**ARIC Manuscript Proposal #2876**

**PC Reviewed: 11/08/16** **Status: _____** **Priority: 2**
**SC Reviewed: _____** **Status: _____** **Priority: ____**

**1.a. Full Title**: **Particulate Matter Air Pollution and DNA Methylation**

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

**2. Writing Group**: WHI-EMPC & 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 late 2016

## 4. Rationale:

Ambient particulate matter (PM) air pollution is a modifiable exposure that has been consistently associated with cardiovascular disease (CVD) morbidity and mortality, partly through autonomic imbalance, oxidative stress and pro-inflammatory pathways.[1] Although air quality regulations in the US have led to declines in ambient PM, lower levels of exposure still pose a CVD risk.[2] Despite the ubiquity of air pollution exposure and the continued population burden of PM, the molecular mechanisms underlying PM-associated cardiovascular health effects have not been completely described.

DNA methylation (DNAm), a heritable but dynamic epigenetic modification that can influence gene expression without altering the genome, may be central to pathways by which environmental factors modify CVD risk.[3] In fact, DNAm has been associated with other modifiable risk factors for CVD (e.g. diet,[4,5] smoking,[6] and exercise[7]). Moreover, DNAm near genes related to coagulation and inflammation has been linked with PM exposure, albeit in candidate-gene studies.[8-13] Only one study has agnostically evaluated DNAm associations with PM on an epigenome-wide scale,[14] but no studies have done so in large, sociodemographically and environmentally diverse, well-characterized populations of women.

The proposed study will therefore do so within a minority over-sample of the WHI Clinical Trial (WHI-CT) participants and replicate results in the biracial Atherosclerosis Risk in Communities (ARIC) study. Additional replication cohorts currently include the Cooperative Health Research in the Region Augsburg (KORA) study, and Normative Aging Study (NAS).

## 5. Main Hypothesis/Study Questions:

To leverage the biracial, geographically diverse data within ARIC to replicate the associations between DNAm and ambient concentrations of particulate matter $\leq 2.5$, $\leq 10$, and between 2.5 and 10 micrometers in diameter ($PM_{2.5}$, $PM_{10}$, and $PM_{2.5-10}$) found within a minority-oversample of WHI-CT participants.

## 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 conduct discovery analyses in WHI CT of PM-DNAm associations for each *Cytosine-phosphate-Guanine* (CpG) methylation site on the Illumina 450K Infinium Methylation BeadChip. The WHI CT analyses described herein will be based on DNA methylation data generated by WHI Ancillary Study (AS) #315 entitled, "Epigenetic Mechanisms of PM-Mediated CVD Risk" (WHI-EMPC; R01-ES020836; MPIs – Hou; Baccarelli; Whitsel) and AS #534 entitled, "Longitudinal study of DNA methylation as a mediator between age and cardiovascular risk" (HHSN268201100046C; MPIs – Conneely; Whitsel). CpG sites will be ranked according to statistical significance followed by replication of CpG sites identified as significant or suggestive of significance within ARIC, KORA, and NAS. The replication analyses in ARIC will rely on air pollution data generated as part of the

"*Modification of PM-Mediated Arrhythmogenesis in Populations*" ancillary study (MOPMAP; R01-ES017794; PI – Whitsel).

*Study Population.* The WHI 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 DNA methylation assay, contemporaneous core analyte data, an address in the contiguous 48 U.S., and no conditions that affect the availability or accuracy of DNA methylation measures. 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), allowing replication for the multi-ethnic WHI CT analyses. Of these participants, 200 have a second assay from a subsequent annual visit and 43 have a third assay from the Long Life Study visit, yielding a total of 2,443 observations.

*Primary Outcomes.* DNA methylation (DNAm) at CpG sites as determined by the Illumina 450K Infinium Methylation BeadChip, quantitatively represented by beta (the percentage of methylated cytosines over the sum of methylated and unmethylated cytosines), then quality controlled, batch-corrected, and normalized using Beta-MIixture Quantile (BMIQ) to correct for differences otherwise attributable to Type I and II probes.[15]

*Main Exposures.* Geocoded participant address-specific 2-, 7-, 28-, and 365- day mean concentrations of ambient $PM_{2.5}$, $PM_{10}$, 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.[16] 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.[17] $PM_{2.5\text{-}10}$ for each averaging period was calculated as the difference of $PM_{10}$ and $PM_{2.5}$.

*Covariates.* Demographic covariates (age; center), technical covariates (plate; chip; row; column), Houseman estimates of cell type proportions (CD8-T, CD4-T, B cell, natural killer, monocyte, and granulocyte), principal components for ancestral admixture, randomly assigned treatment group, relevant meteorological covariates, seasonality, and potential confounders of interest (smoking status, alcohol use, body mass index, physical activity, individual-level education, and neighborhood socioeconomic status).

## STATISTICAL ANALYSIS

*Discovery Association Analyses in WHI CT.* For each PM size fraction and exposure averaging period, covariate-adjusted, three-level, linear mixed effects longitudinal models will leverage repeated measures to estimate PM-DNAm associations. There will be a random intercept and slope for time at the participant level and for PM at the WHI center level as well as a random intercept for technical covariates. These analyses will be stratified by race/ethnicity, and fixed-effects inverse variance-weighted meta-analysis will be used to combine the stratum-specific

estimates.

*Sensitivity Analyses.* To assess sensitivity of WHI results, models will then be run on the entire WHI-EMPC population (i.e. without stratification by race/ethnicity) to include additional minority populations (i.e. American Indian or Alaskan Natives; Asian or Pacific Islanders; Others) that were too small to include in stratified analyses. Race/ethnicity will be adjusted for in sensitivity analyses.

*Replication Association Analyses in ARIC and other cohorts*. For each PM size fraction and exposure averaging period assessed, CpG sites identified as significant ($p<1.0x10^{-7}$) or suggestive ($p<1.0x10^{-5}$) for PM-related methylation will be replicated within ARIC, KORA, and NAS. Thresholds of significance for these analyses will be Bonferroni-corrected based on the number of CpG sites carried over for replication. Analyses will be of European Americans in NAS, Europeans in KORA, and stratified by race/ethnicity for ARIC (African Americans and European Americans). Fixed-effects, inverse variance-weighting will again be used to meta-analytically combine the race/ethnic-specific estimates for replication.

*Quality Control & Meta-Analyses.* We will follow established quality control protocols, including review of results by e.g. graphing the observed *P* values for each CpG site against CpG positions in Manhattan plots, by graphing them against the expected values from a theoretical $\chi^2$ distribution in quantile-quantile (Q-Q) plots, and applying genomic control ($\lambda$) if inflation is observed.

*Functional Annotation.* Function annotation for the implicated CpG sites will be assessed using the UCSC Genome Browser[18] and the WashU Epigenome Browser[19] with data from the Encyclopedia of DNA elements (ENCODE)[20] and Roadmap Epigenomics Project.[21]

**CONCLUSIONS**

In this epigenome-wide association study, we will estimate associations between ambient particulate matter air pollution and DNA methylation, 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)? NA**

**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) Brook, Robert D., et al. "Particulate matter air pollution and cardiovascular disease an update to the scientific statement from the American Heart Association." Circulation 121.21 (2010): 2331-2378.
2) Thurston, George D., et al. "Ambient particulate matter air pollution exposure and mortality in the NIH-AARP diet and health cohort." Environmental health perspectives 124.4 (2016): 484.
3) Bollati, Valentina, and Andrea Baccarelli. "Environmental epigenetics."Heredity 105.1 (2010): 105-112.
4) 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.
5) 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.
6) 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.
7) 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.
8) 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.
9) 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.
10) 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.
11) 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).
12) 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.
13) 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.
14) 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).
15) Wu, Michael C., et al. "A systematic assessment of normalization approaches for the Infinium 450K methylation platform." Epigenetics 9.2 (2014): 318-329.
16) Liao, Duanping, et al. "GIS approaches for the estimation of residential-level ambient PM concentrations." Environmental health perspectives (2006): 1374-1380.
17) 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.

18) Kuhn, Robert M., David Haussler, and W. James Kent. "The UCSC genome browser and associated tools." Briefings in bioinformatics (2012): bbs038.
19) Zhou, Xin, et al. "The human epigenome browser at Washington University." Nature methods 8.12 (2011): 989-990.
20) Rosenbloom, Kate R., et al. "ENCODE whole-genome data in the UCSC Genome Browser: update 2012." Nucleic acids research (2011): gkr1012.
21) Bernstein, Bradley E., et al. "The NIH roadmap epigenomics mapping consortium." Nature biotechnology 28.10 (2010): 1045-1048.

**ARIC Manuscript Proposal #2876**

**PC Reviewed: 11/08/16** **Status: _____** **Priority: 2**
**SC Reviewed: _____** **Status: _____** **Priority: ____**

**1.a. Full Title**: **Particulate Matter Air Pollution and DNA Methylation**

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

**2. Writing Group**: WHI-EMPC & ARIC Epigenetics Working Groups
Writing group members: Jan Bressler, Myriam Fornage, Weihua Guan, Ellen Demerath, Jim Pankow, and Kari North ~~Interested members of the working groups with whom this proposal was discussed by conference call.~~

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 late 2016

**4. Rationale**:

Ambient particulate matter (PM) air pollution is a modifiable exposure that has been consistently associated with cardiovascular disease (CVD) morbidity and mortality, partly through autonomic imbalance, oxidative stress and pro-inflammatory pathways.[1] Although air quality regulations in the US have led to declines in ambient PM, lower levels of exposure still pose a CVD risk.[2] Despite the ubiquity of air pollution exposure and the continued population burden of PM, the molecular mechanisms underlying PM-associated cardiovascular health effects have not been completely described.

DNA methylation (DNAm), a heritable but dynamic epigenetic modification that can influence gene expression without altering the genome, may be central to pathways by which environmental factors modify CVD risk.[3] In fact, DNAm has been associated with other modifiable risk factors for CVD (e.g. diet,[4,5] smoking,[6] and exercise[7]). Moreover, DNAm near genes related to coagulation and inflammation has been linked with PM exposure, albeit in candidate-gene studies.[8-13] Only one study has agnostically evaluated DNAm associations with PM on an epigenome-wide scale,[14] but no studies have done so in large, sociodemographically and environmentally diverse, well-characterized populations of women.

The proposed study will therefore do so within a minority over-sample of the WHI Clinical Trial (WHI-CT) participants and replicate results in the biracial Atherosclerosis Risk in Communities (ARIC) study. Additional replication cohorts currently include the Cooperative Health Research in the Region Augsburg (KORA) study, and Normative Aging Study (NAS).

**5. Main Hypothesis/Study Questions**:

To leverage the biracial, geographically diverse data within ARIC to replicate the associations between DNAm and ambient concentrations of particulate matter $\leq 2.5$, $\leq 10$, and between 2.5 and 10 micrometers in diameter ($PM_{2.5}$, $PM_{10}$, and $PM_{2.5-10}$) found within a minority-oversample of WHI-CT participants.

**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 conduct discovery analyses in WHI CT of PM-DNAm associations for each *Cytosine-phosphate-Guanine* (CpG) methylation site on the Illumina 450K Infinium Methylation BeadChip. The WHI CT analyses described herein will be based on DNA methylation data generated by WHI Ancillary Study (AS) #315 entitled, "Epigenetic Mechanisms of PM-Mediated CVD Risk" (WHI-EMPC; R01-ES020836; MPIs – Hou; Baccarelli; Whitsel) and AS #534 entitled, "Longitudinal study of DNA methylation as a mediator between age and cardiovascular risk" (HHSN268201100046C; MPIs – Conneely; Whitsel). CpG sites will be ranked according to statistical significance followed by replication of CpG sites identified as significant or suggestive of significance within ARIC, KORA, and NAS.

The replication analyses in ARIC will rely on air pollution data generated as part of the "*Modification of PM-Mediated Arrhythmogenesis in Populations*" ancillary study (MOPMAP; R01-ES017794; PI – Whitsel).

*Study Population.* The WHI 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 DNA methylation assay, contemporaneous core analyte data, an address in the contiguous 48 U.S., and no conditions that affect the availability or accuracy of DNA methylation measures. 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), allowing replication for the multi-ethnic WHI CT analyses. Of these participants, 200 have a second assay from a subsequent annual visit and 43 have a third assay from the Long Life Study visit, yielding a total of 2,443 observations.

*Primary Outcomes.* DNA methylation (DNAm) at CpG sites as determined by the Illumina 450K Infinium Methylation BeadChip, quantitatively represented by beta (the percentage of methylated cytosines over the sum of methylated and unmethylated cytosines), then quality controlled, batch-corrected, and normalized using Beta-MIixture Quantile (BMIQ) to correct for differences otherwise attributable to Type I and II probes.[15]

*Main Exposures.* Geocoded participant address-specific 2-, 7-, 28-, and 365- day mean concentrations of ambient $PM_{2.5}$, $PM_{10}$, and $PM_{2.5-10}$ 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.[16] 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.[17] $PM_{2.5-10}$ for each averaging period was calculated as the difference of $PM_{10}$ and $PM_{2.5}$.

*Covariates.* Demographic covariates (age; center), technical covariates (plate; chip; row; column), Houseman estimates of cell type proportions (CD8-T, CD4-T, B cell, natural killer, monocyte, and granulocyte), principal components for ancestral admixture, randomly assigned treatment group, relevant meteorological covariates, seasonality, and potential confounders of interest (smoking status, alcohol use, body mass index, physical activity, individual-level education, and neighborhood socioeconomic status).

## STATISTICAL ANALYSIS

*Discovery Association Analyses in WHI CT.* For each PM size fraction and exposure averaging period, covariate-adjusted, three-level, linear mixed effects longitudinal models will leverage repeated measures to estimate PM-DNAm associations. There will be a random intercept and slope for time at the participant level and for PM at the WHI center level as well as a random intercept for technical covariates. These analyses will be stratified by race/ethnicity, and fixed-

effects inverse variance-weighted meta-analysis will be used to combine the stratum-specific estimates.

*Sensitivity Analyses.* To assess sensitivity of WHI results, models will then be run on the entire WHI-EMPC population (i.e. without stratification by race/ethnicity) to include additional minority populations (i.e. American Indian or Alaskan Natives; Asian or Pacific Islanders; Others) that were too small to include in stratified analyses. Race/ethnicity will be adjusted for in sensitivity analyses.

*Replication Association Analyses in ARIC and other cohorts*. For each PM size fraction and exposure averaging period assessed, CpG sites identified as significant ($p<1.0x10^{-7}$) or suggestive ($p<1.0x10^{-5}$) for PM-related methylation will be replicated within ARIC, KORA, and NAS. Thresholds of significance for these analyses will be Bonferroni-corrected based on the number of CpG sites carried over for replication. Analyses will be of European Americans in NAS, Europeans in KORA, and stratified by race/ethnicity for ARIC (African Americans and European Americans). Fixed-effects, inverse variance-weighting will again be used to meta-analytically combine the race/ethnic-specific estimates for replication.

*Quality Control & Meta-Analyses.* We will follow established quality control protocols, including review of results by e.g. graphing the observed *P* values for each CpG site against CpG positions in Manhattan plots, by graphing them against the expected values from a theoretical $\chi^2$ distribution in quantile-quantile (Q-Q) plots, and applying genomic control ($\lambda$) if inflation is observed.

*Functional Annotation.* Function annotation for the implicated CpG sites will be assessed using the UCSC Genome Browser[18] and the WashU Epigenome Browser[19] with data from the Encyclopedia of DNA elements (ENCODE)[20] and Roadmap Epigenomics Project.[21]

**CONCLUSIONS**

In this epigenome-wide association study, we will estimate associations between ambient particulate matter air pollution and DNA methylation, 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)? NA**

**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) Brook, Robert D., et al. "Particulate matter air pollution and cardiovascular disease an update to the scientific statement from the American Heart Association." Circulation 121.21 (2010): 2331-2378.
2) Thurston, George D., et al. "Ambient particulate matter air pollution exposure and mortality in the NIH-AARP diet and health cohort." Environmental health perspectives 124.4 (2016): 484.
3) Bollati, Valentina, and Andrea Baccarelli. "Environmental epigenetics."Heredity 105.1 (2010): 105-112.
4) 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.
5) 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.
6) 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.
7) 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.
8) 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.
9) 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.
10) 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.
11) 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).
12) 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.
13) 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.
14) 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).
15) Wu, Michael C., et al. "A systematic assessment of normalization approaches for the Infinium 450K methylation platform." Epigenetics 9.2 (2014): 318-329.
16) Liao, Duanping, et al. "GIS approaches for the estimation of residential-level ambient PM concentrations." Environmental health perspectives (2006): 1374-1380.

17) 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.
18) Kuhn, Robert M., David Haussler, and W. James Kent. "The UCSC genome browser and associated tools." Briefings in bioinformatics (2012): bbs038.
19) Zhou, Xin, et al. "The human epigenome browser at Washington University." Nature methods 8.12 (2011): 989-990.
20) Rosenbloom, Kate R., et al. "ENCODE whole-genome data in the UCSC Genome Browser: update 2012." Nucleic acids research (2011): gkr1012.
21) Bernstein, Bradley E., et al. "The NIH roadmap epigenomics mapping consortium." Nature biotechnology 28.10 (2010): 1045-1048.