ARIC Manuscript Proposal #3949

PC Reviewed: 10/12/21	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Social Determinants of Health-Based Cognitive Normative Data

b. Abbreviated Title (Length 26 characters): Non-Race Based Cognitive Norms

2. Writing Group:

Andrea L.C. Schneider (First Author) (University of Pennsylvania)
Rebecca F. Gottesman (NINDS)
Lisa Wruck (Duke University)
James Russell Pike (University of North Carolina)
Roland Faigle (Johns Hopkins University)
Keenan Walker (NIA)
Anna Kucharska-Newton (University of North Carolina)
James Henegan (University of Mississippi Medical Center)
Nicole Williams (NINDS)
Josef Coresh (Johns Hopkins University)
Michael Griswold (University of Mississippi Medical Center)
Thomas Mosley (Senior Author) (University of Mississippi Medical Center)
Others Welcome

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

First author: **Andrea Lauren Christman Schneider, MD, PhD**
Address: 51 North 39th Street, Andrew Mutch Building 416
Philadelphia, Pennsylvania 19104

Phone: 443-827-2352
E-mail: Andrea.Schneider@pennmedicine.upenn.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: **Thomas Mosley, PhD**
Address: 2500 North State Street
Jackson, Mississippi 39216

Phone: 601-984-4467
E-mail: tmosley@umc.edu

3. Timeline: Data for analyses are currently available. Data analysis and manuscript preparation and submission will take place over one year from manuscript proposal acceptance (2021-2022).

4. Rationale:

In order to diagnose cognitive impairment, an individual's performance on cognitive testing is compared to estimated baseline cognitive abilities, which are derived from the test performance of healthy individuals (i.e., normative data). Importantly, these normative data adjust for demographic factors that are associated with performance (e.g., age, education) by either comparing an individual to persons who are demographically similar or by adjusting the normative data based on associations with these demographic factors. Historically, normative data also included adjustment for race (e.g., Heaton norms[1]) as a way to reduce harms that might result from the over identification of cognitive impairment among Black, as compared to White individuals. Accordingly, the cognitive normative data currently used in the ARIC Study are also race-specific[2].

The recent National Football League (NFL) Player's Concussion Injury Litigation case has brought national attention to the weaknesses of using race-adjusted cognitive normative data[3]. When race-adjusted normative data is used, the assumption is that Black men and women start at a lower cognitive baseline than White men and women. Therefore, a Black individual with the same cognitive score as a White individual is assumed to have experienced less cognitive impairment. It is now widely accepted that race is a crude surrogate for lifetime social experiences and prior studies have shown that adjusting cognitive test performance for social determinants of health significantly reduces variance explained by race[3]. In this setting, it has been proposed that race-based normative data should be replaced by regression-based normative approaches that adjust for social determinants of brain health, including education, literacy, psychosocial stress, occupation, economic/financial status, residential characteristics, language, and nativity/acculturation[3].

To this end, we propose to develop social determinants of health-based cognitive normative data as an alternative to race-based normative data in the ARIC Study. Specifically, we will develop social determinants of health-based norms for cognitive function assessed at ARIC Visit 5 using social determinants of health, rather than race, as the adjustment factors and will then compare the performance of these social determinants of health-based norms to the performance of the race-based norms using cognitive domain and algorithm-based cognitive status data from ARIC Visit 6.

5. Main Hypothesis/Study Questions:

1. To develop 5 different sets of cognitive normative data at ARIC Visit 5:
    a. <u>Race-based cognitive normative data</u>:
        i. Derived using age, education, and race (Black/White)
    b. <u>Social determinants of health-based cognitive normative data</u>:
        i. Derived using age and education

ii. Derived using age, education, and the Wide Range Achievement Test (WRAT-3)
iii. Derived using age, education and WRAT-3, and area deprivation index (ADI)
iv. Derived using age, education, and ADI

2. To compare the performance of the social determinants of health-based norms to the performance of the race-based norms at ARIC Visit 6:
    a. We will compare the agreement for meeting criteria for "cognitive domain failure" in each cognitive domain (i.e., having a domain Z score relative to the normative sample of worse than -1.5) using race-based normative data versus each of the social determinants of health-based normative data.
    b. We will compare the agreement for the algorithm-based cognitive status (normal, MCI, or dementia) using race-based normative data versus each of the social determinants of health-based normative data.

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 Population and Inclusion/Exclusion Criteria:
Race-based and social determinants of health-based cognitive norms will be derived using ARIC Visit 5 data. The comparison of the performance of social determinants of health-based versus race-based norms will be performed using ARIC Visit 6 data.

Visit 5: The study population for the derivation of cognitive normative data analysis will be created using similar exclusions to our prior ARIC paper on race-based cognitive normative data[2] and will match the population used for the creation of the Visit 5 race-based normative data. After applying the following exclusion criteria, our normative population will have a sample size of 2,609 (per ARIC Visit 6 Manual 17, Page 9):

1. Exclusions due to clinical neurologic disease:
    a. History of stroke hospitalization as of Visit 5
    b. History of neurologic disease at or before Visit 5 (multiple sclerosis, Parkinson's disease, brain tumor)
2. Exclusions due to diagnosed or self-reported memory problems or factors affecting cognition at Visit 5:
    a. Using medications for dementia at Visit 5
    b. Low MMSE (pro-rated MMSE <22; scored at 30*[number correct] / [30-number not answered])
    c. Self-report memory problems at Visit 5 identified on the Subjective Memory Form (SMF) responses to questions 1 and 3 is "often" (3) or "very often" (4)
    d. Dementia ICD codes prior to Visit 5 (290.x, 294.0x, 294.1x, 294.2x, 294.9x, 331.0x, 331.1x, 331.2x, 331.7x, 331.9x, 331.8, 331.82, 331.83, 331.89)
    e. Diagnosis of Level 3 definition dementia at Visit 5
    f. Depression (CES-D ≥8)
    g. Two APOE ε4 alleles

    h. Substantial decline on Delayed Word Recall Test (DWRT), Digit Symbol Substitution Test (DSST), or Word Fluency Test (WFT) (change defined at Visit 5 score minus the mean of the Visits 2 and 4 scores, excluded if change score was in worse 10^{th} percentile on any one test or between 10^{th} and 20^{th} percentile on at least 2 tests – DWRT 10^{th} percentile: -3.5, 20^{th} percentile: -3 – DSST 10^{th} percentile: -18.5, 20^{th} percentile: -14.5 – WFT 10^{th} percentile: -11.5, 20^{th} percentile: -8).
    i. Diagnosis of MCI or unknown cognitive status at Visit 5
3. Exclusions based on information collected after Visit 5
    a. Semi-annual follow-up General Interview Version A (GEN) response to question 1a (Alzheimer's disease), 1c (memory loss or cognitive impairment), 1d (dementia, vascular dementia, or hardening of the arteries of the brain) of "Yes"
    b. Dementia death code (F00, F00.0, F00.1, F00.2, F00.9, F01, F01.1, F01.2, F01.3, F01.5, F01.50, F01.51, F01.8, F01.9, F02, F02.0, F02.1, F02.3, F02.4, F02.8, F02.80, F02.81, F03, F03.9, F03.90, F03.91, F05.1, F06.7, G31.0, G31.1, G31.09, G31.83, G31.84, G30, G30.0, G30.1, G30.8, G30.9)
    c. Positive dementia surveillance using the Six-item Screener (SIS) and the Ascertain Dementia 8-Item Questionnaire (AD8)
    d. Memory problems self-reported at annual follow-up after Visit 5 (MCU 13a (Alzheimer's disease), 13c (memory loss or cognitive impairment), 13d (dementia, vascular dementia, or hardening of the arteries of the brain)
4. Other exclusions:
    a. Non-white/Non-black Race
    b. Missing education
    c. Missing Wide Range Achievement Test 3^{rd} edition (WRAT-3) or score <10

Visit 6: For the analysis comparing the performance of the social determinants of health-based norms to the performance of the race-based norms, we will use the ARIC Visit 6 population excluding individuals who are non-white/non-black race and blacks at the Minnesota or Maryland field centers.

Social Determinants of Health Variables:
In addition to our race-based cognitive normative data (created using age, education and race), we propose to create 4 unique sets of non-race based cognitive normative data using age plus different combinations of the following social determinants of health variables:

- Education: Education was assessed at ARIC Visit 1 and will be categorized in 3 groups: < high school versus high school, general educational development (GED), or vocational school versus some college, college, graduate, or professional school. We will also explore the 6-group education categorization in our analyses (grade school or no education versus high school no degree versus high school graduate versus vocational school versus college versus graduate or professional school).
- Area Deprivation Index (ADI): The ADI[4, 5] was originally created by the U.S. Health Resources and Services Administration. It is composed of 17 education, employment, housing-quality, and poverty measured drawn from the long-form Census data and updated to incorporate American Community Survey data. The ADI is currently freely available to download on the University of Wisconsin's Neighborhood Atlas website

(https://www.neighborhoodatlas.medicine.wisc.edu). This data has already been cross-linked with the ARIC cohort by investigators of the ARIC outcomes research group, represented herein by Dr. Anna Kucharska-Newton who is included as co-author on this manuscript proposal. We will use the ADI data from ARIC visit 5.

- Wide Range Achievement Test 3rd Edition (WRAT-3): The WRAT-3[6] is a single-word reading test of literacy, which is widely considered a measure premorbid cognitive functioning. WRAT-3 (score range 0-57) was administered at ARIC visit 5.

Cognitive Domains and Component Tests:
We will create race-based and non-race based normative data at Visit 5 for the cognitive domains of memory, language, and executive function (domains defined previously[7]). We will then compare the performance of social determinants of health-based and race-based normative data at ARIC Visit 6 (compare the agreement for meeting criteria for "failure" in each cognitive domain by each normative definition (i.e., having a domain Z score relative to the normative sample of worse than -1.5).

The memory domain is comprised of the following tests: the Delayed Word Recall Test (DWRT)[8], Logical Memory (LM) I and II[9], and Incidental Learning Digit-Symbol Pairs[10]. The language domain is comprised of the following tests: The Word Fluency Test (WFT)[11], Animal Naming[12], and the Boston Naming Test[13]. The executive function domain is comprised of the following tests: Trail Making Test Parts A and B[14], and the Digit Symbol Substitution Test[10].

Cognitive Status Algorithmic Diagnosis:
In addition to comparing the performance of social determinants of health-based and race-based normative data at ARIC Visit 6 by evaluating the agreement for meeting criteria for "failure" in each cognitive domain, we will also compare performance by evaluating the agreement for algorithmic diagnosis (normal, MCI, dementia). The computer algorithmic diagnoses are found on pages 18-19 of ARIC Visit 6 Manual 17. In this analysis we will use a modified algorithm, using only stratum 3-16 (excluding the race-based pro-rated MMSE score and a prior dementia diagnosis at visit 5 [stratum 1-2]) in order to best test our different normative definitions:

Stratum	Decline	Failed Domain	CDR Sum of Boxes	FAQ	Algorithm Diagnosis
3	N	Any	uncollected	uncollected	Normal
4	Y or Y due to missing	0	uncollected	uncollected	Normal
5	Y or Y due to missing	1 failed or at least 1 missing	0, missing	≤5, missing	MCI
6	Y or Y due to missing	1 failed or at least 1 missing	0	>5	Prob MCI
7	Y or Y due to missing	1 failed or at least 1 missing	>0 but ≤3	≤5, missing	MCI

8	Y or Y due to missing	1 failed or at least 1 missing	>0 but ≤3	>5	Prob MCI
9	Y or Y due to missing	1 failed or at least 1 missing	>3	≤5	Prob Dementia
10	Y or Y due to missing	1 failed or at least 1 missing	>3	>5, missing	Prob Dementia
11	Y or Y due to missing	>1	0, missing	≤5, missing	MCI
12	Y or Y due to missing	>1	0	>5	Prob MCI
13	Y or Y due to missing	>1	>0 but ≤3	≤5	MCI
14	Y or Y due to missing	>1	>0 but ≤3	>5, missing	Prob MCI
15	Y or Y due to missing	>1	>3	≤5	Prob Dementia
16	Y or Y due to missing	>1	>3	>5, missing	Dementia

Statistical Analyses:
Visit 5 characteristics will be shown overall for our normative analytic sample (N=2,609) using means (SDs) for continuous variables and n (%) for categorical variables. Using the Visit 5 normative population, we will develop the race-based and 4 different social determinants of health-based cognitive normative data. The race-based norms will use age, education and race and the 4 different sets of social determinants of health-based norms will use age plus the following variables: 1) education only, 2) education and WRAT-3, 3) education and WRAT-3, and ADI, and 4) education and ADI. For each definition, adjusted linear regression models will be run for each domain score (memory, language, executive function). Age will be continuous, race will be binary (white, black), education will be categorical (< high school, high school or equivalent, > high school), WRAT-3 will be continuous, and ADI will be continuous. All continuous variables will be centered at the median of the distribution. Regression coefficients and RMSE will be tabulated for each model.

Visit 6 characteristics will be shown for the "test" population in which we will compare the performance of the social determinants of health-based norms to the performance of the race-based norms. Each participant's Visit 6 domain scores (memory, language, executive function), normed to visit 5, will be calculated using a formula of weighted sums based on factor analysis of Visit 5 data, as described previously[7] and as shown in ARIC Visit 6 Manual 17, page 12 (coefficients derived without regard to race). These calculated domain Z scores will then be converted to Z scores relative to the normative sample as Z score minus predicted mean from normative sample divided by RMSE from the linear regression model adjusted for the variables in each set of norms. We will assess correlations between and scatterplots with localized regression lines for domains Z-scores derived from each set of normative data. The Z scores

relative to each normative sample will then be compared to a cut-point of -1.5 to determine "cognitive domain failure" for each domain. We will calculate percent agreement between race-based and each social determinants of health-based normative definition. In addition to examining the percent agreement for domain failure between race-based and each social determinants of health-based normative definition, we will additionally calculate Cohen's Kappa. Next, we will apply the modified algorithmic diagnoses using each set of normative data and will compare performance of race-based and each social determinants of health-based normative definition by evaluating the percent agreement for algorithmic diagnosis (normal, MCI, dementia). In addition to examining the percent agreement for algorithmic diagnoses between race-based and each social determinants of health-based normative definition, we will additionally calculate Cohen's Alpha (will treat normal versus dementia and MCI versus dementia as equally incorrect) and Krippendorff's Alpha[15] (takes into account that MCI versus dementia is more correct than normal versus dementia). We will additionally consider risk reclassification analyses (e.g., net reclassification index and integrated discrimination improvement).

In sensitivity analyses, we will consider creating normative data for each cognitive test separately (rather than for each cognitive domain). We can additionally consider re-creating our normative data excluding individuals from our normative sample who were later found to have dementia at visit 6. In supplemental analyses, we will also consider the incorporation of other measures of life-course socioeconomic status[16] (instead of the composite ADI) in our normative definitions (including data on occupation, occupational role, home ownership, family income during childhood, young adulthood, and middle/older adulthood).

Limitations: One important limitation of the proposed social determinants of health-based cognitive test normative data that we will create in this manuscript is the age range to which they are applicable. The age range at ARIC Visit 5 is 66 years to 90 years. Using this population, we will not be able to create normative data for younger or older ages. A second limitation is that our race-based norms are inherently geography-based norms due to the ARIC Study recruitment of whites at MD, MN, and NC and blacks at NC and MS. Another limitation is the generalizability of our cognitive normative data to other population as our population consists of individuals who voluntarily agreed to be participants in the ARIC Study and are likely different from members of the general community in several important ways (e.g., they may have more interest in their health, be more educated, and/or be in better general health).

7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? ____ Yes _X_ No

b. If Yes, is the author aware that the current derived consent file ICTDER05 must be used to exclude persons with a value RES_OTH and/or RES_DNA = "ARIC only" and/or "Not for Profit" ? ____ Yes ____ No
(The 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
APOE ε4 genotype

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the current derived consent file ICTDER05 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/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)?

- MSP #2109: Normative data for eight neuropsychological tests for blacks and whites from the Atherosclerosis Risk in Communities Study (Andrea Schneider)
- MSP 2021B: Mid-life vascular risk factors for Mild Cognitive Impairment in the ARIC NCS Study (David Knopman)
- MSP 2120C: Incidence of Dementia and its relationship to midlife vascular risk factors in ARIC (Rebecca Gottesman)

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 https://sites.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.

Understood.

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.

Understood.

References

1. Heaton RK, Psychological Assessment Resources Inc. Revised comprehensive norms for an expanded Halstead-Reitan battery : demographically adjusted neuropsychological norms for African American and Caucasian adults, professional manual, Updated ed. Lutz, Fla.: Psychological Assessment Resources, 2004.
2. Schneider AL, Sharrett AR, Gottesman RF, et al. Normative data for 8 neuropsychological tests in older blacks and whites from the atherosclerosis risk in communities (ARIC) study. Alzheimer Dis Assoc Disord 2015;29:32-44.
3. Possin KL, Tsoy E, Windon CC. Perils of Race-Based Norms in Cognitive Testing: The Case of Former NFL Players. JAMA Neurol 2021;78:377-378.
4. Kind AJH, Buckingham WR. Making Neighborhood-Disadvantage Metrics Accessible - The Neighborhood Atlas. N Engl J Med 2018;378:2456-2458.
5. Zuelsdorff M, Larson JL, Hunt JFV, et al. The Area Deprivation Index: A novel tool for harmonizable risk assessment in Alzheimer's disease research. Alzheimers Dement (N Y) 2020;6:e12039.
6. Wilkinson G. The Wide Range Achievement Test 3rd Edition Manual. Wilmington, DE: Wide Range, 1993.
7. Gross AL, Power MC, Albert MS, et al. Application of Latent Variable Methods to the Study of Cognitive Decline When Tests Change over Time. Epidemiology 2015;26:878-887.
8. Knopman DS, Ryberg S. A verbal memory test with high predictive accuracy for dementia of the Alzheimer type. Arch Neurol 1989;46:141-145.
9. Wechsler D. Wechsler Memory Scale Revised Manual. San Antonio: Psychological Corp, 1987.
10. Wechsler D. Wechsler Adult Intelligence Scale (WAIS) Manual – Third Edition. New York: Psychological Corp, 1997.
11. Benton AL, Hamsher K. Multilingual Aphasia Examination, 2nd Edition. Iowa City: AJA Associates, 1989.
12. Benton A. Differential Behavioral Effects in Frontal Lobe Disease. Neuropsychologia 1968;6:53-60.
13. Williams BW, Mack W, Henderson VW. Boston Naming Test in Alzheimer's disease. Neuropsychologia 1989;27:1073-1079.
14. Reitan RM. Validity of the Trail Making Test as an Indicator of Organic Brain Damage. Percept Mot Skills 1958;8:271-276.
15. Hayes AF, Krippendorff K. Answering the Call for a Standard Reliability Measure for Coding Data. Communication Methods and Measures 2007;1:77-89.
16. George KM, Lutsey PL, Kucharska-Newton A, et al. Life-Course Individual and Neighborhood Socioeconomic Status and Risk of Dementia in the Atherosclerosis Risk in Communities Neurocognitive Study. Am J Epidemiol 2020;189:1134-1142.