ARIC Manuscript Proposal # 3675 (Revised)

PC Reviewed: 10/13/20	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Pre-statistical Harmonization of Cognitive Measures across Six Population-Based Cohorts: ARIC, CARDIA, CHS, FHS, MESA, and NOMAS

b. Abbreviated Title (Length 26 characters): Harmonization of Cognitive Measures

2. Writing Group:

Writing group members: Alden Gross, Ph.D., Johns Hopkins, Baltimore, Maryland, USA, agross14@jhu.edu; Bruno Giordani, Ph.D., University of Michigan, Ann Arbor, MI, USA, giordani@med.umich.edu; Jennifer Manly, PhD., Columbia University, New York City, New York, USA, jjm71@columbia.edu; Rebecca Gottesman, MD., Ph.D., Johns Hopkins, Baltimore, Maryland, USA, rgottesm@jhmi.edu; Mitchell S.V. Elkind, MD, Columbia University, New York City, New York, USA, mse13@cumc.columbia.edu; Steve Sidney, MD, MPH, Kaiser Permanente, Oakland, CA, steve.sidney@kp.org; Stephanie Hingtgen, MPP, University of Michigan, Ann Arbor, MI, USA, smhing@med.umich.edu; Ralph L. Sacco, MD, MS, University of Miami, Miami, FL, USA, RSacco@med.miami.edu; Clinton Wright, MD, MS, wright.clinton@gmail.com; Kristine Yaffe, MD, Kristine.Yaffe@ucsf.edu; Deborah A. Levine, MD., University of Michigan, Ann Arbor, MI, USA, deblevin@med.umich.edu

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

First author: Emily M. Briceño
University of Michigan Medical School
Department of Physical Medicine & Rehabilitation
Rehabilitation Psychology & Neuropsychology
Address: 325 E. Eisenhower
Ann Arbor, MI 48108-1605

Phone: 734-936-7052 Fax: 734-936-7048
E-mail: emilande@med.umich.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: **Alden Gross**

Address: Epidemiology
Center on Aging and Health, Rm 2721
2024 E. Monument St.
Baltimore, MD 21205

Phone: 443-287-7196 Fax: 410-614-9625
E-mail: agross14@jhu.edu

Name: **Rebecca Gottesman**
Address: Neurology
Phipps 446D
600 North Wolfe Street
Baltimore, MD 21287

Phone: 410-614-2381 Fax:
E-mail: rgottesm@jhmi.edu

3. Timeline: Manuscript has been drafted and will be submitted pending approval of this proposal.

4. Rationale:

Data harmonization methods provide powerful approaches for combining multiple large, population-based data sets, and have been increasingly applied to cognitive data (Gross et al., 2014). Pooling cognitive data from longitudinal population-based cohorts facilitates improved power and novel capabilities to investigate biomedical, lifestyle, and cultural factors that may affect cognition across the lifespan (2,3). Harmonization endeavors require pre-statistical harmonization steps to identify common tests and/or test items across datasets.

Pre-statistical harmonization (Griffith et al., 2013) is a complicated, qualitative component of the harmonization process that precedes pooling of data from different studies. It involves a careful review of cohort characteristics (e.g., subject selection procedures, demographic factors such as race/ethnicity, socioeconomic status, language) and of available variables of interest. Pre-statistical harmonization also involves review of cognitive instruments

and associated data in individual datasets to identify common and unique test items across datasets, as well as candidate items that might be made comparable with minimal transformation.

A critical challenge in harmonizing large population-based datasets is the considerable variability that exists in the instruments used to measure cognition, and the procedures for administering them. Even seemingly equivalent cognitive instruments may exhibit differences in implementation across studies that could impact test score interpretation and suitability for data pooling. Variability such as differences in test version, administration and scoring rules, and selection of summary scores for specific instruments may contribute to such differences. Despite these challenges, this critical step of the harmonization process is rarely described. We aim to describe our procedures and findings from this critical step of the harmonization process.

5. Main Hypothesis/Study Questions:

The goal of the present paper is to describe the procedures and findings from the pre-statistical harmonization of cognitive instruments for six longitudinal population-based cohorts: the Atherosclerosis Risk in Communities (ARIC) study, the Coronary Artery Risk Development in Young Adults study (CARDIA), Cardiovascular Health Study (CHS), Framingham Offspring Study (FOS), Multi-Ethnic Study of Atherosclerosis (MESA), and the Northern Manhattan Study (NOMAS). Specifically, we will: 1) describe findings from our detailed review of the administration, scoring procedures, and score ranges of cognitive instruments across cohorts to determine their degree of equivalence and suitability for pooling; 2) examine the impact of procedural differences on test score equivalence across cohorts; and 3) offer recommendations for the pre-statistical harmonization process for future studies.

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).

Cohort studies: Six NIH-funded, longitudinal cohort studies will be included in the present study: The Atherosclerosis Risk in Communities Study (ARIC (5)); The Coronary Artery Risk Development in Young Adults Study (CARDIA (6)), the Cardiovascular Health Study (CHS (7)); the Framingham Offspring Study (FOS (8)), the Multi-Ethnic Study of Atherosclerosis (MESA (9)) and the Northern Manhattan Study (NOMAS (10)). A description of each cohort study is available (11).

Procedure: Neuropsychologists (EB, BG) will identify cognitive instruments in each of the six cohort studies from documentation provided by each cohort, from available published papers, and from scrutiny of provided datasets. Cohort study investigators will be contacted to request unpublished administration, scoring, and procedural details of cognitive test batteries. Cognitive tests will be reviewed and categorized into relevant cognitive domains (e.g., memory, executive functioning) by neuropsychologists (EB, BG). All cognitive instruments will be reviewed for inclusion in the harmonization, with the exception of instruments that measure crystallized abilities that are known to be insensitive to age-related cognitive decline (e.g., measures of literacy and premorbid intellectual functioning) in light of study aims. For cognitive instruments for which composite test items are available (e.g., individual test items for the Mini Mental State Examination (MMSE); Telephone Interview for Cognitive Status (TICS), Modified Mini Mental State Examination (3MSE), or Montreal Cognitive Assessment (MoCA)), each test item will be reviewed. Common and unique instruments and test items across cohorts will be identified.

After available cognitive instruments were identified, procedural details will be gleaned from documentation provided by each cohort study, including test forms, data entry forms, and administration and scoring instructions. Procedural details extracted from this process will include the published test version, administration and scoring details (e.g., stopping rules; acceptable responses for specific items), possible raw score ranges (based upon the instrument structure and number of items), and metrics available for each instrument (e.g., individual item data, raw and standardized summary scores). Available raw data for each instrument will also be reviewed for score ranges and distributions. When procedural or distributional differences across cohorts in common test items are identified, each item will be reviewed for possible data transformation.

Data analysis: The findings from this detailed review, including procedural differences amongst common cognitive instruments, will be summarized in tables. We will select a representative example of an instrument that is administered across two or more cohorts and will illustrate the impact of procedural differences on test score distributions.

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.cscc.unc.edu/aric/mantrack/maintain/search/dtSearch.html

__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)?

Co-author Rebecca Gottesman has published articles on BP and cognition using ARIC data. Co-author Alden Gross has published articles on harmonization of cognitive measures using ARIC data.

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* 2008.06, ARIC-NCS, PI Coresh)

___ 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 https://www2.cscc.unc.edu/aric/approved-ancillary-studies

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

1. Gross AL, Jones RN, Fong TG, Tommet D, Inouye SK. Calibration and validation of an innovative approach for estimating general cognitive performance. *Neuroepidemiology*. 2014;42(3):144-153. doi:10.1159/000357647.
2. Hofer SM, Piccinin AM. Toward an Integrative Science of Life-Span Development and Aging. *Journals Gerontol Ser B Psychol Sci Soc Sci*. 2010;65B(3):269-278.

doi:10.1093/geronb/gbq017.
3. McArdle JJ, Grimm KJ, Hamagami F, Bowles RP, Meredith W. Modeling life-span growth curves of cognition using longitudinal data with multiple samples and changing scales of measurement. *Psychol Methods*. 2009;14(2):126-149. doi:10.1037/a0015857.
4. Griffith L, van den Heuvel E, Fortier I, et al. Harmonization of Cognitive Measures in Individual Participant Data and Aggregate Data Meta-Analysis. In: *Methods Research Report. (Prepared by the McMaster University Evidence-Based Practice Center under Contract No. 290-2007-10060-I.) AHRQ Publication No.13-EHC040-EF. Rockville, MD: Agency for Healthcare Research and Quality.* ; 2013.
5. The ARIC Investigators. The atherosclerosis risk in communities (ARIC) study: Design and objectives. *Am J Epidemiol*. 1989;129(4):687-702. doi:10.1093/oxfordjournals.aje.a115184.
6. Friedman GD, Cutter GR, Donahue RP, et al. CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol*. 1988;41(11):1105-1116. doi:10.1016/0895-4356(88)90080-7.
7. Fried LP, Borhani NO, Enright P, et al. The Cardiovascular Health Study: Design and rationale. *Ann Epidemiol*. 1991;1(3):263-276. doi:10.1016/1047-2797(91)90005-W.
8. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham offspring study. Design and preliminary data. *Prev Med (Baltim)*. 1975;4(4):518-525. doi:10.1016/0091-7435(75)90037-7.
9. Bild DE, Bluemke DA, Burke GL, et al. Multi-Ethnic Study of Atherosclerosis: Objectives and design. *Am J Epidemiol*. 2002;156(9):871-881. doi:10.1093/aje/kwf113.
10. Sacco RL, Boden-Albala B, Gan R, et al. Stroke incidence among white, black, and Hispanic residents of an urban community: The Northern Manhattan Stroke Study. *Am J Epidemiol*. 1998;147(3):259-268. doi:10.1093/oxfordjournals.aje.a009445.