ARIC Manuscript Proposal # 1453

PC Reviewed: 11/11/08	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Development of a Risk Score for Predicting Atrial Fibrillation in a Bi-racial Cohort: The Atherosclerosis Risk in Communities (ARIC) Study

b. Abbreviated Title (Length 26 characters): AF Risk Prediction

2. Writing Group:

Writing group members: Alanna Chamberlain, MPH
Sunil K. Agarwal, MD, MPH
Marietta Ambrose, MD
Aaron Folsom, MD
Richard Crow, MD
Elsayed Z. Soliman, MD, MSc, MS
Lloyd E. Chambless, PhD
Alvaro Alonso, MD, PhD
Others welcome

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

First author: Alanna Chamberlain
Address: 1300 S. 2nd Street, Suite 300
Minneapolis, MN 55454

Phone: 612-625-5352 Fax: 612-624-0315
E-mail: gram0119@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: Alvaro Alonso
Address: 1300 S. 2nd Street, Suite 300
Minneapolis, MN 55454

Phone: 612-626-8597 Fax: 612-624-0315
E-mail: alonso@umn.edu

3. Timeline:

Statistical Analysis:	December 2008 – March 2009
Manuscript Preparation:	March 2009
Manuscript Revision:	April 2009
Manuscript Submission:	May 2009

4. Rationale:

Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia in clinical practice, and is associated with increased stroke and cardiovascular morbidity and mortality[1]. AF currently affects more than 2.2 million Americans, and the lifetime risk for development of AF in men and women over 40 years of age is 1 in 4[2]. AF is an important risk factor for incident stroke[3-5] and heart failure[4, 6], and also carries poor prognosis in heart failure patients[7].

Risk factors for AF include increasing age[8, 9], male gender[8, 10], obesity[11], hypertension[12, 13], diabetes[14, 15], metabolic syndrome[16], obstructive sleep apnea[17, 18], chronic obstructive pulmonary disease[19, 20], and heavy alcohol intake[21-23]. Additionally, recent cardiac surgery[19, 20] and structural abnormalities, such as left ventricular hypertrophy, left atrial enlargement[24], and restrictive left ventricular diastolic filling[25], increase the risk of incident AF.

Although some of the previously mentioned risk factors have been well studied in relation to incident AF on a population level, information on a person's individual risk of developing AF is lacking. This information may aid in risk stratification and in the selection of appropriate candidates for preventive therapies. Risk scores have been successfully developed for coronary heart disease[26], stroke[27], and heart failure[28]. Additionally, a risk score has been developed for predicting stroke or death with new-onset AF[29]; however, a risk score for the prediction of AF has not been developed. Therefore, we propose to develop a risk score for predicting AF incidence using the Atherosclerosis Risk in Communities study for both black and white men and women.

5. Main Hypothesis/Study Questions:

We aim to determine a risk score for atrial fibrillation based on a combination of known risk factors that maximizes discrimination, i.e., the ability to predict who will have an incident event and who will not using Cox regression and receiver operating characteristic (ROC) curves.

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).

Individuals without ECG or who had diagnosed AF or atrial flutter at baseline will be excluded from analyses. We will additionally exclude those who did not fast for at least 8 hours before the baseline exam. We will censor individuals who develop atrial

flutter without AF at their atrial flutter diagnosis date. The independent variables in this analysis include, but are not limited to, age, sex, race, education, income, BMI/obesity, waist circumference, height, systolic and diastolic blood pressure, pulse pressure, hypertension treatment, total cholesterol, LDL cholesterol, HDL cholesterol, triglycerides, lipid lowering medication, diabetes, smoking, alcohol intake, physical activity, myocardial infarction, heart failure, and ECG-based left ventricular hypertrophy. The dependent variable is incident AF. Incident cases of atrial fibrillation will be identified through hospital discharge codes (ICD-9 code 427.31), death certificates (underlying cause of death ICD-9 code I48 or 427.3), and ECG's performed during follow-up visits. Individuals who develop both atrial flutter and AF during follow-up will be considered as having an event and will be censored at the earliest date of either AF or atrial flutter.

We will follow the methods described by Sullivan, et al[30] to determine a risk score for AF. The 10-year risk of developing AF will be estimated, and we may also consider estimating the 5- and 15-year risks. First, we will stratify each independent variable into categories and run Cox regression analyses to obtain the regression coefficients for the categories of each variable. Second, we will determine the reference values for each category. We will use the midpoint value of each category as the reference, except for categories that may have extreme outliers. For example, if $\geq$160 is the highest category for systolic blood pressure (SBP), we will use the midpoint between 160 and the 99th percentile of the observed SBP as the reference. Third, we will pick one category for each risk factor as the referent or "base" category. This category will be assigned 0 points and the other categories for that variable will be assigned positive numbers if they reflect a less healthy state and negative numbers if they reflect a more healthy state than the base category. Fourth, we will calculate how far each category is from the base category in regression units by the following equation: $\beta_i(W_{ij} - W_{iREF})$, where β_i is the regression coefficient, W_{ij} is the reference value of each category, and W_{iREF} is the base category for a particular variable. Fifth, a constant will be set corresponding to the number of regression units that equate to 1 point. For example, we would multiply the regression coefficient for a continuous age variable by 5 if the age categories were set up in 5-year increments. Sixth, we will divide the $\beta_i(W_{ij} - W_{iREF})$ for each category by the constant to determine the points associated with each category. Finally, we will determine the risk of developing AF associated with each possible point total by multiplying the constant by each point total (if we use the continuous age regression coefficient to determine the constant, we will also need to add this back in our calculations).

Once the risk score for AF is completed, we will use ROC curves and estimate the area under the curve (AUC) to assess the performance of our risk score. To conduct the ROC curve analysis, we will use the methodology developed by Chambless and Diao, which takes into account the presence of censoring in survival data[31]. For our risk score determination, we will consider each of the above mentioned independent variables. However, we may find that some of these variables do not contribute to the risk score (by not increasing the AUC compared to a risk score without that variable), and these will be excluded when appropriate. By comparing ROC curves, we will determine whether the risk scores should be computed separately by sex and/or race, or if a pooled risk score is as predictable as these separate risk scores. We will use bootstrapping to determine a

confidence interval around our ROC curve, taking at least 1000 samples with replacement, for internal validation of our risk score. We will additionally compare our risk score to the Cox regression equation to see if our risk score predicts AF as well as plugging in specific values for each variable into a regression equation. Finally, we will use the ARIC and calibrated Framingham CHD risk scores to predict AF, in order to determine whether or not a separate risk score for AF better predicts incident events than a validated CHD risk score.

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

__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)?

MS # 1351: Alonso, et al. Incidence of atrial fibrillation in a bi-racial cohort: the ARIC study

MS # 1156: Soliman, et al. Ethnic distribution of electrocardiographic predictors of atrial fibrillation and its impact on understanding the ethnic distribution of ischemic stroke in the ARIC Study

MS # 1389: Chamberlain, et al. Metabolic Syndrome and Risk of Incident Atrial Fibrillation among Whites and Blacks 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? __X__ Yes ____ No

11.b. If yes, is the proposal

_X__ A. primarily the result of an ancillary study (list number* __2008.09_)

___ 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.

References:

1. Benjamin EJ, Wolf PA, D'Agostino RB, Silbershatz H, Kannel WB, Levy D. Impact of atrial fibrillation on the risk of death: The framingham heart study. *Circulation*. 1998;98:946-952.

2. Rosamond W, Flegal K, Furie K, et al. Heart disease and stroke statistics--2008 update: A report from the american heart association statistics committee and stroke statistics subcommittee. *Circulation*. 2008;117:e25-146.

3. Miyasaka Y, Barnes ME, Gersh BJ, et al. Time trends of ischemic stroke incidence and mortality in patients diagnosed with first atrial fibrillation in 1980 to 2000: Report of a community-based study. *Stroke*. 2005;36:2362-2366.

4. Stewart S, Hart CL, Hole DJ, McMurray JJ. A population-based study of the long-term risks associated with atrial fibrillation: 20-year follow-up of the Renfrew/Paisley study. *Am J Med*. 2002;113:359-364.

5. Wolf P, Abbott R, Kannel W. Atrial fibrillation as an independent risk factor for stroke: The framingham study. *Stroke*. 1991;22:983-988.

6. Miyasaka Y, Barnes ME, Gersh BJ, et al. Incidence and mortality risk of congestive heart failure in atrial fibrillation patients: A community-based study over two decades. *Eur Heart J*. 2006;27:936-941.

7. Ahmed A, Perry GJ. Incident atrial fibrillation and mortality in older adults with heart failure. *European Journal of Heart Failure,*. 2005;7:1118-1121.

8. Go AS, Hylek EM, Phillips KA, 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;285:2370-2375.

9. Psaty BM, Manolio TAMHS, Kuller LHDH, et al. Incidence of and risk factors for atrial fibrillation in older adults. *Circulation*. 1997;96:2455-2461.

10. Friberg J, Scharling H, Gadsboll N, Jensen GB. Sex-specific increase in the prevalence of atrial fibrillation (the copenhagen city heart study). *Am J Cardiol*. 2003;92:1419-1423.

11. Wanahita N, Messerli FH, Bangalore S, Gami AS, Somers VK, Steinberg JS. Atrial fibrillation and obesity--results of a meta-analysis. *Am Heart J*. 2008;155:310-315.

12. Kannel WB, Abbott RD, Savage DD, McNamara PM. Epidemiologic features of chronic atrial fibrillation: The framingham study. *N Engl J Med*. 1982;306:1018-1022.

13. Benjamin EJ, Levy D, Vaziri SM, D'Agostino RB, Belanger AJ, Wolf PA. Independent risk factors for atrial fibrillation in a population-based cohort. the framingham heart study. *JAMA*. 1994;271:840-844.

14. Movahed M, Hashemzadeh M, Mazen Jamal M. Diabetes mellitus is a strong, independent risk for atrial fibrillation and flutter in addition to other cardiovascular disease. *International Journal of Cardiology,*. 2005;105:315-318.

15. Aksnes TA, Schmieder RE, Kjeldsen SE, Ghani S, Hua TA, Julius S. Impact of new-onset diabetes mellitus on development of atrial fibrillation and heart failure in high-risk hypertension (from the VALUE trial). *The American Journal of Cardiology*. 2008;101:634-638.

16. Watanabe H, Tanabe N, Watanabe T, et al. Metabolic syndrome and risk of development of atrial fibrillation: The niigata preventive medicine study. *Circulation*. 2008;117:1255-1260.

17. Gami AS, Pressman G, Caples SM, et al. Association of atrial fibrillation and obstructive sleep apnea. *Circulation*. 2004;110:364-367.

18. Gami AS, Hodge DO, Herges RM, et al. Obstructive sleep apnea, obesity, and the risk of incident atrial fibrillation. *J Am Coll Cardiol*. 2007;49:565-571.

19. Mathew JP, Fontes ML, Tudor IC, et al. A multicenter risk index for atrial fibrillation after cardiac surgery. *JAMA*. 2004;291:1720-1729.

20. Buch P, Friberg J, Scharling H, Lange P, Prescott E. Reduced lung function and risk of atrial fibrillation in the copenhagen city heart study. *Eur Respir J*. 2003;21:1012-1016.

21. Frost L, Vestergaard P. Alcohol and risk of atrial fibrillation or flutter: A cohort study. *Arch Intern Med*. 2004;164:1993-1998.

22. Djoussé L, Levy D, Benjamin EJ, et al. Long-term alcohol consumption and the risk of atrial fibrillation in the framingham study. *The American Journal of Cardiology,*. 2004;93:710-713.

23. Mukamal KJ, Tolstrup JS, Friberg J, Jensen G, Gronbaek M. Alcohol consumption and risk of atrial fibrillation in men and women: The copenhagen city heart study. *Circulation*. 2005;112:1736-1742.

24. Vaziri SM, Larson MG, Benjamin EJ, Levy D. Echocardiographic predictors of nonrheumatic atrial fibrillation. the framingham heart study. *Circulation*. 1994;89:724-730.

25. Tsang TS, Gersh BJ, Appleton CP, et al. Left ventricular diastolic dysfunction as a predictor of the first diagnosed nonvalvular atrial fibrillation in 840 elderly men and women. *J Am Coll Cardiol*. 2002;40:1636-1644.

26. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97:1837-1847.

27. Wolf PA, D'Agostino RB, Belanger AJ, Kannel WB. Probability of stroke: A risk profile from the framingham study. *Stroke*. 1991;22:312-318.

28. Kannel WB, D'Agostino RB, Silbershatz H, Belanger AJ, Wilson PWF, Levy D. Profile for estimating risk of heart failure. *Arch Intern Med*. 1999;159:1197-1204.

29. Wang TJ, Massaro JM, Levy D, et al. A risk score for predicting stroke or death in individuals with new-onset atrial fibrillation in the community: The framingham heart study. *JAMA*. 2003;290:1049-1056.

30. Sullivan LM, Massaro JM, D'Agostino RB,Sr. Presentation of multivariate data for clinical use: The framingham study risk score functions. *Stat Med*. 2004;23:1631-1660.

31. Chambless LE, Diao G. Estimation of time-dependent area under the ROC curve for long-term risk prediction. *Stat Med*. 2006;25:3474-3486.