ARIC Manuscript Proposal #2592

PC Reviewed: 8/11/15 **Status: A** **Priority: 2**
SC Reviewed: _____ **Status: _____** **Priority: ____**

1.a. Full Title: **Dairy Consumption and Body Mass Index (BMI): Mendelian Randomization and Gene-diet Interaction Analyses**

b. Abbreviated Title (Length 26 characters): **Dairy, BMI, Gene, Interaction**

2. Writing Group: **This proposal will include data from the CHARGE consortium, ARIC, and eight other cohorts: [Please insert the name of the cohorts].**

Lu Qi , Tao Huang (lead)
ARIC co-authors and collaborators from other cohorts are listed below.

Cohort	Collaborators
ARIC	Shelly-Ann Love Misa Graff Kari North Gerardo Heiss
NHS	Tao Huang, Ming Ding
HPFS	Tao Huang, Ming Ding
Women Genome Health Initiative	Chu, Audrey Y.,Ph.D.
The Cardiovascular Health Study (CHS)	Rozenn Lemaitre, PhD MPH
The Rotterdam Study	M. Carola. Zillikens，Trudie Voortman
The Family Heart Study	Mary Wojczynski
the Malmö Diet and Cancer study,	Ulrika Ericson
Young Finns Study	Mika Helminen
Framingham	Mary Wojczynski , Adrienne Cupples
MESA	Lekki Wood
InCHIANTI	Tosh
GLACIER	Frida Renstrom
Raine Study: young birth cohort in Australia	Carol Wang
the Danish Diet, Cancer and Health cohort (Danish part of the EPIC study	Tuomas Oskari Kilpeläinen
Danish cohort called Inter99	Camilla Sandholt
the Copenhagen City Heart Study, CCHS	Christina Ellervik
the Copenhagen General Population Study, CGPS	Christina Ellervik
the Danish General Suburban Population Study, GESUS	Christina Ellervik
DESIR: Epidemiological Study on the Insulin Resistance Syndrome cohort	frederic.fumeron
the PREDIMED-Valencia study	M. Dolores Corella Piquer
GOLDN	Smith, Caren E
BPRHS	Smith, Caren E
Health ABC	Denise Houston
CARDIA	Marilyn Cornelis

Note: More coauthors will be included in this study based on contribution.

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

First author: **Tao Huang & Ming Ding**
Address: build 2 room 343, 677 Huntington Ave, Boston, MA 02115
Phone: 4103401626
Fax:
E-mail: nhtah@channing.harvard.edu & mingding517@gmail.com

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 E. North, PhD**
Address: Professor Department of Epidemiology
Carolina Center for Genome Sciences
University of North Carolina Chapel Hill, NC 27599
919-966-2148 (voice)
919-966-9800 (fax)
Kari_north@unc.edu

Name: **Gerardo Heiss, M.D., M.Sc., Ph.D.**
Address: Gillings School of Global Public Health, UNC-CH
Department of Epidemiology
137 East Franklin Street, Suite 306
Chapel Hill, NC 27514
Phone: (919) 962-3253
Email: gerardo_heiss@unc.edu

3. Timeline:
Each cohort will complete their Mendelian randomization and gene-diet interaction analyses and provide their results to Tao Huang as soon as possible. Subsequently, an analyst from Harvard University will conduct the meta-analysis.

4. Rationale:
Dairy consumption has been consistently related to decreased body weight (1,2). However, it is not known whether this association reflects causality, since confounding from lifestyle and socioeconomic factors are difficult to fully take into account in classical observational epidemiological studies. Mendelian Randomization is a newly-developed analytical method addressing causality inference by combining genetic and epidemiological approaches (3).
In order to assess the causal relation between dairy consumption and body weight, we plan a Mendelian Randomization analysis.

1) Chen M, Pan A, Malik VS, Hu FB. Effects of dairy intake on body weight and fat: a meta-analysis of randomized controlled trials. Am J Clin Nutr. 2012 Oct;96(4):735-47. Epub 2012 Aug 29.
2) Abargouei AS, Janghorbani M, Salehi-Marzijarani M, Esmaillzadeh A. Effect of dairy consumption on weight and body composition in adults: a systematic review and meta-analysis of randomized controlled clinical trials. Int J Obes (Lond). 2012 Dec;36(12):1485-93. doi: 10.1038/ijo.2011.269. Epub 2012 Jan 17.
3) Qi L. Mendelian randomization in nutritional epidemiology. Nutr Rev. 2009 Aug;67(8):439-50. doi: 10.1111/j.1753-4887.2009.00218.x.

5. Main Hypothesis/Study Questions:
The main aim of the proposed investigation is to examine the causal effect of dairy consumption on body weight using an established SNP (rs4988235) as the instrumental variable.

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 Criteria

- Sample size: ≥500
- Follow-up time: ≥2 years (*No limitation to maximal follow-up time, please use 10 years as maximal follow-up time if there are repeated measures over time*).

The following analysis plan is for **additive model** for dairy-SNP.
Please also repeat the analysis in **dominant model** (CC vs CT+TT), and **recessive model** (CC+CT vs TT).

Part 1: Mendelian Randomization
Exposure, Outcome, and Instrumental Variable

- **Outcome:** BMI at endpoint (kg/m^2) (follow-up>=2 years)
- **Exposure:** baseline total dairy consumption
    - Total dairy products included 'skim/low fat milk', 'whole milk', 'ice cream', 'yogurt', 'cottage/ricotta cheese', 'cream cheese', 'other cheese', and 'cream'.
- **Instrumental variable:** SNP (rs4988235)
    - SNP rs4988235 Code: TT=2, CT=1, CC=0; T allele is associated with lactase persistence. Please treat SNP as continuous variable.
- **Covariates:** sex, ethnicity, region, years of follow-up, and other baseline covariates if available (age, BMI, smoking status (current vs. former/never), physical activity, total energy intake (kcal), and alcohol consumption).

Note: for covariate region: if the study includes several countries, or USA study includes several states, please control region.

1) **Association between SNP and outcome of BMI**
   Multivariate-adjusted linear model:
   BMI ~ SNP + covariates (age, sex, ethnicity, region)

2) **Association between SNP and baseline dairy intake as an outcome**
   Multivariate-adjusted linear model:
   Total dairy consumption ~ SNP + covariates (age, sex, ethnicity, region)

3) **Association between diary consumption and outcome of BMI**
   Multivariate-adjusted linear model:
   BMI ~ total dairy consumption + covariates

Covariates: sex, ethnicity, region, years of follow-up, and other baseline covariates if available (age, BMI, smoking status, physical activity, total energy intake, and alcohol intake).

Part 2. Gene-Diet Interaction

In this part, we propose to analyze the interaction of rs4988235 and dairy consumption on BMI.

Exposure, Outcome, and Instrumental Variable

- **Outcomes:** BMI at endpoint (kg/m^2) (<u>follow-up>2 years</u>)
- **Diet:** Total dairy consumption [3 categories (tertile1, tertile2, and tertile3) and continuous]
- <u>Total dairy products</u> included 'skim/low fat milk', 'whole milk', 'ice cream', 'yogurt', 'cottage/ricotta cheese', 'cream cheese', 'other cheese', and 'cream'.
- **Genetic variable:** SNP (rs4988235): TT=2, CT=1, CC=0; T allele is associated with lactase persistence. ***(<u>please treat SNP as continuous variable in all models</u>)***
- **Covariates:** sex, ethnicity, region, years of follow-up, and other baseline covariates if available (age, BMI, smoking status, physical activity, total energy intake, and alcohol intake).

Step 1. Interaction model for BMI as an outcome

BMI ~ SNP + Dairy consumption + SNP*Dairy consumption + covariates

- Dairy consumption: continuous variable.
  - SNP (rs4988235): TT=2, CT=1, CC=0; T allele is associated with lactase persistence. Please treat SNP as continuous variable.
- <u>Covariates</u>: sex, ethnicity, region, years of follow-up, and other baseline covariates if available (age, BMI, smoking status, physical activity, total energy intake, and alcohol intake).

Step 2. Stratified analysis on the association between dairy consumption tertiles and the outcome of BMI by SNP (rs4988235)

Please split the data into three subgroups based on tertiles of dairy consumption.

BMI ~ SNP + covariates

- SNP (rs4988235): TT=2, CT=1, CC=0; T allele is associated with lactase persistence. Please treat SNP as continuous variable.
- <u>Covariates</u>: sex, ethnicity, region, years of follow-up, and other baseline covariates if available (age, BMI, smoking status, physical activity, total energy intake, and alcohol intake).

Step 3. Stratified analysis on the association between SNP and the outcome of BMI by diary consumption

Please split the data into three subgroups based on the status of SNP rs4988235 (TT, CT, CC)

BMI ~ Dairy consumption + covariates

- Dairy consumption: (continuous)
- <u>Covariates</u>: sex, ethnicity, region, years of follow-up, and other baseline covariates if available (age, BMI, smoking status, physical activity, total energy intake, and alcohol intake).

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 (We use the genotypic data for SNP rs4988235)

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)?

MS 1005: Associations of dietary calcium, dairy foods and calcium supplementation with anthropometry (lead: E. Nowicki)

MS 1248r: Does the UCP2 Ala55Val Polymorphism Influence the Relation between Dairy Consumption and Weight and Diabetes Risk in a Bi-Ethnic Sample of Adults from the Atherosclerosis Risk in Communities (ARIC) Study (lead: G. George)

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 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.

In one year

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.