ARIC Manuscript Proposal #2090

PC Reviewed: 3/12/13	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Particulate Matter-Gene Interactions and QT Interval Duration

b. Abbreviated Title (Length 26 characters): PM-Gene GWAS of QT

2. Writing Group:

Writing group members: Candidates currently include co-investigators who have been involved in the planning, execution of, or assembly of data for ARIC AS#2009.08 and its WHI clinical trial sister study AS#264: Whitsel EA, Avery CL, Li Y, Yan S, Liao D, Lin D, North KE, Smith RL, Tinker L, Vernon M, Wilhelmsen K, Wu M, Kabisa S, and Zhang Z

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

First author: Melanie D. Napier
Address: University of North Carolina – Chapel Hill
Department of Epidemiology
Cardiovascular Disease Program
Bank of America Center, Suite 306-E
137 East Franklin Street
Chapel Hill, NC 27514
Phone: (919) 966-3168
Fax: (919) 966-9800
E-mail: mdnapier@email.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 – Chapel Hill
Departments of Epidemiology and Medicine
Cardiovascular Disease Program
Bank of America Center, Suite 305-C
137 East Franklin Street
Chapel Hill, NC 27514
Phone: (919) 966-3168
Fax: (919) 966-9800
E-mail: eric_whitsel@unc.edu

3. Timeline:
First pass analyses will begin as soon as requisite electrocardiographic, environmental, genetic and covariate data have been assembled. Anticipated completion is six months once data are available.

4. Rationale:

Repolarization is an active process by which differences in cation concentrations across cell membranes return to their resting levels. The duration of the electrocardiographic QT interval (QT) on the resting standard twelve-lead electrocardiogram (ECG) provides useful information about its temporal dimensions in the ventricular myocardium. QT prolongation has been associated with increased risk of sudden cardiac death.[1] Environmental and genetic causes of QT prolongation and its clinical manifestations have been identified. For example, many epidemiologic studies have suggested that elevated ambient particulate matter (PM) concentrations are associated with increased risk of cardiovascular disease (CVD) morbidity and mortality,[2, 3] including ventricular arrhythmias[4, 5] and QT prolongation.[6] Recent studies have suggested that the effect of PM on ventricular repolarization is one of the underlying mechanisms linking air pollution to increased risk of CVD.[7-13] The National Human Genome Research Institute (NHGRI) has also catalogued the discovery of at least fifteen known genetic variants on eleven loci influencing the distribution of QT, including *NOS1AP*,[14-24] *KCNQ1*,[14, 15, 22] *KCNH2*,[15, 22] *PLN*,[14, 15, 22] *LITAF*,[15, 22] *NDRG4*,[14, 22] and *ATP1B1*[14, 15, 24] in individuals of European and African ancestry. However, few studies have used genome-wide association methods to help identify genetic variants that may modify the QT-prolonging effects of environmental factors like ambient PM. Additional research on PM interactions with QT is therefore warranted.

The proposed manuscript directly addresses the paucity of information on the complex interplay of SNPs and PM on ventricular repolarization. The well-characterized ARIC and WHI CT cohorts are ideal for GWA analyses designed to leverage major, existing, and funded resources. Genomic, environmental and electrocardiographic data from these two cohorts can support a well-powered, rigorous examination of the degree to which SNPs modify PM-QT associations, as described in the power section below. Filling this gap in knowledge will (1) advance understanding of genetic susceptibility to and the pathophysiological mechanisms underlying PM-mediated arrhythmogenesis in ethnically and geographically diverse populations and (2) inform epidemiologists, environmental health scientists, and federal regulators responsible for evaluating air quality standards in terms of their ability to protect cardiovascular health.

5. Main Hypothesis/Study Questions:
We therefore propose a GWA study to (1) examine gene-by-PM effects as they relate to the QT interval and (2) evaluate the consistency of these interactive effects among race and gender groups.

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 proposed study will be conducted in the ARIC and WHI CT cohorts, based on the foundation provided by ARIC Ancillary Study #2009.08, "*Modification of PM-Mediated Arrhythmogenesis in Populations*" (R01- ES017794; Whitsel, PI) and two WHI Ancillary Studies: #140 "*The Environmental Epidemiology of Arrhythmogenesis in WHI*" (R01-ES012238; Whitsel, PI) and #264 "*Genetic Modification of PM-Mediated Arrhythmogenesis*" (R01-ES017794; Whitsel, PI).

Study population. The focus will be on approximately 34,000 uniformly well-characterized and consenting participants living in the contiguous 48 U.S. states (U.S. Environmental Protection Agency Regions 1-10) within 500 miles of their exam site who had one or more high quality ECGs between 1987 and 1999 and consented to use of DNA for genetic research. The population will include seven distinct subpopulations: (1) black women, (2) Hispanic women, and (3) white women in the WHI CT, including ~4000 WHI participants in GARNET (Genome-wide Association Studies of Treatment Response in Randomized Clinical Trials) and ~4000 WHI participants in WHIMS (WHI Memory Study); and (4) black women, (5) white women, (6) black men, and (7) white men in ARIC. These seven subpopulations will be used to independently identify gene-by-PM interactions for QT interval between approximately 2.5 million SNPs using imputed data from the Affymetrix 6.0 platform and daily mean ambient PM concentrations spatially interpolated at geocoded participant addresses.

Environmental exposure: Particulate matter. The proposed work will focus on location-specific daily mean concentrations of $PM_{2.5}$ and PM_{10} at geocoded addresses of ARIC participants between 1987 and 1999. We estimated the concentrations using a practical approach to spatial interpolation developed and validated by our group.[25-28]

Genetic exposure. SNPs were imputed using MACH (version 1.0.16) and the publicly available phased haplotypes from HapMap (release 22, build 36) as a reference panel.

Phenotypes. Ventricular repolarization, defined as maximum QT interval duration (ms) across all twelve leads. This phenotype will be analyzed linearly as an interval-scale measure.

Inclusions. Consenting participants, GWAS data, QT data, PM data

Exclusions. Poor quality ECG, electronic pacemaker, Wolff–Parkinson–White syndrome, QRS ≥ 120 ms, atrial fibrillation, 2° or 3° AV block, heart failure, or anti-arrhythmic medication use

Additive Genetic Model.

$Y_{ij} = \beta_0 + \beta_1 SNP_i + \beta_2 PM_{ij} + \beta_3 SNP_i * PM_{ij} + \beta_4 C_i$, where

Y_{ij} is QT (ms) for the i^{th} participant at the j^{th} visit

β_0 is the intercept

PM_{ij} is PM ($\mu g/m^3$)

SNP_i is the dosage of the genetic variant
C_{ij} is the vector of covariables

To identify interactive effects, we propose a stratified, longitudinal analysis of QT interval duration. The initial strategy is to longitudinally model the average of repeated outcomes and thereby facilitate estimation of effects by increasing power using methods we have identified, tested and applied to large-scale genomic data in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Pharmacogenetics Working Group over the last two years under ARIC AS#2009.10. The methods are easily extended to estimate gene-by-PM interactions using conventional generalized estimation equations (GEE). An identity link and an independence working correlation will be used in this context. Although other structures can be accommodated, the latter will ensure consistency of the GEE estimates in the presence of time-varying covariates and protect against potential bias related to the putative effects of past QT prolongation on future PM exposure,[29, 30] Pan and Wall's small-sample GEE extension[31] of Satterthwaite's method of approximating the degrees of freedom[32] associated with the *t* reference distribution will be implemented in R using the bossWithdf package. We assume separate analyses for the seven subpopulations that are then combined by meta-analysis.

More powerful tests of gene-by-PM interaction across ancestral populations will be based on an extension of kernel machine regression methods[33-35] that aggregate SNP-level score test statistics within genes. Such methods are particularly useful in genetically diverse populations where different SNPs may be in linkage disequilibrium with causal SNP(s). They also can accommodate complex SNP interactions, permit covariate adjustment, and do not penalize SNPs with opposing associations within a gene.

The analysis plan described above relies on identical electrocardiographic, genetic and environmental data in the ARIC study and WHI CT, i.e. the same ECG measures estimated by the same ECG reading center using the same methods; approximately 2.5 million SNPs imputed to the same HapMap reference panel; and daily mean PM concentrations estimated at all geocoded participant addresses using the same kriging methods. Although issues related to phenotype and genotype harmonization have been minimized by design, opportunities for replication of genes implicated in gene-by-PM interaction analyses outside the ARIC study and the WHI CT may be limited to those populations with comparable measures, e.g. the Multiethnic Study of Atherosclerosis.

Adjustments. We will include age (year) and median RR interval duration across all twelve leads (ms) as covariables. Analyses will be stratified by study, race, and within the ARIC study, gender. All analyses will be appropriately adjusted for both ancestral admixture (using race-specific principal components in statistical models) and multiple comparisons. To control for potential influences of season, day of week, time of day, health, and weather in models examining genetic modification of the PM-ectopy association, we will also add temporal, sociodemographic, clinical and weather co-variables to models examining the SNP-PM interactions.

Power. In ARIC manuscript proposal #2078 we illustrated having at least 80% power to detect ORs of 1.4-2.1 and 1.1-1.4 for SNP-$PM_{2.5}$ and SNP-PM_{10} interactions as they relate to ventricular ectopy. We therefore expect greater power to detect small

interactions in this context because QT is an interval-scale outcome measure and this analysis includes ~8000 more participants from GARNET and WHIMS. Indeed, depending on the number of risk alleles and assuming a minor allele frequency (MAF) of 0.05, we will have at least 80% power to detect percent decrements in QT of 1.2%-1.7% and 3.9%-5.6% for SNP-PM_{10} and SNP-$PM_{2.5}$ interactions.

Meta-analysis. We will examine heterogeneity among subpopulations as a function of study, race, and gender. Although we expect modest power to detect heterogeneity, meta-analytic methods will be used to combine gene-level results across subpopulations, when appropriate. We also highlight our exploration of potentially more powerful gene-based tests of association, as facilitated by KMR methods referenced above. These methods will allow us to harness the benefits of gene-based tests (e.g. fewer statistical tests and accommodation of a multiplicity of variants) while avoiding the potential for allelic heterogeneity and population-specific patterns of linkage disequilibrium possible when meta-analytically combining summary results across race/ethnicity.

Genome-wide Significance Level. 1 ÷ number of tests

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)?

#1940 – Whitsel, *Modification of PM-Associated Heart Rate Variability in AS #2009.08.* The lead author of that proposal (Whitsel) is a co-author on this proposal. The focus of #1940 is on gene-PM interactive effects related to HRV, while this proposal focuses on gene-PM interactive effects QT interval duration

#2078 – Napier, *Genome-wide Association Study of Particulate Matter and Ventricular Ectopy* in AS #2009.08. I am the lead author and the focus is on ectopy.

#1643 – Avery, *Genetic association of ventricular repolarization and heart rate in the multi-ethnic cohorts of the PAGE Consortium*. The lead author of that proposal (Avery) is a co-author of this proposal. The focus on #1643 is on examining whether mature genetic variants generalize across multi-ethnic populations. There is no consideration of the effects of PM.

Manuscript proposal #1152 (Post, *Genomic Predictors of Sudden Cardiac Death*) and #1434 (Arking, "*GWA and candidate gene studies for sudden cardiac death*) focus on sudden cardiac death and main effects of selected SNPs. Neither focuses on PM-gene interactions as they relate to the QT interval.

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.

References

1. Newton-Cheh C, and Shah R. Genetic determinants of QT interval variation and sudden cardiac death. Curr Opin Genet Dev. 2007, Jun;17(3):213-21.
2. US Environmental Protection Agency. Integrated Science Assessment for Particulate Matter (Final Report). United States Environmental Protection Agency, Washington, DC, EPA/600/R-08/139F. 2009;
3. Franklin M, Zeka A, and Schwartz J. Association between PM2.5 and all-cause and specific-cause mortality in 27 US communities. J Expo Sci Environ Epidemiol. 2007, May;17(3):279-87.
4. Dockery DW, Luttmann-Gibson H, Rich DQ, Link MS, Mittleman MA, Gold DR, et al. Association of air pollution with increased incidence of ventricular tachyarrhythmias recorded by implanted cardioverter defibrillators. Environ Health Perspect. 2005, Jun;113(6):670-4.
5. Ljungman PL, Berglind N, Holmgren C, Gadler F, Edvardsson N, Pershagen G, et al. Rapid effects of air pollution on ventricular arrhythmias. Eur Heart J. 2008, Dec;29(23):2894-901.
6. Van Hee VC, Szpiro AA, Prineas R, Neyer J, Watson K, Siscovick D, et al. Association of long-term air pollution with ventricular conduction and repolarization abnormalities. Epidemiology. 2011, Nov;22(6):773-80.
7. Campen MJ, McDonald JD, Reed MD, and Seagrave J. Fresh gasoline emissions, not paved road dust, alter cardiac repolarization in ApoE-/- mice. Cardiovasc Toxicol. 2006;6(3-4):199-210.
8. Ghelfi E, Rhoden CR, Wellenius GA, Lawrence J, and Gonzalez-Flecha B. Cardiac oxidative stress and electrophysiological changes in rats exposed to concentrated ambient particles are mediated by TRP-dependent pulmonary reflexes. Toxicol Sci. 2008, Apr;102(2):328-36.
9. Henneberger A, Zareba W, Ibald-Mulli A, Rückerl R, Cyrys J, Couderc JP, et al. Repolarization changes induced by air pollution in ischemic heart disease patients. Environ Health Perspect. 2005, Apr;113(4):440-6.
10. Liao D, Shaffer ML, Rodriguez-Colon S, He F, Li X, Wolbrette DL, et al. Acute adverse effects of fine particulate air pollution on ventricular repolarization. Environ Health Perspect. 2010, Jul;118(7):1010-5.
11. Lux RL, and Pope CA. Air pollution effects on ventricular repolarization. Res Rep Health Eff Inst. 2009, May;(141):3-20; discussion 21-8.
12. Samet JM, Rappold A, Graff D, Cascio WE, Berntsen JH, Huang YC, et al. Concentrated ambient ultrafine particle exposure induces cardiac changes in young healthy volunteers. Am J Respir Crit Care Med. 2009, Jun 1;179(11):1034-42.
13. Yue W, Schneider A, Stölzel M, Rückerl R, Cyrys J, Pan X, et al. Ambient source-specific particles are associated with prolonged repolarization and increased levels of inflammation in male coronary artery disease patients. Mutat Res. 2007, Aug 1;621(1-2):50-60.
14. Avery CL, Sethupathy P, Buyske S, He Q, Lin DY, Arking DE, et al. Fine-mapping and initial characterization of QT interval loci in African Americans. PLoS Genet. 2012;8(8):e1002870.

15. Smith JG, Avery CL, Evans DS, Nalls MA, Meng YA, Smith EN, et al. Impact of Ancestry and Common Genetic Variants on QT Interval in African Americans. Circ Cardiovasc Genet. 2012, Dec 1;5(6):647-655.
16. National Institutes of Health. National Human Genome Research Institute. Office of Population Genomics. A Catalog of Genome-Wide Association Studies [Internet]. Available from: http://www.genome.gov/26525384
17. Aarnoudse AJ, Newton-Cheh C, de Bakker PI, Straus SM, Kors JA, Hofman A, et al. Common NOS1AP variants are associated with a prolonged QTc interval in the Rotterdam Study. Circulation. 2007, Jul 3;116(1):10-6.
18. Arking DE, Pfeufer A, Post W, Kao WH, Newton-Cheh C, Ikeda M, et al. A common genetic variant in the NOS1 regulator NOS1AP modulates cardiac repolarization. Nat Genet. 2006, Jun;38(6):644-51.
19. Lehtinen AB, Newton-Cheh C, Ziegler JT, Langefeld CD, Freedman BI, Daniel KR, et al. Association of NOS1AP genetic variants with QT interval duration in families from the Diabetes Heart Study. Diabetes. 2008, Apr;57(4):1108-14.
20. Post W, Shen H, Damcott C, Arking DE, Kao WHL, Sack PA, et al. Associations between Genetic Variants in the *NOS1AP* (CAPON) Gene and Cardiac Repolarization in the Old Order Amish. Human Heredity. 2007;64(4):214-219.
21. Tobin MD, Kähönen M, Braund P, Nieminen T, Hajat C, Tomaszewski M, et al. Gender and effects of a common genetic variant in the NOS1 regulator NOS1AP on cardiac repolarization in 3761 individuals from two independent populations. Int J Epidemiol. 2008, Oct;37(5):1132-41.
22. Newton-Cheh C, Eijgelsheim M, Rice KM, de Bakker PI, Yin X, Estrada K, et al. Common variants at ten loci influence QT interval duration in the QTGEN Study. Nat Genet. 2009, Apr;41(4):399-406.
23. Nolte IM, Wallace C, Newhouse SJ, Waggott D, Fu J, Soranzo N, et al. Common genetic variation near the phospholamban gene is associated with cardiac repolarisation: meta-analysis of three genome-wide association studies. PLoS One. 2009;4(7):e6138.
24. Pfeufer A, Sanna S, Arking DE, Müller M, Gateva V, Fuchsberger C, et al. Common variants at ten loci modulate the QT interval duration in the QTSCD Study. Nat Genet. 2009, Apr;41(4):407-14.
25. Liao D, Peuquet DJ, Duan Y, Whitsel EA, Dou J, Smith RL, et al. GIS approaches for the estimation of residential-level ambient PM concentrations. Environ Health Perspect. 2006, Sep;114(9):1374-80.
26. Liao D, Peuquet DJ, Lin H, Duan Y, Whitsel EA, Smith RL, and Heiss G. National Kriging Exposure Estimation: Liao et al. Respond. Environmental Health Perspectives. 2007, Jul;115(7):A338-39.
27. Liao D, Peuquet D, Duan Y, Whitsel E, Dou J, Smith R, et al. GIS approaches for estimation of resident-level ambient PM concentrations. Epidemiology. 2005;16(5):S28.
28. Liao D, Peuquet D, Duan Y, Whitsel E, Dou J, Smith R, et al. Estimation of resident-level ambient PM concentrations from the US EPA's air quality monitoring database. Epidemiology. 16(5):S27-28.
29. Pepe MS, and Anderson GL. A cautionary note on inference for marginal regression models with longitudinal data and general correlated response data. Communications in Statistics-Simulation and Computation. 1994;23(4):939-951.

30. Pan W, Louis TA, and Connett JE. A note on marginal linear regression with correlated response data. The American Statistician. 2000, Aug;54(3):191-195.
31. Pan W, and Wall MM. Small-sample adjustments in using the sandwich variance estimator in generalized estimating equations. Stat Med. 2002, May 30;21(10):1429-41.
32. Satterthwaite FE. An approximate distribution of estimates of variance components. Biometrics. 1946, Dec;2(6):110-4.
33. Lee S, Wu MC, and Lin X. Optimal tests for rare variant effects in sequencing association studies. Biostatistics. 2012;13(4):762-775.
34. Wu MC, Lee S, Cai T, Li Y, Boehnke M, and Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet. 2011, Jul 15;89(1):82-93.
35. Wu MC, Kraft P, Epstein MP, Taylor DM, Chanock SJ, Hunter DJ, and Lin X. Powerful SNP-set analysis for case-control genome-wide association studies. The American Journal of Human Genetics. 2010;86(6):929-942.