ARIC Manuscript Proposal #2398

PC Reviewed: 8/12/14	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title:
Metabolomics and Cognitive Function in Middle-Aged African American Adults: The Atherosclerosis Risk in Communities (ARIC) Study

b. Abbreviated Title (Length 26 characters): Metabolomics and Cognition

2. Writing Group:
Writing group members: Jan Bressler, Bing Yu, Thomas H. Mosley, David S. Knopman, Rebecca F. Gottesman, Alvaro Alonso, Richey A. Sharrett, Lisa Wruck, Eric Boerwinkle (other investigators welcome)

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

First author: **Jan Bressler**
Address: Human Genetics Center
Division of Epidemiology, Human Genetics, and Environmental Sciences
University of Texas School of Public Health
1200 Pressler Street
Houston, TX, 77030
Phone: 713-500-9919 Fax: 713-500-0900
E-mail: jan.bressler@uth.tmc.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
Division of Epidemiology, Human Genetics, and Environmental Sciences
University of Texas School of Public Health
1200 Pressler Street
Houston, TX, 77030
Phone: 713-500-9816 Fax: 713-500-0900
E-mail: Eric.Boerwinkle@uth.tmc.edu

3. Timeline:

Statistical analyses: July 2014 – September 2014
Manuscript preparation: October 2014 – December 2014
Manuscript revision: January 2015

4. Rationale:

Alzheimer's disease (AD) is the most common form of dementia[1] and is characterized by significant impairment in memory, behavioral changes, and gradual loss of autonomy. The prevalence of AD is as high as 20-30% in persons aged 75-84 years, and up to 50% in individuals 85 years of age or older.[2] When 2,800 subjects who were free of dementia were followed for 29 years in the Framingham Heart Study, the lifetime risk for dementia was reported to be 1 in 5 for women and 1 in 10 for men.[3] There is currently no known cure or preventive intervention. Cognitive function is influenced by both genetic and environmental factors.[4,5] The human metabolome is a reflection of the interaction between genes and the environment, and studies examining the relationship between metabolomic profiles and cognitive function may lead to the development of biomarkers used to detect AD and cognitive decline before clinical diagnostic criteria for impairment are met. In this context, Mapstone et al. have recently reported that a set of ten lipids identified in a metabolomics screen in peripheral blood could be used with 90% accuracy to predict conversion from normal cognitive status to amnestic mild cognitive impairment (MCI) or AD over a 2-3-year period in adults 70 years or older.[6] Several other investigators have also found significant alterations in metabolic profiles in comparisons of patients with MCI and AD to cognitively normal subjects.[7-11] The goal of this study is to determine whether metabolites measured in serum in middle-aged African-American adults are associated with cognitive function and cognitive change in the ARIC study. As this racial group is affected disproportionately with AD, this investigation may also provide insight into the biological basis of this health disparity.

5. Main Hypothesis/Study Questions:

This proposed research will use both a cross-sectional and longitudinal design. The following hypotheses will be examined:

1. Baseline cognitive function is associated with metabolites detectable in the current metabolomic profile.
2. Change in cognitive function over a 6-year period and/or a 20-year period is associated with metabolites detectable in the current metabolomics profile.
3. Incident dementia is associated with metabolites detectable in the current metabolomics profile.

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 and Sample

This is a longitudinal study that consists of African American ARIC participants at the baseline (visit 1), visit 2, visit 4, and visit 5 examinations. Metabolomics data was obtained at visit 1. Cognitive function was assessed by three neurospsychological tests (Delayed Word Recall Test (DWRT) [12] the Digit Symbol Substitution Test (DSST),[13] and the Word Fluency Test (WFT)[14,15] in the entire cohort at visits 2, visit 4, and visit 5. All analyses will be performed separately for each cognitive test. A preliminary analysis indicated that there are 1,626 participants with both metabolomics data and cognitive data at visit 2 before the application of any exclusions.

Cognitive change over a 6-year period proximate to the time of measurement of the serum metabolites will be assessed by using change in each of the 3 domain-specific cognitive test scores as a quantitative outcome variable in linear regression models. To assess 20-year cognitive change between visit 2 and visit 5, 3 test-specific z-scores will be calculated for each visit by subtracting the mean test score at visit 2 from each participant's test score and dividing by the standard deviation at visit 2. A global z-score will also be generated by averaging the z-scores of the three tests, subtracting the global mean at visit 2 from each participant's test score, and dividing by its standard deviation at visit 2.[16] Individuals with adjudicated dementia/MCI at visit 5 will initially be considered as cases in the analyses of incident dementia. Other case definitions can also be considered such as combining individuals hospitalized for dementia through 2011 identified using ICD-9 codes (Alzheimer's disease (331.0); vascular dementia (290.4); or other forms of dementia (290.0, 290.1., 290.2, 290.3, 290.9, 294.1, 294.2, 294.8, 294.9, 331.1, 331.2, 331.8, 331.9)[17], and/or study participants identified through telephone interview, with dementia/MCI cases adjudicated at visit 5.

.

Exclusions:

- Non-African Americans
- Missing outcome or covariates information
- History of stroke or transient ischemic attack prior to visit 2, and incident stroke after visit 2 for longitudinal analyses.
-

Outcome Variables:

- Baseline cognitive test scores
- Six-year cognitive change (visit 4 test score – visit 2 test score)
- Twenty-year cognitive change (test specific z-scores)
- Incident dementia

Covariates:

- Age
- Sex
- The kidney filters all metabolites from the blood. Molecules of <10,000 Da molecular weight are freely filtered by the kidney and subsequently metabolized (reabsorbed, catabolized and/or secreted). eGFR at each examination will be calculated based on serum creatinine using the Chronic Kidney Disease

Epidemiology Collaboration (CKD-EPI) equation as follows: $eGFR_{CKD\text{-}EPI} = 141\times$ (minimum of Standard Serum Creatinine [mg/dL]/κ or 1)$^{\alpha}$ × (maximum of Standard Serum Creatinine [mg/dL]/κ or 1)$^{-1.209}$ × 0.993^{age} × (1.018 if female) × (1.159 if black), where κ is 0.7 if female and 0.9 if male; α is -0.329 if female and -0.411 if male.[18]

- Education level

Metabolomics data:
Based on both practical and theoretical considerations, we have placed each measured metabolite into three groups by reliability coefficient (RC; from either the medium-term reliability study or the blind duplicate study) and missing percentage (MS#1847 Zheng Y, et al.).

- Group 1 contains metabolites (n=187) that are reliably measured (RC ≥0.60) and have missing values in fewer than 50% of the sample. The metabolites are to be treated as continuous variables during data analysis with the missingness of metabolites replaced by the lowest measured value.
- Group 2 contains metabolites (n=17) that are reliably measured (RC ≥0.60) but have a moderate amount of missing data (values missing in 50- 80% of the sample). For this group, we consider missing values as category 1. For the measured (non-missing) values, we consider values below the median as category 2 and values above the median as category 3. An ordinal variable is to be used during data analysis.
- Group 3 contains metabolites (n=398) that have >80% missing data or RC <0.6; this group is not included in data analysis.

Statistical Methods:
Linear models will be applied for each metabolite to estimate its association with baseline cognitive function; linear models and Cox proportional hazard models will be used to assess its relationship to 6-year change in cognitive function and incident dementia, respectively. To evaluate the association of each metabolite with 20-year cognitive change, generalized estimating equations (GEE) will be used to account for the correlation between cognitive test scores at repeated assessments assuming an unstructured correlation. The GEE models will include each metabolite, time on study, and their interaction term as well as the covariates listed above. In secondary analyses, regression models will be adjusted for vascular risk factors (diabetes, hypertension, body mass index, low density lipoprotein cholesterol, current smoking, and current alcohol consumption) and *APOE* genotype in addition to the covariates listed above. Statistical significance for the metabolomic data will be pre-specified with an experiment-wise α=0.05 (2 tailed) and a modified Bonferroni procedure will be used to consider the correlations among metabolites.[19,20]

To address possible selection bias where attrition between visit 2 and visits 4 and 5 may be related to cognitive status, baseline characteristics and clinical outcomes will be compared for those who attended the later of the two visits (visit 4 or visit 5) and those who did not. To correct for attrition at visit 5, inverse probability of attrition weighting

(IPAW)[16,21] will be explored following the recommendations of the ARIC Neurocognitive Study (NCS) analysis working group.

All of the analyses described for Aims 1-3 will be performed by Jan Bressler and Bing Yu under the supervision of Eric Boerwinkle; a signed data distribution agreement has been completed.

Limitations:

The limitations of the current study include assessment of the metabolomics data at visit 1 while the three cognitive tests were administered in the entire cohort only at later clinical examinations (visits 2, 4, and 5). Similarly, identification of study participants with incident hospitalized dementia based on hospital discharge diagnoses began at visit 1 and continued through 2011 with the consequence that there will be a single measurement of the metabolomic profile obtained up to 23 years earlier for some subjects. In support of this approach, Yan Zheng *et al.* have previously reported detection of novel biomarkers and pathways significantly associated with incident hypertension[22] and incident heart failure[23] during a 10-year or 20-year follow-up period, respectively, using the same dataset. A second caveat is that the metabolomic profile is available for a subset of African-American participants at visit 1 (n = 1,977) with 1,626 having cognitive data at baseline (visit 2), 1,071 having cognitive data at both visits 2 and 4, and 709 having cognitive data at both visits 2 and 5. Since there may be low power to detect an association due to the sample size confirmatory investigations in other cohorts may be needed.

References:

1. Thies, W., Bleiler, L. & Alzheimer's, A. 2013 Alzheimer's disease facts and figures. *Alzheimers Dement* **9**, 208-45 (2013).
2. Fitzpatrick, A.L. *et al.* Incidence and prevalence of dementia in the Cardiovascular Health Study. *J Am Geriatr Soc* **52**, 195-204 (2004).
3. Seshadri, S. *et al.* The lifetime risk of stroke: estimates from the Framingham Study. *Stroke* **37**, 345-50 (2006).
4. McClearn, G.E. *et al.* Substantial genetic influence on cognitive abilities in twins 80 or more years old. *Science* **276**, 1560-3 (1997).
5. Swan, G.E. *et al.* Heritability of cognitive performance in aging twins. The National Heart, Lung, and Blood Institute Twin Study. *Arch Neurol* **47**, 259-62 (1990).
6. Mapstone, M. *et al.* Plasma phospholipids identify antecedent memory impairment in older adults. *Nat Med* **20**, 415-8 (2014).
7. Trushina, E., Dutta, T., Persson, X.M., Mielke, M.M. & Petersen, R.C. Identification of altered metabolic pathways in plasma and CSF in mild cognitive impairment and Alzheimer's disease using metabolomics. *PLoS One* **8**, e63644 (2013).

8. Kaddurah-Daouk, R. *et al.* Alterations in metabolic pathways and networks in Alzheimer's disease. *Transl Psychiatry* **3**, e244 (2013).
9. Sato, Y. *et al.* Identification of a new plasma biomarker of Alzheimer's disease using metabolomics technology. *J Lipid Res* **53**, 567-76 (2012).
10. Han, X. *et al.* Metabolomics in early Alzheimer's disease: identification of altered plasma sphingolipidome using shotgun lipidomics. *PLoS One* **6**, e21643 (2011).
11. Oresic, M. *et al.* Metabolome in progression to Alzheimer's disease. *Transl Psychiatry* **1**, e57 (2011).
12. Knopman, D.S. & Ryberg, S. A verbal memory test with high predictive accuracy for dementia of the Alzheimer type. *Arch Neurol* **46**, 141-5 (1989).
13. Wechsler, D. *The Adult Intelligence Scale - Revised*, (Psychological Corporation, New York, 1981).
14. Lezak, M. *Neuropsychological Assessment*, (Oxford University Press, Oxford 1995).
15. Tombaugh, T.N., Kozak, J. & Rees, L. Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Arch Clin Neuropsychol* **14**, 167-77 (1999).
16. Gottesman, R.F. *et al.* Impact of differential attrition on the association of education with cognitive change over 20 years of follow-up: the ARIC neurocognitive study. *Am J Epidemiol* **179**, 956-66 (2014).
17. Alonso, A. *et al.* Risk of dementia hospitalisation associated with cardiovascular risk factors in midlife and older age: the Atherosclerosis Risk in Communities (ARIC) study. *J Neurol Neurosurg Psychiatry* **80**, 1194-201 (2009).
18. Levey, A.S. *et al.* A new equation to estimate glomerular filtration rate. *Ann Intern Med* **150**, 604-12 (2009).
19. Sankoh, A.J., Huque, M.F. & Dubey, S.D. Some comments on frequently used multiple endpoint adjustment methods in clinical trials. *Stat Med* **16**, 2529-42 (1997).
20. Blakesley, R.E. *et al.* Comparisons of methods for multiple hypothesis testing in neuropsychological research. *Neuropsychology* **23**, 255-64 (2009).
21. Weuve, J. *et al.* Accounting for bias due to selective attrition: the example of smoking and cognitive decline. *Epidemiology* **23**, 119-28 (2012).
22. Zheng, Y. *et al.* Metabolomics and incident hypertension among blacks: the atherosclerosis risk in communities study. *Hypertension* **62**, 398-403 (2013).
23. Zheng, Y. *et al.* Associations between metabolomic compounds and incident heart failure among African Americans: the ARIC Study. *Am J Epidemiol* **178**, 534-42 (2013).

7. a. Will the data be used for non-CVD analysis in this manuscript? _X_ Yes _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?__X_ 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**

8.c. If yes, is the author aware that the participants with RES_DNA = 'not for profit' restriction must be excluded if the data are used by a for profit group?
__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

Yes. There is no overlap between this proposal and other ARIC manuscript proposals or published manuscripts.

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#1847 Zheng Y et al. Role of the Human Metabolome in Incident Heart Failure Etiology among African Americans in the Atherosclerosis Risk in Communities (ARIC) Study

MS#1853 Yu B et al. Genome-Wide Association Study of Heart Failure-related Human Metabolite Profiles among African Americans in the Atherosclerosis Risk in Communities (ARIC) Study.

MS#1882 Yu B et al. A Longitudinal Study of Metabolomics and Kidney Function among African Americans in the Atherosclerosis Risk in Communities (ARIC) Study

MS#2056 Zheng Y et al. A Medium-Term Reliability Study of the Human Serum Metabolome: the Atheroslcerosis Risk in Communities (ARIC) Study

MS#1918 Zheng Y et al. Associations of the Human Metabolome with Blood Pressure, Prevalent and Incident Hypertension among African-Americans in the Atherosclerosis Risk in Communities (ARIC) Study

MS#1982 Gottesman R et al. Estimation of cognitive change from repeat measures in observational studies; associations with education: the ARIC NCS

MS# 2120 Knopman D et al. Prevalence of Mild Cognitive Impairment and Dementia and Their Relationship to Diabetes and Hypertension in ARIC

MS#2145 Dearborn J et al. Nutrition, Healthy Diet and 21-year Cognitive Change

MS#2354 Alonso A et al. Metabolomics and Incident Atrial Fibrillation in African-Americans: the ARIC study

11. a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? Yes

11.b. If yes, is the proposal

__x__ A. primarily the result of an ancillary study (list number*) 2008.16 Metabolomics and Heart Failure: A Novel Approach to Biomarker Discovery

___ 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/

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.

Yes, the lead author is aware that manuscript preparation is expected to be completed in 1-3 years, and if this expectation is not met, the manuscript proposal will expire.