Supplementary Online Content

Bancks MP, Kershaw K, Carson AP, Gordon-Larsen P, Schreiner PJ, Carnethon MR. Association of modifiable risk factors in young adulthood with racial disparity in incident type 2 diabetes during middle adulthood. *JAMA*. doi:10.1001/jama.2017.19546

This supplementary material has been provided by the authors to give readers additional information about their work.

eAppendix. Supplementary Methods, Results, and References

Supplementary Methods

Fasting glucose, post-challenge glucose, and HbA1c were assessed at various examinations. Fasting glucose was measured in examination years 0 (Y0) and Y7-30. Post-challenge glucose was measured in Y10, Y20, and Y25. HbA1c was assessed in Y20 and Y25. Self-reported use of diabetes medication from the medication inventory was available for all examination years.

We assessed the sensitivity of our estimates to possible racial differences in health and medical care services. At CARDIA examination year 7, participants were queried about access to medical care services. Participants were asked to report whether they had a usual source of medical care in the occasion of a check-up or if they are ill and to identify the source of such care (e.g., private/personal physician, walk-in clinic, clinic by appointment, hospital emergency room, other).

To understand the impact of excluding individuals from our analysis who were missing baseline risk factor data we estimated HRs when using multiple imputation to impute missing baseline risk factor data. We used fully conditional specification (FCS) methods for multiple imputation of missing baseline risk factor data, drawing from the full CARDIA cohort, creating five multiple imputation datasets. FCS methods do not rely on the assumption of multivariate normality for all the variables in the imputation model and are ideal for imputing missing data for continuous and categorical variables.[1] FCS methods use data from available cases and each variable with missing values is imputed using a regression model conditional on all of the other variables specified in the imputation model, shown to be a reliable approach given sufficient sample size.[1,2] We estimated HRs for diabetes in each of the five datasets individually and combined estimates according to Rubin's Rules.[3]

Supplementary Results

During follow-up, 17 of 189 (9%) incident cases among whites were ascertained by diabetes medication use data alone, while among blacks 33 of 313 (11%) incident cases were ascertained by diabetes medication data alone.

Among whites, 78% of women and 71% of men reported having a usual source of medical care at exam 7. For blacks, 72% of women and 63% of men reported having a usual source of medical care. The prominent source of medical care for each race-sex groups was private/personal physician, 67% for white women, 56% for black women, 57% for white men, and 40% for black men. When medical care characteristics were included in the final model for time-invariant covariates, the HR for black women compared to white women was 1.24 (95% CI=0.85, 1.81) and for black men compared to white men was 1.50 (95% CI=1.01, 2.22). Neither usual medical care nor source of medical care were significantly ($p>0.05$) associated with incidence of diabetes in this model.

Relative to 4251 included in the primary analysis, the 863 individuals (864 individuals were excluded, but 1 individual withdrew consent to participate and was not analyzed) excluded from the primary analysis were more likely to be black, less educated, have lower parental education, more likely to be a current smoker and heavy consumer of alcohol, have higher fasting glucose and poorer diet score, and more likely to live in a neighborhood with higher than expected black residents and greater percentage of the neighborhood living in poverty. We excluded 162 white women (12% of enrolled white women), 321 black women (22% of enrolled black women), 131 white men (11% of enrolled white men), and 250 black men (22% of enrolled black men) from primary analyses. In analyses using multiple imputation to impute missing baseline risk factor data, black women had a HR of 1.18 (95% CI=0.83, 1.69) for diabetes compared to white women after adjustment for all risk factor groups. In a model adjusted for all risk factor groups, black men had 1.51 times greater risk for diabetes compared their white counterparts (95% CI=1.04, 2.18) when using multiple imputation methods to impute missing baseline risk factor data.

References

1. Liu Y, De A. Multiple Imputation by Fully Conditional Specification for Dealing with Missing Data in a Large Epidemiologic Study. *International journal of statistics in medical research.* 2015;4(3):287-295.
2. Lee KJ, Carlin JB. Multiple Imputation for Missing Data: Fully Conditional Specification Versus Multivariate Normal Imputation. *American journal of epidemiology.* 2010;171(5):624-632.
3. Little R, Rubin D. *Statistical analysis with missing data.* 1st ed. New York, NY: John Wiley; 1987.

eTable 1. Participant Exclusions

CARDIA Recruitment	**5115**
Exclusions	
Withdrew consent to participate	1
Unknown or prevalent diabetes at baseline	35
Unable to determine diabetes status during follow-up	158
Missing baseline body mass index	15
Missing baseline waist circumference	7
Missing baseline fasting glucose	78
Missing baseline physical activity	1
Missing baseline diet history	120
Missing baseline education	2
Missing baseline alcohol consumption	19
Missing baseline smoking status	33
Missing baseline neighborhood segregation	10
Missing baseline forced vital capacity	226
Missing baseline triglycerides	159
Analytic sample	4251

eTable 2. Comparison of Characteristics Between Participants Included in the Analytic Sample and Those Excluded From Analysis

Characteristic	Excluded (n=863)[a]	Included (n=4251)
Age, mean (SD), years	24.2 (3.8)	25.0 (3.6)
Black, n (%)	571 (66)	2066 (49)
Women, n (%)	483 (56)	2304 (54)
Socioeconomic		
Education > 16 years, n (%)	47 (5)	445 (10)
Married, n (%)	163 (19)	974 (23)
Somewhat hard, hard, or very hard to pay for basics, n (%)	364 (42)	1431 (34)
Employed full-time, n (%)	402 (47)	2475 (58)
Mother's education > 16 years, n (%)	34 (4)	338 (8)
Father's education > 16 years, n (%)	56 (6)	533 (13)
Neighborhood		
G-statistic of black segregation, mean (SD)	4.2 (3.7)	3.1 (3.5)
Percentage of census tract living in poverty, mean (SD)	21 (13)	18 (12)
Psychosocial		
CES - Depression, mean (SD), range 0-60, higher is depressive symptoms	12.5 (8.5)	11.0 (8.1)
Behavioral		
Never smoker, n (%)	439 (51)	2417 (57)
Former smoker, n (%)	88 (10)	588 (14)
Current smoker, n (%)	300 (35)	1246 (29)
No daily alcohol, n (%)	335 (39)	1651 (39)
Moderate alcohol, n (%)	398 (46)	2076 (49)
Heavy alcohol, n (%)	109 (13)	524 (12)
Physical activity, mean (SD), exercise units	433 (343)	418 (291)
Diet score, mean (SD), range 0-5, higher is healthier diet	1.1 (1.0)	1.3 (1.0)
Biological		
Fasting glucose, mean (SD), mg/dL	87 (36)	82 (8)
BMI, mean (SD), kg/m^2	24 (6)	25 (5)
Waist circumference, mean (SD), cm	77 (12)	78 (11)
Systolic blood pressure, mean (SD), mm Hg	110 (11)	110 (11)
Use of blood-pressure lowering medications, n (%)	28 (3)	87 (2)
HDL-cholesterol, mean (SD), mg/dL	54 (14)	53 (13)
Triglycerides, mean (SD), mg/dL	72 (52)	73 (48)
Parental history of diabetes (mother or father), n (%)	118 (14)	592 (14)
Forced vital capacity, mean (SD), L	4.2 (1.0)	4.3 (1.0)

[a] 864 individuals were excluded, but 1 individual withdrew consent to participate and was not analyzed. Where categorical smoking and alcohol consumption variable percentages for excluded sample do not sum to 100%, values are missing.
Continuous variables are presented as mean (SD) and categorical as n (%).
Abbreviations: CES, Center for Epidemiologic Studies; BMI, Body Mass Index; HDL, high-density lipoprotein.
SI conversion factors: To convert glucose to mmol/L, multiply values by 0.0556. To convert HDL-cholesterol to mmol/L, multiply values by 0.0259. To convert triglycerides to mmol/L, multiply values by 0.0113.
G-statistic of black segregation: interpretation is similar to a z-score where a value greater than 1.96 would indicate living in a residential area with significantly greater than expected percentage black residents and a value less than -1.96 would represent significantly lower than expected percentage. Values between -1.96 and 1.96 would indicate no appreciable difference in actual percentage of black residents compared to the expected.

eTable 3. Comparison of Characteristics for Participants Who Were not Identified as an Incident Diabetes Case According to Last Examination Visit

Characteristic	Last Examination Visit							
	2	**3**	**4**	**5**	**6**	**7**	**8**	**9**
N	131	97	122	170	184	206	314	2523
Black, n (%)	76 (58)	53 (55)	70 (57)	107 (63)	115 (63)	104 (50)	144 (46)	1082 (43)
Women, n (%)	62 (47)	46 (47)	58 (48)	87 (51)	82 (45)	106 (51)	164 (52)	1430 (57)
Age, mean (SD), years	24.5 (3.7)	23.8 (3.7)	24.3 (3.4)	24.8 (3.8)	24.4 (3.6)	24.7 (3.8)	24.7 (3.8)	25.1 (3.5)
Education >16 years, n (%)	19 (15)	7 (7)	8 (7)	7 (4)	11 (6)	13 (6)	29 (9)	320 (13)
Birmingham, AL, n (%)	39 (30)	16 (16)	25 (20)	32 (19)	42 (23)	48 (23)	75 (24)	513 (20)
Chicago, IL, n (%)	28 (21)	26 (27)	18 (15)	28 (16)	32 (17)	46 (22)	77 (25)	563 (22)
Minneapolis, MN, n (%)	35 (27)	29 (30)	37 (30)	57 (34)	63 (34)	55 (27)	103 (33)	650 (26)
Oakland, CA, n (%)	29 (22)	26 (27)	42 (34)	53 (31)	47 (26)	57 (28)	59 (19)	797 (32)

Continuous variables are presented as mean (SD) and categorical as n (%).
Examination 9 occurred in 2015-2016

eTable 4. Hazard Ratios (95% CIs) for Incident Diabetes in Blacks Compared With Whites and Percent Reduction in Parameter Estimates for Women and Men Combined According to Sequential Adjustment for Each Risk Factor Domain When Including Updated Risk Factor Information for Model Adjustment and When Adjusting for Baseline Measurement of Risk Factors

	Black	**White**
N	2066	2185
Cases/Person-years	315/48337	189/56009
Risk difference/1000 people[a]	66 (95 % CI: 45, 87)	
Updated risk factor information adjustment		
Model	Race HR (95% CI)	Percent Reduction in β[b]
M1: Age and field center	2.22 (1.84, 2.67)	Reference
M2: M1 + Biological[c]	1.03 (0.82, 1.30)	96%
M3: M2 + Neighborhood[d]	0.97 (0.75, 1.26)	104%
M4: M3 + Psychosocial[e]	0.96 (0.74, 1.24)	105%
M5: M4 + Socioeconomic[f]	0.91 (0.70, 1.19)	112%
M6: M5 + Behavioral[g]	0.90 (0.69, 1.18)	113%
Baseline risk factor adjustment		
Model		
M2: M1 + Biological[c]	1.49 (1.19, 1.87)	50%
M3: M2 + Neighborhood[d]	1.45 (1.13, 1.87)	53%
M4: M3 + Psychosocial[e]	1.45 (1.13, 1.87)	53%
M5: M4 + Socioeconomic[f]	1.32 (1.01, 1.72)	65%
M6: M5 + Behavioral[g]	1.29 (0.99, 1.69)	68%

[a] Risk difference is age, sex, and field center adjusted presenting excess cases in blacks relative to whites.
[b] Percent Reduction in β-estimate: $(\beta_0-\beta_n)/(\beta_0)*100$, where β_0 = Age, sex, sex-race interaction term, and field center adjusted reference model
[c] Biological: Fasting glucose, body mass index, waist circumference, parental history of diabetes, triglycerides-to-high-density lipoprotein cholesterol ratio, forced vital capacity, systolic blood pressure, and blood pressure lowering medication use
[d] Neighborhood: G-statistic for racial segregation and tract-level percentage of population living in poverty
[e] Psychosocial: CES - Depression
[f] Socioeconomic: education, current employment status, paying for basics, marital status, and mother's and father's educational attainment
[g] Behavioral: regular alcohol consumption, smoking status, diet score from AHA Life's Simple 7, and regular physical activity