ARIC Manuscript Proposal #4219

PC Reviewed: 03/14/23	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Serum metabolomic markers of artificially sweetened beverage consumption in the Atherosclerosis Risk in Communities (ARIC) study

b. Abbreviated Title (Length 26 characters): Metabolomics of ASBs

2. Writing Group:

Writing group members: Hejingzi Jia, Lauren Bernard, Jingsha Chen, Shutong Du, Lyn M. Steffen, Kari Wong, Eric Boerwinkle, Bing Yu, Casey M. Rebholz

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

First author: Hejingzi (Monica) Jia
Address: 2024 East Monument Street, Suite 2-500

Phone: 417-771-2672 Fax:
E-mail: hjia11@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 East Monument Street, Suite 2-500
Phone: 410-502-2359 Fax:
E-mail: crebhol1@jhu.edu

3. Timeline: Analysis will begin upon manuscript proposal approval. We anticipate the manuscript will be submitted within approximately 1.5 years from the proposal's approval.

4. Rationale:

Sugar-sweetened beverages have been the predominant type of sweetened beverage consumed worldwide. The association between sugar-sweetened beverages and adverse health outcomes (e.g., obesity, cardiometabolic diseases, and mortality) has been widely reported in previous studies.[1–3] As such, artificially-sweetened beverages have become an appealing

alternative to reduce calorie and sugar intake, potentially mitigating the risk of adverse health outcomes.

Artificially-sweetened beverages are defined as beverages sweetened with agents that are synthetic or artificial in nature. Some examples of artificial sweeteners are sucralose, aspartame, acesulfame-potassium (Ace-K), and saccharin. Artificial sweeteners are sometimes also called non-nutritive sweeteners or low-calorie sweeteners. Some studies showed that substituting sugar-sweetened beverages with artificially-sweetened beverages may be associated with improvement in body weight and composition.[4] As a likely consequence of more public awareness for the beneficial health implications of artificially-sweetened beverages, the consumption of artificially-sweetened beverages has increased over time in the United States.[5] Approximately one-fifth of the U.S. population now consumes artificially-sweetened beverages on a daily basis.[5] Given the potential benefits of these beverages and the growing trend of consumption, artificially-sweetened beverages are an important and timely exposure to study.

Self-reported dietary assessments (e.g., 24-hour dietary recalls and food frequency questionnaires) are widely used by studies to ascertain artificially-sweetened beverages intake. These methods are subject to non-random measurement error (e.g., systematic underestimation) and recall bias, which may lead to lower data quality.[6] Some possible sources of errors include misclassification, intrusion, omissions, and portion size misestimation.[7] The portion size misestimation for beverages, especially sugar-free beverages, is considerable even in 24-hour dietary recalls.[7] Therefore, more objective approaches are needed to assess the intake of artificially-sweetened beverages.

Metabolomics is the study of small-molecule metabolites in biospecimen that are responsive to dietary intake.[8,9] An untargeted metabolomics study applies an agnostic approach to biomarker discovery, which is particularly advantageous for the identification of novel markers of dietary intake. There is a paucity of research on biomarkers of artificially-sweetened beverages. One trial of non-nutritive sweeteners has reported several amino acids (e.g., aspartate, serine, N-acetylalanine, ornithine), cofactors (e.g., quinolinate), and nucleotides (e.g., pseudouridine) associated with the consumption of non-nutritive sweeteners.[10] Studies of free-living populations that reflect consumption patterns of U.S. adults are absent from the literature.

Our study aims to identify candidate markers of artificially-sweetened beverages and explore metabolic pathways affected by their consumption. Upon replication, these results may be applied to develop dietary assessment tools (e.g., targeted assays) that will be complementary to existing self-reported methods.

5. Main Hypothesis/Study Questions:

We hypothesize that we will identify novel serum metabolites related to the intake of artificially-sweetened beverages.

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:

We will perform a cross-sectional analysis of artificially-sweetened beverage intake (exposure) and serum metabolites (outcome), which are derived from the food frequency questionnaire and fasting serum samples collected at ARIC visit 1 (1987-1989).

Eligibility criteria:

Participants will be excluded for the following reasons: (1) missing artificially-sweetened beverage intake data; (2) missing metabolomic data; (3) missing covariates information or unrealistic energy intake: <600 or >4200 kcal for men and <500 or >3600 kcal for women. We anticipate that approximately 4,000 participants will be included in our analysis.

Exposure variables:

The exposure is artificially-sweetened beverages consumption (sv/d). Dietary intake data were collected by trained interviewers at ARIC visit 1 (1987-1989) using a modified version of the Willett food frequency questionnaire. The reliability and validity for this food frequency questionnaire are moderate to high.[11,12] Specifically, we will use item #63, which asked intake of low-calorie soft drinks, such as any Diet Coke, Diet Pepsi, Diet 7-Up, to assess the level of exposure. Participants were given the following options to indicate their frequency of artificially-sweetened beverages consumption: (A) greater than 6 glasses per day; (B) 4-6 glasses per day; (C) 2-3 glassed per day; (D) 1 glass per day; (E) 5-6 glasses per week; (F) 2-4 glasses per week; (G) 1 glass per week; (H) 1-3 glasses per month; (I) almost never. We will use a scale to convert these reported values to units of glasses per day. The scale will use the midpoint of each range as the starting value and divide values in week or month frequency (E-H) by 7 and 30, respectively (e.g., response E is converted to 5.5/7=0.79). Responses of almost never (I) will be assigned the most conservative value of 0. We will model the exposure as a continuous variable.

Outcome variables:

The outcome is serum metabolites. We used fasting blood samples collected at baseline (1987-1989) and stored at - 80°C. Approximately 800 metabolites were measured by Metabolon, Inc. (Durham, North Carolina) using an untargeted, gas chromatography/mass spectrometry and liquid chromatography/mass spectrometry-based (LC-MS) protocol.[13] Metabolon profiled samples at two analytic timepoints: 2010 and 2014. In 2010, 374 known metabolites were identified in samples from 1,807 participants that are all African-American from the Jackson, Mississippi center (referred to as “subgroup 1”). In 2014, platform enhancements allowed Metabolon to identify a broader pool of metabolites, 759 metabolites, in samples from 2,004 participants from all four centers (referred to as “subgroup 2”). For this study, we plan to perform metabolomic assessment and data cleaning similar to previously published analyses of ARIC metabolomics data.[14] Our primary outcome will be metabolites that were measured in both subgroup 1 and subgroup 2, in order to internally replicate our candidate metabolomic markers. Our secondary outcome will be metabolites that were measured in only one batch, in order to maximize discovery findings.

Potential covariates:

Baseline covariates of interest include age, sex, race-center, total energy intake, body mass index, physical activity, education level, and dietary intake of sugar-sweetened beverages, coffee, and tea.

Statistical analysis:

Multivariable linear regression models will be used to assess cross-sectional associations between artificially-sweetened beverage consumption (exposure) and serum metabolites (outcome). Model 1 will adjust for age, sex, race-center, and total energy intake. Model 2 will further adjust for body mass index, physical activity, and education level. Model 3 will adjust for dietary intake of sugar-sweetened beverages, coffee, and tea. Race and center will be included as a combined covariate in only subgroup 2.

Analyses will be performed separately in each subgroup, in order to mitigate the possibility of batch effects. We will then meta-analyze the results using fixed effects meta-analysis. We will use Bonferroni correction to minimize the likelihood of false positive associations (e.g., $p = 0.5/360 \text{ metabolites} = 1.4 \times 10^{-3}$). We plan to perform a secondary analysis of metabolites only available in one subgroup. For significant metabolites in the primary and secondary analyses, we plan to assess their correlations with other significant metabolites using Spearman correlation coefficients. Pathway enrichment will be conducted.

We plan to explore non-linearity by modeling the exposure of artificially-sweetened beverage consumption as a categorical variable (quartile) and test for linear trend using the median value of each quartile. C-statistic tests will be used to determine whether significant metabolites predict higher intake of artificially-sweetened beverages, empirically dichotomized as highest quartile versus the lower three quartiles. We will also assess effect modification by diabetes status using likelihood ratio tests and stratification. Analyses will be conducted using Stata and R software.

Anticipated methodological limitations or challenges:

1. Dietary intake was assessed at ARIC visit 1 (1987-1989). These dietary data may not reflect current consumption given that the type and amount of artificial sweeteners and artificially-sweetened beverages has changed over time.[5] Aspartame and saccharin are the most widely used artificial sweeteners that are approved to be used in beverages during our data collection period.[15] However, more different types of artificial sweeteners have since been approved and used in beverages, such as Ace-K and sugar alcohol. We do not have information on packets of artificial sweeteners that are added to food and beverages at the table, which will result in an underestimation of the exposure.
2. Food frequency questionnaires are prone to bias and misclassification.[6,7] Nevertheless, the ARIC food frequency questionnaire has been found to have high reproducibility in previous ARIC publications.[12] In addition, diet beverages are typically purchased in a 12 fl oz can, therefore, the portion size is easily reported.
3. Given the observational study design, residual confounding is possible due to covariates that are either not measured or imprecisely measured.
4. Dietary components are correlated to each other. This can present a challenge in trying to study a single dietary exposure (artificially-sweetened beverages). However, we plan to adjust for other important dietary factors (i.e., sugar-sweetened beverages, coffee, tea), so we anticipate the results to have specificity for the exposure of artificially-sweetened beverages.

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)?

#3969: Metabolomics of Ultra-Processed Food, Incident Chronic Kidney Disease, and Incident Hypertension (lead author: Donghan Su)

This completed nutritional metabolomics project focuses on another dietary factor, ultra-processed food, which is distinct from the present manuscript proposal on artificially-sweetened beverages.

#2034: The Human Metabolome is Associated with Dietary Intake Among African Americans in the Atherosclerosis Risk in Communities Study (lead author: Yan Zheng)

The manuscript based on this proposal has already been published [Zheng Z, Yu B, Alexander D, Steffen LM, Boerwinkle E. Human metabolome associates with dietary intake habits among African Americans. Am J Epidemiol 2014;179(12):1424-1433.]. The paper included a sugar-rich foods and beverages group, which was defined as sugar-sweetened beverages (soda and fruit-flavored drinks), chocolate candy, candy without chocolate, cake, cookies, pie, donuts, Danish pastries, and brownies. They found 23 metabolites that were associated with this group, including 5 metabolites in the 2-hydroxybutyrate–related subpathway, 5 γ-glutamyl dipeptides, and 7 unsaturated long-chain fatty acids. We hypothesize that the biological underpinnings of artificially-sweetened beverages and sugar-rich beverages differ, and thus, the metabolites associated with artificially-sweetened beverages will differ from those identified in this prior work. In addition, this paper only included data from African American participants from the Jackson, Mississippi center, whereas we plan to include a larger biracial sample representative of

all four field centers. Finally, they studied 356 named metabolites. The Metabolon platform has since expanded and metabolite identification has improved, so we are able to study a larger pool of metabolites (~720) and anticipate that we will be able to identify many novel metabolomic associations.

#3228: Metabolomics of coffee consumption and risk of kidney disease in the ARIC study (Lead author: William He)

The manuscript based on this proposal has already been published [He WJ, Chen J, Razavi AC, et al. Metabolites Associated with Coffee Consumption and Incident Chronic Kidney Disease. *Clin J Am Soc Nephrol CJASN*. 2021;16(11):1620-1629. doi:10.2215/CJN.05520421]. This study used a similar untargeted metabolomic approach to discover metabolites that were associated with a beverage (coffee). However, we plan to identify candidate markers of another beverage (artificially-sweetened beverages).

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* _____)

2014.20: Genomics, Metabolomics, and Cardiovascular Disease (PI: Eric Boerwinkle)

2008.16: Metabolomics & Heart Failure: A Novel Approach to Biomarker Discovery (PI: Jennifer Nettleton)

___ 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. Pereira MA. Sugar-Sweetened and Artificially-Sweetened Beverages in Relation to Obesity Risk. *Adv Nutr*. 2014;5(6):797-808. doi:10.3945/an.114.007062

2. de Koning L, Malik VS, Kellogg MD, Rimm EB, Willett WC, Hu FB. Sweetened Beverage Consumption, Incident Coronary Heart Disease, and Biomarkers of Risk in Men. *Circulation*. 2012;125(14):1735-1741. doi:10.1161/CIRCULATIONAHA.111.067017

3. Malik VS, Li Y, Pan A, et al. Long-Term Consumption of Sugar-Sweetened and Artificially Sweetened Beverages and Risk of Mortality in US Adults. *Circulation*. 2019;139(18):2113-2125. doi:10.1161/CIRCULATIONAHA.118.037401

4. McGlynn ND, Khan TA, Wang L, et al. Association of Low- and No-Calorie Sweetened Beverages as a Replacement for Sugar-Sweetened Beverages With Body Weight and Cardiometabolic Risk: A Systematic Review and Meta-analysis. *JAMA Netw Open*. 2022;5(3):e222092. doi:10.1001/jamanetworkopen.2022.2092

5. Fakhouri THI. Consumption of Diet Drinks in the United States, 2009–2010. 2012;(109).

6. Ravelli MN, Schoeller DA. Traditional Self-Reported Dietary Instruments Are Prone to Inaccuracies and New Approaches Are Needed. *Front Nutr*. 2020;7:90. doi:10.3389/fnut.2020.00090

7. Whitton C, Ramos-García C, Kirkpatrick SI, et al. A Systematic Review Examining Contributors to Misestimation of Food and Beverage Intake Based on Short-Term Self-Report Dietary Assessment Instruments Administered to Adults. *Adv Nutr*. Published online August 30, 2022:nmac085. doi:10.1093/advances/nmac085

8. Tzoulaki I, Ebbels TMD, Valdes A, Elliott P, Ioannidis JPA. Design and Analysis of Metabolomics Studies in Epidemiologic Research: A Primer on -Omic Technologies. *Am J Epidemiol*. 2014;180(2):129-139. doi:10.1093/aje/kwu143

9. Guasch-Ferré M, Bhupathiraju SN, Hu FB. Use of Metabolomics in Improving Assessment of Dietary Intake. *Clin Chem*. 2018;64(1):82-98. doi:10.1373/clinchem.2017.272344

10. Suez J, Cohen Y, Valdés-Mas R, et al. Personalized microbiome-driven effects of non-nutritive sweeteners on human glucose tolerance. *Cell*. 2022;185(18):3307-3328.e19. doi:10.1016/j.cell.2022.07.016

11. Stevens J, Metcalf PA, Dennis BH, Tell GS, Shimakawa T, Folsom AR. Reliability of a food frequency questionnaire by ethnicity, gender, age and education. *Nutr Res*. 1996;16(5):735-745. doi:10.1016/0271-5317(96)00064-4

12. Willett WC, Sampsons L, Stampfers MJ, et al. Reproducibility and Validity of a Semiquantitative Food Frequency Questionnaire. *Am J Epidemiol*. 1985;122(1):51-65. doi:10.1093/oxfordjournals.aje.a114086

13. Kim H, Yu B, Li X, et al. Serum metabolomic signatures of plant-based diets and incident chronic kidney disease. *Am J Clin Nutr*. 2022;116(1):151-164. doi:10.1093/ajcn/nqac054

14. Zheng Y, Yu B, Alexander D, et al. Associations between metabolomic compounds and incident heart failure among African Americans: the ARIC Study. *Am J Epidemiol*. 2013;178(4):534-542. doi:10.1093/aje/kwt004

15. Basson AR, Rodriguez-Palacios A, Cominelli F. Artificial Sweeteners: History and New Concepts on Inflammation. *Front Nutr*. 2021;8:746247. doi:10.3389/fnut.2021.746247