ARIC Manuscript Proposal # 1621

PC Reviewed: 3/9/10	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: Marginal Additive Hazards Model for Case-cohort Studies with Multiple Disease Outcomes

b. Abbreviated Title (Length 26 characters): Case-cohort Studies

2. Writing Group:
Writing group members: Jianwen Cai, Sangwook Kang, and Woody Chambless

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

First author: Jianwen Cai
Address: Department of Biostatistics, CB#7420
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599-7420

Phone: 919-966-7788 Fax: 919-966-3804
E-mail: cai@bios.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: Department of Biostatistics, CB#7420
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599-7420

Phone: 919-962-3264 Fax: 919-962-3265
E-mail: Lloyd.Chambless@mail.cscc.unc.edu

3. Timeline: Work can begin as soon as approval is received.

4. Rationale:

Modern analyses of survival data focus on multiplicative models for relative risk using proportional hazards models (Cox, 1972), mostly due to desirable theoretical properties along with a simple interpretation of the results and the wide availability of computer programs. However, epidemiologists often are interested in the risk difference attributed

to the exposure, and the risk difference is known to be more relevant to public health because it translates directly into the number of disease cases that would be avoided by eliminating a particular exposure (Kulich and Lin, 2000). Also, the proportional hazards assumption which is critical for proportional hazards models is often violated in practice. Consequently, the additive hazards model, which model risk differences, has often been suggested as an alternative to the proportional hazards model.

The case-cohort design is originally developed for large cohort studies, where the cost to assemble the exposure and covariates is huge. Such design is appealing especially when the event rate is low and the exposure measurements are expensive to obtain. The design effect needs to be taken into consideration when data from case-cohort studies are analyzed. A key advantage of the case-cohort study design is its capacity to use the same subcohort for several diseases or for several subtypes of disease. In order to compare the effect of a risk factor on different types of diseases, times to different events need to be modeled simultaneously. Valid statistical methods that take the correlations among the outcomes from the same subject into account need to be developed. Recently, Kang and Cai (2009) proposed methods for fitting failure time data from case-cohort studies with multiple disease outcomes under marginal proportional hazards models. However, for the additive hazards model, no methods have been proposed for failure time data from case-cohort studies with multiple disease outcomes. In our research, we will propose an estimating procedure and study the properties of the proposed method. As part of the research, we will also apply our method to a real case-cohort data set to illustrate its use. This is the reason we are requesting data from the ARIC study. These data will only be published as part of a methodological paper with acknowledgement of the ARIC study.

5. Main Hypothesis/Study Questions:

We want to use the proposed method to compare the effects of Lipoprotein-associated Phospholipase A2 and high-sensitivity C-reactive protein on incident coronary heart disease and on stroke using data collected under the case-cohort design while adjusting for possible confounding factors.

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).

Use existing data sets under case-cohort design:

> MS#889: Lipoprotein-associated Phospholipase A2, high-sensitivity C-reactive protein and risk for incident coronary heart disease in middle-aged men and women in the Atherosclerosis Risk in Communities (ARIC) study (published 2004).
>
> MS#940: Lipoprotein-associated phospholipase A2, high-sensitivity C-reactive protein, and risk for incident ischemic stroke in middle-aged men and women in the Atherosclerosis Risk in Communities (ARIC) study (published 2005).

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

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)? This is a proposal to illustrate new statistical methodology using ARIC data, and as such there are no similar papers. Ms889 and ms940 published the original results using this cae-cohort data.

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.