1.a. Full Title: Novel Risk Factors of Diabetes and Their Impact on the Racial Disparity in Risk of Incident Diabetes: The Atherosclerosis Risk in Communities Study

b. Abbreviated Title (Length 26 characters): Racial Disparity in Diabetes Risk

2. Writing Group:
   Hsin Chieh Yeh, PhD,
   Tariq Shafi, MBBS, MHS,
   Elizabeth Selvin, PhD,
   James Pankow, PhD,
   Frederick L. Brancati. MD, MHS,
   others welcome

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. __RC___ [please confirm with your initials electronically or in writing]

First author: Ranee Chatterjee, MD, MPH
Address: 2024 E. Monument St, Suite 2-501, Baltimore, MD 21205

   Phone: 410-614-6441       Fax: 410-955-0476
   E-mail: rchatte2@jhu.edu

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).

   Name: Hsin Chieh Yeh, PhD
   Address: 2024 E. Monument St, Suite 2-600, Baltimore, MD 21205

   Phone: (410) 614-4316       Fax: 410-955-0476
   E-mail: hcyeh@jhsph.edu

3. Timeline: Data analysis and manuscript preparation will be performed over the next eight months.
4. **Rationale:**

African Americans have been found, consistently over the past few decades, to account for a disproportionately high percentage of the diabetes epidemic. Recent estimates from NHANES 2005-2006 found that the prevalence of diabetes among African Americans is 70% higher than that of non-Hispanic whites after adjustment for differences in sex and age (1). Many factors are thought to contribute to the higher prevalence of diabetes seen among African Americans, including differences in socioeconomic status, diet, behavioral factors, as well as related comorbidities, particularly obesity (2). However, not all of the increase in risk of diabetes can be accounted for by traditional risk factors, and there are likely to be other metabolic and genetic factors that contribute to this increased risk.

There have been many studies looking at novel risk factors for diabetes. In one previous ARIC study, traditional and suspected risk factors were evaluated for their role in the racial disparity in diabetes risk. This analysis included smoking, alcohol consumption, and dietary energy as the “suspected diabetes risk factors” but found that obesity, a traditional risk factor, accounted for a substantial portion of the racial disparity in diabetes risk particularly in women (3). Since that time, several studies have looked at other novel risk factors. In a previous ARIC study done by this group, serum potassium was found to be a significant predictor of incident diabetes with low-normal serum potassium associated with higher risk (4). We also found that in models using primarily traditional risk factors of diabetes, serum potassium accounted for almost 20% of the racial disparity in risk of incident diabetes, which was almost equivalent to the effect of BMI on the racial disparity in diabetes risk in these models (unpublished). Most studies of novel risk factors of diabetes have not assessed the impact of these factors on the racial disparity found in risk of incident diabetes. If factors can be identified that have a significant impact on the association between race and risk of diabetes, a better understanding of the biological basis for this disparity might be achieved and interventions could be developed to help reduce this disparity.

5. **Main Hypothesis/Study Questions:**

Using ARIC data, we will study the association between race and risk of incident diabetes. We hypothesize that multivariate models containing both traditional and novel risk factors of diabetes will account for most of the racial disparity in risk of incident diabetes. From these models, we will determine which of these novel risk factors have the greatest impact on the association between race and risk of incident diabetes.

6. **Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).**

**Study design:** Prospective cohort study using Atherosclerosis Risk in Communities (ARIC) data

**Inclusion criteria:** All ARIC participants
Exclusion criteria:
1) Participants with diabetes at the baseline exam
2) Participants with missing information regarding diabetes status at baseline exam
3) Participants with ethnicity other than African American or white
4) Participants with missing information regarding covariates of interest at baseline exam
5) Participants with kidney disease as defined by a serum creatinine > 1.7 mg/dL

Outcome: The main outcome will be incident diabetes, or a new diagnosis of diabetes at Visit 2, 3, or 4, as defined by ARIC, with a participant’s report of a clinical diagnosis and/or use of medications for diabetes, and/or biochemical evidence of diabetes, defined as a fasting glucose ≥ 126 mg/dL or non-fasting glucose ≥ 200mg/dL at any follow-up visit.

Main exposure: Race based on self-report

Covariates: We will include traditional risk factors of diabetes that are thought to confound the association between race and incident diabetes including: age, sex, BMI, waist circumference/waist-to-hip ratio, parental history of diabetes, physical activity level, hypertension, systolic blood pressure, use of antihypertensives (beta-blockers, thiazides, ACE-inhibitors), smoking history, and income.

We will also consider the following novel risk factors, based on ARIC studies as well as studies from other cohorts, as potential mediators of the association between race and incident diabetes: serum potassium (4), serum magnesium (5), actual forced vital capacity (6), leg length (7), hematocrit (8), coffee intake (9), uric acid (10, 11), resting heart rate (12), cereal and total dietary fiber intake (13), factor VIII (14), von Willebrands factor (vWF) (14), activated partial thromboplastin time (aPTT) (14), white blood cell count (15), and albumin (16).

Data analysis:
1) Baseline characteristics of participants will be compared by race—using χ2 tests for categorical variables and students t-tests for continuous variables.

2) We will determine if each covariate meets the criteria of being a potential mediator of the association between race and incident diabetes by testing if 1) the covariate is predicted by race and 2) if the covariate predicts risk of incident diabetes after controlling for race.

3) Cox proportional hazard models will be used to assess the association between race and incident diabetes controlling for those covariates that meet the criteria of being potential mediators of this association.

4) We will calculate the mediation effect of each covariate as the percent change in the coefficient of race in models with and without the covariate of interest. 95% confidence intervals will be calculated using boot-strapping with replacement (1000 samples) (17). Tests of significance will be two-tailed, with an alpha level of 0.05.
5) We will conduct a sensitivity analysis on the residents of Forsyth County, North Carolina, which is the only center which had participants of both races, so as to ensure that the pattern found in the entire cohort is not due to geographic confounding.

Limitations:
We recognize that in calculating mediation effects of covariates, we are assuming a causal relationship, which should not be assumed from an epidemiologic study. However, we think that factors which are found to impact the racial disparity in risk of incident diabetes will deserve further attention and study: 1) to determine the cause of this association and 2) to determine if interventions can be designed to reduce this disparity.

7.a. Will the data be used for non-CVD analysis in this manuscript? ____ Yes __X_ No

b. If Yes, is the author aware that the file ICTDER03 must be used to exclude persons with a value RES_OTH = “CVD Research” for non-DNA analysis, and for DNA analysis RES_DNA = “CVD Research” would be used? _____ Yes _____ No

(This file ICTDER03 has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)

8.a. Will the DNA data be used in this manuscript? ____ Yes __X_ No

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the file ICTDER03 must be used to exclude those with value RES_DNA = “No use/storage DNA”? _____ Yes _____ No

9. The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status. ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: http://www.cscc.unc.edu/ARIC/search.php

____X__  Yes _______ No

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?
11. a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? ____ Yes  ___X___ No

11. b. If yes, is the proposal
   ___ A. primarily the result of an ancillary study (list number* _________)
   ___ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* _________  _________  _________)

*ancillary studies are listed by number at http://www.cscc.unc.edu/aric/forms/

12. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.

   _RC_

References