ARIC Manuscript Proposal #4034

PC Reviewed: 4/12/22 **Status: _____** **Priority: 2**
SC Reviewed: _____ **Status: _____** **Priority: ____**

1.a. Full Title: The Population Burden of Stroke Attributable to Modifiable Risk Factors: Estimates by Race and Ethnicity in the United States

b. Abbreviated Title (Length 26 characters): **Stroke PAFs by Race**

2. Writing Group:

Writing group members: Mark Lee, Kamakshi Lakshminarayan, Behnam Sabayan, Gerardo Heiss, Lin Yee Chen and Pamela L. Lutsey

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

First author: **Mark Lee**
Address: 50 Willey Hall
225 19th Avenue South
Minneapolis, MN 55455
Phone: 763-229-7001
E-mail: leex6611@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: **Pamela Lutsey**
Address: 1300 S 2nd St
300 West Bank Office Building
Minneapolis, MN 55454
Phone: 612-624-5812
E-mail: lutsey@umn.edu

3. Timeline:

Data analysis will begin after approval.

4. Rationale:

In the United States, approximately 795,000 people experience a stroke every year.[1] Although stroke incidence has been declining over the last several decades,[2,3] crude stroke prevalence

among Americans is projected to increase approximately 11% between 2020 and 2030 (from 3.48% to 3.88%).[4] This is because older Americans who are at greater risk for stroke will make up a greater proportion of the population. Consequently, the direct medical costs associated with stroke in the United States is projected to increase from $108 billion in 2020 to $184 billion in 2030—not including the indirect costs of lost productivity among those who have a stroke.[4] Finding effective ways to sustain and enhance stroke rate reductions in the United States is essential to improving population health outcomes and reducing health care costs as the population continues to age.

Encouragingly, previous research has suggested that a large fraction of stroke cases worldwide are attributable to modifiable risk factors. Using data from the INTERSTROKE study that examined risk factors for stroke using a case-control approach in 32 countries, researchers estimated that 90.7% of strokes were attributable to hypertension, physical inactivity, apolipoprotein ratio, poor diet, waist-to-hip ratio, psychosocial factors, smoking, cardiac causes (including atrial fibrillation), excessive alcohol consumption, and diabetes mellitus.[5] This suggests that reducing the prevalence of these risk factors could have a substantial impact on rates of stroke. However, evidence specific to the United States population has not yet been published.

There are racial and ethnic disparities in stroke rates among Americans. Data from population based cohort studies has shown that Black and Hispanic Americans have between 1.5 and 2 times higher risk of incident stroke than non-Hispanic White Americans.[6–8] To design interventions that have an equitable health impact, it is important to investigate whether the risk factors producing the largest attributable fraction of strokes differ across racial and ethnic groups.

The goal of this study is to estimate the percentage of total strokes (both ischemic and hemorrhagic) among non-Hispanic White, non-Hispanic Black, Hispanic, and Asian Americans attributable to well-established and potentially modifiable risk factors for stroke. The risk factors we will consider include atrial fibrillation plus the American Heart Association's Life's Simple 7: blood pressure, glucose, cholesterol, body mass index, physical activity, smoking, and diet. We will also calculate the proportion of stroke cases for each race that could be prevented from a 10% proportional reduction in the prevalence of each risk factor. Additionally, we will calculate the preventable number of stroke cases per 100,000 person years with a 10% proportional reduction in the risk factor.

To calculate PAFs, we will use risk ratios from the most recent meta-analyses along with race-specific prevalence estimates from nationally representative data (e.g., NHANES). Because many of the risk factors co-occur within individuals, we will adjust PAF estimates for communality to calculate the proportion of stroke cases attributable to all risk factors combined. To do this, we will conduct factor analyses using ARIC data, in which all risk factors have been measured. Methods for doing this have previously been described.[9]

5. Main Hypothesis/Study Questions:

Study questions: What are the PAFs for risk of stroke among Black, White, Hispanic, and Asian Americans overall, and attributable to atrial fibrillation and LS7 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).

Study design: To calculate PAFs, we will use risk ratios from most recent meta-analyses combined with race-specific prevalence estimates of each risk factor based on nationally representative data (e.g., NHANES). Using this information, we will calculate PAFs for each risk factor. We will also weight each risk factor's PAF based on communality (derived using ARIC data) to calculate the fraction of stroke cases attributable to all risk factors combined. Note that ARIC is only used as inputs calculation the background communality data and no actual ARIC data will be presented in the published manuscript.

ARIC "background" Communality Weight Calculation

Inclusion/Exclusion: To calculate communality weights, we will exclude ARIC participants with missing data on the relevant risk factors (see below).

Risk Factors Measured in ARIC Visit 1:

- Hypertension (SBP ≥ 140 mm Hg or DBP ≥ 90 mm Hg or antihypertensive medication use)
- High cholesterol (total cholesterol ≥ 240)
- Obesity (BMI ≥30 kg/m^2)
- Smoking
- Physical inactivity (not meeting guidelines of 75 min/week of vigorous or 150 min/week of moderate activity)
- Diabetes (defined by fasting glucose, A1c, and/or medication)
- Diet (meeting ≤ 1 of the AHA LS7 healthy diet components)[10]
- Atrial fibrillation

Statistical analysis:

First, to calculate the unweighted PAFs for each risk factor, we will use the standard formula $P_{e,r}(RR_e-1)/(1+P_{e,r}[RR_e-1])$ in which $P_{e,r}$ is the prevalence of exposure *e* for race group *r* and RR_e is the relative risk of stroke associated with that exposure. This does not involve use of ARIC data.

Next, using ARIC data, we will calculate the communality weights for each risk factor. To do this, we will conduct a principal component analysis of ARIC data to identify latent variables that explain the variance in the observed risk factors. Weights for each risk factor will be calculated as one minus the sum of the square of all factor loadings (i.e., how much each unobserved component explained each measured variable). Then, we will use these weights to calculate the adjusted proportion of stroke cases for each race group attributable to all risk

factors combined. We will use the formula $PAF_{combined} = 1-[(1-w*PAF_1)(1-w*PAF_2)(1-w*PAF_3)\ldots]$ where PAF_1 is the unweighted PAF for the first risk factor, and so on. This method is adopted from a recent Lancet Commission report examining the PAF of dementia.[9]

Next, to evaluate the impact of a proportional 10% reduction in risk factor prevalence, we will calculate potential impact factors (PIFs) using the formula $PIF = \frac{\sum_{i=1}^{n} p_i RR_i - \sum_{i=1}^{n} p'_i RR_i}{\sum_{i=1}^{n} p'_i RR_i}$ where p_i is the observed proportion of cases at the *i*th exposure level and p'_i is the counterfactual proportion of cases at that level given a 10% reduction. We will use the results of the PIF calculation to estimate how many cases of stroke per 100,000 could be prevented with a 10% reduction in risk factor prevalence for each race group.

Limitations:

This analysis has several limitations. First, the risk ratios used in our PAF calculations are drawn from meta-analyses in which the exposure variable was not operationalized and measured in a uniform way. There will also likely be some differences in how the exposure was operationalized and measured for estimating the RR and how it was operationalized and measured in NHANES (especially for diet and physical activity). Additionally, absent reliable evidence to the contrary, we will assume that the relative risk of stroke associated with each risk factor does not vary across racial and ethnic groups. For some risk factors (e.g., atrial fibrillation), it may be difficult to get race-specific prevalence estimates. In this case, we will either drop these risk factors from our analysis or include them only for groups with available data. Our interpretation of the measures of impact will be cautious, acknowledging the general assumptions of measures of impact (e.g., the exposure is causal).[11]

7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? ____ Yes __X__ No

b. If Yes, is the author aware that the current derived consent file ICTDER05 must be used to exclude persons with a value RES_OTH and/or RES_DNA = "ARIC only" and/or "Not for Profit" ? ____ Yes ____ No
(The 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 current derived consent file ICTDER05 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/mantrack/maintain/search/dtSearch.html

___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)?

#3849 – Racial and Ethnic Differences in the Population Burden of Dementia Attributable to Modifiable Risk Factors Among Americans.

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

_X__ A. primarily the result of an ancillary study (list number* 2008.06)

___ 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 https://sites.cscc.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.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. Virani SS, Alonso A, Aparicio HJ, et al. Heart Disease and Stroke Statistics—2021 Update. *Circulation*. 2021;143(8):e254-e743. doi:10.1161/CIR.0000000000000950

2. Aparicio HJ, Himali JJ, Satizabal CL, et al. Temporal Trends in Ischemic Stroke Incidence in Younger Adults in the Framingham Study. *Stroke*. 2019;50(6):1558-1560. doi:10.1161/STROKEAHA.119.025171

3. Koton S, Sang Y, Schneider ALC, Rosamond WD, Gottesman RF, Coresh J. Trends in Stroke Incidence Rates in Older US Adults: An Update From the Atherosclerosis Risk in Communities (ARIC) Cohort Study. *JAMA Neurol*. 2020;77(1):109-113. doi:10.1001/jamaneurol.2019.3258

4. Ovbiagele B, Goldstein LB, Higashida RT, et al. Forecasting the Future of Stroke in the United States. *Stroke*. 2013;44(8):2361-2375. doi:10.1161/STR.0b013e31829734f2

5. O'Donnell MJ, Chin SL, Rangarajan S, et al. Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (INTERSTROKE): a case-control study. *Lancet Lond Engl*. 2016;388(10046):761-775. doi:10.1016/S0140-6736(16)30506-2

6. Gardener H, Sacco RL, Rundek T, Battistella V, Cheung YK, Elkind MSV. Race and Ethnic Disparities in Stroke Incidence in the Northern Manhattan Study. *Stroke*. 2020;51(4):1064-1069. doi:10.1161/STROKEAHA.119.028806

7. Howard VJ, Kleindorfer DO, Judd SE, et al. Disparities in stroke incidence contributing to disparities in stroke mortality. *Ann Neurol*. 2011;69(4):619-627. doi:10.1002/ana.22385

8. Morgenstern LB, Smith MA, Sánchez BN, et al. Persistent ischemic stroke disparities despite declining incidence in Mexican Americans. *Ann Neurol*. 2013;74(6):778-785. doi:10.1002/ana.23972

9. Livingston G, Sommerlad A, Orgeta V, et al. Dementia prevention, intervention, and care. *The Lancet*. 2017;390(10113):2673-2734. doi:10.1016/S0140-6736(17)31363-6

10. Folsom AR, Yatsuya H, Nettleton JA, Lutsey PL, Cushman M, Rosamond WD. Community Prevalence of Ideal Cardiovascular Health, by the AHA Definition, and Relation to Cardiovascular Disease Incidence. *J Am Coll Cardiol*. 2011;57(16):1690-1696. doi:10.1016/j.jacc.2010.11.041

11. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. *Am J Public Health*. 1998;88(1):15-19. doi:10.2105/AJPH.88.1.15