ARIC Manuscript Proposal #4149

PC Reviewed: 11/8/22 **Status: _____** **Priority: 2**
SC Reviewed: _____ **Status: _____** **Priority: ____**

1.a. Full Title: ***Validation of Proteomic Risk Factors for Heart Failure Among Participants of the Chronic Renal Insufficiency Cohort***

b. Abbreviated Title (Length 26 characters): **Validation CRIC HF Proteomics**

2. Writing Group [please provide a middle name if available; EX: Adam Lee Williams]:

Writing group members:
Ruth Dubin, MD
Peter Ganz, MD
Amil Shah, MD
Bing Yu, MD
Vicky Arthur, PhD

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

First author [please provide a middle name if available; EX: Adam Lee Williams]:
Address: 5323 Harry Hines Blvd
Dallas TX 75390
Phone: (510) 847 4955 Fax:
E-mail: ruth.dubin@utsouthwestern.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, MD**
E-mail: ashah11@rics.bwh.harvard.edu

3. Timeline: Analyses should be complete by January 2023

4. Rationale:

We have assayed all Chronic Renal Insufficiency Cohort(CRIC) participant samples from Year 1 for proteomics on the SOMAscan platform. Using machine learning, we have developed proteomic risk prediction model for incident heart failure. In tandem, we have also devised novel clinical models for incident heart failure composed of clinical risk factors, as well as hybrid models consisting of proteins and clinical factors. We have identified high-risk and low-risk protein markers and used Mendelian randomization potential causal mediators.

The same SOMAscan assay has been run on plasma samples from Atherosclerosis Risk in Communities(ARIC) participants, and we plan to request validation of our results in CRIC. We have requested that protein models, novel clinical models, and select individual proteins be replicated in ARIC among 1160 participants

at Visit 5 with estimated glomerular filtration rate (eGFR) < 60ml/min/1.73m^2. For separate outcomes of incident heart failure, heart failure with reduced ejection fraction (HFrEF), and heart failure with preserved ejection fraction (HFpEF), there are three risk models: 1) novel clinical model, 2) novel protein model and 3) novel hybrid clinical – protein model. For each outcome there are 30 individual proteins to replicate: 20 proteins of high risk, 10 proteins of low risk.

Results of the replication analyses in ARIC will be integrated into the main CRIC manuscript.

5. Main Hypothesis/Study Questions:

Although it is well-known that incident heart failure is common and associated with significant morbidity and mortality among patients with CKD, we lack methods of risk stratification for incident heart failure in the CKD population. In the United States Renal Data System, prevalence of heart failure is about 4 x higher in participants with CKD, compared to those without CKD. Among those with CKD stage 3, prevalence of heart failure is 28%, and among those with CKD stages 4-5, prevalence is 41%. In comparison, heart failure prevalence is about 6% among individuals without CKD. Among those with CKD, roughly equal proportions of cases are HFpEF vs HFrEF.[1] Among CRIC participants, heart failure is about twice as common as myocardial infarction.[2] In a recent CRIC manuscript by Bansal et al, over 8 years of follow up, there are 477 incident heart failure events. For 356 of these heart failure events, ejection fraction is available either during the hospitalization for the event or within 1 year of hospitalization. Out of these 356 incident heart failure events, 200 are HFpEF, 156 are HFrEF.[3] In our current analyses of proteomic predictors of heart failure, censoring outcomes at ESRD, we have 14 years of follow-up, including 390 heart failure events. Among these events, 165 identified as HFrEF and 137 HFpEF.

The development of the SOMAscan proteomics assay affords the opportunity to screen nearly 5000 soluble plasma proteins in search of novel, potentially modifiable risk factors for CKD progression. We have conducted a longitudinal analysis of incident heart failure, HFrEF, and HFpEF, in CRIC. We utilized large-scale proteomics performed at Year 1 and applied machine learning to devise proteomics-based risk models for incident heart failure. We compared the prognostic utility of these proteomic models to clinical models for heart failure and examine individual proteins for associations with the three heart failure outcomes. For this proposal, we outline a plan to validate individual proteins, novel clinical models, and novel proteomic models from our CRIC study, among participants among 1160 ARIC participants at Visit 5 with estimated glomerular filtration rate (eGFR) < 60ml/min/1.73m^2.

Hypotheses:

1. Novel protein based risk models for incident heart failure, HFrEF and HFpEF will have better discrimination than refit pooled cohort equations.
2. Hybrid clinical-protein models will have better discrimination than clinical models.
3. Novel protein, clinical and hybrid clinical-protein models will replicate in ARIC.
4. Protein ‘top-hits’ found in CRIC will replicate in ARIC at statistical significance corrected for 30 multiple tests for each heart failure outcome.

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).

Risk model factors and coefficients will be supplied to the statistician in R code. C-statistics (95%CI) and calibration plots will be created for each risk model. Individual proteins will be evaluated for HR (95%CI) with the respective heart failure outcome, in Cox regression adjusted for eGFR. Sample size for individual protein associations is show below, considering α=0.05, Bonferroni corrected α=0.05/5000 = 0.00001. We will likely report statistical significance as false discovery rate of 0.05, for which power lies between that power calculated for α=0.05 and α=0.00001. Events are totaled over median (IQR) 6.96 [5.04, 7.73] years. C-statistics (95%CI) and calibration plots will be created for each risk model (three models for each HF outcome, including protein model, clinical model, and hybrid model for HF, HFpEF, HFrEF). Thirty individual proteins for each outcome will be evaluated for

HR (95%CI) with the respective heart failure outcome, in Cox regression adjusted for eGFR. Statistical criterion for a protein to validate in ARIC will be a HR in the same direction as in CRIC, with significance < 0.05/30.

	N ARIC participants	N events	Minimum HR per SD of protein assuming ***α=0.05***	***Minimum HR per SD of protein assuming α=0.00001***
Heart Failure	1160	149	1.27	1.56
HFpEF	1160	76	1.38	1.83
HFrEF	1160	52	1.47	2.1

7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? ____ Yes __X__ No

b. If Yes, is the author aware that the current derived consent file ICTDER05 must be used to exclude persons with a value RES_OTH and/or RES_DNA = "ARIC only" and/or "Not for Profit" ? ____ Yes ____ No
(The 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 current derived consent file ICTDER05 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? ____ 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 https://sites.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. United States Renal Data System, Vol. 1. Ch. 4. 2020.
2. Scialla JJ, Xie H, Rahman M, Anderson AH, Isakova T, Ojo A, Zhang X, Nessel L, Hamano T, Grunwald JE, Raj DS, Yang W, He J, Lash JP, Go AS, Kusek JW, Feldman H, Wolf M and Chronic Renal Insufficiency Cohort Study I. Fibroblast growth factor-23 and cardiovascular events in CKD. *J Am Soc Nephrol*. 2014;25:349-60.
3. Bansal N, Zelnick L, Go A, Anderson A, Christenson R, Deo R, Defilippi C, Lash J, He J, Ky B, Seliger S, Soliman E, Shlipak M, dagger CSI and Investigatorsdagger CS. Cardiac Biomarkers and Risk of Incident Heart Failure in Chronic Kidney Disease: The CRIC (Chronic Renal Insufficiency Cohort) Study. *J Am Heart Assoc*. 2019;8:e012336.