ARIC Manuscript Proposal #4094

PC Reviewed: 8/9/22	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Investigating relationships between an MRI measure of brain aging with proteomics and cognition.

b. Abbreviated Title (Length 26 characters): **Brain aging and proteomics**

2. Writing Group:

Barnard Ryan T.
Casanova Ramon
Justice Jamie N.
Hughes Timothy M.
Kritchevsky Stephen B.
Lu Lingyi
Prizment Anna
Sedaghat Sanaz
Wagenknecht Lynne E.
Walker Keenan A.

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

First author: Ramon Casanova, PhD
Address: Department of Biostatistics and Data Science
Wake Forest School of Medicine
Winston Salem, NC, 27157
Phone: 336 716 8309
Email: casanova@wakehealth.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: **Lynne Wagenknecht.**
Address: Wake Forest School of Medicine
Division of Public Health Sciences
Wake Forest School of Medicine
Winston Salem, NC, 27157

Phone: 336 716 8309
Phone: Fax:
E-mail: lynne.wagenknecht@wakehealth.edu

3. Timeline:

Expectations are for analyses to be completed within 1 year.

4. Rationale:

There is an increasing body of work based on the use of machine learning and artificial intelligence[1-4] to estimate chronological age using neuroimaging data. The **gap between chronological age and estimated brain age (BAG)** is then used as a measure of brain accelerated/resilient aging. These estimates are often based on structural MRI because it is more available, less invasive and cheaper compared to other brain imaging modalities such as Positron Emission Tomography (PET). Accelerated brain aging calculated in this fashion has been shown to be associated with smoking and alcohol consumption[5] and cognitive impairment and progression to AD[6, 7] among other examples. Cole et al have reported that accelerated brain aging was associated with measures of physical function, health status and mortality[8].

On the other hand, there is an increased interest in proteomics to investigate human aging[9-11] and brain disease[12]. Proteomic clocks have been devised[9, 10, 13, 14] which facilitate the biological interpretation of the results when compared to epigenetic clocks. Ferruci and Tanaka used the SomaScan proteomics platform to measure 1306 proteins in 240 healthy men and women between 22 and 93 years[14]. They found 197 proteins to be positively associated and 20 proteins negatively associated with age. Sathyan et al. also based on elastic net regression models, created a proteomic signature of age based on relative concentrations of 76 proteins that highly correlated with chronological age ($r = 0.94$)[15]. By utilizing the SomaScan® proteomic platform in 1,025 participants of the LonGenity cohort (age range: 65-95, 55.7% females), they found that 754 of 4,265 proteins were associated with chronological age. Pleiotrophin, WNT1-inducible-signaling pathway protein 2, chordin-like protein 1, transgelin and R-spondin-1 were the proteins most significantly associated with age. The correlation between proteomic age prediction based on elastic net regression and chronological age was 0.8 ($p < 2.2E-16$). Johnson et al. proposed a novel proteomic aging clock comprised of proteins that were reported to change with age in plasma in three or more different studies. Using a large patient cohort comprised of 3,301 subjects (aged 18-76 years), they demonstrated that this clock is able to accurately predict human age[9].

Several groups have linked proteomics data to brain structural MRI. Harris and colleagues found that, in the Lothian Birth Cohort, that the associations between EDA2R and general fluid cognitive ability were mediated by total brain volumes[16]. Shi and colleagues identified 17 plasma proteins related to hippocampal volume and 2 correlated with white matter

hyperintensities in cognitively normal individuals using spearman correlations after correcting for multiple comparisons [17]. Based on data from the Atherosclerosis Risk in Communities study (ARIC), Walker and colleagues reported thirty-eight proteins to be associated with incident dementia after Bonferroni correction[12]. Of these, 16 were also associated with late-life dementia risk when measured in plasma collected nearly 20 years earlier, during mid-life. In addition, they found associations of several proteins with MRI derived measures of neurodegeneration and brain disease (e.g. total brain volume, white matter hyperintensities and a meta-ROI based on regions sensitive to AD). We recently investigated associations of an MRI measure called the AD Pattern Similarity scores [18-20] also using ARIC proteomic data (GeroScience, under review). We found the GDF-15 and pleiotrophin to be associated with the AD-PS scores after a Bonferroni correction. Our study was informed by a panel of 32 proteins reported to be associated with aging in the literature in 5 studies or more[9]. However, it is rare linking proteomics to an MRI measure of accelerated/resilient brain aging as we proposed here. This type of analyses can potentially lead to the identification of proteins related to the process of brain aging, understand mechanisms that lead to it and future clinical targets.

Here we capitalize on the rich phenotypic characterization of the ARIC participants to deploy an MRI measure of accelerated/resilient aging and investigate its cross-sectional and longitudinal associations with proteomics, cognitive function and mortality. We will use machine learning to determine associations of proteomic data with accelerate/resilient brain aging.

5. Main Hypothesis/Study Questions:

Our main objective is to deploy a new measure of brain aging and investigate its associations with proteomics, cognitive function and mortality.

Our main hypotheses are:

Hypothesis 1 (Cross-sectional): The measure of accelerated/resilient brain aging implemented using MRIs at visit 5 will be associated with cognitive status at visit 5.

Hypothesis 2 (Longitudinal): The measure of accelerated/resilient brain aging implemented using MRIs at visit 5 will be predictive of incident cognitive impairment and mortality within 5 and 8 years of follow-up respectively.

Hypothesis 3 (Imaging - Proteomics): Based on the full Soma logic panel collected at visits 3-5 and machine learning methods, we will determine proteomics signatures associated with our brain aging measure suggesting their involvement in accelerated/resilient brain aging.

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).

Design: Cross-sectional and longitudinal study design with follow-up through visit 7.

Outcome sets:

Visit 3

Proteomic data

Visits 5-7
Mortality

Visit 6-7

Cognitive status

Datasets:

Visits 5:

MRI
Proteomic data
Physical function data
Demographics
Comorbidities
Clinical data

Analyses:

Hypothesis 1 (Cross-sectional): The measure of accelerated/resilient brain aging implemented using MRIs at visit 5 will be associated with cognitive status at visit 5.

We will use similar methods to those we used to create the AD-PS scores[18, 19] to produce estimates of age from the MRI scans collected at visit 5 but instead of solving a high-dimensional classification problem a regression problem will be solved using the elastic net. The training dataset will be composed by 592 MRI images from cognitively normal participants (55-90 years old) in the Alzheimer's Neuroimaging Initiative Study. Once the elastic net regression model is estimated, MRI data from ARIC visit 5 is provided to the model as input to obtain the estimated ages for each individual. The gap between estimated and chronological age (Brain Aging Gap – BAG) will be used as measure of accelerated/resilient aging. We will use logistic regression analyses to evaluate BAG associations with cognitive status at visit 5. In

addition, receiver operating characteristic (ROC) curve analyses will be used to evaluate discrimination between CN and cognitively impaired individuals (MCI + dementia cases).

Hypothesis 2 (Longitudinal): The measure of accelerated/resilient brain aging implemented using MRIs at visit 5 will be predictive of incident cognitive impairment, and mortality within 5 and 8 years of follow-up respectively.

We will use logistic and Cox regression analyses to evaluate BAG associations with incident cognitive impairment and mortality.

Hypothesis 3 (Imaging - Proteomics): Based on the full Soma logic panel collected at visits 3-5 and machine learning methods, we will determine proteomics signatures associated with our brain aging measure suggesting their involvement in accelerated aging or resilience.

Proteins (visit 3) - Accelerated/resilient Brain Aging visit 5 (Longitudinal) - We will use the elastic net (LASSO) to identify proteomic signatures that predict biological brain aging at visit 5 in individuals that have both proteomics at visit 3 and MRIs at visit 5.

Proteins visit 5 - Accelerated/resilient Brain Aging (visit 5) (Cross-sectional) – Similar to above but cross-sectional at visit 5. We will identify proteins that overlap across the two analyses.

Standard statistical analyses will be adjusted by age, race-center, sex, education, smoking, hypertension, and diabetes. Analyses will be stratified by race and sex.

Limitations/Challenges

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"? _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)?

We have included several ARIC colleagues whose research is related.

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? _x___ Yes ___ No

SomaScan

11.b. If yes, is the proposal

_x_ A. primarily the result of an ancillary study (list number* _2008-06_____)

___ 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.

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. Franke, K. and C. Gaser, *Ten Years of BrainAGE as a Neuroimaging Biomarker of Brain Aging: What Insights Have We Gained?* Front Neurol, 2019. **10**: p. 789,PMID: 31474922.
2. Cole, J.H., *Neuroimaging-derived brain-age: an ageing biomarker?* Aging (Albany NY), 2017. **9**(8): p. 1861-1862.
3. Cole, J.H. and K. Franke, *Predicting Age Using Neuroimaging: Innovative Brain Ageing Biomarkers.* Trends Neurosci, 2017. **40**(12): p. 681-690,PMID: 29074032

4. Niu, X., et al., *Improved prediction of brain age using multimodal neuroimaging data.* Hum Brain Mapp, 2020. **41**(6): p. 1626-1643.
5. Ning, K., et al., *Association of relative brain age with tobacco smoking, alcohol consumption, and genetic variants.* Sci Rep, 2020. **10**(1): p. 10,PMID: 32001736.
6. Gaser, C., et al., *BrainAGE in Mild Cognitive Impaired Patients: Predicting the Conversion to Alzheimer's Disease.* PLoS One, 2013. **8**(6): p. e67346.
7. Gaser, C., et al., *BrainAGE in Mild Cognitive Impaired Patients: Predicting the Conversion to Alzheimer's Disease.* PLoS One, 2013. **8**(6): p. e67346,PMID: 23826273.
8. Cole, J.H., et al., *Brain age predicts mortality.* Mol Psychiatry, 2018. **23**(5): p. 1385-1392,PMID: 28439103.
9. Johnson, A.A., et al., *Systematic review and analysis of human proteomics aging studies unveils a novel proteomic aging clock and identifies key processes that change with age.* Ageing Res Rev, 2020. **60**: p. 101070.
10. Moaddel, R., et al., *Proteomics in aging research: A roadmap to clinical, translational research.* Aging Cell, 2021. **20**(4): p. e13325.
11. Ubaida-Mohien, C., et al., *Proteomics and Epidemiological Models of Human Aging.* Front Physiol, 2021. **12**: p. 674013.
12. Keenan A. Walker, et al., *Large-scale plasma proteomic analysis identifies proteins and pathways associated with dementia risk.* Nature Aging, 2021. **1**: p. 473-489.
13. Lehallier, B., et al., *Data mining of human plasma proteins generates a multitude of highly predictive aging clocks that reflect different aspects of aging.* Aging Cell, 2020. **19**(11): p. e13256.
14. Tanaka, T., et al., *Plasma proteomic signature of age in healthy humans.* Aging Cell, 2018. **17**(5): p. e12799.
15. Sathyan, S., et al., *Plasma proteomic profile of age, health span, and all-cause mortality in older adults.* Aging Cell, 2020. **19**(11): p. e13250.
16. Harris, S.E., et al., *Neurology-related protein biomarkers are associated with cognitive ability and brain volume in older age.* Nat Commun, 2020. **11**(1): p. 800.
17. Shi, L., et al., *Identification of plasma proteins relating to brain neurodegeneration and vascular pathology in cognitively normal individuals.* Alzheimers Dement (Amst), 2021. **13**(1): p. e12240.
18. Casanova, R., et al., *Using high-dimensional machine learning methods to estimate an anatomical risk factor for Alzheimer's disease across imaging databases.* Neuroimage, 2018. **183**: p. 401-411.
19. Casanova, R., et al., *Comparing data-driven and hypothesis-driven MRI-based predictors of cognitive impairment in individuals from the Atherosclerosis Risk in Communities (ARIC) study.* Alzheimers Dement, 2022. **18**(4): p. 561-571.