

ARIC Manuscript Proposal #3231

PC Reviewed: 9/11/18
SC Reviewed: _____

Status: _____
Status: _____

Priority: 2
Priority: _____

1.a. Full Title: Novel blood DNA methylation sites predict risk of death in a longitudinal study of 12,300 individuals

b. Abbreviated Title (Length 26 characters): DNAm & mortality (CHARGE)

2. Writing Group:

Writing group members:

This paper will present meta-analysis results from the CHARGE consortium. Participating cohorts include ARIC, FHS, InCHIANTI, KORA, LBC1921, LBC1936, NAS, TwinsUK, and WHI. Author representatives from each cohort are still being determined.

First author on the paper is Elena Colicino, Icahn School of Medicine at Mount Sinai (elena.colicino@mssm.edu)

ARIC co-authors:

James Pankow
Weihua Guan
Wen Zhang
Jan Bressler

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

First author:

Phone: _____ Fax: _____
E-mail: _____

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: Jim Pankow (ARIC cohort contact)
Address: University of Minnesota
1300 S. 2nd St., Suite 300
Minneapolis, MN 55454

Phone: 612-624-2883 Fax: _____
E-mail: panko001@umn.edu

3. Timeline: Submission for publication in late summer 2018

4. Rationale:

The human epigenome contains DNA methylation marks that change progressively as we age, in part due to risk factors for chronic disease. Previous studies have constructed aging clocks based on blood DNA methylation marks and linked those measures to mortality. However, the clocks have been constructed to predict current chronological age, but not to predict mortality. Indeed, no large-scale analysis has yet been conducted to identify variations in DNA methylation at individual 5'-cytosine-phosphate-guanosine-3' (CpG) sites associated with future risk of death.

5. Main Hypothesis/Study Questions:

Results from epigenome-wide methylation analysis of 12 American and European studies will be meta-analyzed to determine whether site-specific DNA methylation predicts all-cause mortality, independently of age, lifestyle, and previous morbidity.

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).

ARIC will contribute results for African Americans only.

Each cohort will independently run a common pre-specified statistical analysis in R. We will evaluate the association between locus-by-locus blood DNA methylation levels and all-cause mortality in each cohort with a Cox-regression model. Each cohort will adjust for two sets of harmonized covariates; age (categories for decades), gender, technical covariates (plate, chip, row and column) in the basic model; the previous set plus education level, self-reported recreational physical activity, smoking status and cumulative smoking (pack-years), body mass index, alcohol intake, hypertension, diabetes and any personal history of cancer in the fully adjusted model. All cohorts will independently estimate cell type proportions by using the reference-based Houseman method extended by Horvath.

We will perform inverse variance-weighted fixed-effects meta-analysis to combine results across cohorts. We will account for multiple testing by controlling at 5% both the Bonferroni correction and the false discovery rate (FDR) using the Benjamini-Hochberg procedure.

We will evaluate whether the CpG sites associated with mortality were enriched with genomic features provided in the Illumina annotation file. We will also test each gene mapped to the newly identified CpGs for tissue specific expression using data from the Genotype Tissue Expression (GTEx) project as integrated by the Functional Mapping and Annotation (FUMA) tool, which allows us to extract and interpret the relevant biological information from publicly available repositories and provide interactive figures for prioritized genes. To functionally interpret the genomic information identified from the FDR-significant CpGs, we will use the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database, which links genomic

information with higher order functional information. In KORA, a total of 998 individuals had both valid methylation and blood gene expression data, which will be used to assess whether DNA methylation was correlated to gene expression.

7.a. Will the data be used for non-CVD analysis in this manuscript? 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? 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 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.csc.c.unc.edu/aric/mantrack/maintain/search/dtSearch.html>

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)?

None identified

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? Yes 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 <https://www2.csc.c.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.csc.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.