



Racial Differences in the Performance of Existing Risk Prediction Models for Incident Type 2 Diabetes: The CARDIA Study

Diabetes Care 2016;39:285–291 | DOI: 10.2337/dc15-0509

Mary E. Lacy,^{1,2} Gregory A. Wellenius,¹ Mercedes R. Carnethon,³ Eric B. Loucks,¹ April P. Carson,⁴ Xi Luo,⁵ Catarina I. Kiefe,⁶ Annie Gjelsvik,¹ Erica P. Gunderson,⁷ Charles B. Eaton,¹ and Wen-Chih Wu^{1,2,8}

OBJECTIVE

In 2010, the American Diabetes Association (ADA) added hemoglobin A_{1c} (A1C) to the guidelines for diagnosing type 2 diabetes. However, existing models for predicting diabetes risk were developed prior to the widespread adoption of A1C. Thus, it remains unknown how well existing diabetes risk prediction models predict incident diabetes defined according to the ADA 2010 guidelines. Accordingly, we examined the performance of an existing diabetes prediction model applied to a cohort of African American (AA) and white adults from the Coronary Artery Risk Development Study in Young Adults (CARDIA).

RESEARCH DESIGN AND METHODS

We evaluated the performance of the Atherosclerosis Risk in Communities (ARIC) diabetes risk prediction model among 2,456 participants in CARDIA free of diabetes at the 2005–2006 exam and followed for 5 years. We evaluated model discrimination, calibration, and integrated discrimination improvement with incident diabetes defined by ADA 2010 guidelines before and after adding baseline A1C to the prediction model.

RESULTS

In the overall cohort, re-estimating the ARIC model in the CARDIA cohort resulted in good discrimination for the prediction of 5-year diabetes risk (area under the curve [AUC] 0.841). Adding baseline A1C as a predictor improved discrimination (AUC 0.841 vs. 0.863, $P = 0.03$). In race-stratified analyses, model discrimination was significantly higher in whites than AA (AUC AA 0.816 vs. whites 0.902; $P = 0.008$).

CONCLUSIONS

Addition of A1C to the ARIC diabetes risk prediction model improved performance overall and in racial subgroups. However, for all models examined, discrimination was better in whites than AA. Additional studies are needed to further improve diabetes risk prediction among AA.

In 2010, the American Diabetes Association (ADA) modified the diagnostic guidelines for type 2 diabetes to include hemoglobin A_{1c} (A1C) (1,2). However, existing models for predicting diabetes risk were developed before the widespread adoption of A1C as a diagnostic test for diabetes (3–5). Thus, established diabetes risk

¹Department of Epidemiology, School of Public Health, Brown University, Providence, RI

²Center of Innovation in Long-Term Services and Supports, Providence Veterans Affairs Medical Center, Providence, RI

³Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL

⁴Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham, AL

⁵Department of Biostatistics, School of Public Health, Brown University, Providence, RI

⁶Department of Qualitative Health Sciences, University of Massachusetts Medical School, Worcester, MA

⁷Cardiovascular and Metabolic Conditions Section, Division of Research, Kaiser Permanente Northern California, Oakland, CA

⁸Division of Cardiology, Providence Veterans Affairs Medical Center, Providence, RI

Corresponding author: Mary E. Lacy, mary_lacy@brown.edu.

Received 11 March 2015 and accepted 18 October 2015.

This article contains Supplementary Data online at <http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc15-0509/-/DC1>.

© 2016 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered.

prediction models do not include A1C. Additionally, most existing risk prediction models were developed in populations with few or no African Americans (AA), despite the fact that AA are at increased risk of type 2 diabetes and vascular complications (6). Three of the more commonly used risk prediction models for incident type 2 diabetes were developed in the Atherosclerosis Risk in Communities (ARIC) study ($N = 7,916$; 85% white, 15% AA), the Framingham Offspring Study ($N = 3,140$; 99% white), and the San Antonio Heart Study ($N = 3,004$; 61% Mexican American, 39% white) (7–9). An article comparing the validity of these three models in a multiethnic cohort reported good discrimination (area under the curve [AUC] 0.78–0.81) for all three models, though model discrimination was lower in AA than whites across all three models (10). Importantly, A1C was not included in any of these risk prediction models.

Findings from a number of studies have established that AA have higher A1C than whites, with estimates of the absolute A1C difference between AA and whites ranging from 0.40 to 0.65% after adjustment for glucose levels (11–13). Despite consistent evidence of higher A1C values in AA, the clinical significance of this difference in A1C is unclear. No racial differences were found for the association of A1C with incident coronary heart disease, stroke, or chronic kidney disease in a prospective study of older AA and whites; however, a cross-sectional study found that the prevalence of retinopathy was elevated in AA versus whites at the same A1C (14,15). These findings suggest that the benefit of A1C as a potential predictor of incident diabetes should be further explored in different racial groups. The objectives of this study were as follows: 1) to examine the performance of an existing risk prediction model in a biracial cohort of AA and white adults from the Coronary Artery Risk Development Study in Young Adults (CARDIA) using ADA 2004 guidelines, 2) to examine model performance when diabetes status is ascertained using ADA 2010 guidelines, and 3) to examine change in model performance with the addition of baseline A1C as a predictor of diabetes risk.

With the recent changes to diagnostic guidelines for type 2 diabetes, we hypothesized that model discrimination

would be lower using ADA 2010 diagnostic guidelines compared with ADA 2004 guidelines (the guidelines that were in use when existing models were developed). Additionally, we hypothesized that including baseline A1C as a predictor in the ADA 2010 model would improve prediction of diabetes in both whites and AA. However, given the increased risk of diabetes in AA and A1C differences in AA compared with whites, we hypothesized that model discrimination would be significantly lower in AA than whites when A1C was incorporated into the prediction models.

RESEARCH DESIGN AND METHODS

Study Population

Details regarding the CARDIA study design have previously been published (16). In brief, CARDIA is an ongoing, multicenter longitudinal study of the determinants of cardiovascular disease in 5,115 adults aged 18–30 years at the baseline assessment in 1985–1986. A stratified sample of 2,637 AA and 2,478 white men and women were recruited from Minneapolis, MN; Chicago, IL; Birmingham, AL; and Oakland, CA. Participants had follow-up examinations at 2, 5, 7, 10, 15, 20, and 25 years after enrollment. In CARDIA, fasting glucose was measured at baseline and years 7, 10, 15, 20, and 25; 2-h 75-g oral glucose tolerance test was performed in years 10, 20, and 25; and A1C was measured at years 20 and 25. For construction of a diabetes definition based on the current ADA diagnostic guidelines (which include A1C), analyses were restricted to follow-up examinations at which A1C was measured, i.e., years 20 and 25 of follow-up.

Therefore, the year 20 exam (2005–2006) was the baseline for the current study and incident diabetes was determined at the year 25 exam (2010–2011). A total of 3,549 participants completed the year 20 exam (74% of the surviving cohort). Participants with prevalent diabetes at the year 20 exam ($n = 332$) and those who were missing diabetes status at year 20 ($n = 41$) or year 25 ($n = 379$) or who were missing covariate information ($n = 341$) were excluded, resulting in 2,456 participants included in current study. Because fasting glucose levels were used as a covariate and in the outcome definition, analyses were restricted to participants who were fasting.

Data Collection

CARDIA data were collected according to standardized protocols across the four study sites as previously published in detail (16). All covariates were measured at the year 20 examination, except parental history of diabetes, which was measured at the year 10 examination by self-report. Interviewers collected data on participants' self-reported race, sex, and date of birth at the baseline examination and verified these data at each subsequent examination. Self-reported medication use was ascertained by trained interviewers at the year 20 assessment. Height and weight were measured with participants wearing light clothing and no shoes. Body weight was measured to the nearest 0.2 kg using a calibrated balance-beam scale, and height was measured with a vertical ruler to the nearest 0.5 cm; BMI was calculated as weight in kilograms divided by the square of height in meters. Waist circumference was measured to the nearest 0.5 cm at the minimum abdominal girth with participants standing upright; the average of two waist circumference measurements was used. Three seated blood pressure measurements were taken for each participant after a 5-min rest using an automated blood pressure monitor, with the average of the last two measurements used to determine systolic and diastolic blood pressure. Lipid assays were used to measure total, HDL, and LDL cholesterol and triglycerides. At the year 20 and 25 follow-up, fasting and 2-h postload glucose were measured by the hexokinase ultraviolet method and A1C was assessed using a Tosoh G7 high-performance liquid chromatography instrument. The coefficient of variation for all assays was <6%.

Risk Prediction Models

We calculated the 5-year predicted probability of developing type 2 diabetes for CARDIA participants based on the existing type 2 diabetes prediction model derived in the ARIC study (7). The ARIC model was developed on 7,915 participants (85% non-Hispanic white, 15% AA) using ~9 years of follow-up starting in 1987–1989. The ARIC model includes the following predictors: age, race (AA vs. White), waist circumference, height, parent history of type 2 diabetes, systolic blood pressure, HDL cholesterol, triglycerides, and fasting glucose. We selected this model as our primary exposure, as it

is the only existing model that explicitly included an indicator variable for AA race (3–5,7). The ARIC model predicts incident diabetes over 9 years of follow-up. To account for the different length of follow-up available for CARDIA participants, we divided each CARDIA participant's predicted probability by the number of years of follow-up used in the prior study (e.g., in ARIC, 9 years of follow-up) and multiplied this number by 5 to obtain the 5-year predicted probability of developing diabetes; this method assumes a constant risk of diabetes (10).

In sensitivity analyses, we examined the performance of two additional published type 2 diabetes prediction models from the Framingham Offspring Study and the San Antonio Heart Study (8,9). In contrast to the ARIC model, neither Framingham nor the San Antonio Heart Study had any AA participants. Thus, the diabetes prediction models from these two cohorts provide an interesting comparison. Model components and selected cohort characteristics are described in Table 2.

Assessment of Diabetes

We evaluated incident type 2 diabetes using two definitions: 1) an outcome definition adapted from the ADA 2004 diagnostic guidelines, which included reported use of antidiabetes medication, fasting glucose ≥ 126 mg/dL, or 2-h oral glucose tolerance test ≥ 200 mg/dL (1), and 2) an outcome definition adapted from the ADA 2010 diagnostic guidelines, which included reported use of antidiabetes medication or fasting glucose ≥ 126 mg/dL, 2-h oral glucose tolerance test ≥ 200 mg/dL, or A1C $\geq 6.5\%$ (2). Our definition of diabetes required only one elevated measure to define diabetes; confirmatory testing was not available.

In a sensitivity analysis, we examined an alternate definition of type 2 diabetes using an outcome definition adapted from the ADA 2010 diagnostic guidelines as indicated above with the addition of self-reported physician diagnosis of diabetes (i.e., reported use of antidiabetes medication or fasting glucose ≥ 126 mg/dL, 2-h oral glucose tolerance test ≥ 200 mg/dL, or A1C $\geq 6.5\%$ or self-reported physician diagnosis of diabetes).

Statistical Analysis

Using the data from the CARDIA year 20 clinical assessment (the baseline for this study), we examined the distribution

of demographics, anthropometrics, medical history, and clinical covariates overall and by race. We also examined the 5-year incidence of type 2 diabetes in the overall population and stratified by race using both the ADA 2004 and ADA 2010 guidelines.

For each risk prediction model, we ran two sets of logistic regression models. The first set of models (model 1) used CARDIA data and the regression coefficients from the original published ARIC model to calculate the predicted probability of developing diabetes for each participant using ADA 2004 and ADA 2010 diagnostic guidelines. In the second set of models (model 2), we refit the regression models using the same predictors from the original published ARIC model and estimated new regression coefficients using the CARDIA data. We also examined change in model performance after adding baseline A1C as a predictor of risk. All regression models were estimated for the overall cohort and in subgroups stratified by race. To determine whether A1C results varied by other sociodemographic factors, we calculated the odds of developing type 2 diabetes in subgroups defined by sex, age (< 45 years old vs. ≥ 45 years old), and education level (did not complete high school vs. graduated high school) in addition to race (AA vs. white). Models were evaluated using three criteria to assess model fit: 1) model discrimination was evaluated using area under the receiver operating curve, a measure of how well the model ranked individuals who developed diabetes as at higher risk than those who did not (17,18); 2) model calibration, which assesses how close the predicted risks are to the observed risks (summarized using the Hosmer-Lemeshow goodness-of-fit test) (19); and 3) integrated discrimination improvement (IDI) after addition of baseline A1C as a predictor (20–22). IDI is a measure of the separation in predicted probabilities for events and nonevents across the “old” (ARIC predictors only) and “new” (ARIC predictors + A1C) models. Relative IDI compares the relative contribution of the new predictor (A1C) with the average contribution of the predictors from the original model. The original ARIC model has nine predictors; with the assumption that each variable contributes equally to the discrimination

slope, the average contribution of each predictor is 11.1%. Estimates and 95% CIs for absolute and relative IDI findings were estimated using 999 bootstrap replications with replacement (18,23). Analyses were conducted using SAS 9.3 statistical software (SAS Institute, Cary, NC).

RESULTS

Of the 2,456 participants included in analyses, 40.7% were AA and 59.3% were white. Whites tended to be older than AA, were more likely to be male, and had higher mean levels of triglycerides and fasting glucose. AA were more likely to have a parent with a history of type 2 diabetes, were more likely to be prediabetic, and had significantly higher mean values of BMI, waist circumference, systolic blood pressure, and A1C (Table 1).

The 5-year cumulative incidence of type 2 diabetes differed substantially when using the ADA 2004 diagnostic guidelines compared with the ADA 2010 diagnostic guidelines. In the overall sample, the 5-year incidence of type 2 diabetes was 3.0% ($n = 74$) under the ADA 2004 diagnostic guidelines versus 5.1% ($n = 124$) using the ADA 2010 guidelines. AA were significantly more likely than whites to develop diabetes using either guideline: under ADA 2004 diagnostic guidelines, 4.5% ($n = 45$) of AA developed type 2 diabetes versus 2.0% ($n = 29$) of whites ($P < 0.0001$); under ADA 2010 guidelines, 7.6% ($n = 76$) of AA developed type 2 diabetes versus 3.3% ($n = 48$) of whites ($P < 0.0001$).

Components for the ARIC model are presented in Table 2. With use of the previously published regression coefficients, the ARIC model yielded very high discrimination (AUC 0.846) for incident diabetes defined according to the ADA 2004 diagnostic guidelines (Table 3) (model 1a). Model discrimination was slightly lower when incident diabetes was defined according to the ADA 2010 diagnostic guidelines (model 1b) (AUC 0.822 vs. 0.846, $P = 0.48$ for difference between the two AUCs). With use of the same set of predictors from the published ARIC model and re-estimation of the regression equation in CARDIA, the prediction model achieved similar discrimination (AUC 0.841) for diabetes defined according to the ADA 2010 diagnostic guidelines (model 2a). Adding baseline A1C as a covariate in the prediction model improved discrimination significantly (model 2b)

Table 1—Characteristics of CARDIA participants at year 20 (2005–2006)

Characteristic	Overall (N = 2,456)	AA (n = 999)	Whites (n = 1,457)	P†
Age (years)	45.29 ± 3.56	44.59 ± 3.81	45.78 ± 3.30	<0.0001
Male (%)	43.04	38.24	46.33	<0.0001
BMI (kg/m ²)	28.80 ± 6.39	30.55 ± 6.81	27.60 ± 5.78	<0.0001
Waist circumference (cm)	90.42 ± 14.35	92.37 ± 14.18	89.08 ± 14.31	<0.0001
Parental history of diabetes (%)	17.79	22.82	14.34	<0.0001
Systolic blood pressure (mmHg)	114.58 ± 14.07	118.62 ± 15.43	111.81 ± 12.32	<0.0001
HDL cholesterol (mg/dL)	54.87 ± 16.65	55.24 ± 16.16	54.62 ± 16.98	0.37
Triglycerides (mg/dL)	106.16 ± 73.15	94.32 ± 68.28	114.27 ± 75.27	<0.0001
Fasting glucose (mg/dL)	94.33 ± 9.35	93.72 ± 10.37	94.74 ± 8.56	<0.01
A1C (%)	5.34 ± 0.37	5.46 ± 0.41	5.26 ± 0.31	<0.0001
A1C (mmol/mol)	35 ± 4.0	36 ± 4.5	34 ± 3.4	<0.0001
Prediabetes (%)‡	39.05	46.55	33.91	<0.0001

Data are means ± SD unless otherwise indicated. †Comparison by χ^2 for categorical variables and unpaired *t* tests for continuous variables. ‡Prediabetes defined as fasting glucose 100–125 mg/dL, A1C 5.7–6.4%, or oral glucose tolerance test 140–199 mg/dL.

Interaction terms examining the effect of A1C on diabetes risk in subgroups defined by sex, age, education level, and race indicated that the effect of A1C did not differ in any of these subgroups except for race (*P* values for interaction terms: sex * A1C, *P* = 0.74; age * A1C, *P* = 0.45; education * A1C, *P* = 0.26; and race * A1C, *P* = 0.05).

Race-specific analyses revealed important differences in model discrimination. Across all models, model discrimination was higher among whites than AA, though these differences were not all statistically significant. With use of the ADA 2004 diagnostic guidelines (model 1a), model discrimination was higher in whites than in AA, though differences did not reach statistical significance (AUC in AA 0.802 vs. AUC in whites 0.887, *P* = 0.10). For all models that defined diabetes according to the ADA 2010 diagnostic guidelines (models 1b, 2a, and 2b), model discrimination was significantly higher in whites than in AA (model 1b, AUC in AA 0.778 vs. AUC in whites 0.860, *P* = 0.04; model 2a, AUC in

(AUC 0.841 vs. 0.863, *P* = 0.03 for differences between the two AUCs). Hosmer-Lemeshow goodness-of-fit tests revealed

no evidence of model misspecification for any of the models in the overall sample (all *P* values >0.20).

Table 2—Summary of select existing type 2 diabetes prediction models and characteristics of original cohort

Study/version	N	Average follow-up (years)	Baseline for model development	Population sample	Variables	Outcome ascertainment
ARIC: clinical model plus fasting glucose and lipids	7,915	9	1987–1989	85% white, 15% AA	Age, race, waist circumference, height, parental history of diabetes, HDL cholesterol, triglycerides, fasting glucose, systolic blood pressure	Elevated 2-h glucose, elevated fasting glucose, diabetes medications, or report of a clinical diagnosis during follow-up
Framingham: multivariate prediction with continuous variables	3,140	7	Mid-1990s	99% white, 0% AA	Age, sex, BMI, waist circumference, parent history of diabetes, HDL cholesterol, triglycerides, fasting glucose, systolic blood pressure	Diabetes medications or elevated fasting glucose
San Antonio Heart Study: clinical model with no 2-h glucose	3,004	7.5	1979–1982, 1984–1988	61% Mexican American, 39% white, 0% AA	Age, sex, Mexican American ethnicity, BMI, family history of diabetes, HDL cholesterol, fasting glucose, systolic blood pressure	Diabetes medications, elevated fasting glucose, elevated 2-h glucose, or self-report of physician diagnosis

Prediction models used to calculate probability (diabetes) = $\exp(x) / [1 + \exp(x)]$. ARIC study logistic regression model: $x = -9.9808 + 0.0173 * \text{age in years} + 0.4433 * \text{AA race} + 0.4981 * 1$; if parent history of diabetes is present, $+ 0.0880 * \text{fasting glucose in mg/dL} + 0.0111 * \text{systolic blood pressure in mmHg} + 0.0273 * \text{waist circumference in cm} - 0.0326 * \text{height in cm} - 0.0122 * \text{HDL cholesterol in mg/dL} + 0.00271 * \text{triglycerides in mg/dL}$. Framingham Study logistic regression model: $x = -18.607 - 0.0101 * \text{age in years} - 0.4308 * \text{sex (1 if male, 0 if female)} + 0.4383 * 1$ if parent history of diabetes; if present, $+ 0.03922 * \text{BMI} + 0.001 * \text{systolic blood pressure in mmHg} - 0.0488 * \text{HDL in mg/dL} + 0.0488 * \text{waist circumference in cm} + 0.1398 * \text{fasting glucose in mg/dL}$. San Antonio Heart Study logistic regression model: $x = -13.415 + 0.028 * \text{age in years} + 0.661 * \text{sex (1 if female, 0 if male)} + 0.412 * 1$ if Mexican American (all 0 for this study) $+ 0.079 * \text{fasting glucose in mg/dL} + 0.018 * \text{systolic blood pressure in mmHg} - 0.039 * \text{HDL in mg/dL} + 0.070 * \text{BMI} + 0.481 * 1$ if family history of diabetes is present.

Table 3—Discrimination of 5-year incident diabetes using the ARIC risk prediction model: the CARDIA study (2005–2011)

	Overall AUC (95% CI)	AA AUC (95% CI)	Whites AUC (95% CI)	P*
Model 1: previously published regression coefficients†				
Model 1a: ADA 2004 diagnostic guidelines	0.846 (0.794, 0.898)	0.802 (0.721, 0.884)	0.887 (0.827, 0.947)	0.10
Model 1b: ADA 2010 diagnostic guidelines	0.822 (0.782, 0.862)	0.778 (0.716, 0.840)	0.860 (0.814, 0.906)	0.04
P‡	0.48	0.64	0.48	
Model 2: regression equations re-estimated in CARDIA using ADA 2010 diagnostic guidelines§				
Model 2a: original predictors	0.841 (0.806, 0.876)	0.796 (0.737, 0.854)	0.875 (0.830, 0.920)	0.04
Model 2b: original predictors + A1C	0.863 (0.832, 0.894)	0.816 (0.763, 0.869)	0.902 (0.867, 0.936)	0.008
P	0.03	0.14	0.08	

*P for unpaired receiver operating characteristic comparison of models estimated in AA vs. whites. †Models 1a and 1b used the following published ARIC prediction model to calculate diabetes risk: probability (diabetes) = $[\exp(x)/(1+\exp(x))]/9 * 5$, where $x = -9.9808 + 0.0173 * \text{age in years} + 0.4433 * \text{AA race} + 0.4981 * 1 \text{ if parent history of diabetes is present} + 0.0880 * \text{FPG in mg/dL} + 0.0111 * \text{SBP in mmHg} + 0.0273 * \text{waist circumference in cm} - 0.0326 * \text{height in cm} - 0.0122 * \text{HDL cholesterol in mg/dL} + 0.00271 * \text{triglycerides in mg/dL}$. ‡P value for unpaired receiver operating characteristic comparison of models estimated within overall cohort, AA, and whites comparing change in model discrimination when diagnostic guidelines are updated from ADA 2004 to ADA 2010 diagnostic guidelines. §Model 2a included the following predictors from the published ARIC prediction model: age, parent history of diabetes, fasting glucose, systolic blood pressure, waist circumference, height, HDL cholesterol, and triglycerides. Model 2b included the same predictors as model 2a with the addition of A1C. ||P value for paired receiver operating characteristic comparison of models estimated within overall cohort, AA, and whites comparing change in model discrimination when diagnostic guidelines are updated to include baseline A1C as a predictor.

AA 0.796 vs. AUC in whites 0.875, $P = 0.04$; and model 2b, AUC in AA 0.816 vs. AUC in whites 0.902, $P = 0.008$. Hosmer-Lemeshow goodness-of-fit tests revealed no evidence of model misspecification for any of the models in race-specific analyses (all P values >0.20).

In the overall sample, the difference in mean predicted probabilities of type 2 diabetes between participants who developed diabetes and those who did not was 14.2% (model 2a); this increased to 17.3% with the addition of predictor A1C (model 2b), resulting in a statistically significant absolute IDI of 0.031 (95% CI 0.014, 0.047; $P < 0.0001$) (Fig. 1) and a relative IDI of 21.8% (95% CI 9.4, 34.3; $P < 0.0001$). In AA, the absolute IDI was 0.029 (95% CI 0.006, 0.049; $P < 0.0001$)

and the relative IDI was 21.7% (95% CI 4.3, 38.2; $P < 0.0001$). In whites, the absolute IDI was 0.031 (95% CI 0.005, 0.055; $P < 0.0001$) and the relative IDI was 22.8% (95% CI 4.3, 43.4; $P < 0.0001$).

Findings from sensitivity analyses evaluating discrimination of the Framingham and San Antonio Heart Study models among CARDIA participants were similar to findings with the ARIC model (Supplementary Table 1). For each of these models, model discrimination was consistently higher in whites than in AA. Results from a sensitivity analysis that examined an alternate outcome definition (which included self-reported physician diagnosis of diabetes) were similar to results from our main analysis (Supplementary Table 2).

CONCLUSIONS

In the current study, a previously published type 2 diabetes risk prediction model maintained high levels of discriminative validity when applied to a modern, biracial cohort of U.S. adults. We found that model discrimination was high when using the ADA 2004 diagnostic guidelines among CARDIA participants. Defining incident diabetes using the ADA 2010 diagnostic guidelines, which include A1C, resulted in a decrease in model discrimination, which was reversed with the addition of baseline A1C as a predictor of type 2 diabetes risk. In the overall sample, there was no evidence of lack of model calibration. The addition of predictor A1C yielded a statistically significant increase in IDI and a relative IDI of 21.8%, a contribution that is well above the average contribution of the other nine predictors in the ARIC model (11.1%). These findings suggest that including A1C in the prediction model results in significant improvement in model performance.

When racial subgroups in CARDIA were analyzed separately, model performance was better among whites than AA. For all models that used ADA 2010 diagnostic guidelines, model discrimination was significantly higher in whites than AA. The addition of baseline A1C improved model discrimination in whites and AA to a similar degree, and IDI analyses suggested that A1C significantly improved model performance in both whites and AA. Results from the current

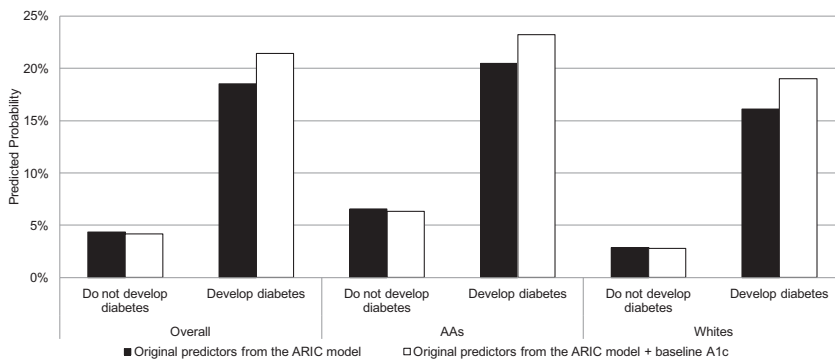


Figure 1—Change in 5-year predicted probability of developing diabetes when updating from the ARIC model predictors (age, parent history of diabetes, fasting glucose, systolic blood pressure, waist circumference, height, HDL cholesterol, triglycerides) to the ARIC model predictors plus predictor A1C in the overall cohort and in racial subgroups. All models defined diabetes using ADA 2010 diagnostic guidelines.

analysis confirm our hypothesis that when type 2 diabetes prediction models are updated to include A1C, a reflection of current clinical practice for the diagnosis of type 2 diabetes, there exists a racial divide in model performance. This difference in model performance may reflect racial differences in diagnostic practices among other factors. Therefore, we also examined other subgroups defined by sex, age, and education level to determine whether racial differences in model performance were being influenced by other key sociodemographic factors. These subgroup analyses indicated that the effect of A1C did not differ in any of these subgroups (*P* values for interaction term >0.20 for all subgroups).

Potential limitations of this study included the availability of A1C measurements only at the year 20 and 25 follow-up. This short follow-up time (5 years) may have affected our study's power to detect significant differences in model performance, particularly among the racial subgroups. Also, our study defined diabetes based on modified diagnostic guidelines using reported use of antidiabetes medication, a single measurement of fasting glucose, or 2-h glucose or A1C; we did not have confirmatory testing available for fasting glucose. The diabetes definition used in our study differed slightly from the definition used in the ARIC study: both studies included fasting glucose and use of antidiabetes medications, but our definition included 2-h glucose where ARIC did not, and ARIC included self-report of diabetes diagnosis where our study did not. Because incident diabetes was assessed at the follow-up exam 5 years after baseline, our analyses assume a constant risk of diabetes over the study period, and we were unable to assess whether the risk of diabetes varied at specific time points. Additionally, while it is not surprising that the inclusion of baseline A1C improved model performance, our findings highlight the extent to which addition of baseline A1C improves diabetes prediction in this biracial population and provide evidence of differential prediction by race in models with and without A1C. Lastly, the existing diabetes risk prediction models did not include lifestyle factors, such as diet and physical activity, so these variables were not evaluated in our study. Strengths of our

study included the use of a large biracial cohort from four cities across the U.S. with detailed clinical and metabolic data available, including all of the recommended tests for assessing diabetes status (i.e., fasting glucose, 2-h glucose, and A1C).

Overall, an existing type 2 diabetes risk prediction model derived from the ARIC study maintained relatively strong discriminatory power in a biracial cohort comprised of AA and whites from four geographically diverse cities across the U.S. Existing type 2 diabetes risk prediction models should be updated to incorporate ADA 2010 diagnostic criteria, as these criteria reflect current clinical guidelines. Findings from our analysis support the inclusion of A1C as a predictor of incident type 2 diabetes for AA and whites, yet suggest that, specifically in AA, there is a need for closer examination of the optimal prediction of diabetes risk in AA. Further investigation in additional cohorts that include racial and ethnic minorities is warranted.

Acknowledgments. The authors gratefully acknowledge the staff and participants of CARDIA for their contributions.

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

Funding. CARDIA is supported by the National Heart, Lung, and Blood Institute (NHLBI) (contracts HHSN268201300025C, HHSN268201300026C, HHSN268201300027C, HHSN268201300028C, HHSN268201300029C, and HHSN268200900041C), the National Institute on Aging (NIA) (Intramural Research Program), and an intra-agency agreement between NIA and NHLBI (AG0005). M.E.L. was supported by National Institute of Diabetes and Digestive and Kidney Diseases grant F31-DK-105791. M.E.L. and W.-C.W. were supported by the Center of Innovation in Long-Term Services and Supports, Providence Veterans Affairs Medical Center. A.P.C. was supported by National Institute of Diabetes and Digestive and Kidney Diseases grant K01-DK-095928.

Author Contributions. M.E.L. researched data and wrote the manuscript. G.A.W. and W.-C.W. contributed to the study design, analysis plan, interpretation of results, and critical revisions of the manuscript. M.R.C. contributed to study design and interpretation of results. E.B.L. contributed to study design and reviewed and edited the manuscript. A.P.C. contributed to interpretation of results and reviewed and edited the manuscript. X.L. contributed to the analytic plan and provided statistical expertise. C.I.K. reviewed and edited the manuscript and contributed to the discussion. A.G. contributed to the interpretation of results and reviewed and edited the manuscript. E.P.G. contributed to the discus-

sion. C.B.E. contributed to the analytic plan and the discussion. M.E.L. 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.

References

1. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2004;27(Suppl. 1):S5–S10
2. American Diabetes Association. Standards of medical care in diabetes—2010. *Diabetes Care* 2010;33(Suppl. 1):S11–S61
3. Buijsse B, Simmons RK, Griffin SJ, Schulze MB. Risk assessment tools for identifying individuals at risk of developing type 2 diabetes. *Epidemiol Rev* 2011;33:46–62
4. Collins GS, Mallett S, Omar O, Yu LM. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. *BMC Med* 2011;9:103
5. Noble D, Mathur R, Dent T, Meads C, Greenhalgh T. Risk models and scores for type 2 diabetes: systematic review. *BMJ* 2011;343:d7163
6. National Center for Chronic Disease Prevention and Health Promotion, Division of Diabetes Translation. National Diabetes Statistics Report, 2014 [Internet], 2014. Atlanta, Centers for Disease Control and Prevention. Available from <http://www.cdc.gov/diabetes/pubs/statsreport14/national-diabetes-report-web.pdf>. Accessed 15 October 2014
7. 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
8. Wilson PWF, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RB Sr. Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. *Arch Intern Med* 2007;167:1068–1074
9. Stern MP, Williams K, Haffner SM. Identification of persons at high risk for type 2 diabetes mellitus: do we need the oral glucose tolerance test? *Ann Intern Med* 2002;136:575–581
10. Mann DM, Bertoni AG, Shimbo D, et al. Comparative validity of 3 diabetes mellitus risk prediction scoring models in a multiethnic US cohort: the Multi-Ethnic Study of Atherosclerosis. *Am J Epidemiol* 2010;171:980–988
11. Herman WH. Do race and ethnicity impact hemoglobin A1c independent of glycemia? *J Diabetes Sci Technol* 2009;3:656–660
12. Ziemer DC, Kolm P, Weintraub WS, et al. Glucose-independent, black-white differences in hemoglobin A1c levels: a cross-sectional analysis of 2 studies. *Ann Intern Med* 2010;152:770–777
13. Kirk JK, D'Agostino RB Jr, Bell RA, et al. Disparities in HbA1c levels between African-American and non-Hispanic white adults with diabetes: a meta-analysis. *Diabetes Care* 2006;29:2130–2136
14. Selvin E, Rawlings AM, Bergenstal RM, Coresh J, Brancati FL. No racial differences in the association of glycated hemoglobin with kidney disease and cardiovascular outcomes. *Diabetes Care* 2013;36:2995–3001

15. Tsugawa Y, Mukamal KJ, Davis RB, Taylor WC, Wee CC. Should the hemoglobin A1c diagnostic cutoff differ between blacks and whites? A cross-sectional study. *Ann Intern Med* 2012; 157:153–159
16. Friedman GD, Cutter GR, Donahue RP, et al. CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol* 1988;41:1105–1116
17. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29–36
18. Gonen M. *Analyzing Receiver Operating Characteristic Curves with SAS*. Cary, NC, SAS Institute, 2007
19. Hosmer DW, Lemeshow SA. A goodness-of-fit test for the multiple logistic regression model. *Commun Stat* 1980;9:1043–1069
20. Pencina MJ, D’Agostino RB Sr, D’Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;27:157–172
21. Pencina MJ, D’Agostino RB Sr, Pencina KM, Janssens AC, Greenland P. Interpreting incremental value of markers added to risk prediction models. *Am J Epidemiol* 2012;176:473–481
22. Kennedy KF, Pencina MJ. A SAS macro to compute added predictive ability of new markers predicting a dichotomous outcome. In *Southeast SAS Users Group 2010 Proceedings, 2010*. Savannah, GA
23. Cassell DL. Don’t be loopy: re-sampling and simulation the SAS way. *SAS Global Forum 2007, Statistics and Data Analysis, 2007*. Cary, NC