

ARIC Manuscript Proposal # 3377

PC Reviewed: 4/9/19

Status: _____

Priority: 2

SC Reviewed: _____

Status: _____

Priority: _____

1.a. Full Title: Improving efficiency for stratified case-cohort studies with multiple disease outcomes using augmentation

b. Abbreviated Title (Length 26 characters): augmentation for case-cohort

2. Writing Group:

Writing group members: Soyoung Kim, Kwang Woo Ahn, and Youngjoo Cho

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

First author: Soyoung Kim

Address: 8701 W Watertown Plank Rd, Milwaukee, WI 53226

Phone: 414-955-8271

Fax: 414-955-0176

E-mail: skim@mcw.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: David Couper

Address: 123 W. Franklin Street, Suite 450, CB #8030, Chapel Hill, NC 27516

Phone: 919-962-3229

Fax:

E-mail: david_couper@unc.edu

3. Timeline: plan to submit by 12/30/2019

4. Rationale:

The case-cohort study design has been widely used to reduce cost in large cohort studies with rare diseases. A case-cohort data consists of a random sample, called the subcohort, and all cases of interest. Case-cohort data is one type of missing data. Missing data can reduce statistical power because of information loss. Therefore, it is important and beneficial to improve

efficiency for the case-cohort study when estimating risk effects for outcomes of interest. If we are able to obtain consistent estimator with smaller standard error, it produces more accurate and reliable results and increase statistical power.

5. Main Hypothesis/Study Questions:

We proposed to improve efficiency for stratified case-cohort studies with multiple diseases using augmented inverse probability weighting (AIPW) method.

Case-cohort data is one type of missing data. One commonly used method to handle missing data is inverse probability weighting method (IPW). This method is to adjust complete case method (i.e. excluding missing data) by using the inverse of the selection probability as weights. It is known that IPW estimators are inefficient. To improve efficiency, we propose AIPW estimator for case-cohort studies with multiple diseases. By augmenting information who of observation who are censored, we can obtain consistent and more efficient estimator than IPW estimators.

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 aim is to investigate the association between PTGS1 polymorphisms and risk of incident coronary heart disease (CHD) and stroke.

Incident CHD is defined as definite or probable myocardial infraction, electrocardiographic evidence of silent myocardial infraction, definite CHD death, or coronary revascularization procedure. Incident stroke is defined as a definite or probable ischemic stroke. We will exclude subjects with missing genotype data or any covariates. Other covariates are age, gender, center, current smoking status, diabetes, hypertension, and race.

Anticipated results are estimated effects of PTGS1 using proposed method are more efficient than those using IPW method.

7.a. Will the data be used for non-CVD analysis in this manuscript? ____ Yes ☒ ____ 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? ☒ ____ 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"? ☒ ____ 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.csc.unc.edu/aric/mantrack/maintain/search/dtSearch.html>

☒ 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)?

Cyclooxygenase Polymorphisms and Risk of Cardiovascular Events: The Atherosclerosis Risk in Communities (ARIC) Study

Improving the efficiency of estimation in the additive hazards model for stratified case-cohort design with multiple diseases

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? ☐ Yes ☒ 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 <https://www2.csc.unc.edu/aric/approved-ancillary-studies>

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.csc.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.