ARIC Manuscript Proposal #3398

PC Reviewed: 5/14/19	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Proteomics and the Risk of Incident Atrial Fibrillation in the Elderly: The Atherosclerosis Risk in Communities (ARIC) study

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

2. Writing Group:
Writing group members: Faye L. Norby, Lin Y. Chen, Alvaro Alonso, Pamela L. Lutsey, James S. Pankow, Christie M. Ballantyne, Eric Boerwinkle, Josef Coresh, Weihong Tang, Aaron R. Folsom, others welcome

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

First author: **Faye Norby**
Address: Division of Epidemiology & Community Health
School of Public Health, University of Minnesota
1300 S 2nd St, Suite 300
Minneapolis, MN 55454
Phone: 612-626-9096
E-mail: flopez@umn.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: **Aaron R. Folsom**
Address: Division of Epidemiology & Community Health
School of Public Health, University of Minnesota
1300 S 2nd St, Suite 300
Minneapolis, MN 55454
Phone: 612-626-8862
E-mail: folso001@umn.edu

3. Timeline: Statistical analysis: 3 months
Manuscript preparation: 6 months

4. Rationale:

Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia, with a lifetime risk of 1 in 3 among whites and 1 in 5 among African Americans.[1] The prevalence of AF increases with older age, from 0.1% among people younger than 55 years to 9% among people 80 years or older.[2] The risk of AF increases with advancing age, taller height, European ancestry, smoking, higher weight, higher blood pressure, blood pressure medication use, diabetes, history of myocardial infarction, and history of heart failure.[3, 4] In addition to the traditional clinical risk factors listed above, various biomarkers have been identified as risk factors for incident AF including markers of inflammation,[5-8] oxidative stress,[9] myocardial necrosis,[5, 10, 11] myocardial stress,[5, 12-17] and mineral metabolism.[18, 19] Despite the relatively high prevalence and incidence of AF, our understanding of its pathobiology and precipitants remains superficial. Identification of novel biomarkers can advance our understanding of AF mechanisms, enhance opportunities for risk prediction, and potentially provide targeted preventive strategies for AF.

Proteomic profiling enables systematic high-throughput analysis of proteins and may substantially accelerate novel biomarker discovery. Relatively unbiased proteomics approaches have the advantage of allowing simultaneous screening for large numbers of proteins involved in different biological pathways. Recently, 3 longitudinal cohort studies have reported proteomic profiling and the risk of new-onset AF. [6, 17, 20] The first study used a proximity extension assay (Olink Proseek Multiplex Cardiovascular 96 x 96 kit) to screen 92 proteins in 2 community-based cohorts of older adults in Sweden with a total of 271 incident AF cases in 1703 participants over a median follow-up of around 9 years.[17] They identified 7 proteins that were associated with incident AF after adjustment for age and sex. Two proteins, NT-proBNP and IL-6, remained significantly associated with incident AF after multivariable adjustment and Bonferroni correction.[17] The second cohort study used a community-based sample from Italy and focused on 75 inflammatory marker proteins identified from proximity extension assays (the Olink Proseek Multiplex CVD I 96 x 96 and the Proseek Multiple Inflammation I 96 x 96 kits).[6] There were 117 new AF cases among 880 participants during a 20-year follow-up. The Italian study reported the results of 75 inflammatory biomarkers including FGF-23, fatty acid binding protein 4, and IL-6, none of which were associated with AF after adjustment for age and sex.[6] The third study, from Framingham, used single-stranded DNA-based aptamers as affinity reagents (measured by the

SOMAscan platform) to screen for 1373 proteins.[20] This study included 1885 participants with 349 incident AF cases during a mean follow-up of 18 years. In this study, Ko *et al.* identified 8 proteins associated with AF after adjustment for age and sex, and after further adjustment for AF risk factors, 2 proteins (ADAMTS13 and NT-proBNP) remained associated with new-onset AF.[20]

The ARIC study provides the opportunity to examine the relationship between proteomics and incident AF in an elderly cohort (mean age at visit 5 = 76). A recent collaboration between SomaLogic and the ARIC study allows for proteomic assay using SOMAscan version 4.0 (~5,000 proteins) in ~5,000 participants (~25% African American) at visit 5. At this time, ARIC has approximately 5 years of follow-up data and nearly 500 incident AF events after visit 5, making this analysis feasible in this cohort. The ARIC cohort has power to detect novel associations by having a larger sample size and more AF cases than the 3 previous studies, along with a shorter follow-up time to potentially detect acute associations. ARIC also allows us to examine associations by black and white race.

5. Main Hypothesis/Study Questions:

The main objective of this proposal is to evaluate aptamer-based proteomic profiles with incident AF in an elderly cohort of black and white men and women.

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 from visit 5 until the end of 2017 or most recent data. Follow-up is approximately 5 years.

<u>Study population</u>

- Inclusion criteria: ARIC participants with proteomic measures at visit 5
- Exclusion criteria: Prevalent AF at visit 5 (from visit 5 ECG or hospitalization codes / ECG prior to visit 5), missing or low quality proteomic data, missing or indeterminate ECG

measures at visit 5, race other than white or black and non-whites in the Minneapolis and Washington County field centers (due to low numbers), and those missing covariates.

Ascertainment of AF

The outcome will be incident AF. Incident AF will be ascertained death certificates or hospitalization discharge diagnosis codes (ICD-9-CM: 427.3, 427.31 or 427.32 or ICD-10: I48 in any position) through the end of 2017. Prevalent AF at visit 5 will additionally be identified by the study visit ECG.

Proteomics Profiling

Our exposure will be individual measures of protein levels from plasma samples (~5000 proteins). In brief, EDTA-plasma was obtained from blood samples that were collected during visit 5 and stored at -80. Protein levels in the plasma samples were measured by the SOMAscan platform, which uses single-stranded DNA-based aptamers to capture conformational protein epitopes. Protein levels are measured in relative fluorescent units (RFU). Standard SomaLogic quality control and normalization process have been applied to the protein measures.

Covariates

Covariates will be obtained at visit 5 and the following variables will be considered for inclusion in our models: age, sex, race/site, cigarette smoking status, height, weight, systolic and diastolic blood pressure, anti-hypertensive medication use, diabetes, prevalent myocardial infarction, prevalent heart failure, diuretic use, liver disease (from ICD hospitalization codes), and estimated glomerular filtration rate (eGFR).

Statistical analysis

Our primary analysis will test the association between protein level and incident AF.

We will examine distributions of protein levels and log-transform where appropriate. We may standardize levels to mean=0 and standard deviation=1 for comparison purposes.

We will use Cox proportional hazards models to relate each protein level to incident AF (censored at the last follow-up time, death, or the end of 2017).

- Model 1 will adjust for age, sex, and race/center.
- Model 2 will adjust for Model 1 and additionally adjust for the remaining AF risk factors from the CHARGE-AF score[4] including current cigarette smoking, height, weight, systolic and diastolic blood pressure, the use of hypertension medications, diabetes, prevalent myocardial infarction and prevalent heart failure.
- Model 3 will additionally adjust for factors that could influence measured protein levels including eGFR, diuretic use, and liver disease.

The P-value threshold for significance will be defined using Bonferroni correction to account for the number of analyzed proteins.

We will explore interactions by sex, race, and age.

Sensitivity analysis: For any proteins that are significant after adjustment in the above models, we will run several exploratory sensitivity analyses. We will exclude unreliable protein data (low quality data or data with low reliability in ARIC QC and pilot study which include both analytic and physiologic variability). We will flag outliers (>4 SD from the mean) and will winsorize them to avoid undue influence and violations of assumptions. Finally, we will restrict the analysis to those who were fasting at visit 5.

We will consider an initial analysis that attempts to replicate the results from the Framingham and Swedish cohorts.

Additional analysis: Finally, we will add any significant proteins from this project to an AF prediction model to determine if the proteins enhance the model discrimination and risk stratification. This new AF prediction model is being developed simultaneously in “ARIC #3236: AF prediction in the elderly – Norby”, using participants and data from visit 5. Full methods can

be found in that proposal. Briefly, we will add any significant proteins to the AF prediction model and calculate the added predicted value of the complex model versus the simple model with the increment in C-statistic and the categorical net reclassification improvement (NRI). We will be lacking a comparative external cohort for validation of our prediction equation results, at which point we will use internal validation with bootstrapping methods.[21]

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 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 __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)?
The most relevant manuscripts include:
#3057: Change in proteome – Tin
#3236: AF prediction in the elderly - Norby
#1578: CHARGE-AF risk score – Alonso

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

11.b. If yes, is the proposal
__x_ A. primarily the result of an ancillary study (2017.14, 2017.27, 2018.13)

___ 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/

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.cscc.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.

References:

1. Mou, L., et al., *Lifetime Risk of Atrial Fibrillation by Race and Socioeconomic Status: ARIC Study (Atherosclerosis Risk in Communities).* Circ Arrhythm Electrophysiol, 2018. **11**(7): p. e006350.
2. Go, A.S., et al., *Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study*. JAMA, 2001. **285**(18): p. 2370-5.
3. Alonso, A., et al., *Incidence of atrial fibrillation in whites and African-Americans: the Atherosclerosis Risk in Communities (ARIC) study.* Am Heart J, 2009. **158**(1): p. 111-7.
4. Alonso, A., et al., *Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium.* J Am Heart Assoc, 2013. **2**(2): p. e000102.
5. Rienstra, M., et al., *Relation between soluble ST2, growth differentiation factor-15, and high-sensitivity troponin I and incident atrial fibrillation.* Am Heart J, 2014. **167**(1): p. 109-115 e2.
6. Willeit, K., et al., *Association Between Vascular Cell Adhesion Molecule 1 and Atrial Fibrillation.* JAMA Cardiol, 2017. **2**(5): p. 516-523.
7. Guo, Y., G.Y. Lip, and S. Apostolakis, *Inflammation in atrial fibrillation.* J Am Coll Cardiol, 2012. **60**(22): p. 2263-70.
8. Schnabel, R.B., et al., *Relation of multiple inflammatory biomarkers to incident atrial fibrillation.* Am J Cardiol, 2009. **104**(1): p. 92-6.
9. Samman Tahhan, A., et al., *Association between oxidative stress and atrial fibrillation.* Heart Rhythm, 2017. **14**(12): p. 1849-1855.
10. Filion, K.B., et al., *High-sensitivity cardiac troponin T and the risk of incident atrial fibrillation: the Atherosclerosis Risk in Communities (ARIC) study.* Am Heart J, 2015. **169**(1): p. 31-8 e3.
11. Hussein, A.A., et al., *Serial measures of cardiac troponin T levels by a highly sensitive assay and incident atrial fibrillation in a prospective cohort of ambulatory older adults.* Heart Rhythm, 2015. **12**(5): p. 879-85.
12. Sinner, M.F., et al., *B-type natriuretic peptide and C-reactive protein in the prediction of atrial fibrillation risk: the CHARGE-AF Consortium of community-based cohort studies.* Europace, 2014. **16**(10): p. 1426-33.
13. Schnabel, R.B., et al., *Relations of biomarkers of distinct pathophysiological pathways and atrial fibrillation incidence in the community.* Circulation, 2010. **121**(2): p. 200-7.
14. Patton, K.K., et al., *N-terminal pro-B-type natriuretic peptide is a major predictor of the development of atrial fibrillation: the Cardiovascular Health Study.* Circulation, 2009. **120**(18): p. 1768-74.
15. Patton, K.K., et al., *N-terminal pro-B-type natriuretic peptide as a predictor of incident atrial fibrillation in the Multi-Ethnic Study of Atherosclerosis: the effects of age, sex and ethnicity.* Heart, 2013. **99**(24): p. 1832-6.
16. Kara, K., et al., *B-type natriuretic peptide for incident atrial fibrillation-The Heinz Nixdorf Recall Study.* J Cardiol, 2015. **65**(6): p. 453-8.
17. Lind, L., et al., *Discovery of new biomarkers for atrial fibrillation using a custom-made proteomics chip.* Heart, 2017. **103**(5): p. 377-382.
18. Mathew, J.S., et al., *Fibroblast growth factor-23 and incident atrial fibrillation: the Multi-Ethnic Study of Atherosclerosis (MESA) and the Cardiovascular Health Study (CHS).* Circulation, 2014. **130**(4): p. 298-307.
19. Alonso, A., et al., *Circulating fibroblast growth factor-23 and the incidence of atrial fibrillation: the Atherosclerosis Risk in Communities study.* J Am Heart Assoc, 2014. **3**(5): p. e001082.
20. Ko, D., et al., *Proteomics Profiling and Risk of New-Onset Atrial Fibrillation: Framingham Heart Study.* J Am Heart Assoc, 2019. **8**(6): p. e010976.
21. Steyerberg, E.W., et al., *Internal validation of predictive models: efficiency of some procedures for logistic regression analysis.* J Clin Epidemiol, 2001. **54**(8): p. 774-81.