

ARIC Manuscript Proposal # 3066

PC Reviewed: 11/14/17
SC Reviewed: _____

Status: _____
Status: _____

Priority: 2
Priority: _____

- 1.a. **Full Title:** Role of obesity in metabolite-associated CVD
b. **Abbreviated Title (Length 26 characters):** Obesity, metabolites, and CVD

2. **Writing Group:**

Writing group members:

Ellen Demerath (University of Minnesota), Bing Yu (University of Texas Health Sciences Center at Houston), Penny Gordon-Larsen (University of North Carolina), Danyu Lin (University of North Carolina), Annie Green Howard (University of North Carolina), Susan Cheng (Brigham and Women's Hospital), Kris Young (University of North Carolina), Casey Rebholz (Johns Hopkins University) *others welcome*

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

First author: Christy Avery
Address: 137 E. Franklin St., Suite 306
Chapel Hill, NC 27514
Phone: 919/966-2148 Fax: 919/966-9800
E-mail: Christy_avery@unc.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: Kari North
Address: 137 E. Franklin St., Suite 306
Chapel Hill, NC 27514
Phone: 919/966-2148 Fax: 919/966-9800
E-mail: kari_north@unc.edu

3. **Timeline:** Analyses will begin once study-specific permissions are obtained.

4. **Rationale:** Cardiovascular diseases (CVDs) and their antecedents impose major societal burdens, being leading causes of morbidity, mortality, and early disability.¹⁻⁵ For example, approximately 660,000 incident cases of CHD occurred in 2013,⁶ putting large segments of the population at increased heart failure, cardiac arrhythmia, and sudden death risk,^{7,8} and imparting significant (and rising) health care costs projected to double by 2030.⁹ Alarming, the decline in CVD mortality evident over the past four decades may have substantially slowed, or even reversed.¹⁰⁻¹² One highly plausible reason for changing CVD mortality patterns are the parallel

increases in obesity and diabetes. For example, over the past 30 years, the prevalence of obesity has increased markedly in the U.S., doubling among adults (35% in 2011-12)¹³ along with onset earlier in adulthood.¹⁴⁻¹⁶

Evolving CVD risk factor burdens emphasize the need to better understand the evolution of CVD and CVD risk factors in the context of constant and unremitting metabolic stress.^{17,18} The “expressed genome”, i.e. factors beyond DNA, offer innovative means by which to interrogate underlying mechanisms influencing CVD and CVD risk factors.¹⁹ For example, systematic interrogations of small-molecule metabolites produced by biochemical and cellular processes that reflect gene and protein activity have identified plasma and urinary correlates of CVD and CVD risk factors,¹⁹⁻²⁵ most notably the novel molecular marker branched chain amino acids (BCAAs), which predict diabetes incidence, weight loss, and improvement in insulin resistance after weight loss.^{75, 94-98} Interestingly, dietary supplementation with BCAAs in rats fed high-fat diets promoted insulin resistance regardless of weight gain,^{26,27} demonstrating that under conditions of high fat intake (i.e., typical of the American diet), BCAA contributes to development of obesity-associated insulin resistance. Yet, there is still a tremendous amount to learn about the role of metabolites as CVD biomarkers and the modifying role of obesity. For example, no studies to the best of our knowledge have evaluated the modifying role of obesity on metabolite-associated CVD risk or changes in CVD risk factors despite strong associations of CVDs with metabolic perturbations including hyperglycemia and hyperlipidemia and the established modifying role of obesity in said perturbations.²⁴

5. Main Hypothesis/Study Questions:

Hypothesis: Metabolites in pathways that lead to CVD are exacerbated by obesity.

Aims: Identify metabolites and major metabolic pathways to CVD risk factors (blood lipids, blood pressure, and glycemic phenotypes) and CVD incidence (coronary heart disease [CHD], heart failure [HF], stroke, and type 2 diabetes [DM]) ***that are modified by obesity. Thus we have a very specific hypothesis about the interaction of obesity on the relationship between plasma metabolites and CVD. We also are extremely happy to include all persons interested in the main effects of metabolites on CVD/CVD risk factors as coauthors on this proposal.***

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).

Event definitions. Follow-up for all outcomes (CHD, HF, and stroke; baseline visit through 12/31/14 or the latest date for which outcome data are available) was accomplished through a combination of active surveillance of local hospital discharge lists, annual participant interviews querying hospitalizations, examination of vital records, and interviews with decedent’s next of kin. Incident HF was defined by hospitalization or death certificate codes listing 428 or I-50 in any position, shown previously to have high levels of accuracy.²⁸ Incident CHD was defined as a validated definite or probable hospitalized myocardial infarction (MI), a definite CHD death, or an unrecognized MI identified by electrocardiograph. The criteria for definite or probable hospitalized MI were based on combinations of chest pain symptoms, electrocardiogram changes, and cardiac enzyme levels.²⁹ Physician-adjudicated stroke was identified based on the presence of ICD-9 codes 430 to 438 and neurological signs and symptoms. Differences in associations by stroke subtype will be evaluated in sensitivity analysis to the degree possible given modest

numbers of non-ischemic stroke. As of 12/31/2014, ARIC investigators identified 2,431, 2,867, and 1,332 cases of incident CHD, HF, and stroke, respectively over a mean of 20 years follow-up. Diabetes is identified through fasting glucose, self-report, and medication inventory.

CVD risk factor measurement protocols are described for select lipid, blood pressure, and glycemic phenotypes. Briefly, fasting LDL-C was calculated using the Friedwald formula;³⁰ medications were assessed through medication inventory and will be used for adjustment or correction consistent with the literature.³¹ Three seated resting blood pressure readings were obtained using random zero sphygmomanometers; the mean of the last two measures is retained. Like LDL-C, blood pressure lowering medications are assessed by medication inventory and will be used for adjustment or correction consistent with the literature.³¹ Fasting plasma glucose will be evaluated in participants without diabetes, as defined above.

Plasma metabolites: Plasma metabolites were measured (Metabolon, Inc., Durham, NC) at baseline in a random sample of Caucasians (n=1,553) and African Americans (n=2,479), standardized, and were: a) excluded when >75% samples had missing values or values below the detection limit (BDL); b) analyzed as a continuous variable when missing/BDL<25%, with data imputed to half of the lowest value; and c) analyzed as an ordinal variable when 25-75% of the data were missing/BDL. We will primarily focus on the ~200 metabolites with limited missing values, reasonable reliability, and present in both batches.

Statistical methods. For all continuous outcomes, normality will be assessed and appropriate transformations will be applied, as required. Confounders will be assessed through literature review and construction of directed acyclic graphs (DAGs).³² We also will include “phase” and measures of kidney function in all models given previously identified strong relationships with metabolites. Although we present our approach for the entire population, we also will examine evidence for effect modification by sex and race/ethnicity as secondary exploratory extensions to each aim.

We will evaluate how obesity, e.g. measured as body mass index (BMI) or waist hip ratio (WHR), modifies cross-sectional or longitudinal associations between individual metabolites and CVD/risk factors using generalized linear mixed models (continuous or categorical outcomes) and an extension of proportional hazards models (time-to-event) both which will allow for time varying exposure measures.^{33,34} Equation 1 below provides a general framework for CVD risk factors, where $g()$ is a link function (identity for continuous variables, binary for categorical variables, e.g. diabetes), Z_{ij} and $obesity_{ij}$ denote respectively the CVD risk factor and the obesity measure for individual i at the j th time point and Y_i denotes the metabolite marker measured at study baseline. As with previous models, within-individual correlation will be accounted through the use of the random effect for the individual (b_{0i}) and all models will adjust for covariates (e.g., phase, kidney function, race, age, sociodemographics, lifestyle health behaviors etc.) as identified by our DAG analysis.

$$\begin{aligned}
 \text{Eq. 1} \quad & g(Z_{ij}) \\
 & = \beta_0 + \beta_1 Y_{ij} + \beta_2 time \\
 & + \beta_3 (obesity_{ij}) + \beta_4 (time)(obesity_{ij}) + \beta_5 (time) Y_{ij} + \beta_6 Y_{ij} (obesity_{ij}) \\
 & + \beta_7 Y_{ij} (time)(obesity_{ij}) + b_{0i} + e_{ij}
 \end{aligned}$$

By adding the three-way interaction in these models, we will be able to not only look at see if obesity modifies the association between metabolites and obesity but also how this modification changes over time. Therefore, our primary hypothesis of interest involves testing simultaneously

whether there is a change over time in the effect modification (i.e. $\beta_7 = 0$) and whether there is effect modification (i.e. $\beta_6 = 0$).

Our time-to-event models will be of a similar structure as they will include baseline metabolite measures and time-varying covariates (specifically longitudinal obesity measures and interactions between longitudinal obesity measures and baseline metabolite measures). However, due to the structure of these models, the focus will be strictly on the interaction term between metabolites and obesity, in order to determine whether obesity modifies the association between a particular metabolite and the time to CVD event.

Mendelian Randomization. *Potential* causal effects of metabolites on CVD/risk factors within strata of obesity will be examined using MR. Briefly, MR is a form of instrumental variable analysis based on the concept that if exposure X (i.e. metabolite) affects outcome Y (i.e. CVD), factors affecting X (i.e. genetic variants) must also effect Y.¹⁰⁷ Genetic variant K is therefore utilized as an ‘instrumental variable’ (IV) to examine the potential causal nature of the X-Y association. Strengths of MR include K-Y associations that are robust to confounding from variables other than ancestry, which can be accurately measured and accounted for,^{108,109} given random assignment of genotypes at conception. Such random assignment also enables assessment of temporality of the K-Y and by extension the X-Y association, thereby eliminating the potential for reverse causation.¹⁰⁸ Although application of MR is dependent on critical assessment of assumptions, which will be evaluated, the approach has been successfully used to assess the potential causality of various exposures, such as cholesterol levels, ethanol intake, BMI, and C-reactive protein on diverse outcomes such as CVD Alzheimer’s disease, exercise levels and depression.

Although the above narrative pertains to MR of main effects, the approach for which readers are most familiar, semiparametric methods for estimation of non-linear exposure-outcome relationships using IVs were recently described.³⁵ Thus, while we expect statistical innovations like these will allow us to extend MR methods developed for main effects analyses to evaluate the *potential* causality of metabolites on obesity associated CVD risk/changes in risk factors, we remain aware of the challenges. ***Thus, we consider this component of our study aims to be exploratory.***

Anticipated methodologic limitations or challenges. We will adjust the significance level by the Bonferroni method to account for multiple comparisons (number of metabolites evaluated and number of phenotypes examined). Statistical power also is may be a challenge, particularly when evaluating interaction hypotheses in the context of a generalized linear mixed effects model. As part of a parallel effort in the CARDIA study, we examined statistical power to detect BMI-metabolite interactions. Briefly, using the statistical framework described, if the metabolite \times obesity interaction effect explains 1% of the variance ($R^2=0.01153$, $f^2=0.01166$; $\alpha=0.05/(200 \text{ metabolites} \times 3 \text{ risk factors})$, i.e. $\alpha=8.3 \times 10^{-5}$) in a CVD risk factor among $N=3,000$ subjects (conservative given $n=4,032$ ARIC participants have metabolomics characterization), we achieve 90% power to detect the metabolite \times obesity interaction effect. We thus anticipate excellent power to examine how obesity modifies associations between plasma metabolites and CVD/risk factors.

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 ___ 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>

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)?

MS #2867 (Yu), MS#2756 (Rebholz). Dr. Yu is a close collaborator on this project, having led many ARIC metabolomics efforts. We also have invited Dr. Rebholz to collaborate.

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* _____)

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.csc.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.csc.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.

13. Per Data Use Agreement Addendum, approved manuscripts using CMS data shall be submitted by the Coordinating Center to CMS for informational purposes prior to publication. Approved manuscripts should be sent to Pingping Wu at CC, at pingping_wu@unc.edu. I will be using CMS data in my manuscript ____ Yes No.

Literature Cited

1. Abbasi, F., Brown, B.W., Jr., Lamendola, C., McLaughlin, T. & Reaven, G.M. Relationship between obesity, insulin resistance, and coronary heart disease risk. *J Am Coll Cardiol* **40**, 937-43 (2002).
2. Ferrannini, E. *et al.* Insulin resistance in essential hypertension. *N Engl J Med* **317**, 350-7 (1987).
3. Reaven, G.M. Insulin resistance/compensatory hyperinsulinemia, essential hypertension, and cardiovascular disease. *J Clin Endocrinol Metab* **88**, 2399-403 (2003).
4. Flegal, K.M., Graubard, B.I., Williamson, D.F. & Gail, M.H. Excess deaths associated with underweight, overweight, and obesity. *JAMA* **293**, 1861-7 (2005).
5. Romero-Corral, A. *et al.* Association of bodyweight with total mortality and with cardiovascular events in coronary artery disease: a systematic review of cohort studies. *Lancet* **368**, 666-78 (2006).
6. Statistics, N.C.f.H. Mortality multiple cause micro-data files. 2013. *Public use data file and documentation. NHLBI tabulations.: National Center for Health Statistics* (2011).
7. Marmot, M.G. & Elliott, P. *Coronary heart disease epidemiology: from aetiology to public health*, (Oxford University Press, USA, 2005).
8. Mozaffarian, D. *et al.* Heart disease and stroke statistics--2015 update: a report from the American Heart Association. *Circulation* **131**, e29-322 (2015).
9. Heidenreich, P.A. *et al.* Forecasting the future of cardiovascular disease in the United States: a policy statement from the American Heart Association. *Circulation* **123**, 933-44 (2011).
10. Sidney, S. *et al.* Recent Trends in Cardiovascular Mortality in the United States and Public Health Goals. *JAMA Cardiol* **1**, 594-9 (2016).
11. Lloyd-Jones, D.M. Slowing Progress in Cardiovascular Mortality Rates: You Reap What You Sow. *JAMA Cardiol* **1**, 599-600 (2016).
12. Mensah, G.A. *et al.* Decline in Cardiovascular Mortality: Possible Causes and Implications. *Circ Res* **120**, 366-380 (2017).
13. Ogden, C.L., Carroll, M.D., Kit, B.K. & Flegal, K.M. Prevalence of childhood and adult obesity in the United States, 2011-2012. *JAMA* **311**, 806-14 (2014).
14. The, N.S., Richardson, A.S. & Gordon-Larsen, P. Timing and duration of obesity in relation to diabetes: findings from an ethnically diverse, nationally representative sample. *Diabetes Care* **36**, 865-72 (2013).
15. The, N.S., Suchindran, C., North, K.E., Popkin, B.M. & Gordon-Larsen, P. Association of adolescent obesity with risk of severe obesity in adulthood. *JAMA* **304**, 2042-7 (2010).
16. Dutton, G.R. *et al.* 25-year weight gain in a racially balanced sample of U.S. adults: The CARDIA study. *Obesity (Silver Spring)* **24**, 1962-8 (2016).
17. Poirier, P. *et al.* Obesity and cardiovascular disease: pathophysiology, evaluation, and effect of weight loss: an update of the 1997 American Heart Association Scientific Statement on Obesity and Heart Disease from the Obesity Committee of the Council on Nutrition, Physical Activity, and Metabolism. *Circulation* **113**, 898-918 (2006).
18. Scherer, P.E. & Hill, J.A. Obesity, Diabetes, and Cardiovascular Diseases: A Compendium. *Circ Res* **118**, 1703-5 (2016).

19. Musunuru, K. *et al.* The Expressed Genome in Cardiovascular Diseases and Stroke: Refinement, Diagnosis, and Prediction: A Scientific Statement From the American Heart Association. *Circ Cardiovasc Genet* **10**(2017).
20. Shah, S.H., Kraus, W.E. & Newgard, C.B. Metabolomic profiling for the identification of novel biomarkers and mechanisms related to common cardiovascular diseases: form and function. *Circulation* **126**, 1110-20 (2012).
21. Ussher, J.R., Elmariah, S., Gerszten, R.E. & Dyck, J.R. The Emerging Role of Metabolomics in the Diagnosis and Prognosis of Cardiovascular Disease. *J Am Coll Cardiol* **68**, 2850-2870 (2016).
22. Marcinkiewicz-Siemion, M., Ciborowski, M., Kretowski, A., Musial, W.J. & Kaminski, K.A. Metabolomics - A wide-open door to personalized treatment in chronic heart failure? *Int J Cardiol* **219**, 156-63 (2016).
23. Nowak, C. *et al.* Effect of Insulin Resistance on Monounsaturated Fatty Acid Levels: A Multi-cohort Non-targeted Metabolomics and Mendelian Randomization Study. *PLoS Genet* **12**, e1006379 (2016).
24. Shah, S.H. & Newgard, C.B. Integrated metabolomics and genomics: systems approaches to biomarkers and mechanisms of cardiovascular disease. *Circ Cardiovasc Genet* **8**, 410-9 (2015).
25. Wurtz, P. *et al.* Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. *Circulation* **131**, 774-85 (2015).
26. Newgard, C.B. *et al.* A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab.* **9**(2009).
27. Newgard, C.B. *et al.* A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab* **9**, 311-26 (2009).
28. Folsom, A.R. *et al.* Community prevalence of ideal cardiovascular health, by the American Heart Association definition, and relationship with cardiovascular disease incidence. *J Am Coll Cardiol* **57**, 1690-6 (2011).
29. NHLBI. Atherosclerosis Risk in Communities (ARIC) Study Operations Manual No. 3: Surveillance Component Procedures. Version 1.0. (ed. ARIC Coordinating Center) (University of North Carolina, Chapel Hill, NC, 1987).
30. McNamara, J.R., Cohn, J.S., Wilson, P.W. & Schaefer, E.J. Calculated values for low-density lipoprotein cholesterol in the assessment of lipid abnormalities and coronary disease risk. *Clin Chem* **36**, 36-42 (1990).
31. Tobin, M.D., Sheehan, N.A., Scurrah, K.J. & Burton, P.R. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. *Stat Med* **24**, 2911-35 (2005).
32. Greenland, S., Pearl, J. & Robins, J.M. Causal diagrams for epidemiologic research. *Epidemiology* **10**, 37-48 (1999).
33. Cox, D.R. *Regression models and life-tables*, (Springer, New York, 1995).
34. Therneau, T.M. & Grambsch, P.M. *Extending the Cox Model*, (2000).
35. Staley, J.R. & Burgess, S. Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization. *Genet Epidemiol* **41**, 341-352 (2017).

