ARIC Manuscript Proposal #4184

PC Reviewed: 1/10/23	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Longitudinal Modeling of Cardiac Function to Identify Trajectories with Differential Risk of Heart Failure

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

2. Writing Group:
Writing group members: Anne Marie Reimer Jensen, James Ross, Victoria Arthur, Michael Hall, Kunihiro Matsushita, Brandon Lennep, Pamela Lutsey, Tor Biering-Sørensen, Amil M Shah, OTHERS WELCOME

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

First author: **Anne Marie Reimer Jensen**
Address: Gentofte Hospital
Hjertesygdomme, forskning 2
Gentofte Hospitalsvej 8, 3.th
2900 Hellerup, Denmark

Phone: Fax:
E-mail: annemarie.reimerjensen@gmail.com
anne.marie.reimer.jensen@regionh.dk

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 M. Shah**
Address: Brigham and Women's Hospital
Cardiovascular Division
75 Francis Street, PBB-1
North Boston, MA, 02115

Phone: Fax:
E-mail: ashah11@bwh.harvard.edu

3. Timeline: Analysis will begin following proposal approval with anticipated manuscript completion within 12 months. All analyses will be performed at the Brigham and Women's Hospital under the supervision of Amil Shah MD MPH.

4. Rationale:

Heart failure (HF) causes substantial morbidity, mortality, and resource utilization[1] and the prevalence of HF is projected to increase to 3% of the US population in 2030[2]. HF differentially burdens older adults who demonstrate the highest prevalence and incidence of HF[2]. Impairments in left ventricular (LV) systolic and diastolic function underlie HF risk and are common precursors of clinical HF. Age-related changes in cardiac structure and function are well recognized and include increasing LV mass[3], smaller LV chamber size[4], greater LV ejection fraction (LVEF)[5,6], and lower tissue Doppler imaging (TDI) relaxation velocities (e')[7]. However, existing studies of age-related cardiac changes are largely cross-sectional, while those with serial imaging have primarily evaluated average changes in individual cardiac measures using mixed linear models. While informative in describing overall trends, these studies are not able to capture the heterogeneity in patterns of change in both systolic and diastolic function over time which may contribute to differences in HF risk. Specifically, little data exist regarding trajectories defined simultaneously by changes in systolic and diastolic function through mid-to-late-life in the community, and the extent to which such trajectories may differentially relate to risk of HF generally, and HF with preserved LVEF (HFpEF) or HF with reduced LVEF (HFrEF) specifically.

Bayesian nonparametric mixture modeling is a data driven approach to trajectory modeling that has previously been used to characterize prototypical lung function trajectories[8,9]. This modeling algorithm uses multiple, longitudinal data points per subject, and relies on Bayesian nonparametrics which enable the best number of subgroups to be determined automatically from the data during inference. Importantly, the algorithm was formulated to identify subgroups by simultaneously modeling *multiple* outcome measures as a function of predictor variables. Trajectory subgroups are therefore defined by identifying groups of individuals who have similar progression patterns across an array of measures. An additional advantage of this approach is that there is no assumption of balanced or time-structured data, so that individuals can differ with respect to the number of longitudinal data points collected.

In this proposal, we will utilize this novel modelling approach to identify trajectories of cardiac function based on LVEF (a measure of systolic function) and E/A ratio (a measure of diastolic function) in mid-to-late-life in Black American participants in both the Jackson Heart Study (JHS) and the Atherosclerosis Risk in Communities (ARIC) study who had two or more echocardiograms over an 18-year period between JHS Visit 1 (2000-2004) and ARIC Visit 7 (2018-2019). We will then apply the identified trajectories to ARIC-only participants who attended Visit 5 and assess the association of trajectory membership with risk of incident HF hospitalization, HFpEF, and HFrEF. To better understand the potential biologic underpinnings of

the identified trajectories, we will assess for potential differential association of trajectories with proteomic measures from ARIC Visit 5.

5. Main Hypothesis/Study Questions:

1. Using a Bayesian nonparametric mixture modeling approach with longitudinal echocardiographic data in a community-based cohort, we will identify trajectories of cardiac function based on LVEF and E/A ratio that associate differentially with risk of HF

2. Using single-timepoint echocardiographic data in a separate set of participants, we will be able to predict trajectory membership, and predicted trajectory membership will demonstrate similar associations with risk of incident HF and will demonstrate differential risk of incident HFpEF versus HFrEF.

3. Predicted trajectories of cardiac function will demonstrate differential associations with plasma proteins

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).

Hypothesis 1: trajectories of cardiac function in Black Americans from shared ARIC-JHS population

Study population: All shared ARIC-JHS participants who underwent protocol echocardiography at JHS Visit 1 (2000-04) and either ARIC Visit 5 (2011-13) or Visit 7 (2018-19). Those with missing echocardiographic data (LVEF or E/A-ratio) or prevalent heart failure or atrial fibrillation at JHS Visit 1 will be excluded.

Echo measures: LVEF (calculated from LV dimensions using Teichholz formula) and E/A-ratio at JHS visit 1, ARIC Visit 5, and ARIC Visit 7. Echocardiographic measures from JHS Visit 1 studies will be recalibrated against ARIC Visit 5 measures based on re-analysis of 400 representative studies JHS Visit 1 studies at the ARIC (and JHS Visit 4) Echocardiography Reading Center at Brigham and Women's Hospital. Consistent with recalibration approaches previously applied to serial laboratory measures,[10] using the original and remeasured values on these 400 JHS Visit 1 studies, for each echocardiographic measure, we will first perform an iterative outlier removal process whereby studies with differences between the original and re-measured values that are >3 standard deviations from the mean difference are removed. Means and standard deviations of differences will then be recalculated the process repeated until no

outliers remain. We will then use Deming regression of the original versus re-measured valued, and recalibration equations will be derived from the Deming regression coefficients.

Trajectory of cardiac function: We will use a Bayesian nonparametric mixture model to identify trajectories based on LVEF and E/A ratio in shared JHS-ARIC participants with ≥2 echocardiogram. Age will be included as covariate in the model. Trajectory modeling assumes that there are distinct subpopulations of individuals with distinct progression characteristics and latent risk for adverse events. The goal of trajectory modeling is to identify the number of subgroups, the progression characteristics of each subgroup, and to which subgroup each individual in a data sample is most likely to be a member. Dr Ross (writing group member) has developed a Bayesian nonparametric trajectory algorithm that uses multiple, longitudinal data points per subject. To determine the most suitable number of subgroups, the algorithm relies on Bayesian nonparametrics which enable the best number of subgroups to be determined automatically from the data during inference. The user has control over a probabilistic prior to indicate whether there should be more or less subgroups. Importantly, the algorithm was formulated to identify subgroups by simultaneously modeling multiple outcome measures (e.g. LVEF, E/A in our proposal) as a function of predictor variables (e.g. age)[8]. Trajectory subgroups are therefore defined by identifying groups of individuals who have similar progression patterns across an array of measures. An additional advantage of this approach is that there is no assumption of balanced or time- structured data, so that individuals can differ with respect to the number of longitudinal data points collected. This technique has been successfully applied to multidimensional emphysema subtype data from CT scans[8] and to spirometry measures to improve understanding of lung function and COPD heterogeneity[9].

Incident HF events: Incident HF events after JHS Visit 1 will be ascertained based on the ongoing active surveillance for hospitalizations (including HF events) since 2005 and deaths. Endpoint of interest is adjudicated overall HF.

Statistical analysis: We will categorize shared ARIC-JHS participants by trajectory of cardiac function and use multivariable Cox proportional hazard regression models to assess association between trajectory membership and risk of HF after JHS Visit 1.

- **Model 1**: adjust for age and sex
- **Model 2**: adjust for age, sex, obesity, hypertension, diabetes, chronic kidney disease and coronary heart disease

Hypothesis 2: Predicted trajectories of cardiac function in ARIC-only participants

Study population: ARIC-only study participants who attended Visit 5 (2011-13) and underwent echocardiography. Participants who participated in JHS Visit 1 and those with missing echocardiographic data will be excluded.

Trajectory of cardiac function: We will use the Bayesian model generated in Hypothesis 1 analyses above to assign ARIC-only participants to the most likely trajectory based on LVEF, E/A ratio and age at ARIC Visit 5.

Incident HF events: Incident HF events post-Visit 5 will be ascertained based on the ongoing active surveillance for hospitalizations (including HF events) and deaths. Endpoints of interest include incident adjudicated HF overall (adjudicated HF events classified as A-C) and HF subtypes - HFrEF and HFpEF. The HF subtypes are defined based on abstracted ejection fraction (EF) at the time of the incident event and classified as HFpEF (EF ≥50%) or HFrEF (EF <50%).

Echo measures: We will describe the association of predicted trajectories with Visit 5 measures of LV structure (chamber dimensions and volumes, wall thickness, mass, LV remodeling), LV systolic (LVEF, strain), and diastolic function (E wave, TDI e', E/e' ratio, LA size, PASP), RV function, and LA size. For those who attended both Visits 5 (2011-13) and 7 (2018-2019), we will also describe the changes in these measures between study visits by predicted trajectory.

Statistical analysis: We will categorize shared ARIC-only participants by trajectory of cardiac function. Clinical and echocardiographic characteristics at ARIC Visit 5 will be described by assigned trajectory. Multivariable Cox proportional hazard regression models will be employed to assess the association between trajectory membership and risk of HF and HF type (with preserved [HFpEF] or with reduced ejection fraction [HFrEF]) after ARIC visit 5. We will use the following models:

- **Model 1**: age and sex
- **Model 2**: age, sex, obesity, hypertension, diabetes, chronic kidney disease and coronary heart disease

Hypothesis 3: Identify proteins associated with predicted trajectories of cardiac function in ARIC-only participants

Study population: This analysis will use the same population as defined in Hypothesis 2 above. Those with missing proteomics data at Visit 5 will not be included in this analysis.

Statistical analysis: To identify proteins associated with trajectories, we will conduct multinominal logistic regression adjusted for age and sex with the most common trajectory as the reference. Missing covariate data will be imputed using multiple imputation using chained equations. Protein levels and age will all be centered and scaled to a mean of 0 and standard deviation of 1 to allow for comparison between the different protein models. Multiple testing correction will be conducted using the false discovery rate (FDR) method.

For proteins significantly associated with trajectory membership, we will apply a two-sample Mendelian randomization (MR) approach to assess potential causal relationships between these

proteins and cardiac structure and function (LVEDV, LVESV, LVEF). To minimize bias, each analysis will use summary statistics from two independent European samples for exposure and outcome data, respectively. Instrumental variables (IVs) for protein quantitative trait loci (pQTLs) were obtained from three published studies: the INTERVAL study (n=3,301),[1] which included summary statistics for 2,994 proteins measured by SOMAscan, the AGES cohort study (n=5,368),[2] with summary statistics for 4,782 SOMAscan measured proteins, and Fenland study individuals with European descent (n=10,708),[3] including summary statistics for 4,775 SOMAscan protein targets. Summary statistics for cardiac structure/function outcomes (LVEDV, LVESV and LVEF) were obtained from UK Biobank (n=36,041).[4] The significance thresholds for MR association will be determined using Bonferroni correction for the number of independent sets of IVs tested.

Anticipated limitations:

Trajectories are modelled using data on LVEF from ARIC Visit 5 and possibly Visit 7. Therefore, for the JHS/ARIC overlap analysis relating trajectories to HF risk, HF events may occur concurrent with the defined trajectories between JHS Visit 1 and ARIC Visit 5 (i.e. using future information to predict incident events). However, the primary object of the JHS/ARIC overlap analysis is to derive trajectory groups while the analysis relating trajectories to HF risk is exploratory. We believe this important limitation to the incident HF analysis is mitigated by our analysis in the ARIC-only participants, where trajectory membership is assigned solely based on the Visit 5 echocardiograms and all incident HF events occur post-Visit 5. LVEF is preserved in the major of participants in this community-based sample, which may limit its utility in identifying prognostically relevant trajectories of change over time and it use in predicting trajectory membership. However, we have previously demonstrated the prognostic importance of differences in LVEF values within the 'normal' range.

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)?

- Longitudinal Tracking of Left Ventricular Mass and Systolic Function over the Adult Life Course in African Americans #2400 – Ervin Fox, Ramachandran Vasan. Evaluated change in LV mass and LV systolic function in shared ARIC-JHS participants between JHS exam 0 (conducted as part of ARIC exam 3), JHS visit 1 (2000-2004) and ARIC exam 5 (2010-2013).
- Proteomic Profiling and Cardiac Structure and Function in the Atherosclerosis Risk in Communities (ARIC) Study #3539 – Amil Shah, Bing Yu. Identify individual circulating proteins and protein networks that predict longitudinal changes in cardiac function in late life
- Association of cardiac troponin T and N-terminal pro-B-type Natriuretic Peptide with longitudinal changes in left ventricular structure and function: The Atherosclerosis Risk in Communities Study #3542 – Peder L Myhre, Amil Shah. Evaluated association between hs-cTnT and NT-proBNP at Visit 5 and change in echocardiographic parameters between Visit 5 and Visit 7.
- Longitudinal changes in left ventricular diastolic function in late life - the ARIC study #3557 – Li Zhao, Amil Shah. Evaluated change in diastolic function between Visit 5 and Visit 7.
- Association of stable CHD with longitudinal changes in cardiac structure and function: the ARIC study #3881 – Khaled Shelbaya, Amil Shah. Evaluated impact of CHD at Visit 5 on longitudinal changes in cardiac structure and function between Visit 5 and Visit 7.

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

___ A. primarily the result of an ancillary study (list number* _____)
_X_ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* __2015.34__ __2019.13__ _____)

*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. Heidenreich PA, Bozkurt B, Aguilar D, et al. 2022 AHA/ACC/HFSA Guideline for the Management of Heart Failure: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation*. 2022;145(18):E895-E1032. doi:10.1161/CIR.0000000000001063

2. Tsao CW, Aday AW, Almarzooq ZI, et al. Heart Disease and Stroke Statistics-2022 Update: A Report from the American Heart Association. *Circulation*. 2022;145(8):E153-E639. doi:10.1161/CIR.0000000000001052

3. Lieb W, Xanthakis V, Sullivan LM, et al. Longitudinal tracking of left ventricular mass over the adult life course: clinical correlates of short- and long-term change in the framingham offspring study. *Circulation*. 2009;119(24):3085-3092. doi:10.1161/CIRCULATIONAHA.108.824243

4. Cheng S, Xanthakis V, Sullivan LM, et al. Correlates of echocardiographic indices of cardiac remodeling over the adult life course: Longitudinal observations from the framingham heart study. *Circulation*. 2010;122(6):570-578. doi:10.1161/CIRCULATIONAHA.110.937821

5. Reimer Jensen AM, Zierath R, Claggett B, et al. Association of Left Ventricular Systolic Function with Incident Heart Failure in Late Life. *JAMA Cardiol*. 2021;6(5):509-520. doi:10.1001/jamacardio.2021.0131

6. Cheng S, Fernandes VRS, Bluemke DA, McClelland RL, Kronmal RA, Lima JAC. Age-related left ventricular remodeling and associated risk for cardiovascular outcomes the multi-ethnic study of atherosclerosis. *Circ Cardiovasc Imaging*. 2009;2(3):191-198. doi:10.1161/CIRCIMAGING.108.819938

7. Shah AM, Claggett B, Kitzman D, et al. Contemporary Assessment of Left Ventricular Diastolic Function in Older Adults: The Atherosclerosis Risk in Communities Study. *Circulation*. 2017;135(5):426-439. doi:10.1161/CIRCULATIONAHA.116.024825

8. Ross JC, Castaldi PJ, Cho MH, et al. A Bayesian Nonparametric Model for Disease Subtyping: Application to Emphysema Phenotypes. *IEEE Trans Med Imaging*. 2017;36(1):343-354. doi:10.1109/TMI.2016.2608782

9. Ross JC, Castaldi PJ, Cho MH, et al. Longitudinal modeling of lung function trajectories in smokers with and without chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2018;198(8):1033-1042. doi:10.1164/rccm.201707-1405OC

10. Parrinello CM, Grams ME, Couper D, et al. Recalibration of blood analytes over 25 years in the Atherosclerosis Risk in Communities Study: Impact of recalibration on chronic kidney disease prevalence and incidence. *Clin Chem*. 2015;61(7):938-947. doi:10.1373/clinchem.2015.238873