

ARIC Manuscript Proposal #2847

PC Reviewed: 09/13/16
SC Reviewed: _____

Status: _____
Status: _____

Priority: 2
Priority: _____

1.a. Full Title: Metabolomics and incident AF: the ARIC study

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

2. Writing Group

Writing group members: Alvaro Alonso, Bing Yu, Wesley O'Neal, Yan Sun, Elsayed Soliman, Lin Y Chen, Eric Boerwinkle

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

First author: Alvaro Alonso

Address: Dept of Epidemiology, Rollins School of Public Health
Emory University
1518 Clifton Rd NE, Room 3051
Atlanta, GA 30322

Phone: 4047278714

E-mail: alvaro.alonso@emory.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

Address: Human Genetics Center
1200 Herman Pressler, Suite E447
Houston, TX 77030

Phone: 713-500-9803

E-mail: Eric.Boerwinkle@uth.tmc.edu

3. Timeline:

Manuscript to be completed in <1 year after approval

4. Rationale:

Application of metabolomics to the study of atrial fibrillation (AF), a common arrhythmia, has been limited. Metabolomic approaches could potentially identify novel biomarkers of AF and provide new clues regarding the etiopathogenesis of this arrhythmia. Only three prior publications have studied the metabolomics of AF. Two of them used metabolomics to assess the effects of AF in cardiac tissues.^{1,2} The third one used ARIC data to systematically examine the

prospective association of molecules identified through metabolomic profiling with the incidence of AF.³ In this ARIC analysis, which included 1919 African-American participants and 183 cases of AF, we found that two conjugated bile acids (glycolithocholate sulfate and glycochenolate sulfate) were significantly associated with AF risk after multivariable adjustment and correction for multiple comparisons. We propose to extend our previous work taking advantage of the new ARIC metabolomics data obtained in ~2000 additional study participants and the inclusion of European-Americans.

5. Main Hypothesis/Study Questions:

The primary aim of this analysis is to study the association of molecules measured through untargeted metabolomic profiling with incident AF. We hypothesize that (1) concentrations of glycolithocholate sulfate and glycochenolate sulfate, previously associated with AF in 1919 African-American ARIC participants, will remain associated with AF in a larger sample; and that (2) the analysis will identify novel biomarkers of AF risk.

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).

We will conduct a prospective study of participants with metabolomic data and without AF at visit 1. Eligible participants will be followed up through the end of 2013.

Assessment of metabolites

Metabolite profiling in baseline serum samples was done in 1,977 ARIC participants from the Jackson field center in 2010 and in an additional ~2,000 ARIC participants in 2014 using an untargeted, gas chromatography-mass spectrometry and liquid chromatography-mass spectrometry-based metabolomics quantification protocol. The second sample set includes primarily European-Americans.

Based on both practical and theoretical considerations, we will limit the analysis to those metabolites with acceptable reliability ($r > 0.6$), no obvious batch effect, low missing rate ($< 25\%$ in both batches) and shared between African Americans and European Americans. It is estimated that there are to be 245 named metabolites that satisfy these criteria.

Metabolites with missing/below detection limit in $< 25\%$ of the samples will be analyzed as continuous variables, assigning the lowest detected value for that metabolite to those with missing/below detection limit.

Main outcome variable

Incident atrial fibrillation will be identified as previously described.^{4, 5} Specifically, we will identify new cases of AF from study ECGs, hospital discharge codes, and death certificates. Follow-up through the end of 2013 will be considered. We expect to include 300-400 incident AF events in this analysis.

Statistical analysis

We will conduct two separate analyses. The first one will aim to replicate the results from our previous publication. Specifically, we will assess the association of glycolithocholate sulfate and

glycocholate sulfate with AF incidence in the ~2000 participants, mainly of European ancestry, with new metabolomic data. The second analysis will include all participants with metabolomic data in a single population and explore associations of the 245 selected metabolites with AF incidence.

We will assess the association of each metabolite with incident AF using Cox regression. Hazard ratios will be calculated per 1-standard deviation difference in the concentrations of each metabolite. An initial model will adjust for age, race-center-batch, and sex. A second model will additionally adjust for traditional risk factors of AF (smoking, body mass index, systolic blood pressure, antihypertensive medication, prevalent diabetes, prevalent CHD, and prevalent HF). A final model will additionally adjust for eGFR. The manuscript will present models with and without adjustment for eGFR, and our interpretation of the findings will carefully evaluate results from the different models. We will apply a modified stepwise Bonferroni procedure to correct for multiple comparisons, as previously described.⁶ In exploratory analyses, we will explore differences by race.

7.a. Will the data be used for non-CVD analysis in this manuscript? ☐ Yes ☒ 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 ICTDER 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 ☒ 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.csc.unc.edu/ARIC/search.php>

☒ 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)?

MS #2354 Metabolomics and AF (Alonso)—previous publication including data from 1919 participants. The new proposal will extend the findings from MS#2354 with a larger sample size and into the European-American population.

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? ____ Yes ____ No

11.b. If yes, is the proposal

☒ **A. primarily the result of an ancillary study (list number* 2008.16 and 2014.20)**

____ **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.csc.unc.edu/aric/forms/>

12a. 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.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is **your responsibility to upload manuscripts to PubMed Central** whenever the journal does not and be in compliance with this policy. Four files about the public access policy from <http://publicaccess.nih.gov/> are posted in <http://www.csc.unc.edu/aric/index.php>, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.

13. Per Data Use Agreement Addendum, approved manuscripts using CMS data shall be submitted by the Coordinating Center to CMS for informational purposes prior to publication. Approved manuscripts should be sent to Pingping Wu at CC, at pingping_wu@unc.edu. I will be using CMS data in my manuscript ____ Yes ☒ No.

REFERENCES

1. Mayr M, Yusuf S, Weir G, et al. Combined metabolomic and proteomic analysis of human atrial fibrillation. *J Am Coll Cardiol*. 2008;51:585-594.
2. De Souza AI, Cardin S, Wait R, et al. Proteomic and metabolomic analysis of atrial profibrillatory remodelling in congestive heart failure. *J Mol Cell Cardiol*. 2010;49:851-863.
3. Alonso A, Yu B, Qureshi WT, et al. Metabolomics and incidence of atrial fibrillation in African Americans: the Atherosclerosis Risk in Communities (ARIC) Study. *PLoS One*. 2015;10:e0142610.
4. Alonso A, Agarwal SK, Soliman EZ, et al. Incidence of atrial fibrillation in whites and African-Americans: the Atherosclerosis Risk in Communities (ARIC) study. *Am Heart J*. 2009;158:111-117.
5. Chamberlain AM, Agarwal SK, Ambrose M, Folsom AR, Soliman EZ, Alonso A. Metabolic syndrome and incidence of atrial fibrillation among blacks and whites in the Atherosclerosis Risk in Communities (ARIC) Study. *Am Heart J*. 2010;159:850-856.
6. Zheng Y, Yu B, Alexander D, et al. Associations between metabolomic compounds and incident heart failure among African Americans: the ARIC Study. *Am J Epidemiol*. 2013;178:534-542.