ARIC Manuscript Proposal #2447

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

1.a. Full Title: Comparing stroke incidence and survival in American Indians, Blacks, and Whites: the Strong Heart Study, Atherosclerosis Risk in Communities Study, and the Cardiovascular Health Study

b. Abbreviated Title (Length 26 characters): Racial differences in stroke

2. Writing Group:

Writing group members: Clemma Muller; Richard MacLehose, Alvaro Alonso, Jean Forster, David Vock, Rebecca Gottesman, Wayne Rosamond (?); we welcome other suggestions from the ARIC P&P committee.

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

F**irst author**: Clemma Muller
Address: West Bank Office Building
1300 S. Second St., Ste 300
Minneapolis, MN 55454

Phone: 612-327-7237 Fax: 612-624-0315
E-mail: jaco0484@umn.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: Alvaro Alonso
Address: Same as first author

Phone: 612-624-1818 Fax: 612-624-0315
E-mail: alonso@umn.edu

3. Timeline: To be completed no later than May 31, 2015

4. Rationale: In 2010 the prevalence of self-reported stroke was 6% among American Indians, higher than for all other racial and ethnic groups, including Blacks (4%) and Whites (2%).[1] American Indians also have lower mean age at stroke onset than Whites,[2] younger age at death from stroke,[3] and increasing burdens of many stroke risk factors including hypertension, smoking, obesity, and type 2 diabetes.[4-7] The Strong Heart Study (SHS), a population-based cohort study of 4,549 American Indians aged 45-74 at baseline, documented higher stroke incidence for men and women (707 and 653 per 100,000 person years) than observed for Blacks and Whites in other large cohorts.[8] However, no prospective studies exist that allow direct comparison of stroke incidence or post-stroke survival in American Indians vs. other racial groups, and the inferential value of cross-study comparisons erodes as data are parsed into multiple categories defined by age or sex.

The SHS used similar methods and was conducted on a similar timeline as the Atherosclerosis Risk in Communities Study (ARIC) and the Cardiovascular Health Study (CHS).[9-11] The ARIC and CHS enrolled Black and White participants aged 45-64 (ARIC) and ≥ 65 years old (CHS) at baseline, and have generated extensive public health literature on stroke. Accordingly, we propose to combine SHS, ARIC, and CHS data to allow direct comparison of stroke incidence and post-stroke survival in American Indians vs. Blacks and Whites. Although the chronology of enrollment and data collection differs slightly across studies, the similarities in timing, design, and implementation warrant pooling for a combined analysis and there is precedent for pooling ARIC and CHS data for stroke outcomes.[12-18] We will evaluate any stroke as the primary outcome, with secondary analyses separately for ischemic and hemorrhagic stroke risk.

5. Main Hypothesis/Study Questions:

1) Compare stroke incidence in American Indians vs. Blacks and Whites across categories defined by sex and <u>baseline age</u> (45-54, 55-64, 65-69, and 70-74 years old) using Cox regression. *We expect that American Indians will have higher stroke incidence than Blacks or Whites, though these differences may attenuate for older baseline age categories.*

2) Compare stroke incidence in American Indians vs. Blacks and Whites by sex and <u>attained age</u> up to death or censoring using adjusted survival curves. *We expect that American Indians will have higher stroke incidence than Blacks or Whites, with less attenuation for older ages than in the Cox regression analysis of Aim 1.*

3) Among people with incident stroke, compare post-stroke survival for American Indians vs. Blacks and Whites. *We expect that American Indians will exhibit higher fatality rates than Blacks or Whites.*

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).

Design: This will be a cohort analysis, primarily using data collected at baseline and from event surveillance through December 31, 2011. We will pool the SHS, ARIC, and CHS cohorts to allow time-to-event analyses comparing stroke incidence among all participants who were free of stroke at baseline, and post-event survival among the subset of participants who experienced stroke during follow-up.

Inclusion/Exclusion: Analysis for Aims 1 and 2 will include all participants from the three cohorts who did not have prevalent stroke at baseline. We will include prevalent transient ischemic attack as a potential confounder in the adjusted analysis. For Aim 1, CHS participants will be restricted to people who were ≤ 74 years old at baseline, the maximum baseline age represented in the SHS. Analysis of Aim 3 will be restricted to the subset of participants who were free of stroke at baseline and who had an adjudicated primary stroke event on or before December 31, 2011.

Outcome: The primary outcome for Aims 1 and 2 is any adjudicated incident stroke. Secondary analyses will separately evaluate stroke incidence by type (ischemic, hemorrhagic) and severity (fatal, nonfatal). Time to stroke (Aims 1 and 2) or administrative censoring will be measured in years to 2 significant digits. The primary outcome for Aim 3 is post-stroke mortality, measured as binary indicators of 30-day and 1-year survival. The secondary outcome is time in years to 2 significant digits until post-stroke death or administrative censoring.

Other Baseline Variables: The exposure of interest is race (American Indian, Black, White). Cohort (SHS, ARIC, CHS) will also be used to account for potential study-specific differences between the ARIC and CHS (see Data Analysis plan, below). Other baseline variables will include demographics (age, sex, education, study site), lifestyle (smoking and alcohol use), clinical exam (blood pressure, prevalent hypertension, fasting glucose, 2-hour glucose challenge, prevalent diabetes, blood lipids, albuminuria, body mass index, and waist:hip ratio), cardiovascular disease (coronary heart disease, myocardial infarction, congestive heart failure, and atrial fibrillation), and medication (antihypertensive drugs, aspirin).

Post-Baseline Variables: For people who experienced incident stroke during follow-up (subset for analysis of Aim 3), we will also consider prevalent comorbidities (hypertension, diabetes, albuminuria, cardiovascular disease) and medication use that were assessed at follow-up examinations conducted before the stroke occurred.

Data Analysis: **Aim 1:** We will use Cox regression to estimate differences in time to incident stroke, with American Indians as the reference category and separate coefficients for the hazard ratios associated with Blacks and Whites. This analysis will adjust for sex and baseline age, and will also be performed separately for sex-age categories. Because there is no overlap between baseline age in the ARIC (45-64 years old) and CHS (≥ 65 years old), we will pool each study separately with the SHS for this analysis. We will also estimate racial differences in stroke incidence adjusting for potential confounding by baseline health factors (e.g., hypertension, diabetes, prevalent cardiovascular disease) that were similarly measured across the three cohorts. In these models, confounding is viewed as arising from an unmeasured sociocultural cause of race that also influences health conditions relevant to stroke risk, rather than as the health conditions themselves directly acting to cause a person's race.

Aim 2: We will use survival curves to allow evaluation of stroke incidence on the absolute scale (risk difference) in addition to the multiplicative scale (hazard ratio) that is estimated by Cox regression. Survival curves do not rest on the assumption of proportional hazards over follow-up time, and they allow estimation of the underlying risk, or hazard function, that is not directly estimated in Cox regression. We will use a method of covariate adjustment for survival curves that has been previously described.[19] The time scale for this analysis will be attained age (45-49, 50-54, 55-59, etc.), and participants from each study will contribute information to the model for each category in which they survived stroke-free to the minimum age threshold at any point during follow-up. In addition to the survival curves analysis, we will calculate age-specific stroke rates for each attained age category separately by racial group.

Aim 3: This analysis will be restricted to people who experience incident stroke during follow-up. We will use logistic regression to compare 30-day and 1-year post-stroke survival in Blacks and Whites vs. American Indians. We will use marginal standardization to extract predicted probabilities from the logit models, to allow reporting of comparisons using risk differences and risk ratios. The analysis will adjust for age at stroke onset, and potential confounding by comorbidities that were present at or before the stroke event.

Limitations: This proposal has several limitations. First, we assume the pooled data can be treated as if they derive from the same study, after conditioning on study in the inferential analysis. For the Black and White participants of ARIC and the CHS, this assumption can be partially evaluated by including terms in statistical models to account for study and baseline age differences. Because American Indian race is completely conflated with the SHS, however, it is impossible to disentangle study-specific differences from racial comparisons involving American Indians. By restricting the pooled data set to three cohorts with similar study design and timing, and by not including data from other prospective studies (e.g., the REGARDS cohort which was not launched until 2003[20]), we hope to minimize this concern. Nevertheless, the limitation will be carefully considered when interpreting and discussing results. We will take particular care to consider implications of differences in event surveillance and stroke adjudication. For example, the ARIC study design captured in-hospital stroke, whereas the SHS design also captured out-of-hospital strokes. Secondary analyses that separately evaluate fatal and nonfatal stroke events will also help us consider potential bias in outcome ascertainment.

Second, analyzing race as a categorical exposure assumes that any given race label confers the same health effects on everyone to whom it is applied. We acknowledge that this assumption is unlikely to be met. Instead, the meaning of race and its impact on health likely varies across culture, geography, and time. We will therefore interpret racial differences in stroke outcomes as reflecting overall associations while acknowledging the likelihood that population-level differences may not apply equally at the individual level.

Third, evaluation of racial differences in health outcomes can be contentious from both theoretical and methodological perspectives. Specifically, debate persists about race as a marker for sociocultural forces vs. innate biological differences that influence health. We will discuss these perspectives and articulate the pros and cons of each. In addition, we will present results for racial differences estimated with and without adjustment for covariates so that readers can reach their own conclusions depending on whether they

view other health conditions as confounding (sociocultural perspective) or mediating (biological perspective) the race/stroke association.

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

__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)?

1863: Development and validation of prediction models for hemorrhagic and ischemic stroke risk in the Rotterdam Study, Cardiovascular Health Study, and Atherosclerosis Risk in Communities Study.

1663: Risk factors for hemorrhagic stroke II: a pooled study of CHS and ARIC

1111r: Prognosis after MI and stroke: survival, recurrence, and comorbid events in three large prospective studies. (*Note:* the proposal does not name the studies, but I don't see any indication that they are including the SHS.)

1003: Risk factors for hemorrhagic stroke: a pooled study of CHS and ARIC.

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. Centers for Disease C, Prevention. Prevalence of stroke--United States, 2006-2010. *Mmwr.* 2012;61(20):379-382.
2. Frey JL, Jahnke HK, Bulfinch EW. Differences in stroke between white, Hispanic, and Native American patients: the Barrow Neurological Institute stroke database. *Stroke.* 1998;29(1):29-33.
3. Centers for Disease C, Prevention. Prevalence of stroke--United States, 2005. *Mmwr.* 2007;56(19):469-474.
4. Rhoades DA, Welty TK, Wang W, et al. Aging and the prevalence of cardiovascular disease risk factors in older American Indians: the Strong Heart Study. *J Am Geriatr Soc.* 2007;55(1):87-94.
5. Welty TK, Rhoades DA, Yeh F, et al. Changes in cardiovascular disease risk factors among American Indians: The Strong Heart Study. *Ann Epidemiol.* 2002;12(97-106).
6. Jernigan VB, Duran B, Ahn D, Winkleby M. Changing patterns in health behaviors and risk factors related to cardiovascular disease among American Indians and Alaska Natives. *American journal of public health.* 2010;100(4):677-683.
7. Centers for Disease C, Prevention. Declining prevalence of no known major risk factors for heart disease and stroke among adults--United States, 1991-2001. *Mmwr.* 2004;53(1):4-7.
8. Zhang Y, Galloway JM, Welty TK, et al. Incidence and risk factors for stroke in American Indians: the Strong Heart Study. *Circulation.* 2008;118(15):1577-1584.
9. Lee ET, Welty TK, Fabsitz R, et al. The Strong Heart Study. A study of cardiovascular disease in American Indians: design and methods. *Am J Epidemiol.* 1990;132(6):1141-1155.
10. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol.* 1989;129(4):687-702.
11. Fried LP, Borhani NO, Enright P, et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol.* 1991;1(3):263-276.
12. Sturgeon JD, Folsom AR, Longstreth WT, Jr., Shahar E, Rosamond WD, Cushman M. Risk factors for intracerebral hemorrhage in a pooled prospective study. *Stroke; a journal of cerebral circulation.* 2007;38(10):2718-2725.
13. Sturgeon JD, Folsom AR, Longstreth WT, Jr., Shahar E, Rosamond WD, Cushman M. Hemostatic and inflammatory risk factors for intracerebral hemorrhage in a pooled cohort. *Stroke; a journal of cerebral circulation.* 2008;39(8):2268-2273.
14. Ferket BS, van Kempen BJ, Wieberdink RG, et al. Separate prediction of intracerebral hemorrhage and ischemic stroke. *Neurology.* 2014;82(20):1804-1812.
15. Weiner DE, Tighiouart H, Amin MG, et al. Chronic kidney disease as a risk factor for cardiovascular disease and all-cause mortality: a pooled analysis of community-based studies. *Journal of the American Society of Nephrology : JASN.* 2004;15(5):1307-1315.
16. Weiner DE, Tighiouart H, Vlagopoulos PT, et al. Effects of anemia and left ventricular hypertrophy on cardiovascular disease in patients with chronic kidney disease. *Journal of the American Society of Nephrology : JASN.* 2005;16(6):1803-1810.
17. Folsom AR, Yatsuya H, Mosley TH, Jr., Psaty BM, Longstreth WT, Jr. Risk of intraparenchymal hemorrhage with magnetic resonance imaging-defined leukoaraiosis and brain infarcts. *Annals of neurology.* 2012;71(4):552-559.
18. Wilkins JT, Ning H, Berry J, Zhao L, Dyer AR, Lloyd-Jones DM. Lifetime risk and years lived free of total cardiovascular disease. *JAMA.* 2012;308(17):1795-1801.
19. Hernan MA. The hazards of hazard ratios. *Epidemiology.* 2010;21(1):13-15.
20. Howard VJ, Cushman M, Pulley L, et al. The reasons for geographic and racial differences in stroke study: objectives and design. *Neuroepidemiology.* 2005;25(3):135-143.