ARIC Manuscript Proposal # 1493

PC Reviewed: 04/14/09	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Evaluating various methods to assess improvement when adding risk factors to accepted models for longterm risk

b. Abbreviated Title (Length 26 characters): Evaluating risk prediction

2. Writing Group:

Writing group members: Lloyd Chambless, Gang Cui, Chris Cummiskey, Guoqing Diao

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

First author: Lloyd Chambless

Address: ARIC Coordinating Center
CSCC, Biostatistics Dept
UNC
Chapel Hill NC 27514

Phone: 9199623264 Fax: 9199623265
E-mail: wchambless@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: Lloyd Chambless
Address:

Phone: Fax:
E-mail:

3. Timeline: This would combined work from two finished masters paper, so could be finished by May 1, 2009.

4. Rationale: Several parameters are in use for evaluating the benefits of adding risk factors to accepted models of longterm risk prediction. Some of these have ignored the censoring and time-dependency inherent in the application of these methods to longterm

survival data. The parameters include area under the ROC curve (AUC), an extended AUC suggested by Harrell, proportion of total variance explained by the regression variables (R^2), population attributable risk (PAR) related to having elevated risk score, the ratio of predicted risks in the top and bottom quintiles of risk score, and correlation between risk score and time of event. When traditional risk prediction models are compared with newer extended models, differences in these parameters between the models can be considered. Pencina et al have named the difference in R^2 the integrated discrimination improvement (IDI), and have also introduced the net reclassification improvement (NRI) index. For completeness we will also discuss some statistical tests of goodness-of-fit of the models, the Hosmer Lemeshow chi-squared test and the Gronnesby-Borgan test.

5. Main Hypothesis/Study Questions: The purpose of this paper is to extend the application of these parameters to survival data and to compare estimates of the extended parameters with those ignoring censoring and time-dependency, using both real data from the ARIC study, for prediction of risk of CHD, and from simulated data, in which the true values of the parameters are known. We will also provide SAS macros for computation of the extended parameters.

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).

The study design is that of a longterm cohort study, with risk factors measured at baseline used to predicted incident event over time. The analysis tool will be the Cox proportional hazard model, though the methods are applicable to parametric survival models. Comparison be will be made to parameter estimation that uses logistic regression. For illustration the ARIC cohort data will be used to model incidence of coronary heart disease, through 2004. The risk factors included will be those in traditional risk scores, such as Framingham's or in the ARIC risk prediction papers. Analysis will be separate by sex. Race will be included as a covarariate instead of a stratification variable. No true "novel" risk factors will be included – instead each of the traditional risk factors will be treated as a "novel" factor, to investigate the benefit of its addition to a model excluding it.

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 ICTDER03 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? **____ Yes**
_X___ 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"?
____ Yes ____ No

8.c. If yes, is the author aware that the participants with RES_DNA = 'not for profit' restriction must be excluded if the data are used by a for profit group?
____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
Since this is a methodology paper and includes no novel risk factors, it does not overlap with ongoing research. It will included models that have been presented in published papers, ms611 and ms824.

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)?

None

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.