ARIC Manuscript Proposal #3755

PC Reviewed: 1/12/21 **Status: _____** **Priority: 2**
SC Reviewed: _____ **Status: _____** **Priority: ____**

1.a. Full Title: Protein Biomarkers of the DASH Diet and Hypertension Risk in the ARIC Study

b. Abbreviated Title (Length 26 characters): Proteomics, DASH, and Hypertension

2. Writing Group:

Writing group members:

Shutong Du
Hyunju Kim
Jingsha Chen
Josef Coresh
Nilanjan Chatterjee
Peter Ganz
Eric Boerwinkle
Bing Yu
Casey M. Rebholz

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

First author: Shutong Du
Address: Johns Hopkins Bloomberg School of Public Health
615 N. Wolfe St
Baltimore MD 21205

Phone: 206-601-3274 Fax: N/A
E-mail: sdu12@jhu.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: Casey M. Rebholz
Address: 2024 E. Monument St, Suite 2-519
Baltimore MD 21287

Phone: 410-502-2359 Fax: N/A
E-mail: crebhol1@jhu.edu

3. Timeline:

Data analysis and manuscript preparation is anticipated to take place within one year of approval of this proposal.

4. Rationale:

Proteomics is a high-throughput approach that allows for an unbiased characterization of proteins in biological specimens which reflect dietary intake. The proteome represents biologic function and is a meaningful indicator of metabolic activity related to food consumption. Discovery approaches using untargeted proteomics maximize the opportunity to identify new dietary biomarkers and to elucidate mechanisms of the cardioprotective effects of diet.

The Dietary Approaches to Stop Hypertension (DASH) dietary pattern is a high-quality diet, consisting of high intake of fruits, vegetables, and whole grains, and restriction of red and processed meat and sweets, that reduces blood pressure levels and reduces the risk of cardiovascular disease.[1-4] Nonetheless, the mechanism is poorly understood. Whereas biomarkers exist for single nutrients, there remains an urgent need for objective biomarkers of the overall DASH dietary pattern, as a clinically important diet for heart health. Identifying protein biomarkers of the DASH diet that are also related to risk of hypertension could help to characterize biological pathways underlying the beneficial cardiovascular health effects of this high-quality diet.

5. Main Hypothesis/Study Questions:

Aim 1: To evaluate whether protein biomarkers are associated with usual intake of the DASH diet in an observational cohort study.

Aim 2: To assess the long-term association of DASH diet-related proteins with risk of hypertension.

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 Aim 1, we will investigate whether plasma proteins are associated with the DASH diet. We will use available plasma proteomics data and the DASH diet score calculated based on dietary data collected with a food frequency questionnaire at study visit 3. For Aim 2, we propose to investigate whether DASH diet-related proteins to be identified in Aim 1 are prospectively associated with risk of incident hypertension over ~20 years of follow-up.

Inclusion/Exclusion Criteria:

We will include all participants who have dietary intake data and have plasma proteomic data measured by the SomaScan® (SomaLogic, Inc., Boulder, CO, USA) platform at study visit 3. For Aim 2, we will exclude all individuals with prevalent hypertension at visit 3.

Exposure:
For Aim 1, the primary exposure will be the DASH diet adherence score. For Aim 2, the primary exposure will be proteins that are found to be significantly associated with the DASH dietary pattern in Aim 1.

- **DASH Diet Score**: The DASH diet score will be calculated using dietary data collected with a food frequency questionnaire at study visit 3. The DASH diet score is based on 8 items: high intake of (1) fruits, (2) vegetables, (3) whole grains, (4) low-fat dairy products, (5) nuts and legumes, as well as low intake of (6) sodium, (7) red and processed meats, and (8) sweetened beverages.[5] Each component was scored on a scale of 1 to 5 based on ranked distribution in quintiles. The more frequent the consumption of healthy food components (fruits, vegetables, whole grains, low-fat dairy products, nuts and legumes), the higher the score. The more frequent the consumption of unhealthy food components (sodium, red and processed meats, sweetened beverages), the lower the score.

- **Proteins**: Proteins were measured in plasma collected at study visit 3 using an aptamer-based proteomics assay (SomaScan®).[6]

Outcome:
For Aim 1, the primary outcome will be all proteins that have been identified in the ARIC study. For Aim 2, the primary outcome will be incident hypertension. As secondary outcomes, we will also investigate change in blood pressure and patterns/trajectories of change in blood pressure.

- **Incident Hypertension:** Incident hypertension will be defined as first occurrence of systolic blood pressure ≥140 mm Hg or diastolic blood pressure ≥90 mm Hg or self-reported use of antihypertensive medication, as done previously in the ARIC study.[7] We will also perform a sensitivity analysis using a new definition of incident hypertension defined as first occurrence of systolic blood pressure ≥130 mm Hg or diastolic blood pressure ≥80 mm Hg or self-reported use of antihypertensive medication.[8]

Summary of Data Analysis:
Aim 1:
We will use linear regression models to examine the association between proteins and DASH diet adherence scores. β coefficients along with the corresponding 95% confidence intervals will be calculated per doubling of proteins ($\log_2$-transformed). We will use multivariable regression models to examine the independent association between proteins and DASH diet adherence after accounting for demographic characteristics (age, sex, race), study center, education level as a proxy for socioeconomic status, related health behaviors (smoking status, physical activity level), and health status (history of cardiovascular disease, fasting glucose, body mass index, and eGFR). We will adjust the statistical threshold to account for multiple comparisons using Benjamini-Hockberg (false discovery rate) and reduce the likelihood of detecting false positive findings.[9] We will calculate Pearson's correlation coefficient between the significant proteins to describe their interrelationships.

We will use C statistics to examine the ability of the proteins to improve the prediction of DASH diet adherence beyond established risk factors (i.e., covariates used in the multivariable

regression model). For the purpose of calculating the C statistic, we will dichotomize the DASH diet score empirically (highest quartile vs. lower 3 quartiles of the distribution) given the lack of a clinically meaningful threshold for the DASH diet score, as we have previously done in an analysis of dietary acid load.[10]

We will test for interaction by sex and race using likelihood ratio tests and, if present, stratify the analysis by these factors.

Aim 2:
We will use Cox proportional hazards regression models to evaluate the prospective association between DASH diet-related proteins and time to incident hypertension with censoring for death and end of follow-up. In addition, we will use linear mixed models in order to incorporate repeated measures of blood pressure across follow-up study visits and account for the non-independence of the repeated measures within individuals using random effects. Hazard ratios (for the analysis of incident hypertension) and β coefficients (for the analysis of blood pressure) along with the corresponding 95% confidence intervals will be calculated per doubling of proteins (log_2-transformed). We will also explore non-linear associations between proteins and hypertension risk using cubic and linear splines. We will use latent growth models to identify patterns of change in blood pressure among individuals and determine the association of proteins with each blood pressure trajectory. We will use multivariable regression models to examine the independent association between proteins and hypertension after accounting for known risk factors, including age, sex, race, center, education level as a proxy for socioeconomic status, smoking status, physical activity level, body mass index, history of cardiovascular disease, fasting glucose, total cholesterol, and eGFR.

In order to examine the ability of the significant DASH-related proteins to improve the prediction of hypertension risk beyond established risk factors for elevated blood pressure, we will calculate C statistics for models with and without the statistically significant proteins and test for differences in C statistics. We will adjust the statistical threshold to account for multiple comparisons using Benjamini-Hockberg and reduce the likelihood of detecting false positive findings.[11] We will calculate Pearson's correlation coefficient between the significant proteins to describe their interrelationships. We will test for interaction by sex and race using likelihood ratio tests and, if present, stratify the analysis by these factors. We will employ pathway analysis to determine which metabolic pathways are significantly implicated in the association between the DASH diet and hypertension risk.[12]

<u>Anticipated Methodologic Limitations or Challenges</u>:
Self-reported dietary intake relies on respondents' ability to report frequency and portion size and is prone to biases and errors. Proteins are measured in relative concentrations using the SomaLogic platform. Subsequent work will be conducted to develop quantitative assays of candidate protein biomarkers and replication will be conducted to investigate protein biomarkers in controlled feeding studies.

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

___√___ 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)?

#3398: Proteomics and the Risk of Incident Atrial Fibrillation in the Elderly: The Atherosclerosis Risk in Communities (ARIC) study
Description: This prospective study investigated aptamer-based proteomic profiles with incident Atrial Fibrillation in an elderly cohort of black and white men and women.

#3663: Association of Proteomic Markers with Incident Venous Thromboembolism
Description: This study identified novel proteins and biological pathways that associated with incident Venous Thromboembolism in ARIC.

#3389: Proteomic Profiling and Heart Failure Risk in the Atherosclerosis Risk in Communities (ARIC) Study
Description: This study identified individual circulating proteins and protein networks that associate with prevalent Heart Failure and Heart Failure stages and individual proteins and networks that predict Heart Failure and Heart Failure phenotype.

#3738: Proteomics and risk prediction for incident cardiovascular disease, recurrent cardiovascular disease, and chronic kidney disease progression: A validation analysis
Description: This study evaluated prediction measures of proteomic risk models, derived in the CRIC study, for incident cardiovascular disease, recurrent cardiovascular disease and chronic kidney disease progression.

There are no manuscript proposals on the proteomics of hypertension, and there are no existing manuscript proposals on the proteomics of the DASH diet.

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? ____ Yes __√__ No

11.b. If yes, is the proposal

_√_ A. primarily the result of an ancillary study (list number* 2019.26, 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:

1. Appel LJ, Moore TJ, Obarzanek E, et al. A clinical trial of the effects of dietary patterns on blood pressure. DASH Collaborative Research Group. The New England journal of medicine. 1997;336(16):1117-1124.

2. Fung TT, Chiuve SE, McCullough ML, Rexrode KM, Logroscino G, Hu FB. Adherence to a DASH-style diet and risk of coronary heart disease and stroke in women. Archives of internal medicine. 2008;168(7):713-720.

3. Sacks FM, Svetkey LP, Vollmer WM, et al. Effects on blood pressure of reduced dietary sodium and the Dietary Approaches to Stop Hypertension (DASH) diet. DASH-Sodium Collaborative Research Group. The New England journal of medicine. 2001;344(1):3-10.

4. Schwingshackl L, Hoffmann G. Diet quality as assessed by the Healthy Eating Index, the Alternate Healthy Eating Index, the Dietary Approaches to Stop Hypertension score, and health outcomes: a systematic review and meta-analysis of cohort studies. Journal of the Academy of Nutrition and Dietetics. 2015;115(5):780-800 e785.

5. Fung TT, Chiuve SE, McCullough ML, Rexrode KM, Logroscino G, Hu FB. Adherence to a DASH-style diet and risk of coronary heart disease and stroke in women. Arch Intern Med. 2008;168(7):713-720.

6. Gold L, Ayers D, Bertino J, et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. PLoS ONE, 5(12): e15004.

7. Yao L, Folsom AR, Pankow JS, et al. Parathyroid hormone and the risk of incident hypertension: the Atherosclerosis Risk in Communities study. J Hypertens. 2016;34(2):196-203.

8. Whelton PK, Carey RM, Aronow WS, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Hypertension. 2018; *71*(6), 1269–1324.

9. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B-Statistical Methodology. 1995;57(1):289-300.

10. Rebholz CM, Surapaneni A, Levey AS, et al. The Serum Metabolome Identifies Biomarkers of Dietary Acid Load in 2 Studies of Adults with Chronic Kidney Disease. J Nutr. 2019;149(4):578-585.

11. Kramer A, Green J, Pollard J, Jr., Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics. 2014;30(4):523-530.

12. Appel LJ, Champagne CM, Harsha DW, et al. Effects of comprehensive lifestyle modification on blood pressure control: main results of the PREMIER clinical trial. JAMA : the journal of the American Medical Association. 2003;289(16):2083-2093.