ARIC Manuscript Proposal #4306

PC Reviewed: 8/8/23	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Metabolomic and Proteomic Markers of Vitamin D Supplementation and Clinical Outcomes

b. Abbreviated Title (Length 26 characters): Omics of Vitamin D

2. Writing Group:
Writing group members:

Casey M. Rebholz, Valerie Sullivan, Lauren Bernard, Jingsha Chen, Bing Yu, Erin Michos, Lawrence J. Appel, Josef Coresh

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

First author: Casey M. Rebholz
Address: 2024 East Monument Street, Suite 2-500
Baltimore, Maryland 21287

Phone: 410-502-2359
E-mail: crebhol1@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: Valerie Sullivan
Address: 2024 East Monument Street, Suite 2-500
Baltimore, Maryland 21287
E-mail: vsulliv5@jhmi.edu

3. Timeline: Analyses will begin after this proposal is approved. We anticipate that a first draft of the manuscript will be available within approximately one year of manuscript proposal approval.

4. Rationale:
Vitamin D is a fat-soluble vitamin with well-established causal roles in bone health.[1,2] Its involvement in renin-angiotensin-aldosterone system regulation and inhibition of vascular cell growth, inflammation, and fibrosis give biological plausibility to a hypothesized cardioprotective effect of vitamin D.[1,3] Higher serum concentrations of 25(OH)D, the main circulating form of vitamin D, are generally associated with a lower risk of cardiovascular disease in observational

studies.[4–6] However, vitamin D supplementation in randomized controlled trials has consistently failed to reduce cardiovascular events or cardiovascular mortality.[5,6]

While 25(OH)D is a reliable biomarker of vitamin D supply, it does not directly reflect biologically available vitamin D or its functions in the body.[7] The hormonally active vitamin D molecule, $1,25(OH)_2D$, is not an informative marker of function because circulating concentrations are maintained even in deficiency states through upregulation of kidney function.[7,8] Thus, there is a need for better biomarkers of vitamin D. Modern high-throughput technologies can identify thousands of metabolites and proteins in blood and urine. The metabolome, inclusive of both endogenous and exogenous metabolites, and the proteome, which reflects products of gene expression, can provide information about vitamin D exposure as well as its physiological effects and relation to health outcomes. Characterizing metabolites and proteins associated with vitamin D supplementation, and examining sources of heterogeneity in vitamin D-related metabolite and protein profiles, may help elucidate potential mechanisms underlying the epidemiological associations and yield insights into the null findings in clinical trials.

The overarching objective of this research is to use untargeted metabolomics and untargeted proteomics to discover novel biomarkers of vitamin D supplementation and to characterize metabolic pathways through which vitamin D supplements may promote cardiovascular health.

5. Main Hypothesis/Study Questions:

Aim #1: We hypothesize that we will be able to discover novel metabolomic biomarkers of vitamin D supplementation in free-living, older adults using untargeted serum and urine metabolomics.

Aim #2: We hypothesize that we will identify plasma proteins that are associated with vitamin D supplement use.

Aim #3: We hypothesize that a subset of metabolites and proteins that are positively associated with vitamin D supplement use will be inversely associated with cardiovascular outcomes and mortality. Furthermore, we hypothesize that metabolites and proteins that are inversely associated with vitamin D supplement use will be positively associated with cardiovascular outcomes and mortality.

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 conduct a cross-sectional analysis using visit 5 (2011-13) medication (supplement use) data and metabolomic data. For Aim #2, we will use visit 5 medication data and proteomic data. For Aim #3, we will study metabolites and proteins that are significantly associated with vitamin D supplement use in Aim #1 and Aim #2. We plan to study their associations with cardiovascular disease incidence, recurrence, and mortality over approximately 10 years of follow-up beyond ARIC study visit 5.

Eligibility Criteria: For Aim #1, participants must have provided blood samples for metabolomic analysis. For Aim #2, participants must have provided blood samples for plasma proteomic analyses. For Aim #3, participants must have either metabolomic or proteomic biomarkers measured and have follow-up clinical data available for relevant outcome(s) (i.e., cardiovascular disease incidence, recurrence, or mortality). For all analyses, participants must have completed visit 5, provided medication data, and have covariate information available.

Exposures & Outcomes: For Aim #1 and Aim #2, the exposure will be dietary supplementation of isolated vitamin D (in international units). We will examine differences in omics profiles according to vitamin D dose among vitamin D supplement users, as well as compare profiles in vitamin D supplement users (any dose) to non-users. For Aim #3, the exposure will be metabolites and proteins that are significantly associated with vitamin D supplement use in Aim #1 and Aim #2.

Vitamin D Supplementation Use: Participants were instructed to bring all medications and supplements used within the past 4 weeks to the clinic visit. Supplement bottles were scanned or transcribed by study staff to record their identity and strength. Twelve relevant medication codes have been identified and will be used to ascertain vitamin D supplement use (**Table 1**). Vitamin D in other forms (e.g., capsule, liquid, powder, crystals, drops), vitamin D analogs (e.g., paricalcitol), and nutrient combination supplements (e.g., calcium with vitamin D, multivitamins, vitamins D & K) will be excluded.

Table 1. Vitamin D Supplement Codes from ARIC Visit 5

Medication Name	Code	Dosage
Ergocalciferol Cap 50000 Unit	77202030000110	50000
Cholecalciferol Cap 400 Unit	77202032000105	400
Cholecalciferol Cap 1000 Unit	77202032000110	1000
Cholecalciferol Cap 2000 Unit	77202032000120	2000
Cholecalciferol Cap 5000 Unit	77202032000140	5000
Cholecalciferol Cap 10000 Unit	77202032000160	10000
Cholecalciferol Cap 50000 Unit	77202032000180	50000
Cholecalciferol Tab 400 Unit	77202032000320	400
Cholecalciferol Tab 1000 Unit	77202032000330	1000
Cholecalciferol Tab 2000 Unit	77202032000340	2000
Cholecalciferol Tab 5000 Unit	77202032000350	5000
Cholecalciferol Chew Tab 400 Unit	77202032000520	400

Metabolomic Profiling: Serum samples and urine samples were collected at the visit 5 examination (2011-13). Aliquots were stored at -80˚C until metabolomic profiling was conducted by Metabolon, Inc. on serum samples in January 2023 and urine samples in December 2022. We estimate serum metabolomic profiles are available for ~5,400 participants, and urine metabolomic profiles are available for ~1,600 participants. Paired serum and urine metabolomic profiles (i.e., from the same person) are available for

~1,600 participants. In visit 5 samples, there were 953 serum metabolites and 1104 urine metabolites that were identified.

Proteomic Profiling: Plasma aliquots were stored at -80˚C until proteomic analyses by SomaLogic, Inc. (Boulder, Colorado) in May 2020. Plasma proteins were measured using the SomaScan Assay, a DNA-based aptamer technology that uses chemically modified nucleotides to bind proteins with high affinity and specificity. Proteins were quantified in relative fluorescence units using DNA microarrays, and measurements were normalized according to SomaLogic's standard process.[9] Plasma proteomic profiles are available for ~5,200 ARIC study participants at visit 5. In the visit 5 plasma samples, 4,955 proteins were identified.

Incident Cardiovascular Disease, Recurrent Cardiovascular Events, and All-Cause Mortality: For Aim #3, the primary outcome will be incident cardiovascular disease, which will be a composite outcome of coronary heart disease (hospitalized myocardial infarction or fatal coronary heart disease), stroke, and heart failure. Events were ascertained by annual telephone interviews; active surveillance of hospital discharge records, obituaries, and community death certificates; and linkage to the National Death Index.

As secondary analyses, we plan to investigate each cardiovascular disease subtype (coronary heart disease, stroke, and heart failure) separately, and we plan to ascertain and study recurrent cardiovascular events using an established definition in ARIC.[10] We also plan to investigate all-cause mortality identified by annual telephone interviews with proxies, hospital records, obituaries, community death certificates, or linkage to the National Death Index.[11]

Other Variables of Interest: For the metabolomic and proteomic analyses (Aim #1 and Aim #2, respectively), we will consider adjusting for the following variables: age, sex, education, smoking status, physical activity, body mass index, estimated glomerular filtration rate, factors that affect endogenous vitamin D synthesis (including season when blood was collected,[12] clinical center, and race). For the prospective analyses with clinical outcomes (Aim #3), we will consider additional adjustment for cardiovascular risk factors (fasting glucose and cholesterol) and medication use (blood glucose-lowering and lipid-lowering).

Statistical Analysis:

<u>Aim #1 (metabolites) and Aim #2 (proteins)</u>

We will use linear regression to investigate cross-sectional associations between vitamin D supplementation (independent variable) and log-transformed metabolites (dependent variable) for Aim #1 and between vitamin D supplementation (independent variable) and log-transformed proteins for Aim #2. For both analyses, we will model vitamin D supplementation continuously according to daily dose (in international units) to identify dose-dependent differences in metabolite and protein levels. We will also examine differences in metabolite and protein levels comparing vitamin D supplement users (any dose) to non-users. For Aim #1, the primary analysis will be conducted using serum metabolites given the larger sample size. Secondary analyses will be conducted using

urine metabolites as a complementary and innovative approach, which offers even broader coverage of food and plant-derived compounds.

Multivariable linear regression models for Aim #1 and Aim #2 will adjust for age, sex, education, smoking status, physical activity, body mass index, estimated glomerular filtration rate, and factors that affect endogenous vitamin synthesis (e.g., season when blood was collected, clinical center, and race). To account for multiple statistical tests and reduce type 1 error, we will use an adjusted significance threshold (e.g., Bonferroni or Benjamini-Hochberg) for these analyses. We plan to also examine differences in associations by sex, race, and body mass index using stratified analyses and likelihood ratio tests. For Aim #2, we will additionally perform an enrichment analysis to investigate biological pathways that are overrepresented by proteins significantly associated with vitamin D supplementation.

Aim #3 (cardiovascular outcomes and mortality)
We will use Cox proportional hazards regression models to investigate prospective associations between vitamin D-related metabolites and proteins with incident cardiovascular disease (primary outcome). Secondary outcomes for Aim #3 include recurrent cardiovascular disease, cardiovascular disease subtypes, and all-cause mortality. The time origin will be visit 5, as this is when metabolite and proteins were measured, and the end of follow up will be 10 years after visit 5. We will calculate hazard ratios (95% confidence intervals) per one standard deviation higher log-transformed metabolite or protein level. We will examine non-linear associations between metabolite and protein levels with outcomes (e.g., by modeling levels categorically or using splines).

Cox regression models will adjust for all covariates included in the cross-sectional analyses, as well as cardiovascular risk factors (fasting glucose and cholesterol) and medication use (blood glucose-lowering and lipid-lowering). We will use an adjusted significance threshold to account for multiple tests of metabolite- and protein-outcome associations (e.g., Bonferroni or Benjamini-Hochberg). We will also examine whether associations between metabolites and proteins with outcomes differ by sex, race, and body mass index using stratified analyses and likelihood ratio tests.

Anticipated Methodologic Limitations or Challenges: The frequency and duration of supplement use was not recorded. Given that isolated vitamin D supplements are usually taken once daily, we will assume once daily dosing. Additionally, dietary intake was not assessed at visit 5 in the ARIC study, so we are unable to quantify concurrent dietary intake of vitamin D in our study. Given that average dietary vitamin D intake is low in U.S. adults (mean 4.9 µg/day, or ~200 IU),[13] we do not consider this to be a critical limitation. As a sensitivity analysis, we will adjust for usual dietary vitamin D estimated from ARIC study visits 1 (1987-89) and 3 (1993-95).

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? ____ 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 file ICTDER03 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/search.php

___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)?
MP3774: Replication Request for: Omic signatures of vitamins C and E (Lead author: Young KL)

Authors investigated DNA methylation and gene expression signatures of vitamin C and E intake, assessed using dietary questionnaires (food frequency questionnaires). The ARIC study was one of 4 cohorts used for replication analyses. However, several differences distinguish our study from this proposal, including (1) our focus on vitamin D, (2) assessment of vitamin D supplementation using the medication questionnaire (vitamin D supplementation was not assessed on the food frequency questionnaire), and (3) analysis of metabolomic and proteomic signatures rather than epigenomic correlates.

MP2377: 25-hydroxyvitamin D levels, vitamin D binding protein gene polymorphisms, and vitamin D3 epimer with risk of incident coronary heart disease (CHD) among whites and blacks: the ARIC Study (Lead author: Michos ED)
MP2425: Serum 25-hydroxyvitamin D levels and incidence of atrial fibrillation: the ARIC study (Lead author: Alonso A)
MP2435: Serum 25-hydroxyvitamin D and incident hypertension (Lead author: Lutsey PL)

Authors of these studies examined prospective associations between 25(OH)D and cardiovascular outcomes, which is similar to our proposed Aim 3. Due to limitations of 25(OH)D, our proposal uniquely proposes to examine prospective associations between novel metabolomic and proteomic vitamin D biomarkers with cardiovascular disease.

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

2023.08 Metabolomics and Proteomics of Dietary Supplementation with Vitamin D, Cardiovascular Disease, and Mortality (PI: Casey Rebholz)
2017.26: Proteomic longitudinal ARIC study: SOMAscan of multiple visits (PI: Josef Coresh)
2022.09 Urine Metabolomics and Chronic Kidney Disease (PI: Josef Coresh)
2017.10: Metabolic signatures underlying cardiac structure for heart failure (PI: Bing Yu)

___ 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 http://www.cscc.unc.edu/aric/forms/

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. Bouillon R, Marcocci C, Carmeliet G, et al. Skeletal and Extraskeletal Actions of Vitamin D: Current Evidence and Outstanding Questions. *Endocr Rev*. 2019;40(4):1109-1151. doi:10.1210/er.2018-00126
2. Institute of Medicine (US) Committee to Review Dietary Reference Intakes for Vitamin D and Calcium. *Dietary Reference Intakes for Calcium and Vitamin D*. (Ross AC, Taylor CL, Yaktine AL, Del Valle HB, eds.). National Academies Press (US); 2011. http://www.ncbi.nlm.nih.gov/books/NBK56070/
3. Al Mheid I, Quyyumi AA. Vitamin D and Cardiovascular Disease: Controversy Unresolved. *J Am Coll Cardiol*. 2017;70(1):89-100. doi:10.1016/j.jacc.2017.05.031
4. Zhang R, Li B, Gao X, et al. Serum 25-hydroxyvitamin D and the risk of cardiovascular disease: dose-response meta-analysis of prospective studies. *Am J Clin Nutr*. 2017;105(4):810-819. doi:10.3945/ajcn.116.140392
5. Barbarawi M, Kheiri B, Zayed Y, et al. Vitamin D Supplementation and Cardiovascular Disease Risks in More Than 83 000 Individuals in 21 Randomized Clinical Trials: A Meta-analysis. *JAMA Cardiol*. 2019;4(8):765-776. doi:10.1001/jamacardio.2019.1870

6. Manson JE, Bassuk SS, Buring JE, VITAL Research Group. Principal results of the VITamin D and OmegA-3 TriaL (VITAL) and updated meta-analyses of relevant vitamin D trials. *J Steroid Biochem Mol Biol*. 2020;198:105522. doi:10.1016/j.jsbmb.2019.105522
7. Prentice A, Goldberg GR, Schoenmakers I. Vitamin D across the lifecycle: physiology and biomarkers. *Am J Clin Nutr*. 2008;88(2):500S-506S. doi:10.1093/ajcn/88.2.500S
8. Seamans KM, Cashman KD. Existing and potentially novel functional markers of vitamin D status: a systematic review. *Am J Clin Nutr*. 2009;89(6):1997S-2008S. doi:10.3945/ajcn.2009.27230D
9. Candia J, Cheung F, Kotliarov Y, et al. Assessment of Variability in the SOMAscan Assay. *Sci Rep*. 2017;7(1):14248. doi:10.1038/s41598-017-14755-5
10. Wattanakit K, Folsom AR, Chambless LE, Nieto FJ. Risk factors for cardiovascular event recurrence in the Atherosclerosis Risk in Communities (ARIC) study. *Am Heart J*. 2005;149(4):606-612. doi:10.1016/j.ahj.2004.07.019
11. Magnani JW, Norby FL, Agarwal SK, et al. Racial Differences in Atrial Fibrillation-Related Cardiovascular Disease and Mortality: The Atherosclerosis Risk in Communities (ARIC) Study. *JAMA Cardiol*. 2016;1(4):433-441. doi:10.1001/jamacardio.2016.1025
12. Millen AE, Nie J, Sahli MW, et al. Vitamin D Status and Prevalent Early Age-Related Macular Degeneration in African Americans and Caucasians: The Atherosclerosis Risk in Communities (ARIC) Study. *J Nutr Health Aging*. 2017;21(7):772-780. doi:10.1007/s12603-016-0827-6
13. Herrick KA, Storandt RJ, Afful J, et al. Vitamin D status in the United States, 2011-2014. *Am J Clin Nutr*. 2019;110(1):150-157. doi:10.1093/ajcn/nqz037