ARIC Manuscript Proposal #2987

PC Reviewed: 06/06/17	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Epigenetic mediation of particulate matter-associated decreases in heart rate variability

b. Abbreviated Title (Length 26 characters): **PM, DNAm, and HRV**

2. Writing Group: WHI-EMPC, WHI BAA23, & ARIC Epigenetics Working Groups
Writing group members: Jan Bressler, Myriam Fornage, Weihua Guan, Ellen Demerath, Jim Pankow, and Kari North

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

First author: **Rahul Gondalia, MPH**
Address: University of North Carolina at Chapel Hill
Gillings School of Global Public Health
Department of Epidemiology
Cardiovascular Disease Program
CVS Plaza, Suite 306
137 East Franklin Street
Chapel Hill, NC 27514

Phone: (919) 966-1967 Fax: (919) 966-9800
E-mail: rahgonda@unc.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 A. Whitsel**
Address: University of North Carolina at Chapel Hill
Departments of Epidemiology and Medicine
Cardiovascular Disease Program
CVS Plaza, Suite 301-B
137 East Franklin Street
Chapel Hill, NC 27514
Phone: (919) 966-3168 Fax: (919) 966-9800
E-mail: eric_whitsel@unc.edu

3. Timeline: Primary analyses & draft manuscript to be completed by early 2018

4. Rationale:

Heart rate variability (HRV), an electrocardiographic (ECG) measure of cardiac autonomic control, is inversely associated with incident coronary heart disease and mortality.[1] Additionally, elevated exposure to ambient particulate matter (PM) air pollution has been associated with decreases in HRV,[2, 3] suggesting that cardiac autonomic dysfunction may be a mechanism by which PM contributes to cardiovascular disease. Despite the ubiquity and population burden of ambient air pollution exposure, the molecular mechanisms underlying PM-associated decreases in HRV have not been completely described.

DNA methylation (DNAm) at Cytosine-phosphate-Guanine (CpG) sites, a heritable but dynamic epigenetic modification that can influence gene expression without altering the genome, may be central to pathways by which environmental factors mediate CVD risk.[4] In fact, DNAm has been associated with other modifiable risk factors for CVD (e.g. diet,[5,6] smoking,[7] and exercise[8]). Moreover, DNAm near genes related to coagulation and inflammation has been linked with PM exposure[9-15] and has been implicated as a mediator of the PM-fasting blood glucose association.[16] However, no studies have evaluated DNAm as a potential mediator of PM-associated decreases in HRV.

The proposed study will therefore evaluate the PM-HRV association and its possible mediation by DNAm at PM-sensitive CpG sites in the Women's Health Initiative (WHI) and the Atherosclerosis Risk in Communities (ARIC) study.

5. Main Hypothesis/Study Questions:

1) To examine the PM-HRV association
2) To assess mediation of the PM-HRV association by DNAm

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).

Overview. The general approach is to first assess the PM-HRV association (Objective 1) in the WHI Clinical Trials (CT), WHI Observational Study (OS), and ARIC; then evaluate mediation of the PM-HRV association by DNAm at PM-sensitive CpG sites (Objective 2) previously identified by WHI MS #3188 and ARIC MS #2876.

Study Populations. Objective 1. Association analyses will rely on HRV data in approximately 62,195 participants in the WHI CT, 3,372 participants in the WHI OS Myocardial Ischemia and Migraine Study (MIMS), and 15,697 participants in ARIC. The WHI CT and ARIC designs are detailed elsewhere.[17, 18] MIMS was a ten-center ancillary study in a sample of WHI OS participants recruited at the screening or first follow-up visit.

Objective 2. Mediation analyses will rely on DNAm data generated by WHI Ancillary Study (AS) #315 entitled, "Epigenetic Mechanisms of PM-Mediated CVD Risk" (WHI-EMPC; R01-ES020836; MPIs – Hou; Baccarelli; Whitsel), AS #311, entitled "*DNA Methylation Measured in Prospectively Collected Blood Samples and Risk of Bladder Cancer Among Post-menopausal*

Women" (PI - Bhatti), WHI BAA23 entitled, "*Integrative Genomics for Risk of CHD and Related Phenotypes in WHI*" (MPIs – Horvath; Assimes; Absher), and an ARIC ancillary study, entitled "*Building on GWAS for NHLBI-Diseases: the U.S. CHARGE Consortium*" (PI – Boerwinkle).

AS #315 focuses on the core analytes subpopulation, an exam site- and race-stratified, randomly selected minority oversample of WHI CT participants who had repeated, fasting blood draws and resting, standard, twelve-lead electrocardiograms beginning at baseline. From this population, AS #315 randomly selected 2,200 participants with an available aliquot of DNA between 1993 and 2001 for DNAm assay, contemporaneous core analyte data, an address in the contiguous 48 U.S., and no conditions that affect the availability or accuracy of DNAm measures. Of these participants, 200 have DNAm measures from a subsequent annual visit. Two other WHI populations, BAA23 and AS #311, have similar available data for inclusion. BAA23 is a case-control study of cardiovascular disease among approximately 2,100 WHI CT and OS participants. AS #311 is a case-control study of bladder cancer among approximately 880 WHI CT and OS participants. The ARIC DNAm data are available from a subset of African American participants at visit 2/3 (n=2,850) and will soon be available for a subset of European American participants (n=1,102).

Exposures. The PM-HRV association will be estimated for DNAm-relevant PM size fractions and exposure averaging periods identified by WHI MS #3188 and ARIC MS #2876, which can include geocoded participant address-specific 2-, 7-, 28-, and 365- day mean concentrations of PM ≤ 2.5, ≤ 10, and between 2.5 and 10 um in diameter ($PM_{2.5}$, PM_{10}, and $PM_{2.5\text{-}10}$) regulated under the Clean Air Act by the U.S. Environmental Protection Agency (EPA) according to its National Ambient Air Quality Standards (NAAQS). Concentrations at the time of blood draw were estimated using national-scale, log-normal kriging and EPA Air Quality System monitoring data.[18] Data on $PM_{2.5}$ was not widely available until 1999, so before that year, its concentrations were instead estimated using generalized additive mixed models, the log-transformed ratio of $PM_{2.5}$ to predicted PM_{10}, and geographic information system (GIS)-based predictors.[20] $PM_{2.5\text{-}10}$ for each averaging period was calculated as the difference between PM_{10} and $PM_{2.5}$.

Outcomes. Three reliably estimated HRV indices: the mean RR interval duration (RR, ms), i.e. the unit-corrected inverse of mean heart rate; the standard deviation of normally conducted RR intervals (SDNN, ms); and the square root of mean squared differences in successive, normally conducted RR intervals (RMSSD, ms). In the WHI CT and ARIC, the estimates were based on ten-second, resting, supine, standard twelve-lead electrocardiograms (ECGs) recorded by MAC PCs.[21] In MIMS, the estimates were based on 24-hour, ambulatory three-lead ECGs recorded by a Holter monitor and a Zymed Model 3100–001 digital recorder. For Objective 1, HRV indices from WHI CT, MIMS, and ARIC will be standardized to allow for comparability between ten-second and 24-hour ECGs. For Objective 2, only HRV indices from WHI CT and ARIC will be analyzed given limited availability of DNAm among MIMS participants. Therefore, standardization will not be necessary.

Mediators. DNAm at CpG sites interrogated by the Illumina 450K Infinium Methylation BeadChip and previously identified as PM-sensitive by WHI MS #3188 and ARIC MS #2876. DNAm will be quantitatively represented by beta (the percentage of methylated cytosines over the sum of

methylated and unmethylated cytosines), then quality controlled, Beta-MIixture Quantile (BMIQ)-normalized to correct for differences otherwise attributable to Type I and II probes,[22] and batch-corrected.

Potential covariates. Objective 1. Demographic covariates (age; race/ethnicity; center), relevant meteorological covariates, seasonality, potential confounders of interest (smoking status, alcohol use, body mass index, physical activity, individual-level education, and neighborhood socioeconomic status), randomly assigned treatment group in the WHI CT, and clinical covariates (hypertension, hyperlipidemia, diabetes, chronic lung disease, coronary heart disease, congestive heart failure) will be considered.

Objective 2. In addition to those in Objective 1, technical covariates (array; row; column), Houseman cell type proportions (CD8-T, CD4-T, B cell, natural killer, monocyte, and granulocyte),[23] and principal components for ancestral admixture will be considered.

STATISTICAL ANALYSIS

Objective 1. For each DNAm-relevant PM size fraction and exposure averaging period identified by WHI MS #3188 and ARIC MS #2876, covariate-adjusted, multi-level, linear mixed-effects longitudinal models will leverage repeated measures in the WHI CT to estimate PM-HRV associations. There will be a random intercept and slope for time at the participant level and for PM at the WHI center level. Inverse probability weighting will account for attrition in longitudinal analyses. Multivariate imputation by chained equations (MICE) will be used to impute missing data. Similar models will also be used to estimate cross-sectional associations in MIMS using a random intercept and slope for PM at the WHI center level. Complementary analyses will be conducted in ARIC, and fixed-effects inverse variance-weighted meta-analyses will be used to combine the study-specific test statistics across studies.

Objective 2.*Selecting potential mediators.* For each PM-sensitive CpG site identified by WHI MS #3188 and ARIC MS #2876, covariate-adjusted, two-level, linear mixed-effects longitudinal models will leverage repeated measures to estimate CpG site-specific DNAm-HRV associations. The models will contain a random intercept and slope for time at the participant level. MICE will be used to impute missing data. CpG sites that are significantly associated with HRV after Bonferroni-correction will be investigated using Mendelian Randomization (MR) methods[23] that employ genetic variants as instrumental variables in analyses evaluating causality, direction, and magnitude of DNAm-HRV associations. The proposed study will identify CpG site-specific genetic variants from genome-wide data when they are 1) associated with DNAm at the CpG site; 2) unassociated with HRV except through DNAm; and 3) unassociated with any potential confounders. CpG sites that are considered causally associated with HRV will undergo mediation analyses.

Mediation Analyses. For each PM size fraction and exposure averaging period, causal mediation methods[23-25] will be used to decompose the total effect (TE) between PM and HRV to the natural direct effect (NDE), i.e. the effect of PM on HRV independent of the mediator; and the natural indirect effect (NIE), i.e. the effect of PM on HRV through the mediator, where the sum of NDE and NIE is the TE.

We will first obtain the coefficient for the adjusted association between PM and DNAm at each selected CpG site i ($\boldsymbol{\beta_{1i}}$) using

$$CpG_i\ DNAm = \beta_{0i} + \boldsymbol{\beta_{1i}} PM + \beta_{2i} Confounders$$

Then we will obtain the coefficient for the adjusted association between DNAm at CpG i and HRV ($\boldsymbol{\theta_{2i}}$), and its interaction with PM ($\boldsymbol{\theta_{3i}}$), using

$$HRV = \theta_{0i} + \theta_{1i} PM + \boldsymbol{\theta_{2i}} CpG_i\ DNAm + \boldsymbol{\theta_{3i}} PMxCPG_i\ DNAm + \theta_{4i} Confounders$$

Then, for each CpG site i, NDE and NIE will be estimated as

$$NDE_i = \theta_{1i} + \theta_{3i}(\beta_{0i} + \beta_{2i})$$
$$NIE_i = \beta_{1i}(\theta_{2i} + \theta_{3i})$$

Boostrapping will be used to estimate standard errors for the NDE and NIE estimates.[24-26]

Sensitivity of Mediation Analyses. Causal mediation analyses may not be valid if a variable affected by the exposure under study (e.g. PM concentration) is also a confounder of the mediator-outcome (e.g. DNAm-HRV) association.[23-28]. WHI MS #3251 and ARIC MS #2924 may identify leukocyte composition as such a variable, either when measured via cytometry as part of a complete blood count / differential, or in its absence, when estimated by constraining the sum of the CD8+ T cell, CD4+ T cell, natural killer cell, B cell, monocyte, and granulocyte proportions in whole blood to 100%, then fitting a regression model to DNAm data.[23] We will therefore assess whether mediation results are robust to exposure-induced mediator-outcome confounding by leukocyte composition using methods proposed by VanderWeele[28] and Vansteelandt.[29]

CONCLUSIONS

In this study, we will evaluate the mediation by DNA methylation of ambient particulate matter air pollution-associated decreases in heart rate variability, the nature of which may ultimately affect our understanding of molecular consequences of exposure.

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? _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

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)? 2876, 2924

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* _2009.08_)

___ 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.

13. Per Data Use Agreement Addendum, approved manuscripts using CMS data shall be submitted by the Coordinating Center to CMS for informational purposes prior to publication. Approved manuscripts should be sent to Pingping Wu at CC, at pingping_wu@unc.edu. I will be using CMS data in my manuscript ____ Yes __x__ No.

REFERENCES

1) Dekker, Jacqueline M., et al. "Low heart rate variability in a 2-minute rhythm strip predicts risk of coronary heart disease and mortality from several causes." *Circulation* 102.11 (2000): 1239-1244.
2) Pieters, Nicky, et al. "An epidemiological appraisal of the association between heart rate variability and particulate air pollution: a meta-analysis." *Heart* 98.15 (2012): 1127-1135.
3) Buteau, Stephane, and Mark S. Goldberg. "A structured review of panel studies used to investigate associations between ambient air pollution and heart rate variability." *Environmental research* 148 (2016): 207-247.
4) Bollati, Valentina, and Andrea Baccarelli. "Environmental epigenetics."Heredity 105.1 (2010): 105-112.
5) Heijmans, Bastiaan T., et al. "Persistent epigenetic differences associated with prenatal exposure to famine in humans." Proceedings of the National Academy of Sciences 105.44 (2008): 17046-17049.
6) Tobi, Elmar W., et al. "DNA methylation differences after exposure to prenatal famine are common and timing-and sex-specific." Human molecular genetics 18.21 (2009): 4046-4053.
7) Breton, Carrie V., et al. "Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation." American journal of respiratory and critical care medicine 180.5 (2009): 462-467.
8) Rönn, Tina, et al. "A six months exercise intervention influences the genome-wide DNA methylation pattern in human adipose tissue." PLoS Genet 9.6 (2013): e1003572.
9) Madrigano, Jaime, et al. "Air pollution and DNA methylation: interaction by psychological factors in the VA Normative Aging Study." American journal of epidemiology 176.3 (2012): 224-232.
10) Zhong, Jia, et al. "Cardiac autonomic dysfunction: particulate air pollution effects are modulated by epigenetic immunoregulation of Toll-like receptor 2 and dietary flavonoid intake." Journal of the American Heart Association 4.1 (2015): e001423.
11) Bind, Marie-Abele C., et al. "Beyond the mean: quantile regression to explore the association of air pollution with gene-specific methylation in the normative aging study." Environmental health perspectives 123.8 (2015): 759.
12) Peng, Cheng, et al. "Particulate Air Pollution and Fasting Blood Glucose in Non-Diabetic Individuals: Associations and Epigenetic Mediation in the Normative Aging Study, 2000-2011." Environmental health perspectives (2016).
13) Bind, Marie-Abele, et al. "Air pollution and markers of coagulation, inflammation and endothelial function: Associations and epigene-environment interactions in an elderly cohort." Epidemiology (Cambridge, Mass.) 23.2 (2012): 332.
14) Bind, Marie-Abele, et al. "Air pollution and gene-specific methylation in the Normative Aging Study: association, effect modification, and mediation analysis." Epigenetics 9.3 (2014): 448-458.

15) Panni, Tommaso, et al. "A Genome-Wide Analysis of DNA Methylation and Fine Particulate Matter Air Pollution in Three Study Populations: KORA F3, KORA F4, and the Normative Aging Study." Environmental health perspectives (2016).
16) Peng, Cheng, et al. "Particulate air pollution and fasting blood glucose in nondiabetic individuals: associations and epigenetic mediation in the Normative Aging Study, 2000–2011." Environmental health perspectives 124.11 (2016): 1715.
17) The Women's Health Initiative. "Design of the women's health initiative clinical trial and observational study." Controlled clinical trials 19.1 (1998): 61-109.
18) Aric Investigators. "The Atherosclerosis Risk in Communities (ARIC) study: design and objectives." American journal of epidemiology 129.4 (1989): 687-702.
19) Liao, Duanping, et al. "GIS approaches for the estimation of residential-level ambient PM concentrations." Environmental health perspectives (2006): 1374-1380.
20) Yanosky, Jeff D., et al. "Spatio-temporal modeling of particulate air pollution in the conterminous United States using geographic and meteorological predictors." Environmental Health 13.1 (2014): 1.
21) Schroeder, Emily B., et al. "Repeatability of heart rate variability measures." *Journal of electrocardiology* 37.3 (2004): 163-172.
22) Wu, Michael C., et al. "A systematic assessment of normalization approaches for the Infinium 450K methylation platform." Epigenetics 9.2 (2014): 318-329.
23) Houseman, Eugene Andres, et al. "DNA methylation arrays as surrogate measures of cell mixture distribution." BMC bioinformatics 13.1 (2012): 1.
24) VanderWeele, Tyler. *Explanation in causal inference: methods for mediation and interaction*. Oxford University Press, 2015.
25) Valeri, Linda, and Tyler J. VanderWeele. "Mediation analysis allowing for exposure–mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros." Psychological methods 18.2 (2013): 137.
26) VanderWeele, Tyler, and Stijn Vansteelandt. "Conceptual issues concerning mediation, interventions and composition." *Statistics and its Interface* 2 (2009): 457-468.
27) VanderWeele, Tyler J., and Yasutaka Chiba. "Sensitivity analysis for direct and indirect effects in the presence of exposure-induced mediator-outcome confounders." Epidemiology, biostatistics, and public health 11.2 (2014).
28) VanderWeele, Tyler J., Stijn Vansteelandt, and James M. Robins. "Effect decomposition in the presence of an exposure-induced mediator-outcome confounder." Epidemiology (Cambridge, Mass.) 25.2 (2014): 300.
29) Vansteelandt, Stijn, and Rhian M. Daniel. "Interventional effects for mediation analysis with multiple mediators." *Epidemiology* (2017).

TABLES (Template)

Table 1. Demographic Characteristics of WHI and ARIC participants

		WHI CT	WHI MIMS ….
		Female	Female …
N			
Age	Mean (SD)		
Race	African American		
	European American		
	Hispanic/Latino		
SES			
Body Mass Index	Mean (SD)		
Smoking Status			
Alcohol Use			
Physical Activity			
…			
$PM_{2.5}$			
PM_{10}			
$PM_{2.5-10}$			

Table 2. DNAm-HRV associations at CpG-sensitive sites

				SDNN		…	
CpG site	Chr	Position	Nearest Gene	Coeff	*P*	Coeff	*P*
cg…	1						
cg…	2						

Table 3. Mendelian randomization results showing the mean association between Instrumental Variable (IV) SNPs and HRV

CpG site	Trait	IV SNPs	Mean estimate	(SE)	*P*
cg…	*SDNN*	6			
	…	3			
cg…	*SDNN*	2			
	..	5			

Table 4. Mediation results of the PM-HRV associations for each causal CpG site

CpG	Trait	Total effect (SE)	Direct effect (SE)	Indirect effect (SE)	Proportion mediated (%)
cg…	*SDNN*				
	…				
cg…	*SDNN*				
	..				