1. Full Title: Metabolomic compounds and incident peripheral artery disease: the Atherosclerosis Risk in Communities (ARIC) Study

b. Abbreviated Title (Length 26 characters): Metabolomics & PAD risk

2. Writing Group:
   Writing group members: Kunihiro Matsushita, Bing Yu, Shoshana Ballew, Morgan Grams, Vijay Nambi, Elizabeth Selvin, Josef Coresh, and Eric Boerwinkle; others are welcome.

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

First author: Kunihiro Matsushita
Address: Department of Epidemiology
         Johns Hopkins Bloomberg School of Public Health
         Welch Center for Prevention, Epidemiology, and Clinical Research
         2024 E. Monument St., Suite 2-600, Baltimore, MD 21287

   Phone: (443) 287-8766        Fax: (443) 683-8358
   E-mail: kmatsus5@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: Eric Boerwinkle
Address: Human Genetics Center
         UT School of Public Health
         Houston, TX 77025

   Phone: 713-500-9816 Fax:
   E-mail: eric.boerwinkle@uth.tmc.edu

3. Timeline: Data to be used in this proposal are basically available. Analyses and manuscript preparation will be performed over the next 6-12 months.

4. Rationale:
Lower extremity peripheral artery disease (PAD) affects 8-10 million adults in the US and more than 200 million individuals around the world.\textsuperscript{1-3} Individuals with PAD have 2-to 5-fold higher risk of total mortality, mainly due to cardiovascular disease, compared to those without.\textsuperscript{4} PAD also impacts mobility, induces leg symptoms, and thus reduces quality of life.\textsuperscript{5}

It is often considered that epidemiological understanding of other atherosclerotic diseases (e.g., coronary disease) can be simply extrapolated to PAD.\textsuperscript{6} However, traditional cardiovascular risk factors do not necessarily equally contribute to different types of cardiovascular diseases. Specifically, smoking and diabetes are more strongly associated with PAD than coronary heart disease,\textsuperscript{7} suggesting unique pathophysiological pathways in PAD. Moreover, for critical limb ischemia (CLI), a severe manifestation of PAD including ischemic ulcer, gangrene, or rest pain, small clinical studies suggest the pathophysiological involvement of non-traditional factors, such as microvascular injury, coagulation factors, and edema, in its development.

In this context, modern technologies allow for comprehensive characterization of metabolic networks, i.e., metabolomic profiling, through quantification of a wide range of low molecular weight metabolomic compounds in a given biological sample, which may reveal new pathophysiological pathways for certain diseases. However, to the best of our knowledge, there are no metabolomics studies for PAD. Given that some metabolites, e.g., creatinine, are reported to be associated with PAD risk, extensive investigation is needed. Therefore, using data from the Atherosclerosis Risk in Communities (ARIC) Study with over 25 years of follow-up, we aim to explore metabolomic compounds predicting the risk of PAD and/or its severe form, CLI.

5. Main Hypothesis/Study Questions:
   i. Levels of metabolites will be associated with incident PAD.
   ii. Levels of metabolites will be associated with incident CLI.

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).

Inclusions:
- ~4,000 black and white ARIC participants with metabolomic profiling data using visit 1 blood sample under ancillary study # 2014.20 and 2008.16
- Data available for incident PAD (details summarized below)

Exclusions:
- Missing data on variables of interest

Exposures:
- Metabolomic measures were obtained from stored fasting serum samples by Metabolon, Inc. (Durham, North Carolina) using an untargeted, gas chromatography–mass spectrometry and liquid chromatography–mass spectrometry approach. This untargeted approach identified ~600-800 named and unnamed metabolites (depending on batch). Of
these, in the present study, we will focus on ~200 metabolites with limited missing values and reasonable reliability. Because the analysis plan calls for races combined, the analyses will be limited to those metabolites shared between African-Americans and European-Americans.

**Outcomes:**
PAD-related hospitalizations will be identified according to the following ICD codes based on previous literature: atherosclerosis of native arteries of the extremities, unspecified (440.20); atherosclerosis of native arteries of the extremities with intermittent claudication (440.21); atherosclerosis of native arteries of the extremities with rest pain (440.22); atherosclerosis of native arteries of the extremities with ulceration (440.23); atherosclerosis of native arteries of the extremities with gangrene (440.24); other atherosclerosis of native arteries of the extremities (440.29); atherosclerosis of bypass graft of the extremities (440.3); atherosclerosis of other specified arteries (440.8); peripheral vascular disease, unspecified (443.9); leg artery revascularization (38.1, 39.2, 39.29, 39.50). Of these, 440.22, 440.23, and 440.24 will be considered CLI. Also, we will consider any cases with the above code as CLI when the following codes coexist: leg amputation (84.1), lower extremity ulcer (707.1), and gangrene (785.4).

**Other variables of interest and covariates:**
Sociodemographics: age, race, gender, education
Physical information: blood pressure including ankle-brachial index (ABI) obtained at visit 1 in a randomly selected single leg, body mass index
Lifestyle: smoking status/amount
Comorbidities: diabetes, dyslipidemia, kidney function, coronary heart disease, stroke
Batch: The metabolomic measurements were collected in two batches.

**Statistical analysis plan:**
The primary analysis will use Cox proportional hazards models to quantify the prospective association of each metabolomic compound with incident clinical PAD and CLI. All hazard ratios will be estimated for 1-SD increment of each metabolomic compound. To assess the impact of potential confounders, we will run several models. Model 1 will be minimally adjusted for demographic variables, i.e., age, gender, race, and education and the batch. Model 2 will further include traditional risk factors, smoking, diabetes, hypertension, lipids, and body mass index. Model 3 will additionally account for manifestation of atherosclerotic diseases, i.e., history of coronary heart disease and stroke well as kidney function. Multiple comparisons will be taken into account when we interpret the results. To evaluate whether some metabolomic compounds have uniquely strong associations with CLI, differences in log hazard ratios will be obtained by subtracting log-HRs for CLI from the log-HRs for non-CLI PAD. Standard errors for the differences in log-HRs will be estimated by 1,000 bootstraps of the difference of log-HRs. For metabolomic compounds with significant associations with the risk of PAD and/or CLI, we will repeat the analysis after stratifying the study sample by key demographic and clinical subgroups according to age, gender, race, smoking status and the presence/absence of diabetes, hypertension, reduced kidney function, low ABI, and
history of other cardiovascular diseases at baseline. We will formally test interaction using likelihood ratio test.

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.csc.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)?
    There is no proposal for metabolomics in PAD. In the context of metabolomics, a few proposals target other cardiovascular phenotypes such as heart failure (#1696 & #1847) and atrial fibrillation (#2354).

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, 2008.16, 2014.05)
    ____ 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