

Identifying risk for type 2 diabetes in different age cohorts: does one size fit all?

Maria L Alva,¹ Thomas J Hoerger,² Ping Zhang,³ Edward W Gregg³

To cite: Alva ML, Hoerger TJ, Zhang P, *et al*. Identifying risk for type 2 diabetes in different age cohorts: does one size fit all?. *BMJ Open Diab Res Care* 2017;**5**:e000447. doi:10.1136/bmjdr-2017-000447

► Additional material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjdr-2017-000447>).

Received 22 June 2017
Revised 18 August 2017
Accepted 3 September 2017

ABSTRACT

Objective To estimate age-specific risk equations for type 2 diabetes onset in young, middle-aged, and older US adults, and to compare the performance of simple equations based on readily available demographic information alone, against enhanced equations that require both demographic and clinical information (fasting plasma glucose, high-density lipoprotein, and triglyceride levels).

Research design and methods We estimated the probability of developing diabetes by age group using data from the Coronary Artery Risk Development in Young Adults (for ages 18–40 years), Atherosclerosis Risk in Communities (for ages 45–64 years), and the Cardiovascular Health Study (for ages 65 years and older). Simple and enhanced equations were estimated using logistic regression models, and performance was compared by age group. Thresholds based on these risk equations were evaluated using split-sample bootstraps and calibrating the constant of one age cohort to others.

Results Simple risk equations had an area under the receiver-operating curve (AUROC) of 0.72, 0.79, 0.75, and 0.69 for age groups 18–30, 28–40, 45–64, and 65 and older, respectively. The corresponding AUROCs for enhanced equations were 0.75, 0.85, 0.85, and 0.81. Risk equations based on younger populations, when applied to older cohorts, underpredict diabetes incidence and risk. Conversely, risk equations based on older populations overpredict the likelihood of diabetes in younger cohorts.

Conclusions In general, risk equations are more successful in middle-aged adults than in young and old populations. The results demonstrate the importance of applying age-specific risk equations to identify target populations for intervention. While the predictive capacity of equations that include biomarkers is better than of those based solely on self-reported variables, biomarkers are more important in older populations than in younger ones.

INTRODUCTION

Several risk equations have been developed to identify those at high risk of developing type 2 diabetes^{1,2} using data from the Framingham Heart Study, the National Health and Nutrition Examination Survey (NHANES), Coronary Artery Risk Development in Young Adults (CARDIA), Atherosclerosis Risk in Communities (ARIC), the Cardiovascular Heart Study (CHS), and the San Antonio Heart Study for various follow-up periods (from 5 to 24 years),

Significance of this study

What is already known about this subject?

- Most diabetes type 2 risk equations in the literature focus on predictions for people older than 45 years of age.
- The literature finds good levels of discrimination but high levels of discrimination often arise when studies are not able to exclude at baseline individuals with undiagnosed diabetes from the estimating sample.

What are the new findings?

- One size does not fit all when we want to identify what are the most important risk factors for type 2 diabetes across different age groups. Relative and absolute risks vary by age.
- The predictive capacity of equations based on biomarkers is, on average, better than those based on self-reported variables but information from biomarkers are more important in older populations than in younger ones. We find no significant difference in the area under the receiver-operating curve between simple and enhanced equations in young adults.

How might these results change the focus of research or clinical practice?

- A screening strategy based on self-reported variables in younger populations would be as effective as one that requires collecting clinical samples. For older populations, there is a tradeoff between a simple model that can be applied to more people, and an enhanced model that would be more accurate, but would require costly laboratory tests.

each using different estimation methods. The variables used to generate predicted probabilities are common across many of these studies and include self-reported demographic information, such as age, sex, race, medication use, family history of diabetes, body mass index (BMI) and/or waist circumference, smoking status, alcohol consumption, and food consumption. Enhanced risk equations also use clinical measures, such as systolic blood



CrossMark

¹D Phil Public Health Economics Program, RTI International, Washington, DC, USA

²RTI International, Research Triangle Park, Durham, North Carolina, USA

³Centers for Disease Control and Prevention, Atlanta, Georgia, USA

Correspondence to

Dr Maria L Alva; malva@rti.org

pressure, fasting plasma glucose (FPG), triglyceride, and high-density lipoprotein (HDL) cholesterol levels.

The development of simple yet accurate risk scores is important for risk stratification and prevention by clinical and public health interventions. Similarly, quantifying the absolute and relative risks for diabetes associated with combinations of key risk factors is essential for cost-effectiveness modeling efforts. However, cohort studies in the USA have generally been limited to specific segments of the population age range. The most important sets of risk factors, as well as the relative and absolute risks, may vary considerably by age.

In this analysis, we assembled data from three major US epidemiological studies to develop diabetes risk equations and to estimate separate, age-specific risk equations. Our objectives were to (1) examine whether core risk factors and risk equations vary with age and (2) quantify the performance of simple risk equations, based on self-reported variables (age, sex, race, BMI, smoking status, family history, and binary indicators for high blood pressure and high cholesterol), and enhanced risk equations, which include added clinical variables (blood pressure, cholesterol, FPG, HDL, and triglycerides).

RESEARCH DESIGN AND METHODS

Data

Study data were obtained from three epidemiological studies: CARDIA, ARIC, and CHS. The CARDIA study, initiated in 1985 to investigate lifestyle and other factors that influence the evolution of coronary heart disease (CHD) risk factors during young adulthood, recruited 5116 black and white women and men, aged 18–30 years, in four urban areas: Birmingham, Alabama; Chicago, Illinois; Minneapolis, Minnesota; and Oakland, California.³ Participants were followed for 20 years.

ARIC, initiated in 1987, was conducted in four communities (Washington County, Maryland; Forsyth County, North Carolina; Jackson, Mississippi; and Minneapolis, Minnesota) by randomly selecting a cohort of 15 792 individuals aged 45–64 years. Participants were followed for 9 years. ARIC was designed to investigate the causes of atherosclerosis and its clinical outcomes, and the variation in cardiovascular risk factors, medical care, and disease by race and gender.

The CHS, initiated in 1989, enrolled 5888 men and women aged 65 years and older in four communities: Forsyth County, North Carolina; Sacramento County, California; Washington County, Maryland; and Pittsburgh, Pennsylvania. Eligible participants were sampled from Medicare eligibility lists in each area, and were followed for 7 years. The main objective of the CHS study was to identify factors related to the onset and course of CHD and stroke.

Analysis samples and diabetes variable

Our dependent variable was the first incident/diagnosis of type two diabetes. Therefore, we excluded participants

with a previous diagnosis of type two diabetes at the time of enrollment. Diabetes status was assessed using slightly different methods and at different follow-up intervals in the three studies. In CARDIA, diabetes status was defined by FPG measurements and/or self-report of taking oral diabetes medications with or without insulin injections. In ARIC, diabetes status, based on self-reporting of previous diagnosis and glucose values, was available at triennial follow-up encounters (year 3: 1990–1992; year 6: 1993–1995; year 9: 1996–1998). In years 3 and 6, self-reported diabetes medication use was collected; in year 9, a 2-hour oral glucose tolerance test (OGTT) was also administered. In CHS, diabetes was defined annually by the new use of insulin or oral hypoglycemic medication and/or by FPG values. To create a uniform definition of diabetes across all datasets, we defined diabetes by reported physician diagnosis and/or $\text{FPG} \geq 7.0$ mmol/L (126 mg/dL).

None of the three datasets included adults aged 31–44 years; we therefore created a fourth data set by splitting CARDIA into two samples. Individuals recruited for CARDIA at age 18–30 years were aged 28–40 years at the 10-year follow-up. Thus, our new data set, which we will refer to as CARDIA-10, included CARDIA participants who had not developed diabetes by year 10 as a new sample baseline.

Statistical analysis

To address the question of whether a particular set of variables would predict equally effectively across different age groups, we estimated simple and enhanced risk equations for four age groups (18–30, 28–40, 45–64, and 65 and older) using the same sets of variables (self-reported variables for the simple equation and self-reported plus clinical data for the enhanced equation), and the same statistical method (logistic regression) to isolate the predictive power of individual coefficients for the likelihood of diabetes for each age group. The outcome variable in our models was the cumulative incidence of diabetes throughout the observational period for each sample (10 years in CARDIA and CARDIA-10, 9 years in ARIC, and 7 years in CHS).

To test and compare the performance of the simple and enhanced models, we randomly selected 70% of each sample to develop the models, and used the remaining 30% of individuals for validation. This bootstrap exercise was repeated 1000 times for each of the four datasets to strengthen the validity and generalizability of our findings. We evaluated the diagnostic properties of the simple and enhanced models across the four datasets on the remaining 30% of the sample. Predictive capacity for each continuous factor was assessed using the area under the receiver-operating curve (AUROC). A model with no predictive power has an area equal to 0.5, whereas a perfect model has an area equal to 1. We also assessed sensitivity, specificity, and positive and negative predictive values (PPV and NPV). Because the risk equations were estimated over different time periods for the different

datasets, we computed 1-year probabilities to make the results more comparable.

To explore the consequences of using coefficients from one model to estimate the probability of developing diabetes for individuals from another dataset and thus a different age group (eg, using the risk equation estimated with ARIC data for those aged 45–64 years to estimate risk in the CHS population aged 65 years and older), we applied coefficients from one risk equation to data for different age groups. We did so using a constrained logistic regression where the intercept was allowed to vary (accounting for differences in the absolute risk across age groups) but the coefficients for the other variables were constrained to be equal to the coefficients from the original risk equation (maintaining the same OR as the original equation). Thus, we applied the ARIC equation to the CARDIA, CARDIA-10, CHS data, and so on. By allowing the intercept to be re-estimated, we controlled for differences in the absolute probability of diabetes across age groups, but the constrained coefficients on the variables maintained the OR for the original risk equation.

RESULTS

The initial CARDIA sample consisted of 5116 individuals aged 18–30 years. We excluded 78 with diabetes at baseline and 999 with incomplete data, leaving an analytic sample of 4039. At year 20, 3413 remained in the CARDIA sample, which represents our potential sample for the CARDIA-10 cohort. At years 10 and 20, we had information on self-reported diabetes status and FPG. We excluded those who had diabetes at year 10 (n=266) and used the covariates measured at year 10 as the baseline

year. We excluded participants with incomplete information on the variables of interest at year 10 (n=274). The final CARDIA-10 sample used for estimation was 2873 people aged 28–40 years at baseline (ie, year 10 of CARDIA).

In ARIC, 15 792 individuals, aged 45–64 years were recruited at baseline. We used baseline explanatory variables as predictors of the cumulative incidence rate of diabetes at year 9. We excluded 1163 individuals with diabetes at baseline, 2080 with missing data for the explanatory variables at baseline, and 3674 with missing data for the dependent variable at year 9 follow-up, leaving a final sample of 8875 individuals.

Our starting CHS sample was 5888 individuals aged 65 years or older. We used 7-year follow-up for the purpose of this analysis because that is the latest data in which laboratory values were included in the public use dataset. By year 7, the sample consisted of 4100 participants with laboratory information on FPG. We excluded 501 persons with diabetes at baseline. Our final dataset, excluding individuals without a complete set of covariates, was 3094.

Table 1 shows the baseline characteristics of study participants. Data are presented as means and SD, unless otherwise noted.

Tables 2 and 3 show the results of the simple and enhanced models. BMI was the only variable with significant predictive power in both simple and enhanced equations across all age groups. Parental history had predictive power for younger cohorts in the simple and enhanced models, but less predictive importance for the oldest cohort. In the enhanced models, FPG and triglycerides had the best predictive power (see online supplementary appendix A-1).

Table 1 Characteristics among participants of CARDIA, ARIC, and CHS at baseline included in the regression analyses

Characteristic	CARDIA (n=4039, T=10)		CARDIA-10 (n=2813, T=10)		ARIC (n=8875, T=9)		CHS (n=3094, T=7)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age (years)	24.89	3.60	35.03	3.59	53.81	5.67	72.41	4.95
Black (=1)	48.58%	49.99%	44.80%	49.74%	17.12%	37.67%	3.88%	19.31%
Male (=1)	44.54%	49.71%	45.14%	49.77%	44.37%	49.69%	38.40%	48.64%
BMI	24.35	4.72	27.07	5.87	27.09	4.91	26.13	3.79
Smoker (=1)	13.02%	33.66%	13.76%	34.45%	21.90%	41.36%	51.39%	49.99%
High cholesterol (=1)	2.23%	14.76%	64.62%	47.82%	25.06%	43.34%	28.41%	45.11%
Parental history (=1)	13.52%	34.20%	20.84%	40.62%	24.60%	43.07%	35.36%	47.82%
SBP (mm Hg)	110.08	11.01	109.21	12.16	118.60	16.91	133.99	20.90
FPG (mg/dL)	81.66	8.26	86.52	11.18	98.44	9.13	99.49	9.37
HDL (mg/dL)	53.16	13.23	50.40	13.89	52.74	17.12	55.55	15.72
Triglycerides	72.36	47.16	89.36	68.28	124.39	75.24	135.73	64.48

Age groups: CARDIA: 18–30 years; CARDIA-10: 28–40 years; ARIC: 45–64 years; CHD: 65 years and older. (=1) indicates a binary variable. T indicates maximum follow-up time in the sample in terms of years. High cholesterol=1 if 240 mg/dL and above. BMI=(weight in kg)/(height in meters)².

ARIC, Atherosclerosis Risk in Communities; BMI, body mass index; CARDIA, Coronary Artery Risk Development in Young Adults; CHS, Cardiovascular Heart Study; FPG, fasting plasma glucose; HDL, high-density lipoprotein cholesterol; SBP, systolic blood pressure.

Table 2 ORs and diagnostic accuracy for type 2 diabetes over T years—simple model

	CARDIA (10years)	CARDIA-10 (10years)	ARIC (9years)	CHS (7years)
ORs				
Age group†	1.343*	1.242	1.078	0.848
Black (=1)	0.947	1.408*	1.323***	1.265
Male (=1)	0.384***	1.380*	1.575***	1.512**
BMI	1.087***	1.143***	1.138***	1.144***
Parental history (=1)	1.661***	2.357***	1.871***	1.324
Smoker (=1)	0.878	0.987	1.357***	1.198
High SBP (>140 mm Hg=1)	3.846**	1.094	1.472***	1.887***
High cholesterol (>240 mg/dL=1)	1.539	1.388*	1.002	0.948
Observations	4039	2813	8875	3094
Diabetes (N)	171	188	836	150
Diagnostic statistics				
AUROC (95% CI)	0.72 (0.69–0.76)	0.79 (0.76–0.83)	0.75 (0.73–0.77)	0.69 (0.65–0.73)
PPV	50.00%	31.25%	34.92%	0.00%
NPV	95.79%	93.46%	90.76%	95.15%
Cumulative probability	4.23%	6.68%	9.42%	4.85%
1-year probability‡	0.44%	0.73%	1.04%	0.53%

*** p<0.01, ** p<0.05, * p<0.1

†For CARDIA age (25–30)=1, CARDIA-10 age (35–40)=1, ARIC age (55–64)=1, CHS age (75+)=1.

‡Supplementary appendix A-6 shows how to compute annual probabilities using the Logit model.

T represents the maximum time in the sample. For CARDIA and CARDIA-10, T=10 years. For ARIC, T=9; and for CHD, T=7.

For PPV and NPV the cut-off used is 5.

Supplementary appendix A-5a shows non-exponentiated coefficients.

BMI=(weight in kg)/(height in meters)².

PPV, $p \times \text{sensitivity} / (p \times \text{sensitivity} + (1-p)(1-\text{specificity}))$, where p is the prevalence.

NPV, $\text{specificity} \times (1-p) / (p \times (1-\text{sensitivity}) + (1-p) \times \text{specificity})$.

ARIC, Atherosclerosis Risk in Communities; AUROC, area under the receiver-operating curve; CARDIA, Coronary Artery Risk Development in Young Adults; CHS, Cardiovascular Heart Study; BMI, body mass index; NPV, negative predictive value; PPV, positive predictive value; SBP, systolic blood pressure.

As expected, we found that AUROCs were higher for the enhanced models than for the simple models (tables 2 and 3). The simple and enhanced models show different levels of predictive power across age groups. AUROCs for simple risk equations were 0.72 for CARDIA; 0.79 for CARDIA-10; 0.75 for ARIC; and 0.69 for CHS. However, AUROCs for enhanced equations were 0.75 for CARDIA; 0.85 for CARDIA-10; 0.85 for ARIC; and 0.81 for CHS. Statistically significant differences between simple and enhanced models were present only for older age cohorts (ARIC and CHS).

Online supplementary appendix A-2 shows the performance of the simple and enhanced equations in identifying people that developed diabetes by quintiles of predicted risk thresholds using the split sample approach as a test of internal validity. At the top quintile, few individuals would be above the predicted value cut-off, and thus the sensitivity is low and specificity is high. PPV and NPV depend on the prevalence of diabetes in each sample. In younger cohorts, where the prevalence of diabetes is relatively low, PPV is low and NPV is high. In older cohorts, as the prevalence increase so does the PPV.

Table 4 shows the extent to which the coefficients in each of the models change predictions in the other datasets, by testing what happens when the equation estimated with one dataset was used to estimate the probability of developing diabetes for individuals of an age group from another dataset. We report constrained regressions, forcing all coefficients excluding the constant term to be the same as in the source equations, and thus allowing for the calibration of the constant term. We report the AUROC and forecasted 1-year probabilities. All constant terms (analogous to the baseline hazard rate) in the target and source datasets are statistically different. The further away, in terms of age, the underlying cohort is from the coefficients of the risk model used for constrained regression in the test dataset, the more imprecise the results are compared with the results originating from the same underlying coefficients and data. Irrespective of whether the calibration is done using the simple or the enhanced model, equations based on younger age groups when applied to older cohorts underpredict diabetes incidence. Conversely, risk equations based on older

Table 3 ORs and diagnostic accuracy for type 2 diabetes over T years—enhanced model

	CARDIA (10 years)	CARDIA-10 (10 years)	ARIC (9 years)	CHS (7 years)
ORs				
Age group†	1.320	0.981	0.898	0.911
Black (=1)	1.174***	1.554**	1.898***	1.772
Male (=1)	0.292***	0.650**	0.795**	1.104
Parental history (=1)	1.736***	2.219***	1.670***	1.304
SBP (mm Hg)	0.995	1.020***	1.006**	1.005**
Smoker (=1)	0.918	0.943	1.230**	1.169
BMI	1.059***	1.079***	1.091***	1.055**
FPG (mg/dL)	1.034***	1.080***	1.116***	1.114***
HDL (mg/dL)	0.990	0.980**	0.975***	0.986
Triglycerides (mg/dL)	1.005	1.002*	1.002***	1.004***
Observations	4039	2813	8875	3094
Diagnostic statistics				
Diabetes (N)	171	188	836	150
AUROC (C.I.)	0.75 (0.71–0.78)	0.85 (0.82–0.88)	0.85 (0.84–0.86)	0.81 (0.77–0.85)
PPV	62.04%	73.47%	56.02%	50.00%
NPV	57.26%	94.50%	91.88%	95.24%
Cumulative probability	4.23%	6.68%	9.42%	4.85%
1-year probability‡	0.44%	0.88%	1.13%	0.53%

*** p<0.01, ** p<0.05, * p<0.1

†For CARDIA age (25–30)=1, CARDIA-10 age (35–40)=1, ARIC age (55–64)=1, CHS age (75+)=1.

‡Supplementary appendix A-6 shows how to compute annual probabilities using the Logit model.

T represents the maximum time in the sample. For CARDIA and CARDIA-10, T=10 years. For ARIC, T=9; and for CHD, T=7.

For PPV and NPV the cut-off used is 5.

Supplementary appendix A-5b shows non-exponentiated coefficients.

BMI=(weight in kg)/(height in meters)².

PPV, $p \times \text{sensitivity} / (p \times \text{sensitivity} + (1-p)(1-\text{specificity}))$, where p is the prevalence.

NPV, $\text{specificity} \times (1-p) / (p \times (1-\text{sensitivity}) + (1-p) \times \text{specificity})$.

ARIC, Atherosclerosis Risk in Communities; AUROC, area under the receiver-operating curve; CARDIA, Coronary Artery Risk Development in Young Adults; CHS, Cardiovascular Heart Study; BMI, body mass index; FPG, fasting plasma glucose; HDL, high-density lipoprotein cholesterol; NPV, negative predictive value; PPV, positive predictive value; SBP, systolic blood pressure.

age cohorts overpredict the likelihood of diabetes in younger cohorts.

For each dataset, we show the corresponding predicted probability of developing diabetes, and compare these with the probability of developing diabetes using the equation in which they were developed (ie, using the target data). The graphs in online supplementary appendix A-3 illustrate how one can overpredict or underpredict using the wrong equation, by plotting the information presented in table 4 across the entire distribution. Calibration of the constant term across cohorts helps resolve some of the unaccounted discrepancies across cohorts. Even with calibration however, older cohorts typically overpredict, while younger cohorts underpredict, irrespective of the set of covariates used.

DISCUSSION

Our goal was to generate simple and enhanced age group-specific risk equations to predict the probability of developing type two diabetes, and to determine the extent

to which patient characteristics matter differently across age groups. Often, risk factors are selected from many potential covariates based on the strength of association with the outcomes in a study sample. This study shows which variables matter in predicting the risk of diabetes and how their importance varies depending on age. Based on the rules by Hosmer and Lemeshow⁴ for interpreting AUROC values, we find that simple equations have an acceptable level of discrimination ($0.7 \leq \text{AUROC} < 0.8$), while enhanced equations have very good discrimination ($0.8 \leq \text{AUROC}$) except for the youngest (18–30) age group, which is in the acceptable range. Overall, we find that risk equations have better predictability in middle-aged adults than in young and older populations. Thus, it is not surprising that most risk equations published in the literature focus on predictions for people older than 45 years.

Our study shows that predictions vary markedly and significantly when coefficients derived from one age group are used to predict non-adjacent age groups. This

Table 4 Annual predicted values from constrained regressions across simple and enhanced models

Test dataset	Source of constrained coefficients	AUROC (calibrated constant)	1-year probability (calibrated constant) (%)
CARDIA simple	CARDIA	0.72 (0.69–0.76)	0.44
	CARDIA-10	0.63 (0.58–0.68)	0.45
	ARIC	0.61 (0.56–0.66)	0.49
	CHS	0.60 (0.55–0.65)	0.63
CARDIA-10 simple	CARDIA-10	0.79 (0.76–0.83)	0.73
	CARDIA	0.74 (0.70–0.77)	0.72
	ARIC	0.78 (0.74–0.82)	0.80
	CHS	0.78 (0.74–0.81)	1.03
ARIC simple	ARIC	0.75 (0.73–0.77)	1.04
	CARDIA	0.73 (0.72–0.75)	1.01
	CARDIA-10	0.75 (0.73–0.76)	1.05
	CHS	0.75 (0.73–0.76)	1.47
CHS simple	CHS	0.69 (0.65–0.73)	0.72
	CARDIA	0.59 (0.54–0.63)	0.50
	CARDIA-10	0.69 (0.64–0.73)	0.50
	ARIC	0.69 (0.65–0.73)	0.56
CARDIA enhanced	CARDIA	0.75 (0.71–0.78)	0.44
	CARDIA-10	0.66 (0.61–0.70)	0.46
	ARIC	0.65 (0.60–0.70)	0.53
	CHS	0.62 (0.57–0.67)	0.66
CARDIA-10 enhanced	CARDIA-10	0.85 (0.82–0.88)	0.88
	CARDIA	0.83 (0.80–0.86)	0.92
	ARIC	0.84 (0.81–0.87)	0.97
	CHS	0.83 (0.80–0.86)	1.20
ARIC enhanced	ARIC	0.85 (0.84–0.86)	1.13
	CARDIA	0.84 (0.82–0.85)	1.10
	CARDIA-10	0.85 (0.83–0.86)	1.14
	CHS	0.85 (0.83–0.86)	1.56
CHS enhanced	CHS	0.81 (0.77–0.85)	0.75
	CARDIA	0.77 (0.73–0.82)	0.53
	CARDIA-10	0.81 (0.77–0.85)	0.53
	ARIC	0.81 (0.77–0.85)	0.59

suggests that the covariates have different predictive power of future risk of diabetes for different age groups. For example, while risk increases with age, age has a lower predictive power in older cohorts than in younger cohorts. Race, sex, and parental history are stronger predictors for younger age groups. Younger males are significantly less likely to develop diabetes than younger women, while this relationship does not hold true for older men and older women. This initial difference may be driven by the risk of gestational diabetes among women. BMI is the most consistent statistically significant indicator for diabetes across age groups and for both simple and enhanced equations. However, BMI matters more in the simple model than in the enhanced model.

On average, a one unit increase in BMI increases the distal probability of diabetes by 10% across studies (see online supplementary appendix figure A-1). Biomarkers have much narrower confidence intervals. SBP and HDL matter only for the middle age group. Triglycerides and FPG are statistically significant across all age groups, but they matter marginally more for older cohorts than for younger ones in correctly predicting the likelihood of diabetes.

Online supplementary appendix table A-4 summarizes results from 19 previous studies. Fourteen of these studies use logit models, and five use a proportional hazard model. All studies measuring current prevalence used non-clinical data only, while studies measuring future risk

tended to include information on biomarkers. Approximately half of the studies found an acceptable level of discrimination ($0.7 \leq \text{AUROC} < 0.8$), while the remaining reported very good discrimination ($0.8 \leq \text{AUROC}$). FIND-RISC⁵ and the Diabetes Risk Calculator⁶ had the highest combined sensitivity and specificity, with AUROCs of 0.86 and 0.85, respectively. Cabrera de León *et al*⁷ achieved an AUROC of 0.84 for men and 0.87 for women. The high values in the discrimination ability of the Diabetes Risk Calculator and Cabrera de León *et al* are likely due to sample selection; these studies included individuals with undiagnosed diabetes and prediabetes. Previous models in the literature using the same data sources as our study (CARDIA,⁸ ARIC,^{9,10} and CHS¹¹) achieved comparable sensitivity and specificity to ours, despite these previous models including additional variables (physical activity and diet) that we did not include because they were not uniformly coded across studies.

Enhanced risk equations provide better discrimination than simple risk equations, but the benefit of enhanced equations is less in younger cohorts, and there was no significant difference in the AUROC between simple and enhanced in young adults. This implies that screening strategy based on sex, family history, race, and BMI in younger populations would be nearly as effective as one that requires collecting clinical samples. For older cohorts, there is a tradeoff between a simple model that could be used by more people, and an enhanced risk equation that would be more accurate, but would require costly laboratory tests. It is important to note that the simple and enhanced models did not differ significantly in terms of cumulative predictions; therefore, at the population level, a less expensive model performs as well as the more costly model. At the individual level, however, the costlier model will significantly increase the sensitivity of the estimates.

Four limitations related to the data are important to highlight. First, all three surveys experienced loss during follow-up. Individuals exited the sample as a result of death, relocation, or loss of interest in the study. CARDIA had a follow-up rate of 80%; CARDIA-10 84%, ARIC 75%, and CHS 61%. Loss to follow-up could bias estimates, if it is correlated with the likelihood of having diabetes and individual characteristics. Second, the surveys do not define diabetes through OGTT, but through self-reported questionnaires and FPG; however, this is also a benefit as it more closely reflects common practice. Third, because the surveys used are not nationally representative, it is possible that the differences we attributed to age reflect, in part, geographical variations. Fourth, the surveys began in the 1980s and 1990s, and may not reflect current population characteristics and treatment approaches. However, they may reflect the underlying natural history of diabetes progression in the absence of formal interventions to prevent diabetes.

In summary, we found that risk equations have better predictability in middle-aged adults than in young and old populations. While the predictive capacity of

equations based on biomarkers is, on average, better than those based solely on self-reported variables, information from biomarkers are more reliable and important in older populations than in younger ones. This variability emphasizes the importance of using age-specific risk equations when assessing the need to screen for type two diabetes to improve accuracy of individual-level predictions. Using age-specific risk equations may be especially important for the development of practical risk stratification tools, as well as to provide more precise parameters for cost-effectiveness analyses.

Contributors MLA had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Study concept: MLV, TJH, EWG, PZ. Study design: MLV, TJH, EWG, PZ. Acquisition of data: TJH. Drafting of the manuscript: MLA. Critical revision of the manuscript for important intellectual content: TJH, EWG, PZ. Statistical analysis: MLA. Interpretation of data: MLA, TJH, EWG, PZ. Review and approval of the manuscript: TJH, EWG, PZ, MLA.

Funding This research was supported by Contract Number 20072008727958, Task Order 40 from the Centers for Disease Control and Prevention (CDC) and by RTI International. The opinions in this paper are solely those of the authors and do not necessarily reflect the opinions of CDC or RTI.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

1. Noble D, Mathur R, Dent T, *et al*. Risk models and scores for type 2 diabetes: systematic review. *BMJ* 2011;343:d7163.
2. Brown N, Critchley J, Bogowicz P, *et al*. Risk scores based on self-reported or available clinical data to detect undiagnosed type 2 diabetes: a systematic review. *Diabetes Res Clin Pract* 2012;98:369–85.
3. Friedman GD, Cutter GR, Donahue RP, *et al*. CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol* 1988;41:1105–16.
4. Hosmer DW, Lemeshow S. *Applied logistic regression*. 2nd ed. New York: Wiley, 2000.
5. Lindström J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. *Diabetes Care* 2003;26:725–31.
6. Heikes KE, Eddy DM, Arondekar B, *et al*. Diabetes risk calculator: a simple tool for detecting undiagnosed diabetes and pre-diabetes. *Diabetes Care* 2008;31:1040–5.
7. Cabrera de León A, Coello SD, Rodríguez Pérez MC, *et al*. A simple clinical score for type 2 diabetes mellitus screening in the Canary Islands. *Diabetes Res Clin Pract* 2008;80:128–33.
8. Vassy JL, Durant NH, Kabagambe EK, *et al*. A genotype risk score predicts type 2 diabetes from young adulthood: the CARDIA study. *Diabetologia* 2012;55:2604–12.
9. Schmidt MI, Duncan BB, Bang H, *et al*. Identifying individuals at high risk for diabetes: The Atherosclerosis Risk in Communities study. *Diabetes Care* 2005;28:2013–8.
10. Kahn HS, Cheng YJ, Thompson TJ, *et al*. Two risk-scoring systems for predicting incident diabetes mellitus in U.S. adults age 45 to 64 years. *Ann Intern Med* 2009;150:741–51.
11. Mozaffarian D, Kamineni A, Carnethon M, *et al*. Lifestyle risk factors and new-onset diabetes mellitus in older adults: the cardiovascular health study. *Arch Intern Med* 2009;169:798–807.