ARIC Manuscript Proposal # 3236

PC Reviewed: 9/11/18   **Status: _____**   **Priority: 2**
SC Reviewed: _____   **Status: _____**   **Priority: ____**

1.a. Full Title: Prediction of Atrial Fibrillation in an Elderly Cohort: The Atherosclerosis Risk in Communities (ARIC) study

   b. Abbreviated Title (Length 26 characters): AF prediction in elderly

2.   Writing Group:
    Writing group members: Faye L. Norby, Lin Y. Chen, Alvaro Alonso, Pamela L. Lutsey, Elsayed Z. Soliman, Amil Shah, 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]**

    F**irst 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: 8 months
                 We expect to submit an abstract with preliminary results to the AHA Epi conference (submission deadline Oct 2018).

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] AF is associated with an increased risk of adverse cardiovascular outcomes including stroke,[3] myocardial infarction,[4] and mortality.[5] Additionally, AF is associated with significant costs; an AF patient spends an additional $8,705 per year on medical costs compared to someone without AF, and AF costs the US healthcare system $6 billion annually.[6, 7] The aforementioned complications and financial burden associated with AF underscore the importance of accurate AF risk assessment. A risk score for prediction of AF identifies high-risk individuals for screening, clinical trial enrollment, and targeted preventive strategies. A risk score that remains accurate in elderly populations is important, as AF risk increases dramatically with age and the burden that AF places on the health care system will increase with the expected growth in the number of individuals in older age categories.[2]

The need for the accurate prediction of AF has given rise to the development of several population-based prediction equations.[8] Risk scores for AF have been developed in the Framingham Heart Study (FHS),[9] ARIC,[10] the Women's Health Study,[11] and the Cohorts for Aging and Research in Genomic Epidemiology (CHARGE)-AF consortium.[12] The CHARGE-AF risk score derived a 5-year predictive model that used pooled data from 18,556 participants from ARIC, FHS, and the Cardiovascular Health Study (CHS), and was validated in the Age, Gene/Environment Susceptibility Reykjavik study (AGES) and the Rotterdam Study, and in a separate study was again validated in MESA.[13] Investigators developed a simple model, which incorporated common clinically-measured variables, and an augmented model, which incorporated additional electrocardiogram (ECG) measures and blood tests. The advanced-age cohort CHS (mean age 73) was used in deriving the prediction models, however the model did not perform as well in the elderly AGES cohort (mean age 76) as it did in the pooled derivation sample; c-statistic of the simple model = 0.664 in AGES vs. 0.765 in the pooled derivation cohort. This could be a reflection of differing populations, or perhaps the predictive ability of variables changes depending on participants' age. Furthermore, since the publication of the CHARGE-AF risk score in 2013, additional variables have been associated with an increased

risk of incident AF and should be considered when re-evaluating an AF risk prediction equation. Most notably, these variables include the biomarker N-terminal pro B-type natriuretic peptide (NT-proBNP),[14, 15] several P-wave indices including abnormal P-wave axis,[16, 17] and several echocardiograph variables.[18, 19]

ARIC visit 5 gives us the opportunity to re-evaluate the CHARGE-AF risk score in an elderly cohort (mean age at visit 5 = 76), and also to take advantage of the study visit ECG measures, echocardiograms, and biomarker measures. At this time, ARIC has approximately 5 years of follow-up data and nearly 500 incident AF events after visit 5, making a re-evaluation of the CHARGE-AF risk score feasible in this elderly cohort.

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

The main objective of this proposal is to evaluate, recalibrate, and potentially modify the CHARGE-AF risk score in an elderly, mostly bi-racial cohort.

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

Study population

- Inclusion criteria: ARIC participants at visit 5.
- Exclusion criteria: Prevalent AF at visit 5 (from visit 5 ECG or hospitalization codes / ECG prior to visit 5), 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, those missing covariates.

Ascertainment of AF

The main outcome of interest will be the incidence of AF. AF will be ascertained from study visit ECGs (prevalent only), death certificates, or hospitalization discharge diagnosis codes

(ICD-9-CM: 427.3, 427.31 or 427.32 or ICD10: I48 in any position) through the end of 2016, or 2017 when that data becomes available. The CHARGE-AF score includes atrial flutter as part of the AF definition.

The following variables will be considered. Some clinical variables that were evaluated and not included in the CHARGE-AF score will be re-evaluated in this study as their predictive value may differ in the elderly.

* Variables included in the simple CHARGE-AF model

† Variables included in the augmented model

Clinical variables:

- Age*
- Sex
- Race*
- Height*
- Weight*
- Current cigarette smoking*
- Systolic and diastolic blood pressure*
- Use of antihypertensive medication*
- Diabetes*
- Fasting blood glucose
- Estimated glomerular filtration rate (eGFR) <60 mL/kg/m2
- Total cholesterol
- HDL cholesterol
- Triglycerides
- Alcohol consumption
- Use of lipid-lowering medications
- Heart rate
- History of coronary artery bypass graft (CABG)
- History of heart failure*
- History of myocardial infarction*
- History of stroke

Biomarkers:

- NT-proBNP † (added in Sinner et al, 2014 paper, and used in MESA)[13, 15]
- C-reactive protein † (added in Sinner et al, 2014 paper and used in MESA)[13, 15]
- Troponin

ECG variables

- Left ventricular hypertrophy: gender-specific Cornell voltage criteria (SV3 + RaVL > 2.8mV for men, and >2.2mV for women) †
- QRS duration
- QT interval (Prolonged is assoc. w/ AF)
- Abnormal P-wave axis (<0 or >75)
- P-wave duration > 120 ms
- PR interval (<120; 120-199; > 200) †
- P-wave terminal force > 4000 uV.ms
- Advanced interatrial block

Echo variables

- LA diameter (consider continuous and cut-points: >4.0cm M and > 3.7cm W)
- LA volume index (consider continuous and cut-points: >31mL/m2 M and > 30 mL/m2 W)
- E to E prime lateral ratio (consider continuous and cut-points: >11.5 M and > 13.3 W)
- E/A ratio
- Left atrial global longitudinal strain (may need from Shah/Solomon)
- Left atrial emptying fraction (WILL need from Shah/Solomon)
- LV mass index (consider continuous and cut-points: >45 M and > 41.5 W)
- Ejection fraction (<50%)

Statistical analysis

Performance of CHARGE-AF risk score and simple recalibration
First we will assess the prediction of AF using the simple CHARGE-AF risk score to determine the model performance in an elderly cohort. We will evaluate model performance using the C-statistic,[20] and Nam and D'Agostino's modified Hosmer-Lemeshow chi-square statistic for survival analysis (calibration).[21] We will also re-calibrate the score according to the baseline risk and age of the ARIC visit 5 population and see if that improves the model C-statistic and calibration.

Derive a new predictive model
Next, we will create our own AF risk score in the elderly by deriving a 5-year predictive model. A concern is that the CHARGE-AF risk score could be over-predicting actual AF risk because it doesn't take into account the competing risk of mortality which would bias the results in this older population. Therefore, when we are developing a predictive model we will use methods that allow adjustment for the competing risk of death, similar to those used for developing the FHS 30-year CVD prediction model.[22]

Initially, we will run race-specific, competing-risk Cox proportional hazard models to assess individual predictors of AF after adjustment for age and sex. Hazard ratios (95% CI) will be reported for all of the variables listed above. We will select any variable significantly associated with AF (p<0.05) for inclusion in a multivariable model. We will run a final Cox proportional hazards model including these variables using a bootstrap resampling method for developing predictive models.[23] The final model will also be adjusted for the competing-risk of death. We will test age, sex and race interactions for inclusion in the model.

Once we have our final model, we will develop a risk prediction equation by calculating the model-based individual 5-year risk of AF. Rather than assigning points to each variable, we will use the same methods as the CHARGE-AF paper and develop a continuous score using the beta estimates and reference values from each variable in the final model. Then we will evaluate model performance of this new model using the C-statistic,[20] calibration chi-square, and discrimination slopes.[24]

Similar to the CHARGE-AF methods, we will first develop a simple clinical model, and then develop a more complex model using blood test, ECG, and echo variables. We will 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) using the following risk categories: <2.5%, 2.5 to 5%, > 5%. [24]

Evaluate CHARGE-AF + additional variables

Third, we will try to improve upon the CHARGE-AF risk score by adding any significant variables to the model. We realize it is simpler clinically to add to an accepted model than to totally replace it. Therefore, variables that were significant in the final Cox model step above will be added to the CHARGE-AF score (the base model), and we will evaluate the predictive ability of these additional variables. We will compare the C-statistic, calibration, and the NRI of the base model to the model with CHARGE + additional variables.

Determine the best 5-year AF risk equation

Finally, we will determine the best AF risk equation in an elderly cohort by comparing C-statistics and calibration from our models. We will compare: 1) the original (or re-calibrated) CHARGE score vs. 2) CHARGE + additional variables vs. 3) a new predictive model.

We may be lacking a comparative external cohort sample for validation of our prediction equation results, at which point we will consider internal validation with bootstrapping.[25]

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:
#1453: AF risk score in ARIC – Chamberlain
#1578: CHARGE-AF risk score – Alonso
#1818: Genetic score for AF - Lubitz
#1897: Arterial indices and AF - Chen

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/

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. Wolf, P.A., R.D. Abbott, and W.B. Kannel, *Atrial fibrillation as an independent risk factor for stroke: the Framingham Study.* Stroke, 1991. **22**(8): p. 983-8.
4. Soliman, E.Z., et al., *Atrial Fibrillation and Risk of ST-Segment-Elevation Versus Non-ST-Segment-Elevation Myocardial Infarction: The Atherosclerosis Risk in Communities (ARIC) Study.* Circulation, 2015. **131**(21): p. 1843-50.
5. Miyasaka, Y., et al., *Mortality trends in patients diagnosed with first atrial fibrillation: a 21-year community-based study.* J Am Coll Cardiol, 2007. **49**(9): p. 986-92.
6. January, C.T., et al., *2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on practice guidelines and the Heart Rhythm Society.* Circulation, 2014. **130**(23): p. e199-267.
7. Mozaffarian, D., et al., *Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association.* Circulation, 2016. **133**(4): p. e38-e360.
8. Alonso, A. and F.L. Norby, *Predicting Atrial Fibrillation and Its Complications.* Circ J, 2016. **80**(5): p. 1061-6.
9. Schnabel, R.B., et al., *Development of a risk score for atrial fibrillation (Framingham Heart Study): a community-based cohort study.* Lancet, 2009. **373**(9665): p. 739-45.
10. Chamberlain, A.M., et al., *A clinical risk score for atrial fibrillation in a biracial prospective cohort (from the Atherosclerosis Risk in Communities [ARIC] study).* Am J Cardiol, 2011. **107**(1): p. 85-91.
11. Everett, B.M., et al., *Novel genetic markers improve measures of atrial fibrillation risk prediction.* Eur Heart J, 2013. **34**(29): p. 2243-51.
12. 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.
13. Alonso, A., et al., *Prediction of Atrial Fibrillation in a Racially Diverse Cohort: The Multi-Ethnic Study of Atherosclerosis (MESA).* J Am Heart Assoc, 2016. **5**(2).
14. Li, L., et al., *Association of N-terminal pro B-type natriuretic peptide (NT-proBNP) change with the risk of atrial fibrillation in the ARIC cohort.* Am Heart J, 2018. **204**: p. 119-127.
15. 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.
16. Maheshwari, A., et al., *Refining Prediction of Atrial Fibrillation Risk in the General Population With Analysis of P-Wave Axis (from the Atherosclerosis Risk in Communities Study).* Am J Cardiol, 2017. **120**(11): p. 1980-1984.
17. Magnani, J.W., et al., *P-wave indices and atrial fibrillation: cross-cohort assessments from the Framingham Heart Study (FHS) and Atherosclerosis Risk in Communities (ARIC) study.* Am Heart J, 2015. **169**(1): p. 53-61 e1.
18. Bekwelem, W., et al., *Echocardiographic measures of cardiac structure and function are associated with risk of atrial fibrillation in blacks: the Atherosclerosis Risk in Communities (ARIC) study.* PLoS One, 2014. **9**(10): p. e110111.
19. Olsen, F.J., et al., *Relationship Between Left Atrial Functional Measures and Incident Atrial Fibrillation in the General Population: The Copenhagen City Heart Study.* JACC Cardiovasc Imaging, 2018.
20. Pencina, M.J. and R.B. D'Agostino, *Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation.* Stat Med, 2004. **23**(13): p. 2109-23.
21. D'Agostino R. B., N.B., *Evaluation of the performance of survival analysis models: discrimination and calibration measures.* Amsterdam: Elsevier, 2004. **23**: p. 1-25.
22. Pencina, M.J., et al., *Predicting the 30-year risk of cardiovascular disease: the framingham heart study.* Circulation, 2009. **119**(24): p. 3078-84.
23. Austin, P.C.T., J.V., *Bootstrap Methods for Developing Predictive Models.* The American Statistician, 2004. **58**(2): p. 131-137.

24.   Pencina, M.J., et al., *Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond.* Stat Med, 2008. **27**(2): p. 157-72; discussion 207-12.
25.   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.