:1153–1160
18. Parker ED, Pereira MA, Stevens J, Folsom AR. Association of hip circumference with incident diabetes and coronary heart disease: the Atherosclerosis Risk in Communities Study. Am J Epidemiol 2009;169:837–847
19. Weitzman S, Wang CH, Pankow JS, Schmidt MI, Brancati FL. Are measures of height and leg length related to incident diabetes mellitus? The ARIC (Atherosclerosis Risk in Communities) Study. Acta Diabetol 2010;47:237–242
20. Zhu N, Pankow JS, Ballantyne CM, et al. High-molecular-weight adiponectin and the risk of type 2 diabetes in the ARIC Study. J Clin Endocrinol Metab 2010;95:5097–5104
21. Crawford SO, Hoogeveen RC, Brancati FL, et al. Association of blood lactate with type 2 diabetes: the Atherosclerosis Risk in Communities Carotid MRI Study. Int J Epidemiol 2010;39:1647–1655
22. ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. Am J Epidemiol 1989;129:687–702
23. Rasmussen-Torvik LJ, Alonso A, Li M, et al. Impact of repeated measures and sample selection on genome-wide association studies of fasting glucose. Genet Epidemiol 2010;34:665–673
24. Voight BF, Scott LJ, Steinthorsdottir V, et al.; MAGIC investigators; GIANT Consortium. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 2010;42:579–589
25. 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
26. Chambless LE, Cummiskey CP, Cui G. Several methods to assess improvement in risk prediction models: extension to survival analysis. Stat Med 2011;30:22–38
27. Sundström J, Byberg L, Gedeborg R, Michaëlsson K, Berglund L. Useful tests of usefulness of new risk factors: tools for assessing reclassification and discrimination. Scand J Public Health 2011;39:439–441
28. Cook NR, Ridker PM. Advances in measuring the effect of individual predictors of cardiovascular risk: the role of reclassification measures. Ann Intern Med 2009;150:795–802
29. Mihaescu R, van Zitteren M, van Hoek M, et al. Improvement of risk prediction by genomic profiling: reclassification measures versus the area under the receiver operating characteristic curve. Am J Epidemiol 2010;172:353–361
30. Sun F, Tao Q, Zhan S. An accurate risk score for estimation 5-year risk of type 2 diabetes based on a health screening population in Taiwan. Diabetes Res Clin Pract 2009;85:228–234
31. Chien K, Cai T, Hsu H, et al. A prediction model for type 2 diabetes risk among Chinese people. Diabetologia 2009;52:443–450
32. Sattar N, Wannamethee SG, Forouhi NG. Novel biochemical risk factors for type 2 diabetes: pathogenic insights or prediction possibilities? Diabetologia 2008;51:926–940
33. Willems SM, Mihaescu R, Sijbrands EJ, van Duijn CM, Janssens AC. A methodological perspective on genetic risk prediction studies in type 2 diabetes: recommendations for future research. Curr Diab Rep 2011;11:511–518
34. Talmud PJ, Hingorani AD, Cooper JA, et al. Utility of genetic and non-genetic risk factors in prediction of type 2 diabetes: Whitehall II prospective cohort study. BMJ 2010;340:b4838
35. de Miguel-Yanes JM, Shrader P, Pencina MJ, et al.; MAGIC Investigators; DIAGRAM+ Investigators. Genetic risk reclassification for type 2 diabetes by age below or above 50 years using 40 type 2 diabetes risk single nucleotide polymorphisms. Diabetes Care 2011;34:121–125
36. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461:747–753