ARIC Manuscript Proposal # 1507

PC Reviewed: 5/12/09	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: The clinical utility of multiple single point polymorphisms in reclassifying risk for incident CHD in the ARIC study

b. Abbreviated Title (Length 26 characters): **Genetic variation, risk assessment and reclassifying risk**

2. Writing Group:

Writing group members:
Ariel Brautbar, MD
Christie Ballantyne, MD
Lloyd E. Chambless, PhD
Aaron Folsom, MD, MPH
Vijay Nambi, MD
Alanna Morrison PhD
Maja Barbalic
Eric Boerwinkle, PhD

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

F**irst author**: **Ariel Brautbar**
Address: Ariel Brautbar, Baylor College of Medicine, One Baylor Plaza, M.S. BCM225, Houston, TX 77030

Phone: 713-798-5034 Fax:
E-mail: brautbar@bcm.tmc.edu

Corresponding/senior author (if different from first author correspondence will be sent to both the first author & the corresponding author):
Christie Ballantyne
Address: Baylor College of Medicine, 6565 Fannin, M.S. A656, Houston, TX 77030
Phone: 713-798-5034 Fax: 713-798-3057
E-mail:

3. Timeline: Analysis to start as soon as approval is obtained. Manuscript is to be prepared as soon as analysis is available. We anticipate that the analysis and manuscript preparation will take place within 1 year from approval of the proposal.

4. Rationale: Prediction of coronary heart disease (CHD) is based on well-established and commonly measured risk factors which are also called "traditional" risk factors. A large number of studies have shown association of single nucleotide polymorphisms (SNPs) and the risk for CHD. In the ARIC study, an initial genetic risk score (GRS) was developed that encompassed 10 SNPs for whites and 11 SNPs for blacks (Morrison et al). The GRS, in addition to the conventional risk factors, has been shown to have a modest impact on prediction of CHD in blacks and in some cases in whites. Another SNP, with a prominent effect on CHD, on the 9p21 chromosomal region, has been thoroughly studied in ARIC and proven to modestly improve CHD risk prediction (Brautbar et al *Circulation: Cardiovascular Genetics*. 2009). Recently a number of new SNP's associated with CHD have been reported in the literature in a series of studies (Myocardial Infarction Genetics Consortium, Nature Genetics 2009; Tre´goue et al, Nature Genetics, 2009; Erdmann et al, Nature Genetics, 2009)

We would like to examine the influence of recently described SNP's, with and without the 9p21 SNP, on CHD risk prediction, to study how this influences clinical classification and what would be the practical outcome of such reclassification. In addition, we would examine the cost-effectiveness of such a test.

We hypothesize that addition of all or part of these SNPs will improve risk classification (i.e., refine classification of patients thought to be intermediate, low, or high risk based on "traditional" risk factors using Framingham/ARIC risk scores) and influence medical management strategy based on the Adult Treatment Panel III (ATP III) treatment guidelines. Reclassification would particularly influence treatment decisions in individuals who are considered to be at intermediate risk by traditional risk factors but who are reclassified as high risk or low risk by the addition of the new SNPs.

5. Main Hypothesis/Study Questions:

Hypothesis: Adding recently reported SNPs, to traditional risk scores such as the ARIC risk score (ACRS) and the 9p21 SNP, will improve risk classification of patients in the various risk groups.

Questions to be addressed in a stepwise manner:

1. Will recently reported SNP's addition to the 9p21 SNP improve risk classification beyond traditional risk factors alone and traditional risk factors combined with the 9p21 SNP?

3. How will addition of the new SNPs influence risk reclassification, and then, applying this new risk classification to the ATP III treatment guidelines.

3. How many individuals would actually require a change in therapy based on the data available in ARIC?

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).

After excluding patients with CHD and stroke at baseline and those who have not provided consent for use of genetic information, all the other patients in the ARIC study on whom ACRS can be calculated and who have available relevant SNP testing will be eligible for the analysis. In addition, analysis will be done specific for a particular gender and race.

We would:

1. Define the ACRS at baseline and classify as low (10-year CHD risk ≤5%), intermediate (10-year CHD risk 5–20%), and high (10-year CHD risk >20%). Also classify patients as defined in ATP III, i.e., intermediate risk as 10-year CHD risk of 10–20%.

2. Describe the incident CHD events (cardiovascular death, myocardial infarction, and revascularization) in the different categories of ACRS and then stratify them after the addition of the recently reported SNPs . We will perform analyses separately for African Americans and Caucasians.

3. Using the Cox proportional hazards model, fit models with traditional ARIC risk factors with the following: (1) with the addition of 9p21 SNPs alone; (2) with recently reported SNP's and SNPs in the 9p21 region. We will then examine the effect on reclassifying risk for incident cardiovascular events including cardiovascular death, myocardial infarction, and coronary artery or cerebrovascular revascularization. We will perform analyses separately for African Americans and Caucasians. In light of the low frequency of women who have intermediate risk scores, we will also perform analyses for the population as a whole and then separately for men and women in regard to reclassification. We will then compute the area under the curve (AROC) for the models with and without the defined SNP's. Then, to assess model calibration or how closely the predicted probabilities reflect actual risk, the following strategies will be applied: Calculate the actual observed risk and then compute the Gronnesby-Borgan test comparing the observed and predicted risk using participant's actual follow-up time, with 10 categories based on 2% point increases in predicted risk ranging from <2% to 18% with and without 9p21/other SNPs. Also compute the statistic using decile categories of predicted probabilities. Clinical utility will be estimated by comparing predicted risk estimates based on models using ACRS with and without 9p21/other SNPs and then using weighted *kappa* statistics to compare the predicted probabilities with and without 9p21/other SNPs. Group the predicted probabilities into 10-year risk categories of 0 to <5%, 5 to <10%, 10% to <20%, and ≥20%. Generate a table as below to describe the same:

10 year risk without genetic evaluation 10 year risk with defined SNP's reclassified

	0 -<5%	5-<10%	10-<20%	>20%	
0 to <5% Total participants 10 year risk					
5 to <10% Total participants 10 year risk					
10 to <20% Total participants 10 year risk					
>20% Total participants 10 year risk					

In addition we will show percentages of participants reclassified and their recomputed predicted risk.

ii. Another strategy that will be used to compare observed and predicted risk is to use a Kaplan-Meier curve (not modeling with risk factors) to get a 10-year observed risk estimation for the cells of the table. We can also obtain predicted risk using traditional risk factors. Following this we will obtain a 10 year predicted risk using the new risk score . We will then compare this to the 10-year observed risk based on the Kaplan-Meier estimate.

iii. We will then examine correct reclassification by calculating the net reclassification index (NRI). We will calculate both the clinical and total NRI and the integrated discrimination improvement (IDI) as previously suggested (Pencina MJ *Stat Med.* 2008; Cook NR. *Stat Med.* 2008).

SNPs' to be examined for the genetics risk score will include the following: rs6922269, rs2943634, rs599839, rs17465637, rs501120, rs17228212, rs4977574, rs646776, , rs1746048, , rs9982601, rs12526453, rs6725887, rs1122608, rs11206510, rs9818870 rs2259816, rs2048327, rs3127599, rs7767084, rs10755578 (based on references 1-4)

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 ICTDER02 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 ICTDER02 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 ICTDER02 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

_____ 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)?
There are two related published papers that were conducted in the ARIC cohort. Both first and last authors on these studies are authors in this proposal (references 5, 6).

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/

12. 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.

Reference:

1. Samani, N.J. et al. Genome wide association analysis of coronary artery disease. N. Engl. J. Med. 357, 443–453 (2007).
2. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants Myocardial Infarction Genetics Consortium. Nat Genet. 2009 Mar;41(3):334-41.
3. Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. Nat Genet. 2009 Mar;41(3):283-5.
4. New susceptibility locus for coronary artery disease on chromosome 3q22.3. Nat Genet. 2009 Mar;41(3):280-2.
5. Morrison AC, Bare LA, Chambless LE, Ellis SG, Malloy M, Kane JP, Pankow JS, Devlin JJ, Willerson JT, Boerwinkle E. Prediction of coronary heart disease risk using a genetic risk score: the Atherosclerosis Risk in Communities Study. Am J Epidemiol. 2007;166:28 –35.
6. Ariel Brautbar; Christie M. Ballantyne; Kim Lawson; Vijay Nambi; Lloyd Chambless; Aaron R. Folsom; James T. Willerson and Eric Boerwinkle. Impact of adding a single allele in the 9p21 locus to traditional risk factors on reclassification of coronary heart disease risk and implications for lipid-modifying therapy in the Atherosclerosis Risk in Communities (ARIC) study. Circulation: cardiovascular genetics. 2009.