ARIC Manuscript Proposal #2599

PC Reviewed: 8/11/15	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Evaluating evidence of shared genetic effects for cardiac conduction metrics

b. Abbreviated Title (Length 26 characters): ECG pleiotropy

2. Writing Group:

Writing group members: Antoine Baldassari[1], Christy L. Avery[1], Eric A. Whitsel[1, 2], Qianchuan He[3], Dan Arking[4], Elsayed Soliman[5]

[1] Department of Epidemiology, University of North Carolina at Chapel Hill
[2] Department of Medicine, University of North Carolina at Chapel Hill
[3] Public Health Science Division, Fred Hutchinson Cancer Research Center
[4] Center for Complex Disease Genomics, McKusick-Nathans Institute of Genetics Medicine, Johns Hopkins University School of Medicine
[4] Epidemiological Cardiology Research Center (EPICARE), Wake Forest School of Medicine

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. __ARB__

F**irst author**: Antoine Baldassari
Address: University of North Carolina Department of Epidemiology
137 E. Franklin St., Suite 306 CB #8050
Chapel Hill, NC 27514
Phone: 919/966-4312 Fax: 919/966-9800
E-mail: baldassa@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: Christy Avery
Department of Epidemiology
Address: 137 E. Franklin St, Suite 306 CB #8050
Chapel Hill, NC 27514
Phone: 919/966-4312 Fax: 919/966-9800
E-mail: christy_avery@unc.edu

3. Timeline:

Analyses will begin once the manuscript is approved

4. Rationale:

Abnormal cardiac electrophysiology broadly impacts cardiovascular health; for instance, excessively short or long ventricular depolarization-repolarization sequences (QT interval duration outside 350ms to 440ms) meaningfully increase risk for arrhythmias and sudden cardiac death [1]. Recent efforts in genome-wide association studies (GWAS) have started unraveling the high heritability of QT-interval duration and other electrocardiographic (ECG) traits [2-20], and the identification of genetic factors modulating their normal variation has shed an important light on arrythmogenesis [21, 22]. Despite these promising results, much of the heritability in cardiac conduction remains unexplained.

A number of genetic loci identified to date are implicated in more than one ECG trait; for instance, single nucleotide polymorphisms (SNPs) near *SNC5A* and *SNC10A*, which respectively encode the $Na_v1.5$ and $Na_v1.8$ sodium ion channels, have been repeatedly associated with QT [5, 15, 16, 19, 23-26], PR [6, 8, 12] and QRS interval durations [6, 11, 27]. Further evidence for pleiotropy of ECG effects—where single genes influence multiple traits—is apparent across the GWAS literature (Table 1).

Gene pleiotropy in ECG traits opens the door to the use of multivariate statistical methods jointly assessing associations of SNPs with multiple ECG features. These models leverage correlation structures between traits to increase statistical power, and constitute a promising and computationally feasible yet largely unexplored avenue to detect new sources of genetic variation in ECG traits [28, 29].

5. Main Hypothesis/Study Questions:

We propose to assess multivariate relationships between imputed SNPs (1000 genome) and ECG variables measured on the resting, standard 12-lead electrocardiogram that in combination, exhaustively characterize an average heartbeat in the temporal domain (Figure 1): P-wave duration, PR segment duration, QRS-complex duration, ST segment duration, T-wave duration (including U-wave, when present), and TP segment duration. TP-segment duration, corresponding to the isoelectric span preceding atrial depolarization, will be calculated as the difference between RR duration and the sum of all other temporal measures.

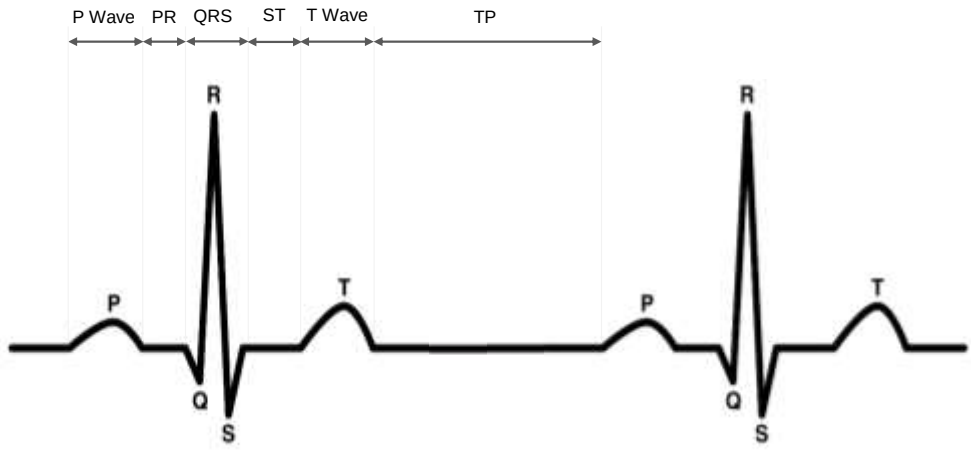

Figure 1. Decomposition of the ECG into interval variables for use in multivariate analyses.

First, race-specific multivariate test statistics constituting global tests for the associations between each SNP and all ECG traits will be computed. Race-stratified and trans-ethnic results combined by meta-analysis will be presented, using our previously described approach [28]. Where significant in the multivariate step, univariate analyses will be conducted to identify the ECG traits responsible for the multivariate signal.

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).

Exclusion criteria

Participants meeting any of the following criteria will be excluded from all analyses: current use of type I or III anti-arrhythmic drugs; paced rhythm; prevalent heart failure (defined by Gothenburg criteria), prevalent coronary heart disease; atrial fibrillation/flutter, ectopy, Wolf-Parkinson-White pattern, second or third degree atrioventricular block, or bundle branch block.

Covariates

We will stratify by race and adjust analyses for age at baseline, sex, ARIC study site, and for population substructure using the first ten principal components from EIGENSTRAT [30].

SNP criteria

In addition to standard QC practices (e.g. excluding all SNPs with oevar_imp < 0.3), we will apply a race/ethnic-specific expected heterozygosity filter of 30 (expected heterozygosity $=$ $2 \times$ minor allele frequency $\times (1 -$ minor allele frequency$)$ $\times$ oevar_imp $\times N$).

Multivariate association test-statistic

We will use a general framework for association tests with multivariate traits that we recently developed [28]. The power of these multivariate models to detect association increases with the extent of the correlations between outcomes, making our approach particularly attractive for the evaluation of ECG traits. Compared with other multivariate methods, this general framework has

the advantages that it allows covariate adjustment, accommodates family data, and is computationally efficient.

Briefly, these methods relate a vector of K traits (in our case, ECG traits) to a set of covariates and a SNP G within a generalized linear model. The score statistic U_k for the null hypothesis that that there is no association between G and the k^{th} trait can be computed. The score vector U is asymptotically K-variate normal with mean 0, and its covariance matrix V can be calculated accordingly.

From there, we test the global null hypothesis that the tested SNP is associated with *none* of the K traits. Most simply, the global test statistic can be calculated as the quadratic form $= U^T V^{-1} U$, which follows a chi-square distribution with K degrees of freedom.

Two other tests statistics T and T' can be calculated, which have greater statistical power under the strong assumptions that the effects of the SNP on all K traits will be similar; however, these can have substantially less power than the Q statistic in the absence of such similarity of effects, and we intend to use the Q statistic in our analyses of ARIC cohort data.

Replication

We are investigating numerous avenues for replication. For instance, we will seek to collaborate with other studies that measured ECG traits and have 1000 genome imputed data. We namely have established relationships with the HCHS/SOL and WHI CT studies, both of which are promising candidates.

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? __X__ 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)?

The most related manuscript proposals in ARIC are those performing GWAS for any of the aforementioned ECG phenotypes. We have attempted to minimize overlap by inviting leaders of the main ECG GWAS to join our writing group.

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? ____ Yes __X__ 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 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 for the Use of Linked ARIC CMS Data, approved manuscripts using linked ARIC 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.

Table 1. Genome-wide significant (10^{-8}) loci associated with multiple cardiac conduction phenotypes.

Chromosome Region	Mapped gene	PR Interval	PR Segment	P wave	QRS Duration	QT Interval	RR	References
11q12.2	FADS2				1		1	[14, 24]
12p12.1	KNOP1P1-RPL21P102	1	1					[8, 17]
12q24.21	TBX3-UBA52P7		1	1				[8, 14]
	TBX5		3			2		[10-13]
13q22.1	KLF12					2	2	[11, 24]
2p14	MEIS1		3	1				[8, 12, 14]
2p22.1	LOC101929667		2				2	[5-7, 24]
3p22.2	SCN5A-SCN10A					1	1	[11, 24]
	SCN10A		6	1	1	3	1	[2, 6, 8-14, 24]
	SCN5A		3			3	3	[2, 8, 10-13, 15, 16, 24]
4q21.23	ARHGAP24		3	1				[8, 10, 13, 14]
6q22.31	CEP85L					1	2	[2, 15, 24]
	SLC35F1	1					1	[17, 24]
7q31.2	CAV1		4	1			1	[6, 8, 10, 13, 14, 24]

References

1. Cubeddu, L.X., *QT prolongation and fatal arrhythmias: a review of clinical implications and effects of drugs.* American journal of therapeutics, 2003. **10**(6): p. 452-457.
2. Ritchie, M.D., et al., *Genome- and phenome-wide analyses of cardiac conduction identifies markers of arrhythmia risk.* Circulation, 2013. **127**(13): p. 1377-85.
3. Seoane, J.A., et al., *Canonical correlation analysis for gene-based pleiotropy discovery.* PLoS Comput Biol, 2014. **10**(10): p. e1003876.
4. Arking, D.E., et al., *Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization.* 2014. **46**(8): p. 826-36.
5. Kim, J.W., et al., *A common variant in SLC8A1 is associated with the duration of the electrocardiographic QT interval.* Am J Hum Genet, 2012. **91**(1): p. 180-4.
6. Hong, K.W., et al., *Identification of three novel genetic variations associated with electrocardiographic traits (QRS duration and PR interval) in East Asians.* Hum Mol Genet, 2014. **23**(24): p. 6659-67.
7. Sano, M., et al., *Genome-wide association study of electrocardiographic parameters identifies a new association for PR interval and confirms previously reported associations.* Hum Mol Genet, 2014. **23**(24): p. 6668-76.
8. Pfeufer, A., et al., *Genome-wide association study of PR interval.* Nat Genet, 2010. **42**(2): p. 153-9.
9. Chambers, J.C., et al., *Genetic variation in SCN10A influences cardiac conduction.* Nat Genet, 2010. **42**(2): p. 149-52.
10. Holm, H., et al., *Several common variants modulate heart rate, PR interval and QRS duration.* Nat Genet, 2010. **42**(2): p. 117-22.
11. Sotoodehnia, N., et al., *Common variants in 22 loci are associated with QRS duration and cardiac ventricular conduction.* Nat Genet, 2010. **42**(12): p. 1068-76.
12. Smith, J.G., et al., *Genome-wide association studies of the PR interval in African Americans.* PLoS Genet, 2011. **7**(2): p. e1001304.
13. Butler, A.M., et al., *Novel loci associated with PR interval in a genome-wide association study of 10 African American cohorts.* Circ Cardiovasc Genet, 2012. **5**(6): p. 639-46.
14. Verweij, N., et al., *Genetic determinants of P wave duration and PR segment.* Circ Cardiovasc Genet, 2014. **7**(4): p. 475-81.
15. Newton-Cheh, C., et al., *Common variants at ten loci influence QT interval duration in the QTGEN Study.* Nat Genet, 2009. **41**(4): p. 399-406.
16. Pfeufer, A., et al., *Common variants at ten loci modulate the QT interval duration in the QTSCD Study.* Nat Genet, 2009. **41**(4): p. 407-14.
17. Eijgelsheim, M., et al., *Genome-wide association analysis identifies multiple loci related to resting heart rate.* Hum Mol Genet, 2010. **19**(19): p. 3885-94.
18. Deo, R., et al., *Common genetic variation near the connexin-43 gene is associated with resting heart rate in African Americans: a genome-wide association study of 13,372 participants.* Heart Rhythm, 2013. **10**(3): p. 401-8.
19. Nolte, I.M., 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): p. e6138.
20. Marjamaa, A., et al., *A common variant near the KCNJ2 gene is associated with T-peak to T-end interval.* Heart Rhythm, 2012. **9**(7): p. 1099-103.
21. Lundby, A., et al., *Annotation of loci from genome-wide association studies using tissue-specific quantitative interaction proteomics.* Nature methods, 2014.
22. van den Boogaard, M., P. Barnett, and V.M. Christoffels, *From GWAS to function: Genetic variation in sodium channel gene enhancer influences electrical patterning.* Trends in Cardiovascular Medicine, 2014. **24**(3): p. 99-104.
23. Arking, D.E., et al., *A common genetic variant in the NOS1 regulator NOS1AP modulates cardiac repolarization.* Nat Genet, 2006. **38**(6): p. 644-51.

24. Arking, D.E., et al., *Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization.* Nat Genet, 2014. **46**(8): p. 826-36.
25. Marroni, F., et al., *A genome-wide association scan of RR and QT interval duration in 3 European genetically isolated populations: the EUROSPAN project.* Circ Cardiovasc Genet, 2009. **2**(4): p. 322-8.
26. Smith, J.G., et al., *Impact of ancestry and common genetic variants on QT interval in African Americans.* Circ Cardiovasc Genet, 2012. **5**(6): p. 647-55.
27. Jeff, J.M., et al., *Generalization of variants identified by genome-wide association studies for electrocardiographic traits in African Americans.* Ann Hum Genet, 2013. **77**(4): p. 321-32.
28. He, Q., C.L. Avery, and D.Y. Lin, *A general framework for association tests with multivariate traits in large-scale genomics studies.* Genet Epidemiol, 2013. **37**(8): p. 759-67.
29. Avery, C.L., et al., *A Phenomics-Based Strategy Identifies Loci on <italic>APOC1</italic>, <italic>BRAP</italic>, and <italic>PLCG1</italic> Associated with Metabolic Syndrome Phenotype Domains.* PLoS Genet, 2011. **7**(10): p. e1002322.
30. Price, A.L., et al., *Principal components analysis corrects for stratification in genome-wide association studies.* Nature genetics, 2006. **38**(8): p. 904-909.