ARIC Manuscript Proposal #3458

PC Reviewed: 9/10/19	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title:

Novel biomarkers for dietary acid load and incident CKD

b. Abbreviated Title (Length 26 characters):

Metabolomics of dietary acid load

2. Writing Group:

Anam Tariq, Bing Yu, Eric Boerwinkle, Josef Coresh, Morgan E. Grams, Casey M. Rebholz, *others welcome*

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

First author: Anam Tariq
Address: 1830 E. Monument Street, Suite 416
Baltimore, MD

Phone: 410-810-4005
Fax: 410-367-2258
E-mail: tanam1@jhmi.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 Street, Suite 2-500
Baltimore, MD 21287
Office: (410) 502-2359
E-mail: crebhol1@jhu.edu

3. Timeline:

Analyses will begin immediately after the manuscript proposal has been approved. The first author will complete the analysis and prepare a manuscript as a Master's thesis project by the end of the 2019-2020 academic year. We anticipate that a manuscript will be submitted to the ARIC Publications Committee within one year of approval of the proposal.

4. Rationale:

Dietary acid load is a clinically important aspect of dietary intake. Higher dietary acid load comes from the consumption of animal sources of food, e.g. cheese, meat, and eggs[1-3]. In contrast, diets that are lower in dietary acid load consist of plant-derived foods, e.g. fruits and vegetables[4,5]. In the ARIC study, higher dietary acid load was associated with higher risk for kidney disease independent of demographic characteristics, socioeconomic status, total energy intake, lifestyle factors, comorbid conditions, antihypertensive medication use, and baseline kidney function[6]. The observed association may be partly due to the fact that the kidneys play an important role in the excretion of acid and other toxins to maintain acid-base balance. The degree of acid load may be burdensome and may promote decline of estimated glomerular filtration rate (eGFR).

The measurement of dietary acid load often relies upon the use of self-reported dietary intake. To address this issue, candidate blood biomarkers of dietary acid load were recently reported using untargeted metabolomics data in two studies of kidney disease patients: the African American Study of Kidney Disease and Hypertension (AASK) and the Modification of Diet in Renal Disease (MDRD) study[7]. Recently, thirteen metabolites were found to be associated with dietary acid load. These metabolites were: S-methylmethionine, indolepropionylglycine, indolepropionate, N-methylproline, N-δ-acetylornithine, threonate, oxalate, chiro-inositol, methyl glucopyranoside, stachydrine, catechol sulfate, hippurate, and tartronate. However, these metabolites have not yet been shown to be related to dietary acid load in a general population and the relationship between the metabolites and kidney outcomes is still unknown.

5. Main Hypothesis/Study Questions:

1. We hypothesize that higher dietary acid load will be associated with lower serum levels of 13 metabolites: S-methylmethionine, indolepropionylglycine, indolepropionate, N-methylproline, N-δ-acetylornithine, threonate, oxalate, chiro-inositol, methyl glucopyranoside, stachydrine, catechol sulfate, hippurate, and tartronate.
2. We hypothesize that lower serum levels of these 13 metabolites will be associated with higher risk of incident chronic kidney disease.

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 hypothesis #1, we will conduct a cross-sectional analysis of dietary acid load and serum metabolites at study visit 1 (1987-1989).
- For hypothesis #2, we will conduct a prospective analysis of dietary acid load-related metabolites (measured at study visit 1, 1987-1989) and incident chronic kidney disease.

Inclusion:

- Those with dietary intake data collected at study visit 1 (baseline, 1987-1989)
- Those with metabolomics data at study visit 1

Exclusion:

- Participants with baseline eGFR $\leq$60 mL/min/1.73 m^2 or prevalent CKD at baseline
- Those with extreme values for energy intake representing implausible estimates of dietary intake

Exposure: Dietary intake was assessed using a semi-quantitative 66-item food frequency questionnaire (FFQ), modified from the Willett questionnaire. As detailed in a previous ARIC publication [6], dietary acid load will be estimated using the equation for potential renal acid load (PRAL) by Remer and Manz: PRAL= 0.49 * protein + 0.037 * phosphorus - 0.021 * potassium – 0.026 * magnesium - 0.013 * calcium [8]. We will also calculate the net endogenous acid production (NEAP), where NEAP = 54.5 * (protein/potassium) – 10.2 [9].

Outcome: The primary outcome is incident CKD defined as eGFR <60 mL/min/1.73 m^2 accompanied by ≥25% eGFR decline from baseline, an International Classification of Diseases (ICD), Ninth/Tenth Revision code for a kidney disease-related hospitalization or death, or ESRD (dialysis or transplantation) identified by linkage to the USRDS registry between baseline (study visit 1, 1987-1989) and December 31, 2017[10,11].

Covariates:

- In multivariable regression models, we will adjust for the following baseline covariates: age (continuous), sex (binary), race (binary), center (categorical), education level (categorical), physical activity (continuous), smoking status (categorical), body mass index (continuous), systolic blood pressure (continuous), diabetes status (binary), total energy intake (continuous), and eGFR (continuous).
- Baseline kidney function will be assessed using the 2009 CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration) creatinine equation for eGFR. As a sensitivity analysis, we will use the cystatin C-based equation for eGFR[12,13].

Data analysis:

- Baseline characteristics of the study participants will be described according to tertiles of PRAL and NEAP using descriptive statistics (mean, standard deviation, frequency, proportion). We will test for differences in baseline characteristics using chi-squared and analysis of variance (ANOVA) for continuous and categorical variables, respectively.
- For hypothesis #1, linear regression models will be used to examine the cross-sectional association between dietary acid load (exposure) and metabolites (outcome).
- For hypothesis #2, Cox proportional hazards regression models will be used to examine the prospective association between metabolites (exposure) and incident chronic kidney disease (outcome).
- In addition to analyzing individual metabolites, we will create a dietary acid load metabolite score. We will use the coefficients from the analysis of individual metabolites as weights for the score. We will calculate the AUC as a measure of the cumulative ability of the dietary acid load metabolite score in addition to the covariates to classify participants in the highest quartile compared with lower quartiles of dietary acid load. We

will also analyze the prospective association between the dietary acid load metabolite score with incident chronic kidney disease. In order to explore potential causal mechanisms, we will conduct causal mediation analysis. Causal mediation analysis estimates the indirect effect (IE) of dietary acid load on chronic kidney disease risk that was mediated through metabolites as well as the proportion of the effect mediated.

- We will analyze the two batches of metabolomics data separately.
- In addition to adjusting for race as a confounder, we will explore potential race-specific associations in a secondary analysis through the stratification by race.

Limitations:

- There is the potential for measurement error with self-report of dietary intake. The purpose of this study is to strengthen the validity of proposed biomarkers of dietary acid load which directly addresses this limitation.
- Another limitation is the long-time interval between dietary assessment (visit 1) and the end of follow-up (2017) considering that diet may change over time. However, in the ARIC study, we have demonstrated that there is little change in diet quality between visit 1 and visit 3 [14].

7.a. Will the data be used for non-CVD analysis in this manuscript? ____ Yes _X___ No

b. If Yes, is the author aware that the file ICTDER03 must be used to exclude persons with a value RES_OTH = "CVD Research" for non-DNA analysis, and for DNA analysis RES_DNA = "CVD Research" would be used? ____ Yes ____ No
(This 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? _X___ Yes ____ No

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the file ICTDER03 must be used to exclude those with value RES_DNA = "No use/storage DNA"? __X__ 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)?

ARIC manuscript proposal #2325: "Relationship of dietary features related to acid load and subsequent kidney disease: Atherosclerosis Risk in Communities study"

The paper based on proposal #2325 has already been published and provides strong rationale for pursuing the present proposal[6].

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 at https://www2.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. Gannon RH, Millward DJ, Brown JE, et al. Estimates of daily net endogenous acid production in the elderly UK population: analysis of the National Diet and Nutrition Survey (NDNS) of British adults aged 65 years and over. Br J Nutr 2008;100:615-23.
2. Engberink MF, Bakker SJ, Brink EJ, et al. Dietary acid load and risk of hypertension: the Rotterdam Study. Am J Clin Nutr 2012;95:1438-44.
3. Murakami K, Sasaki S, Takahashi Y, Uenishi K, Japan Dietetic Students' Study for N, Biomarkers G. Association between dietary acid-base load and cardiometabolic risk factors in young Japanese women. Br J Nutr 2008;100:642-51.
4. Appel LJ, Moore TJ, Obarzanek E, et al. A clinical trial of the effects of dietary patterns on blood pressure. DASH Collaborative Research Group. N Engl J Med 1997;336:1117-24.
5. Ausman LM, Oliver LM, Goldin BR, Woods MN, Gorbach SL, Dwyer JT. Estimated net acid excretion inversely correlates with urine pH in vegans, lacto-ovo vegetarians, and omnivores. J Ren Nutr 2008;18:456-65.
6. Rebholz CM, Coresh J, Grams ME, et al. Dietary Acid Load and Incident Chronic Kidney Disease: Results from the ARIC Study. Am J Nephrol 2015;42:427-35.
7. 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;149:578-85.
8. Remer T, Manz F. Potential renal acid load of foods and its influence on urine pH. J Am Diet Assoc 1995;95:791-7.
9. Frassetto LA, Todd KM, Morris RC, Jr., Sebastian A. Estimation of net endogenous noncarbonic acid production in humans from diet potassium and protein contents. Am J Clin Nutr 1998;68:576-83.

10. Group. KDIGOKCW. KDIGO 2012clincial practice guideline for the evaluation and management of chronic kidney disease. . Kidney Int Suppl 2013;3:1-150.
11. Grams ME, Rebholz CM, McMahon B, et al. Identification of incident CKD stage 3 in research studies. Am J Kidney Dis 2014;64:214-21.
12. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med 2009;150:604-12.
13. Inker LA, Eckfeldt J, Levey AS, et al. Expressing the CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration) cystatin C equations for estimating GFR with standardized serum cystatin C values. Am J Kidney Dis 2011;58:682-4.
14. Xu Z SL, Selvin E, Rebholz CM. Diet quality, change in diet quality, and risk of cardiovascular disease and diabetes. . Public Health Nutr 2019.