1.a. Full Title: The Association of a Genetic Predisposition to Non-Alcoholic Fatty Liver Disease with the Development of Metabolic Risk Factors and Incident Cardiovascular Disease: The Atherosclerosis Risk in Communities (ARIC) Study

b. Abbreviated Title (Length 26 characters): NAFLD SNP and CVD risk

2. Writing Group: Chiadi E. Ndumele; Eric Boerwinkle; Kari North; Gerardo Heiss; Mariana Lazo; Christie M. Ballantyne; Linda Kao; Elizabeth Selvin; Josef Coresh; others welcome

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

First author: Chiadi E. Ndumele, MD, MHS
Address: Assistant Professor of Medicine, Johns Hopkins University
         600 North Wolfe Street, Carnegie 568
         Baltimore MD 21287
Phone: 410-502-2319       Fax: 410-614-8222
E-mail: cndumel2@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: Josef Coresh
Address: Professor of Epidemiology, Medicine & Biostatistics
         Johns Hopkins University
         2024 E. Monument Street, Suite 2-600
         Baltimore MD 21287
Phone: 410-955-0495       Fax: 410-955-0476
E-mail: coresh@jhu.edu

3. Timeline: We aim to submit this manuscript to the ARIC publications committee in <12 months from the date of approval of this proposal.
4. Rationale:

Non-Alcoholic Fatty Liver Disease (NAFLD) is an increasingly prevalent condition in industrialized countries (1). Most commonly encountered among individuals with obesity and insulin resistance, NAFLD is caused by excess delivery of free fatty acids to the liver and increased de novo hepatic lipogenesis (2). Previous studies have found associations of NAFLD with each of the components of the metabolic syndrome (3) (low HDL, elevated triglycerides, increased waist circumference, elevated blood pressure and hyperglycemia), with the development of type 2 diabetes mellitus (4) and with increased rates of cardiovascular disease (CVD) (5). However, it is presently unknown whether NAFLD plays a causal role in the development of any of these conditions.

Increasing insight into the heredity underlying NAFLD may also provide us with tools to better understand its metabolic consequences. A recently discovered single nucleotide polymorphism (rs738409[G]) in a gene involved in triglyceride hydrolysis (PNPLA3) has been shown to be associated with a 2-fold increase in hepatic steatosis among homozygotes for the mutation (6). This SNP could be utilized to enhance our understanding of the inter-relationships among NAFLD, insulin resistance syndromes and cardiovascular disease: If significant associations are found between this genetic predisposition to NAFLD and the development of cardiometabolic risk factors and cardiovascular disease, this could indicate a causal relationship. Because obese individuals have the greatest likelihood of NAFLD (7), we anticipate a gene-environment interaction in which a genetic predisposition to NAFLD has the strongest association with future metabolic abnormalities and CVD among individuals with excess weight, as suggested by previous analyses (8).

In this prospective analysis of the Atherosclerosis Risk in Communities Study, we will evaluate the association of a genetic predisposition to NAFLD with the development of the metabolic syndrome, diabetes mellitus and incident cardiovascular disease, and assess whether this association differs among individuals in different weight categories.

5. Main Hypothesis/Study Questions:

Aims:

1) To determine whether a genetic predisposition to NAFLD in PNPLA3 (rs738409[G]) is associated with an increased risk of developing the metabolic syndrome, diabetes mellitus and CVD. We will also examine cross-sectional associations of rs738409[G] with the prevalence of cardiometabolic abnormalities

2) To evaluate whether there is an interaction of weight and weight change with a genetic predisposition to NAFLD in PNPLA3 (rs738409[G]) on the development of cardiometabolic abnormalities.

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 prospective analysis of the association between an allele associated with NAFLD (rs738409[G]) and the incidence of the metabolic syndrome, diabetes mellitus and cardiovascular disease, using ARIC Visit 1 (1987-1989) as the baseline for this analysis. We will also perform cross sectional analyses to assess the associations of rs738409[G] with cardiometabolic risk factors at baseline (ARIC Visit 1) and with abnormalities closely associated with the metabolic syndrome, such as hs-CRP (data from ARIC Visit 4) and LDL and HDL particle characteristics (data from the ARIC Carotid MRI Visit).

**Exposures:** The primary exposures are the genotype for rs738409[G] (allele frequency 0.23 in European Americans and 0.17 in African Americans(6)) and anthropometric measures (body-mass index [BMI] and waist circumference).

**Outcomes:** The primary outcomes of interest are the development of the metabolic syndrome (defined as the presence of 3 of the following factors: abdominal obesity [≥ 102 cm for men or ≥ 88 cm for women], hypertension [systolic blood pressure ≥ 130 mmHg, diastolic blood pressure ≥ 85mmHg, prior physician diagnosis of hypertension, or use of anti-hypertensive medications], elevated fasting glucose [≥ 100 mg/dl], low HDL [< 40 mg/dl for men or < 50 mg/dl for women] and hypertriglyceridemia [≥ 150 mg/dl]); diabetes mellitus (defined as either a self-reported history of diabetes, the use of prescribed hypoglycemic medications, a blood glucose of ≥126 mg/dl after at least 8 hours of fasting, or a blood glucose of ≥200 mg/dl among non-fasting participants); and cardiovascular disease (incident CHD, CHF or stroke) after ARIC Visit 1. Additional secondary outcomes will include metabolic risk factors closely associated with the metabolic syndrome: hs-CRP (measured at Visit 4) and LDL particle size/number and HDL particle number (Carotid MRI Visit). Secondary outcomes will also include the individual components of the metabolic syndrome.

**Exclusions:** We will exclude non-black/non-white participants or those missing data on the primary exposures or covariates of interest. We will also exclude individuals with a history of excess alcohol intake (> 42 drinks per week for men and >28 drinks per week for women). For prospective analyses, we will exclude individuals with the metabolic syndrome, diabetes mellitus and CVD at baseline.

**Covariates:** Age, gender, race, smoking status, alcohol intake, systolic and diastolic blood pressures, blood pressure medication use, LDL cholesterol, physical activity and liver enzymes (ALT, AST and GGT).

**Main Analyses:** After an initial cross sectional analysis, we will focus on prospective associations since very few studies have ARIC’s capability to assess these longitudinal associations. Linear regression, logistic regression and GEE will be used for cross-sectional analyses. Multivariable Poisson and Cox Regression analyses will also be
performed to assess the association of rs738409[G] with incident metabolic syndrome, diabetes and cardiovascular disease, in race and BMI stratified analyses:

1) We will perform univariate comparisons of baseline characteristics among noncarriers, heterozygotes, and homozygotes for the allele rs738409[G], with stratification by race and BMI category (normal weight [BMI 18.5-24.9 kg/m²], overweight [BMI 25-29.9 kg/m²], obese [BMI 30-34.9 kg/m²] and severely obese [BMI >35 kg/m²])

2) In cross sectional analyses, we will use linear and logistic regression to estimate the association of rs738409[G] with cardiometabolic abnormalities at Visit 1. We will also use GEE to estimate the association between rs738409[G] and average values of cardiometabolic risk factors at all 4 ARIC Visits. We will additionally evaluate the cross-sectional relationship between rs738409[G] and inflammatory (hs-CRP) and lipoprotein (LDL particle size and number, HDL particle number) biomarkers closely associated with the metabolic syndrome. Analyses will be stratified by race and BMI category.

3) In prospective analyses, we will use Poisson regression to estimate relative risk ratios and their 95% CIs for the risk of incident metabolic syndrome associated with rs738409[G], with stratification by race and BMI category.

- We will also assess the relationship between rs738409[G] and changes in individual components of the metabolic syndrome, modeled as continuous and categorical (meeting metabolic syndrome criteria) variables

4) We will also use Poisson regression to estimate relative risk ratios and their 95% CIs for the risk of incident diabetes associated with rs738409[G], with stratification by race and BMI category.

5) Cox regression analyses will be used to estimate hazard ratios and 95% CIs for incident CVD (CHD, CHF or stroke) associated with rs738409[G], with stratification by race and BMI category. If rs738409[G] is associated with increased rates of CVD, we will assess the extent to which this is explained by increased rates of the metabolic syndrome and diabetes

6) If differences are observed in the relationship between rs738409[G] and the outcomes of interest across race or BMI categories, we will test interaction terms for statistical significance

7) Finally, we will assess whether the development of metabolic and cardiovascular abnormalities associated with weight gain (from ARIC Visit 1 to Visit 4) differs according to carrier status for rs738409[G]

Sensitivity Analysis: We will perform sensitivity analyses using a higher BMI threshold (≥32 kg/m²) for the definition of obesity among African Americans, as has been suggested by previous analyses.

Limitations:

- Imaging data is not available in ARIC to assess the association between rs738409[G] and hepatic steatosis in this cohort. However, we will perform analyses assessing the association between rs738409[G] and elevated liver
enzymes among individuals without excess alcohol use, as a surrogate for NAFLD.

- We will likely have limited power to assess associations between rs738409[G] and incident cardiovascular events, particularly with stratification by BMI categories

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 ICTDER03 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/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)?

ARIC Manuscript Proposal # 1833: Non-Alcoholic Fatty Liver Disease and the Risk of Incident Cardiovascular Disease: The Atherosclerosis Risk in Communities (ARIC) Study

ARIC Manuscript Proposal # 1824: Nonalcoholic fatty liver disease and myocardial subclinical disease

ARIC Manuscript Proposal # 1789: Elevated Liver Enzymes and Risk of Diabetes
ARIC Manuscript Proposal # 1465: Relationship between single nucleotide polymorphisms previously associated with lipid levels, HDL-C or triglyceride extreme levels, and atherosclerotic events.

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.cscn.unc.edu/aric/forms/](http://www.cscn.unc.edu/aric/forms/)

12. 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.

Reference List


