ARIC Manuscript Proposal #2041

PC Reviewed: 12/11/12	**Status: A**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: The effect of selection bias on the relationship between cardiovascular risk factors and mortality

b. Abbreviated Title (Length 26 characters): Selection bias and mortality

2. Writing Group:
Writing group members:

Hailey Banack, PhD candidate, McGill University
Jay Kaufman, PhD, McGill University
Sam Harper, PhD, McGill University
Wayne Rosamond, PhD, University of North Carolina
Lisa Wruck, PhD, University of North Carolina

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

F**irst author**: Hailey Banack
Address:
Department of Epidemiology, Biostatistics, and Occupational Health
Purvis Hall
1020 avenue des Pins Ouest
Montreal, Quebec, Canada
H3A 1A2

Phone: 514-242-4527
E-mail: hailey.banack@mail.mcgill.ca

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (must be an ARIC investigator).
Name: Wayne Rosamond
Address: 137 E Franklin Street, Suite 306 Campus Box 7435
Chapel Hill, North Carolina 27514

Phone: 919-962-3230
E-mail: wayne_rosamond@unc.edu

3. Timeline:
August-October 2012: Data acquisition **April 2013- August 2013**: Writing manuscript
November 2012-March 2013: Data analysis **August 2013**: Submission for publication

4. Rationale:
Evidence from epidemiologic research has repeatedly supported the link between modifiable cardiovascular risk factors (e.g., hypertension, dyslipidemia, diabetes, and cigarette smoking) and the incidence of CVD[1]. A substantial body of literature focused on cardiovascular risk factors has been developed, yet researchers continue to search for new CVD risk factors[1]. Recently, Canto and colleagues reported an inverse association between the number of cardiovascular risk factors and risk of mortality post-myocardial infarction, speculating that the result could be explained by pathophysiologic processes[2]. An accompanying editorial suggested "novel but as-yet uncharacterized and deadly CVD risk factors"[4] could explain the study results. However, rather than highlighting the possibility of novel CVD risk factors, the study better serves to emphasize a key methodological challenge associated with prospectively studying the effects of cardiovascular risk factors on mortality. In the study by Canto et al., admission into the analysis data set was a function of both the exposure (number of cardiovascular risk factors) and outcome (mortality), since deaths occurring prior to hospitalization and patients with pre-existing cardiovascular disease were excluded[2]. Conditioning on a variable affected by both the exposure and outcome produces a selection bias that will distort estimates of the risk factors' association with mortality[3]. The selection bias is produced by differential selection from the source population into the analysis data set, and in this case, it could be so substantial that it even reverses the direction of the expected association, making a greater number of cardiovascular risk factors appear protective. Studies of risk factors that are themselves associated with substantial risk of mortality, such as hypertension, dyslipidemia, smoking, and diabetes are especially vulnerable to selection bias (Figure 1). To our knowledge, no prior studies have quantified and corrected for the potential influence of selection bias on the estimated relation between cardiovascular risk factors and mortality.

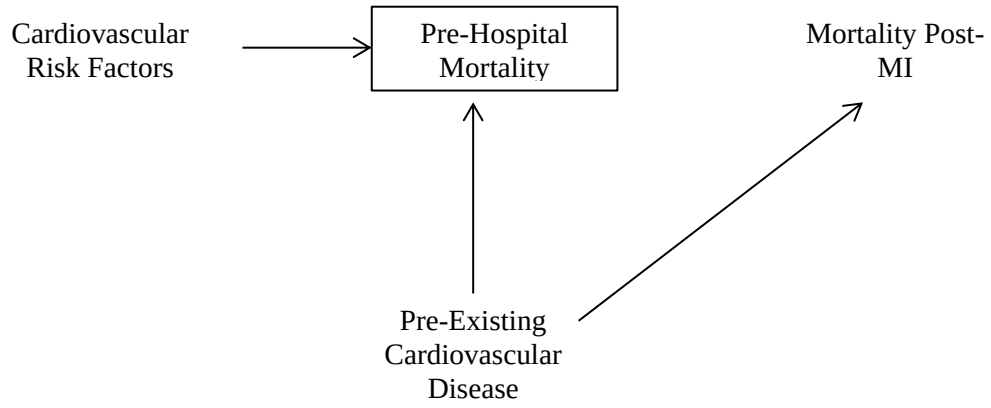

Figure 1. Directed acyclic graph depicting the causal structure underlying selection bias due to pre-hospital mortality in the relation of CVD risk factors and mortality post-MI.

5. Main Hypothesis/Study Questions:
The objective of the proposed study is to quantitatively examine the influence of selection bias due to pre-hospital mortality on the estimated association between cardiovascular risk factors and mortality in ARIC data. Pre-hospital mortality is defined as any death that occurs after an incident MI but prior to hospitalization. We hypothesize that among people hospitalized post myocardial infarction, those with a greater number of cardiovascular risk factors will experience greater mortality (both all-cause and cardiovascular) over the course of the study follow-up. Additionally, contrary to the findings of Canto and colleagues, we expect that after accounting for selection bias due to pre-hospital mortality, there will be a positive correlation between the number of cardiovascular risk factors and mortality post-MI.

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).

Study design
The proposed study will use data collected from visit exams of cohort participates from all four ARIC study sites (Forsyth County, North Carolina; Jackson, Mississippi; Minneapolis, Minnesota; Washington County, Maryland) between 1987 and 1999. Follow-up data on mortality will be used through 2010.

In order quantify the magnitude of selection bias due to pre-hospital mortality, the data set we require for the present study must include the entire ARIC cohort, including information on all MIs (date of MI) between 1987 and 2010, MI hospitalization data, cardiovascular risk factors at all 4 clinic visits, and mortality status up to 2010. Having information on the entire cohort will enable us to quantify pre-hospital mortality due to MI.

Additionally, we require demographic information. See below for further information on the variables required.

Institutional ethics approval has been obtained by the McGill University Research Ethics Board.

Inclusion/exclusion criteria
Inclusion into the ARIC cohort

Outcome variables
Mortality status up to 2010
Date of death
Cause of death (ICD-9 code)

Other variables of interest:

MI hospitalization data (for <u>each</u> MI that occurred in the study period, including repeat MIs):
-Date of hospitalization
-MI type (STEMI/NSTEMI)
-Cardiac procedures (stenting, angioplasty, coronary revascularization, CABG)
-Self reported revascularization procedures from the annual follow up dataset
-Killip Class (S3, pulmonary edema, and shock)

	Visit 1	**Visit 2**	**Visit 3**	**Visit 4**
Participant ID	**X**	**X**	**X**	**X**
Study site	**X**	**X**	**X**	**X**
Date of birth	**X**			
Gender	**X**			
Race	**X**			
Marital Status	**X**	**X**	**X**	**X**
Body Mass Index	**X**	**X**	**X**	**X**
Family History of CVD	**X**	**X**	**X**	**X**
Personal History of CVD	**X**	**X**	**X**	**X**
Systolic Blood Pressure	**X**	**X**	**X**	**X**
Diastolic Blood Pressure	**X**	**X**	**X**	**X**
Blood Lipids (TC, HDL, LDL)	**X**	**X**	**X**	**X**
Smoking	**X**	**X**	**X**	**X**
Diabetes	**X**	**X**	**X**	**X**

Summary of data analysis
Inverse probability weighting will be used to examine the effect of censoring due to pre-hospitalization mortality on the relationship between cardiovascular risk factors and mortality[5]. We will develop models of the probability of remaining alive (i.e., not dying prior to hospitalization) in order to calculate the predicted probabilities of continuation for each observation. We will then compute analytic weights (inverse probability weights) using the predicted probabilities of continuation. The primary exposure variables in the proposed analysis are the four cardiovascular risk factors of interest: hypertension, dyslipidemia, smoking, and diabetes, as well as family history of cardiovascular disease. The primary outcome is mortality. We will examine both all-cause mortality and mortality due to cardiovascular causes. Relevant baseline covariates will be included in the model, including age, sex, race, marital status, and a modified Killip class score.

Methodological limitations and challenges
Participants in the ARIC cohort were receiving screening (at clinic visits) every 3 years, and, as a result, people in the cohort might have a different MI risk than the general population. For example, if an individual had very high blood pressure at the screening

visit, they would be informed of their blood pressure value and possibly seek treatment from their physician, therefore preventing an MI. To verify whether the MI event rate in the ARIC cohort is similar to the general population, we will compare event rates in age strata of the cohort members to event rates in age strata of the community surveillance data.

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 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 __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

_____ Yes _____X__ 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)?

Manuscript #528: Weitzman S, Wang C, Rosamond WD, Chambless LE, Cooper LS, Shahar E, Goff DC. Is diabetes an independent risk factor for mortality after myocardial infarction? The ARIC (Atherosclerosis Risk in Communities) Surveillance Study.. Acta Diabetologica 2004; 41: 77-83.

Manuscript #531: McGinn AP, Rosamond WD, Goff DC Jr, Taylor HA, Miles JS, Chambless L. Trends in prehospital delay time and use of emergency medical services for acute myocardial infarction: experience in 4 US communities from 1987-2000. Am Heart J. 2005 Sep;150(3):392-400.

Manuscript #731: Wattanakit K, Folsom AR, Chambless LE, Nieto FJ. Risk factors for cardiovascular event recurrence in the Atherosclerosis Risk in Communities (ARIC) Study. Am Heart J 2005 Apr; 149(4):606-12.

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.

References:

1. Beaglehole R, Magnus, P. The search for new risk factors for coronary heart disease: occupational therapy for epidemiologists? International Journal of Epidemiology. 2002;31:1117-1122

2. Canto JG, Kiefe CI, Rogers WJ, Peterson ED, Frederick PD, French WJ, Gibson CM, Pollack CV Jr, Ornato JP, Zalenski RJ, Penney J, Tiefenbrunn AJ, Greenland P; NRMI Investigators. Number of coronary heart disease risk factors and mortality in patients with first myocardial infarction. JAMA. 2011 Nov 16;306(19):2120-7.

3. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. Epidemiology. 2004 Sep;15(5):615-25.

4. Peterson ED, Gaziano JM. Cardiology in 2011--amazing opportunities, huge challenges. JAMA. 2011 Nov 16;306(19):2158-9.

5. Weuve J, Tchetgen Tchetgen EJ, Glymour MM, Beck TL, Aggarwal NT, Wilson RS, Evans DA, Mendes de Leon CF. Accounting for bias due to selective attrition: The example of smoking and cognitive decline. Epidemiology; 23(1): 119-128.