ARIC Manuscript Proposal #2382

PC Reviewed: 6/10/14  Status: A  Priority: 2
SC Reviewed: _________  Status: _____  Priority: ____

1.a. Full Title: Examining the Healthy Cohort Effect: Predictors of Attrition in the Atherosclerosis Risk in Communities (ARIC) Study

b. Abbreviated Title (Length 26 characters): ARIC cohort attrition

2. Writing Group:
   Writing group members: Mehul Patel, Anna Kucharska-Newton, Andreea Rawlings, Michael Griswold, Lisa Wruck, Gerardo Heiss, Elsayed Soliman (invited), others welcome

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

First author: Mehul D. Patel, PhD
Address: Department of Epidemiology
         UNC-Chapel Hill
         137 E. Franklin St., Suite 306
         Chapel Hill, NC 27514
         Phone: 919-966-1967  Fax: 919-966-9800
         E-mail: mdpatel@email.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: Anna Kucharska-Newton, PhD, MPH
Address: Department of Epidemiology
         UNC-Chapel Hill
         137 E. Franklin St., Suite 306
         Chapel Hill, NC 27514
         Phone: 919-966-4564  Fax: 919-966-9800
         E-mail: anna_newton@unc.edu

3. Timeline:
The analysis will begin following proposal approval with the aim of completing the manuscript within 1 year.
4. **Rationale:**

Longitudinal studies of aging populations are often complicated by substantial loss of individuals through death and dropout (1). When attrition processes are systematic, estimates of risk factor and outcome associations based on the resulting selected population can be significantly biased (2). This potential selection bias is especially a concern in studies of aging-related outcomes such as cognitive decline since deficits in cognition significantly predict mortality (3, 4) and dropout (5, 6). The biasing effect of attrition is strongest when selection depends on both the outcome and risk factor of interest, either directly or through their predictors. Therefore, it is critical for researchers to consider the sources of attrition and explore systematic reasons to determine their impact on study results. Knowledge of systematic attrition may better inform design strategies to prevent or reduce dropout and statistical methods to account for potential bias.

A previous study noted important differences in socioeconomic status, general health, and cardiovascular risk factors between the Atherosclerosis Risk in Communities (ARIC) participants and those who did not complete the Visit 1 baseline clinic exam (7). Recent Visit 5 examination of ARIC participants provides the opportunity to investigate attrition occurring over more than 25 years of follow-up in a large, population-based cohort. The purpose of this study is to describe the rate of attrition, due to both death and non-death causes, in the ARIC cohort. Further, we will examine sociodemographic characteristics, health behaviors, and physical and mental health states that are predictive of dropout. Lastly, we will explore the vulnerability to bias due to selective attrition in two example outcomes: temporal changes in cognitive performance and electrocardiographic (ECG) traits. Specifically, we will evaluate the degree to which these changes predict attrition, independent of sociodemographics and health characteristics and status.

5. **Main Hypothesis/Study Questions:**

**Hypothesis 1:** Attrition between ARIC Visits 1-4 was largely due to dropout. Due to aging, mortality contributed more than non-participation to attrition occurring from Visits 4 to Visit 5.

**Hypothesis 2:** Conditional on survival, factors most significantly contributing to dropout are old age, low educational attainment, poor cardiovascular risk profile, and poor self-rated health.

**Hypothesis 3:** Given cognitive impairment predicts mortality and study dropout, we expect cognitive decline up to Visit 4 is significantly associated with attrition from Visit 4-5, independent of other predictors of attrition. Alternatively, we anticipate changes in ECG traits (defined by changes in PR, QRS, and QT interval durations up to Visit 4) are less predictive, assuming changes are asymptomatic.
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 population
The ARIC study enrolled 15,792 men and women ages 45-64 years, sampled from 4 U.S. communities: Forsyth County, NC; Jackson, MS; suburban Minneapolis, MN; and Washington County, MD. The first, baseline examination took place from 1987-1989 and was followed by 3 follow-up examinations approximately 3 years apart. The first nine years of cohort follow-up included repeat, comprehensive evaluations of cardiovascular disease, subclinical atherosclerosis, and vascular risk factors (Visits 1-4). In 2011-2013, the fifth ARIC examination (Visit 5) was conducted among 6,538 of the surviving cohort members, ages 70-89 years at the time. Since study onset, cohort members have been monitored for mortality and cardiovascular events and contacted annually to maintain current contact information and health status. Whereas 61% of the 10,748 surviving cohort members participated in the fifth examination (n=6,538), a larger proportion were successfully contacted through annual follow-up, with the overall follow-up response rate at contact year 24 at 83% (Table 1).

Table 1. Participation and follow-up contacts in the ARIC study (1987-2013)

<table>
<thead>
<tr>
<th>Calendar Year</th>
<th>Visit 1</th>
<th>Visit 2</th>
<th>Visit 3</th>
<th>Visit 4</th>
<th>Visit 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of cohort members alive</td>
<td>15,792</td>
<td>15,644</td>
<td>15,267</td>
<td>14,809</td>
<td>10,748</td>
</tr>
<tr>
<td>No. of participants at visit (% of those alive)</td>
<td>(100%)</td>
<td>(91.7%)</td>
<td>(84.4%)</td>
<td>(78.7%)</td>
<td>(60.8%)</td>
</tr>
</tbody>
</table>

*Based on annual follow-up response rates for contact years 4, 7, 10 and 24

The high frequency of clinical exams in the first 9 years of cohort follow-up, compared to the lack of clinical exams in the 15 years between Visit 4 and Visit 5 create two distinct time periods during which mechanisms of attrition may have been different. For that reason, we will examine attrition due to death and to non-death causes separately for the period between Visit 1 and Visit 4 and between Visit 4 and Visit 5.

Sources of attrition
In this study, attrition will be estimated from visit participation and with information collected at annual follow-up (AFU). Study participants’ final statuses, including reasons for contact refusal, are tracked as part of the AFU. Using AFU form versions A-M (administered at contact years 2-24), participants’ final status information will be categorized to contacted and alive, contact and refused, reported alive, reported deceased, and unable to contact or unknown. Attrition will be attributed to death and non-death causes, of which the latter will be further sub-classified into lost to follow-up (alive but not contacted) or non-participation (contacted but refused).
Predictors of non-death attrition
Relevant sociodemographic covariates assessed at Visit 1 will include age, gender, race×center, education level, marital status, occupation status and type, and family income. We will also use baseline prevalence of smoking, alcohol drinking, obesity, hypertension, high cholesterol, and diabetes in examining potential risk factors that may be associated with non-death attrition. Additional factors will be considered, including self-rated general health, objectively measured cognitive and pulmonary function, self-reported health conditions, and distance between participant’s residence and field center. In examining the effect of risk factors on attrition occurring between Visit 4 and Visit 5, we will consider as predictors the prevalence of those risk factors at Visit 4, as well as change in risk factor levels between Visit 1 and Visit 4. Additionally, for continuous, repeated measures, including systolic blood pressure, glucose levels and self-rated health, we will examine trajectories over time from Visit 1 to Visit 4 and use trajectory groups as predictors in examining the risk of attrition.

Example outcomes: changes in cognition and ECG traits
A valid, three-test battery on cognitive function was administered at Visit 2 and 4 participants. We will use these scores to estimate cognitive change leading up to Visit 4 and estimate its effect, as well as the effect of cognitive status at Visit 4, on subsequent attrition to Visit 5. Similarly, we will use repeat ECG data to estimate the influence of changes in ECG traits (PR, QRS, QT interval durations) from Visit 1 to Visit 4 on attrition occurring between Visit 4 and Visit 5.

Statistical analysis
Hypothesis 1: Descriptive analyses will assess the rates of types of attrition (death, loss to follow-up not due to death, non-participation) between Visits 1 and 4 and between Visits 4 and 5, separately. Frequency of visits and age of study participants at the time of Visit 1 as compared to Visit 4, impart a qualitative difference on attrition occurring between Visit 1 and 4 and between Visits 4 and 5. We will therefore perform separate analyses for those two distinct time intervals. We will explore field center effects in stratified analyses.

Hypothesis 2: Multivariable logistic regression will be used to predict non-death attrition, among surviving cohort members, first between Visit 1 and Visit 4 and secondly between Visit 4 and Visit 5 (among those followed to Visit 4). If there are sufficient numbers, we will sub-divide the outcome by types of non-death attrition into a 3-level outcome (not contacted, contacted but refused, participated) and fit multinomial logistic regression models. Predictors with a p-value < 0.1 will be considered statistically significant.

Hypothesis 3: We will model non-death attrition from Visit 4 to 5 on changes in each example outcome (preceding Visit 4) and significant predictors of attrition using multivariable logistic regression. Change in cognitive performance will be defined by Visit 2 and 4 cognitive test scores. In the analysis of ECG traits, we will exclude prevalent and incident CHD and heart failure cases prior to Visit 4. Additional standard exclusions based on medication use and ECG abnormalities will be made.
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.cscu.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)?

MS #1982: Estimation of cognitive change from repeat measures in observational studies; associations with education: the ARIC NCS (Gottesman)

MS #2041: The effect of selection bias on the relationship between cardiovascular risk factors and mortality (Banack)

MS #2115: Sensitivity Analyses with Shared Parameter Models for studying Cognitive Change in the presence of potentially Informative Dropout—the Atherosclerosis Risk in Communities (ARIC) Neurocognitive Study (Griswold)

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data?  ___x__ Yes _____ No

11.b. If yes, is the proposal  ___

A. primarily the result of an ancillary study (list number* __________)  ___x__

B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)*)
To calculate distance to field center, we will use geocoded participants’ addresses collected by AS #1998.02 Life course SES, social context, and CVD (PI: Heiss).

*ancillary studies are listed by number at [http://www.cscc.unc.edu/aric/forms/](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/](http://publicaccess.nih.gov/) are posted in [http://www.cscc.unc.edu/aric/index.php](http://www.cscc.unc.edu/aric/index.php), under Publications, Policies & Forms. [http://publicaccess.nih.gov/submit_process_journals.htm](http://publicaccess.nih.gov/submit_process_journals.htm) shows you which journals automatically upload articles to Pubmed central.

References