ARIC Manuscript Proposal #3389

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

1.a. Full Title:
Proteomic Profiling and Heart Failure Risk in the Atherosclerosis Risk in Communities (ARIC) Study

b. Abbreviated Title (Length 26 characters):
HF and Proteomics

2. Writing Group:
Writing group members:
Bing Yu, Brian Claggett, Adrienne Tin, Ryan Demmer, Ron Hoogeveen, Kuni Matsushita, Elizabeth Selvin, Christie Ballantyne, Gerardo Heiss, Eric Boerwinkle, Joe Coresh and Amil Shah. Others are welcome.

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: 1200 Pressler Street, RAS E-605

Phone: 713-500-9285 Fax:
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: Amil Shah
Address: 75 Francis Street, Boston, MA 02115

Phone: 617-525-9678 Fax:
E-mail: ashah11@rics.bwh.harvard.edu

3. Timeline:
The data collection of proteomics and incident heart failure are already accomplished with baseline at visit 5. When the proposal is approved, data analysis process will start. The manuscript will be prepared when data analysis is done (~ 3-6 months).

4. Rationale:

Heart failure (HF) is common, increasing in prevalence, leads to significant resource utilization, and disproportionately affects the elderly and African Americans.[1] The size of the elderly population is growing, and the large majority of elderly persons have HF risk factors.[2] HF with preserved left ventricular ejection fraction (LVEF; HFpEF), which comprises half of HF cases,[3, 4] accounts for up to 80% of prevalent HF in the elderly.[5] This heterogeneous syndrome is associated with similar impairment in quality of life and rates of hospitalization as HF with reduced LVEF (HFrEF), greater mortality than similarly aged persons without HF, but no efficacious disease specific therapy.[6-9]

Although a number of HF biomarkers have been identified to provide information on diagnosis and prognostication,[10] little is known regarding the biologic pathways underlying HF in the elderly, precluding advances in preventive and therapeutic sciences targeting HFpEF. Prevention of HF is essential given the high prevalence of at risk elderly persons that is growing in size, the heightened morbidity, mortality, and resource utilization once overt HF develops, and the lack of efficacious therapies for the majority of HF in the elderly.

Untargeted proteomics has effectively identified novel pathways relevant to myocardial infarction,[11, 12] and holds promise to similarly identify novel molecular pathways relevant to HFpEF. A recent study has identified an eigenprotein that is positively associated with prevalent and incident HF in Caucasian population.[13] To date, no large studies have employed targeted proteomics to identify molecular pathways predictive of HF and its phenotypes in a biracial population. We propose to evaluate the association between the proteome and HF risk in ARIC African and European Americans to identify novel modifiable biomarkers to improve the understanding of HF pathophysiology and risk prediction.

5. Main Hypothesis/Study Questions:

1. Identify individual circulating proteins and protein networks that associate with prevalent HF and HF stages;

2. Identify individual circulating proteins and protein networks that predict incident HF and HF phenotype (HFpEF vs HFrEF).

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:

For the current analysis, we will focus on ARIC visit 5, consisting of ~5,000 ARIC participants with plasma proteomes and incident HF information followed up to 2017.

We would like to extend our analysis to visits 2 and 3 when proteomics data at those visits are available to examine the protein changes in relation to HF and the difference of HF protein signatures at mid and old ages. This will likely involve new proposals if separate papers are planned.

Exclusion criteria:
Participants will be excluded if they are missing proteome data, HF information or key covariates. Participants with prevalent HF at visit 5 will also be excluded from the incident HF analysis.

Variables:
Outcome variables:

1) Prevalent HF and HF stages at visit 5:
Prevalent HF at Visit 5 will be defined based on adjudicated HF hospitalizations (A through C) since 2005, hospitalizations with HF ICD codes prior to 2005, and HF self-report as codified in the existing ARIC Variable 'v5_prevhf52'.

HF stages are defined using ACC/AHA guideline based on the presence of clinical HF risk factors, cardiac structural and functional abnormalities based on the Visit 5 echocardiogram, and the presence of prevalent HF at Visit 5.[2]

Prevalent HF cases with quantifiable LVEF by echocardiography will be categorized as HFpEF (LVEF ≥ 50%) or HFrEF (LVEF < 50%) based on the Visit 5 echocardiogram.

2) Incident HF followed by 2017:
Incident HF will be defined based on ARIC committee adjudicated HF events post-Visit 5. Incident HFpEF and HFrEF will be defined using LVEF data abstracted from adjudicated hospitalizations.

Exposure variables:
Proteins measured by SOMAscan assay (Somalogic Inc., Boulder, CO). We will investigate the appropriate transformation of the relative florescence units (RFU) of protein levels, such as log 2 and inverse normal transformation. Post-analysis filtering may be applied based on protein calibration factor.

Covariates:
Age, gender, study center, race, body mass index (BMI), smoking, alcohol, systolic blood pressure (SBP), blood pressure lowering medication, estimated glomerular filtration rate (eGFR), prevalent coronary heart disease (CHD), prevalent stroke, prevalent atrial fibrillation, prevalent diabetes, and prevalent chronic kidney disease.

Statistical data analysis:

We will run logistic and Cox proportional hazards regressions for prevalent and incident HF respectively using several models to account for potential confounding factors. Model 1 will be adjusted for age, sex, and race-center. Model 2 will additionally adjust for BMI, SBP and blood pressure lowering medication. Model 3 will further adjust for smoking, alcohol, eGFR, prevalent diabetes and CHD. Model 4 will additionally adjust for prevalent stroke and atrial fibrillation. Model 5 will additionally adjust for NT-pro-BNP and hs-TnT and hs-TnI to detect which protein

signals add information beyond these important targeted assays. The Benjamini–Hochberg procedure (BH step-up procedure) will be used to control false discovery rate (FDR) at level of 0.05 to correct for multiple testing. For incident HFpEF, participants developing HFrEF or HF with unknown LVEF will be censored at the time of the event. For incident HFrEF participants developing HFpEF or HF with unknown LVEF will be censored at the time of the event. We will compare these proteins to identified biomarkers (i.e. C-reactive protein, Troponin T and N-terminal pro-B-type natriuretic peptide) that predict heart failure in the ARIC.[14, 15] We will also explore 1) regularized regression methods (i.e. elastic net with Cox regression[16]) incorporating 10-fold cross validation for parameter tuning and using the one-standard error rule to select a most parsimonious set of protein predictors; 2) statistical measure of discrimination (i.e. area under the receiver operating characteristic curve (AUC)) using 10-fold cross validation to report an average cross-validated C-statistics; [17, 18] and 3) Weighted Gene Co-expression Network Analysis[19] to identify protein modules that relate to HF phenotypes. We will seek external studies (i.e. HUNT3 and/or MESA) for validation to minimize potential false positive findings originated from the large number of models tested.

We already know, NT-pro-BNP and troponin elevations are strong risk factors for HF. Therefore, we will also explore the proteins correlated with these targeted assays to try to understand coordinated pathways, which are altered when these known protein risk factors are altered. Exploratory analysis will also consider discriminant analysis to try to understand the extent to which clusters of proteins can distinguish endophenotypes within HF, such as HFpEF vs. HFrEF but other subgroups with different protein profiles and outcomes may exist as well.

Interpretation of the statistical associations is enhanced by additional information on the biology of the disease and candidate proteins. We will explore protein databases related to their pathophysiology and coordinated regulation and relationship of blood to tissue expression to enhance the interpretation of statistical signals. Finally, we will consider the use of Mendelian Randomization to test if candidate protein associations are likely to be causal.

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/mantrack/maintain/search/dtSearch.html

___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)?

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* __AS2017.27__)

___ 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 https://www2.cscc.unc.edu/aric/approved-ancillary-studies

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. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Jr., Drazner MH, Fonarow GC, Geraci SA, Horwich T, Januzzi JL, Johnson MR, Kasper EK, Levy WC, Masoudi FA, McBride PE, McMurray JJ, Mitchell JE, Peterson PN, Riegel B, Sam F, Stevenson LW, Tang WH, Tsai EJ and Wilkoff BL. 2013 ACCF/AHA guideline for the management of heart failure: executive summary: a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. *Circulation*. 2013;128:1810-52.
2. Shah AM, Claggett B, Loehr LR, Chang PP, Matsushita K, Kitzman D, Konety S, Kucharska-Newton A, Sueta CA, Mosley TH, Wright JD, Coresh J, Heiss G, Folsom AR and Solomon SD. Heart Failure Stages Among Older Adults in the Community: The Atherosclerosis Risk in Communities Study. *Circulation*. 2017;135:224-240.
3. Bhatia RS, Tu JV, Lee DS, Austin PC, Fang J, Haouzi A, Gong Y and Liu PP. Outcome of heart failure with preserved ejection fraction in a population-based study. *N Engl J Med*. 2006;355:260-9.
4. Owan TE, Hodge DO, Herges RM, Jacobsen SJ, Roger VL and Redfield MM. Trends in prevalence and outcome of heart failure with preserved ejection fraction. *N Engl J Med*. 2006;355:251-9.

5. Kitzman DW, Gardin JM, Gottdiener JS, Arnold A, Boineau R, Aurigemma G, Marino EK, Lyles M, Cushman M, Enright PL and Cardiovascular Health Study Research G. Importance of heart failure with preserved systolic function in patients > or = 65 years of age. CHS Research Group. Cardiovascular Health Study. *Am J Cardiol*. 2001;87:413-9.
6. Pitt B, Pfeffer MA, Assmann SF, Boineau R, Anand IS, Claggett B, Clausell N, Desai AS, Diaz R, Fleg JL, Gordeev I, Harty B, Heitner JF, Kenwood CT, Lewis EF, O'Meara E, Probstfield JL, Shaburishvili T, Shah SJ, Solomon SD, Sweitzer NK, Yang S, McKinlay SM and Investigators T. Spironolactone for heart failure with preserved ejection fraction. *N Engl J Med*. 2014;370:1383-92.
7. Massie BM, Carson PE, McMurray JJ, Komajda M, McKelvie R, Zile MR, Anderson S, Donovan M, Iverson E, Staiger C, Ptaszynska A and Investigators IP. Irbesartan in patients with heart failure and preserved ejection fraction. *N Engl J Med*. 2008;359:2456-67.
8. Cleland JG, Tendera M, Adamus J, Freemantle N, Polonski L, Taylor J and Investigators P-C. The perindopril in elderly people with chronic heart failure (PEP-CHF) study. *Eur Heart J*. 2006;27:2338-45.
9. Yusuf S, Pfeffer MA, Swedberg K, Granger CB, Held P, McMurray JJ, Michelson EL, Olofsson B, Ostergren J, Investigators C and Committees. Effects of candesartan in patients with chronic heart failure and preserved left-ventricular ejection fraction: the CHARM-Preserved Trial. *Lancet*. 2003;362:777-81.
10. Chow SL, Maisel AS, Anand I, Bozkurt B, de Boer RA, Felker GM, Fonarow GC, Greenberg B, Januzzi JL, Jr., Kiernan MS, Liu PP, Wang TJ, Yancy CW, Zile MR, American Heart Association Clinical Pharmacology Committee of the Council on Clinical C, Council on Basic Cardiovascular S, Council on Cardiovascular Disease in the Y, Council on C, Stroke N, Council on Cardiopulmonary CCP, Resuscitation, Council on E, Prevention, Council on Functional G, Translational B, Council on Quality of C and Outcomes R. Role of Biomarkers for the Prevention, Assessment, and Management of Heart Failure: A Scientific Statement From the American Heart Association. *Circulation*. 2017;135:e1054-e1091.
11. Jacob J, Ngo D, Finkel N, Pitts R, Gleim S, Benson MD, Keyes MJ, Farrell LA, Morgan T, Jennings LL and Gerszten RE. Application of Large-Scale Aptamer-Based Proteomic Profiling to Planned Myocardial Infarctions. *Circulation*. 2018;137:1270-1277.
12. Ngo D, Sinha S, Shen D, Kuhn EW, Keyes MJ, Shi X, Benson MD, O'Sullivan JF, Keshishian H, Farrell LA, Fifer MA, Vasan RS, Sabatine MS, Larson MG, Carr SA, Wang TJ and Gerszten RE. Aptamer-Based Proteomic Profiling Reveals Novel Candidate Biomarkers and Pathways in Cardiovascular Disease. *Circulation*. 2016;134:270-85.
13. Emilsson V, Ilkov M, Lamb JR, Finkel N, Gudmundsson EF, Pitts R, Hoover H, Gudmundsdottir V, Horman SR, Aspelund T, Shu L, Trifonov V, Sigurdsson S, Manolescu A, Zhu J, Olafsson O, Jakobsdottir J, Lesley SA, To J, Zhang J, Harris TB, Launer LJ, Zhang B, Eiriksdottir G, Yang X, Orth AP, Jennings LL and Gudnason V. Co-regulatory networks of human serum proteins link genetics to disease. *Science*. 2018;361:769-773.
14. Nambi V, Liu X, Chambless LE, de Lemos JA, Virani SS, Agarwal S, Boerwinkle E, Hoogeveen RC, Aguilar D, Astor BC, Srinivas PR, Deswal A, Mosley TH, Coresh J, Folsom AR, Heiss G and Ballantyne CM. Troponin T and N-terminal pro-B-type natriuretic peptide: a biomarker approach to predict heart failure risk--the atherosclerosis risk in communities study. *Clin Chem*. 2013;59:1802-10.
15. Bekwelem W, Lutsey PL, Loehr LR, Agarwal SK, Astor BC, Guild C, Ballantyne CM and Folsom AR. White blood cell count, C-reactive protein, and incident heart failure in the Atherosclerosis Risk in Communities (ARIC) Study. *Ann Epidemiol*. 2011;21:739-48.
16. Simon N, Friedman JH, Hastie T and Tibshirani R. Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. *2011*. 2011;39:13.
17. Taylor JMG, Ankerst DP and Andridge RR. Validation of Biomarker-based risk prediction models. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2008;14:5977-5983.
18. Uno H, Cai T, Pencina MJ, D'Agostino RB and Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med*. 2011;30:1105-17.
19. Langfelder P and Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.