ARIC Manuscript Proposal #3989

PC Reviewed: 1/11/22 **Status: _____** **Priority: 2**
SC Reviewed: _____ **Status: _____** **Priority: ____**

1.a. Full Title: Harnessing the plasma proteome to predict mortality in heart failure subpopulations

b. Abbreviated Title (Length 26 characters): Novel HF prognostic models

2. Writing Group:

Writing group members:

SomaLogic investigators: Jessica Chadwick, Rachel Ostroff, Kelsey Loupy, Michael Hinterberg, Laura Sampson, Natasha Kureshi, Hannah Biegel, Missy Simpson

Study Collaborators: David Lanfear (WSU-SoM/Henry Ford Hospital), Isabella Kardys (Erasmus Medical Center), Nancy Sweitzer (University of Arizona), Peter Ganz (UCSF), Lei Zhao (BMS and Penn Heart Failure Study, may include other Penn Heart Failure investigators).
ARIC Collaborators: Josef Coresh, *others welcome*

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

First author: Jessica Chadwick
Address: 2995 Wilderness Place
Boulder CO 80031

Phone: 760-505-3403 Fax:
E-mail: jchadwick@somalogic.com

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: Josef Coresh
Address: Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health and Welch Center for Prevention, Epidemiology, and Clinical Research,
2024 E. Monument St., Suite 2-600,
Baltimore, MD 21287

Phone: 410-955-0495 Fax:
E-mail: coresh@jhu.edu

3. Timeline: We anticipate that the manuscript will be drafted and submitted to the ARIC Publications Committee within 6 months of the manuscript proposal being approved.

4. Rationale:

Limitations of individual marker measurements and clinical risk scores, currently used in heart failure (HF) prognosis have led an increased interest (within the medical community) in pursuing precision medicine initiatives that seek to accurately stratify HF patients based on risk and aid in personalized treatment decisions (Feldman, 2017; Cresci, 2019). Precision medicine may direct the allocation of resources and improve clinical outcomes, thereby reducing unnecessary workloads placed on health systems and the economic impact on patients and their insurance providers. In driving precision medicine initiatives forward, there is an unmet need to support physicians' clinical decision-making through the creation and utilization of sensitive, unbiased, and personalized prognostic tests that are simple, cost effective, and can be monitored over time.

To that end, we leveraged the SomaScan (v. 4.0) platform to measure >5,000 plasma proteins in over 2,500 persons across three different studies (including ARIC visit 5) in heart failure with reduced ejection fraction (HFrEF) and heart failure with preserved ejection fraction (HFpEF) populations; employing a machine learning approach, we built and validated individual mathematical models predicting six-month and one-year risk of all-cause death based solely on the proteomic signatures of patients with HFrEF or HFpEF, respectively. The proposed manuscript will describe the development and validation of these models (which is completed), as well as assess the performance of these models against current standard of care biomarker NT-proBNP. We will also determine whether the HFrEF model predicts increased risk in longitudinal assessments as a fatal or non-fatal event approach. Our study will provide a novel step forward in precision medicine, demonstrating that proteomic models discovered and validated with the SomaScan platform can accurately predict HF mortality.

5. Main Hypothesis/Study Questions:

We hypothesize:

A. The plasma proteome (as measured by SomaScan) can be used to identify proteins that predict risk of all-cause death in different heart failure populations and that these proteins are dynamic, exhibiting substantial differences in relative abundance as the risk of death or non-fatal events increases.
B. Proteomic risk models for 6-month and 1-year prediction of all-cause mortality in HFrEF and HFpEF patients will be superior to current standard-of-care biomarker NTProBNP.
C. The HFrEF risk will be higher and increase more over time in participants that go on to have an event (all-cause death) vs. those that do not have an event.

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).

The proposed manuscript will describe the development and validation our the proteomic HFrEF and HFpEF models, which is complete. We are currently comparing model performance to current standard of care biomarker: NT-proBNP. We are also applying the HFrEF model to a

new longitudinal observational cohort (Bio-SHiFT) to determine whether the model predicts increased risk in longitudinal assessments as a fatal or non-fatal event approach.

Plasma proteins were measured using the SomaScan assay, version 4, which measures approximately 5000 human proteins (Williams 2019). The Penn Heart Failure Study (PHFS) cohort, ARIC visit 5 cohort and Henry Ford Heart Failure (HFHF) cohort were split as shown in the table below for model training/verification and validation. The three cohorts included all participants with diagnosed HFpEF or HFrEF respectively.

HFrEF Model			
Dataset	Data Use (%)		
	Training	Verification	Validation
PHFS	80%	0%	20%
ARIC visit 5	0%	100%	0%
HFHF	0%	20%	80%
HFpEF Model			
Dataset	Data Use (%)		
	Training	Verification	Validation
PHFS	80%	0%	20%
ARIC visit 5	0%	50%	50%
HFHF	0%	20%	80%

Separate accelerated failure time (AFT) models with Weibull distribution were trained and validated for HFrEF (LVEF <40%) and HFpEF (LVEF ≥50%) patients. The HFrEF and HFpEF models were trained and validated in 2,008 and 1,510 HFrEF and HFpEF patients, respectively, and predict the absolute likelihood of all-cause death within 6 months and 1 year. The HFrEF and HFpEF models include 17 and 14 protein features (protein-binding aptamers), respectively, and both models utilize samples from across 3 independent studies.

The HFrEF and HFpEF models were required to meet predefined performance criteria; a C-statistic and 1-year AUC ≥ 0.7 to ensure the proteomic models were at least equivalent to the previously reported performance of BNP in predicting death (Masson et. al, 2008, Spinar et. al, 2019, and Simpson et. al, 2020). Both models exceeded the prespecified performance requirements in model development (training and verification datasets) and validation. The HFrEF and HFpEF models had a C-Index of 0.75 and 0.761, 6-month AUC of 0.762 and 0.8, and 1-year AUC of 0.781 and 0.807, respectively, in the combined validation datasets (see table below for performance across all datasets).

Model	**Data Set**	**C-Index* (95% CI)**	**6-month AUC (95% CI)**	**1-year AUC (95% CI)**
HFrEF	PHFS (Training)	0.754 (0.73, 0.78)	0.783 (0.75, 0.83)	0.790 (0.76, 0.83)
	ARIC v5 (Verification)	0.757 (0.65, 0.85)	0.905 (0.5, 0.976)	0.894 (0.77, 0.98)

	HFHF (Verification)	0.733 (0.66, 0.81)	0.739 (0.47, 0.91)	0.837 (0.66, 0.97)
	Combined Validation Datasets	0.750 (0.72, 0.78)	0.762 (0.70, 0.83)	0.781 (0.74, 0.82)
	PHFS (Validation)	0.714 (0.66, 0.77)	0.707 (0.50, 0.81)	0.710 (0.61, 0.81)
	HFHF (Validation)	0.755 (0.71, 0.79)	0.775 (0.67, 0.87)	0.834 (0.77, 0.90)
HFpEF	PHFS (Training)	0.834 (0.79, 0.88)	0.839 (0.77, 0.90)	0.850 (0.79, 0.91)
	ARIC v5 (Verification)	0.819 (0.74, 0.90)	0.810 (0.50, 0.90)	0.855 (0.70, 0.98)
	HFHF (Verification)	0.750 (0.69, 0.80)	0.732 (0.50, 1.00)	0.747 (0.62, 0.84)
	Combined Validation Datasets	0.761 (0.73, 0.79)	0.800 (0.73, 0.88)	0.807 (0.73, 0.87)
	PHFS (Validation)	0.796 (0.67 - 0.92)	0.896 (0.73, 1.00)	0.912 (0.79, 0.99)
	ARIC v5 (Validation)	0.701 (0.65, 0.76)	0.626, (0.45, 0.90)	0.682 (0.42, 0.90)
	HFHF (Validation)	0.789, (0.75, 0.83)	0.830 (0.76, 0.90)	0.825 (0.74, 0.88)

7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? Yes

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
(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? 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"? N/A

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

I don't have access to the above link. I have searched public databases and I believe this is the most relevant ARIC publication:
Agarwal S.K. et al. Prediction of Incident Heart Failure in General Practice: The ARIC Study. Circ Heart Fail. 2012 Jul 1; 5(4): 422–429.

This study generated a risk score based on traditional/clinical risk factors, compared to existing HF risk scores and tested the incremental value of including known biomarkers such as NT-proBNP as part of these clinical HF risk scores. This publication has minimal overlap with our work, which utilizes large scale proteomics to agnostically derive novel prognostic heart failure models specifically developed for HFpEF and HFrEF subpopulations, using a mortality endpoint, as the proteomic models do not include clinical risk factors. This work will be referenced in our proposed manuscript.

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)?

I am not aware of related manuscript proposals.

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? Yes

11.b. If yes, is the proposal

_X A. primarily the result of an ancillary study (list number* 2017.27)

___ 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 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

Feldman, A. M., Kontos, C. D., McClung, J. M., Gerhard, G. S., Khalili, K., & Cheung, J. Y. (2017, Jun). Precision Medicine for Heart Failure: Lessons From Oncology. Circ Heart Fail, 10(6). https://doi.org/10.1161/CIRCHEARTFAILURE.117.004202

Cresci, S., Pereira, N. L., Ahmad, F., Byku, M., de Las Fuentes, L., Lanfear, D. E., Reilly, C. M., Owens, A. T., & Wolf, M. J. (2019, Oct). Heart Failure in the Era of Precision Medicine: A Scientific Statement From the American Heart Association. Circ Genom Precis Med, 12(10), 458-485. https://doi.org/10.1161/HCG.0000000000000058

Williams S.A. et al. Plasma protein patterns as comprehensive indicators of health. Nature Med 2019. https://doi.org/10.1038/s41591-019-0665-2