ARIC Manuscript Proposal #2380

PC Reviewed: 6/10/14	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title:
Associations of the Serum Metabolome and Mortality among African Americans in the Atherosclerosis Risk in Communities (ARIC) Study

b. Abbreviated Title (Length 26 characters): Metabolomics and mortality

2. Writing Group:
Writing group members: Bing Yu, Gerardo Heiss, Danny Alexander, Eric Boerwinkle

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

First author: Bing Yu
Address: Division of Epidemiology, Human Genetics & Environmental Sciences
University of Texas Health Science Center
1200 Hermann Pressler, suite E-435
Houston, TX, 77030
Phone: 713-500-9892 Fax: 713-500-0900
E-mail: Bing.Yu@uth.tmc.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: Eric Boerwinkle
Division of Epidemiology, Human Genetics & Environmental Sciences
University of Texas Health Science Center
1200 Herman Pressler, suite E-447
Houston, TX 77030
Phone: 713-500-9816 Fax: 713-500-0900
Email: Eric.Boerwinkle@uth.tmc.edu

3. Timeline:
We expect that the manuscript will be prepared within six months from approval of the analysis plan.

4. Rationale:
All-cause mortality is a traditional measure of disease burden in public health, and currently, cardiovascular disease (CVD) is the leading cause of death for all ages in the United States (1). Thus, early and accurate identification of people with high risk of

death, particularly death of CVD, could assist targeting of preventive therapies in order to improve the overall health. Previous studies have shown that circulating biomarkers have the potential to predict all-cause mortality and cardiovascular mortality (2-5), but there lacks of a systematic way to fully explore the effect of all circulating molecules. Metabolomics, which characterizes small-molecular metabolites produced by a variety of biochemical and cellular processes, is a step closer to this ideal and is one path toward new biomarker discovery. In one recent study 4 out of 116 metabolites were associated with all-cause mortality as well as cardiovascular mortality using targeted metabolomic technology among Estonia and Finland populations (6).

Untargeted metabolomic technology can capture all possible small molecules in a biosample, and thus it provides the opportunity to more fully assess relationships between metabolites and mortality. African-American adults have a higher rate of cardiovascular disease and mortality (1) than other racial groups, and the exact mechanism underlying such disparity is not clear. To date, no study has explored the metabolomic antecedents of all causal mortality and cardiovascular mortality in a large African-American population. In a well-characterized, population-based sample of African Americans from the Atherosclerosis Risk in Communities (ARIC) study, we propose to explore the longitudinal association of serum metabolome quantified by untargeted metabolomic technology with all-cause mortality and cardiovascular mortality.

5. Main Hypothesis:

1. Metabolomic variables measured at baseline in ARIC African Americans are associated with all casual mortality by December 31 2011, independent of traditional risk factors.
2. Metabolomic variables measured at baseline, in ARIC African Americans are associated with cardiovascular mortality by December 31 2011, independent of traditional risk factors.

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 and sample:
This is a study consisting of ARIC African Americans with serum metabolomic data quantified at baseline (visit 1).

Exclusion:

- Persons with no metabolomic data (metabolomics profiling was completed in a random sample of African Americans from the Jackson, MS, field center);
- Persons with missing outcome variables or baseline covariates;

Outcome:

- Death from any cause.

- Cardiovascular mortality was defined as deaths in which the principal cause of mortality was cardiovascular in nature, using ICD-9 codes 390-459, or ICD-10 codes I00–I99.

Follow-up time:

- From visit 1 in 1987-89 to event or to December 31, 2011.

Likely covariates at baseline exam:

- Age (yrs)
- Sex
- Body mass index (kg/m^2)
- The presence or absence of prevalent cardiovascular disease, hypertension and diabetes
- Lipids levels (HDL and total cholesterol)
- Current smoking status
- Estimated Glomerular Filtration Rate ($eGFR_{CKD\text{-}EPI}$, $mL/min/1.73m^2$)

Statistical Methods:

As described in our previous work (7-8), a total of 204 reliable metabolites with reliability coefficients (RC) $\geqslant$ 0.6 and missing values < 80% will be included in the analyses. Metabolites (n=187) with RC $\geqslant$0.6 and have missing values in fewer than 50% of the sample will be treated as continuous variables (standardized prior to the analyses) with the missingness of metabolites are replaced by the lowest measured value. Metabolites (n=17) with RC $\geqslant$0.60 but have a moderate amount of missing data (values missing in 50- 80% of the sample) will be treated as ordinal variables (category 1: missing values, category 2: values below or equal to the median, category 3: values above the median).

Cox proportional hazards models will be used to assess associations between each metabolite and all causal mortality adjusting for traditional risk factors (age, sex, BMI, prevalent diabetes, hypertension, HDL, total cholesterol, current smoking status, eGFR; and prevalence of major cardiovascular disease). In addition, cox proportional hazards models will be used to assess the associations between each metabolite and cardiovascular mortality accounting for the same risk factors. Statistical significance for the metabolomic data will be pre-specified at $p < 1.6 \times 10^{-4}$ using Bonferroni correction (0.05 / 308 metabolites). If more than one metabolite is identified, a MetScore, which sums the quartiles of each identified metabolite, will be used to test the joint effect on mortality prediction of these metabolites.

Reference:

1. Go AS, et al., Heart disease and stroke statistics--2014 update: a report from the American Heart Association. *Circulation*, 2014, 129(3):e28-e292.
2. Wang TJ, et al., Multiple biomarkers for the prediction of first major cardiovascular events and death. *N Engl J Med*. 2006, 355(25):2631-2639.
3. Clarke R, et al., Biomarkers of inflammation predict both vascular and non-vascular mortality in older men. *Eur Heart J*. 2008, 29(6):800-809.

4. Emerging Risk Factors Collaboration, et al., C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. *Lancet*. 2010, 375(9709):132-140.
5. Emerging Risk Factors Collaboration, et al., Diabetes mellitus, fasting glucose, and risk of cause-specific death. *N Engl J Med*. 2011, 364(9):829-841.
6. Fischer K, et al., Biomarker profiling by nuclear magnetic resonance spectroscopy for the prediction of all-cause mortality: an observational study of 17,345 persons. *PLoS Med*. 2014, 11(2):e1001606.
7. Zheng Y, et al., Associations between metabolomic compounds and incident heart failure among African Americans: the ARIC Study. *Am J Epidemiol*. 2013, 178(4):534-542.
8. Zheng Y, et al., Metabolomics and incident hypertension among blacks: the atherosclerosis risk in communities study. *Hypertension*. 2013, 62(2):398-403.

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

8.c. If yes, is the author aware that the participants with RES_DNA = 'not for profit' restriction must be excluded if the data are used by a for profit group?
___ 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
Yes. There is no overlap between this proposal and current proposals/published manuscripts.

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)?

MS#1847 Zheng Y, et al. Role of the Human Metabolome in Incident Heart Failure Etiology among African Americans in the Atherosclerosis Risk in Communities (ARIC) Study.
MS#1882 Yu B, et al. A longitudinal Study of Metabolomics and Kidney Function among African Americans in the Atherosclerosis Risk in Communities (ARIC) Study.

11. a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? Yes
11.b. If yes, is the proposal
__X__ A. primarily the result of an ancillary study (list number* 2008.16**)**
___ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* 2008.16 "Metabolomics & Heart Failure: A Novel Approach to Biomarker Discovery")
*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.
Yes, the lead author is aware that manuscript preparation is expected to be completed in 1-3 years, and if this expectation is not met, the manuscript proposal will expire.