ARIC Manuscript Proposal Form

ARIC Publication Admin Use Only

Publication Committee Review Date: **09/17/24**
ARIC Manuscript Proposal Number: **# 4530**

1.a. Full Title: Determining differences in Stroke Etiology and Stroke Risk by Age: The Atherosclerosis Risk in Communities (ARIC) Study

b. Abbreviated Title (Length 26 characters): Stroke in Older Adults

2. Writing Group [please provide a middle initial if available; EX: Adam L Williams]:

Jing Wang (First), Michelle C. Johansen (Last)

Writing group members (alphabetically): Josef Coresh, Marco Egle, Rebecca F. Gottesman, Kamakshi Lakshminarayan, Chiadi Ndumele, with others welcome

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. _JW__

First author: **Jing Wang**
Address: Phipps 4th floor, suite 446; 600 N Wolfe Street, Baltimore, MD 21287

Phone: 410-955-7144
E-mail: jwang480@jhmi.edu

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (The ARIC author should be involved enough in ARIC to be able to point the lead author to appropriate ancillary study PIs and to be able to search ARIC manuscript proposals if the lead author doesn't have the access needed to do such a search).

Name: **Michelle C. Johansen, MD, PhD**
Address: Phipps 4th floor, suite 477; 600 N Wolfe Street, Baltimore, MD 21287
Phone: 410-955-8210
E-mail: mjohans3@jhmi.edu

3. Timeline: Data cleaning will begin as soon as the proposal is approved, and data has been shared. Data cleaning will take about a month and analysis will begin as soon as data cleaning is done. We anticipate preparing an abstract for submission around Spring 2025 with the full manuscript ready for submission in the fall of 2025.

4. Rationale:

Although the overall prevalence and incidence of stroke is decreasing,[1] temporal trends show that individuals at older ages have a higher incidence of stroke compared to those at younger ages.[2] Approximately 75% of stroke occurs in individuals aged 65 and older and the risk of stroke doubles every decade after age 55.[3] Even though this is the case, most studies related to stroke etiology focus on those aged 65 or younger,[1] which makes those with incident stroke at older age less represented.

Prior work has suggested that stroke subtypes can be age-dependent,[4] so the increased incidence of stroke at older ages may be due to a difference in a particular stroke subtype. For instance, it has been suggested that risk of cardioembolic stroke increases with age, and older adults are more likely to have cardioembolism compared to those at younger ages or those who are middle-aged.[4,5] Data has suggested that large-artery atherosclerosis and small-vessel occlusion are more likely to occur in middle-aged individuals compared to those at younger or older ages, and stroke of other determined causes appears to be more common in younger individuals.[4]

It is likely that risk factors for incident stroke differ between those who have their first stroke at a younger age versus those who have their first stroke, or incident stroke in older age. For example, diabetes increased stroke risk in those at 45-64 years old but this risk increase was not as high among those with incident stroke at older ages.[6] In work by Dr. Egle et al., (MP #4267: *The association between midlife and late life risk factor and chronic disease clustering and stroke incidence and stroke severity* conducted in ARIC that is currently under review), he demonstrated that different risk factor clusters in midlife were only found to be significantly associated with incident stroke among those ≤70 years old. Specifically, there was only one cluster (obesity, hypertension, diabetes and hypertriglyceridemia) that was associated with incident stroke at both younger and older ages. In other work, modifiable risk factors like smoking, sedentary lifestyle, alcohol intake, and obesity appear to increase the risk of acute cerebrovascular events more for those younger than 75 years old, whereas high blood pressure appears more strongly associated with the risk of having an acute cerebrovascular event at 75 and older.[7]

Cardiac disease states are certainly important risk factors for incident stroke at both older and younger ages, but it may be that these cardiac conditions are more important for incident stroke in older age. Atrial fibrillation and left ventricular hypertrophy were found to be significantly associated with stroke risk across age, but appeared more important for incident stroke in the elderly,[6] suggesting that age should be taken into account when assessing the degree of importance of cardiac risk factors in contributing to stroke.

The purpose of understanding these associations would be to tailor stroke prevention efforts for the older adult, and do a better job predicting if a person is at particular high risk for stroke. Regarding the first, it has been shown that older adults do not respond the same to medications as their younger counterparts for reasons that are still being understood. A prime example is the "Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial" (ALLHAT)) which randomized high-risk hypertensive patients to receive hypertension medications to determine whether receiving an angiotensin-converting enzyme inhibitor, a calcium channel blocker, or an alpha-blocker in comparison to diuretics would result in a lower incidence of heart

disease (heart attack) and stroke. It found a significant differential effect by age between combined CHD (defined as coronary revascularization and hospitalized angina) and lisinopril vs. chlorthalidone with those greater than 65 having a higher rate of combined CHD when taking chlorthalidone and those less than 65 having a higher rate of combined CHD when taking lisinopril.[8] This age difference in medication response suggests that specific stroke risk factors might not represent the same disease when studied in the younger versus older adult. Therefore, understanding the association between risk factors and age among individuals with incident stroke is important and further research is needed.

Another important consideration regarding stroke in older age is that males have a higher risk for stroke than females at younger or middle age, but the risk of stroke in females becomes approximately equal to or even higher than in males as they age.[1] This may be due to the fact that females have a longer life expectancy and potentially estrogen loss occurring post-menopause being linked to increased inflammation.[9] Additionally, cardiac risk factors are more common to occur in females at older ages than at younger ages.[10] Prior work highlights that females tend to have a higher risk of having hypertension and atrial fibrillation compared to males as they age, suggesting that there are different risk factors for stroke in females, especially as they age.[11,12] Therefore, determining differences between embolic risk factors and age among those with incident stroke by sex is also critical.

Finally, early identification of individuals at high risk for stroke allows for timely intervention, which can significantly reduce the likelihood of stroke occurrence and its associated morbidity. Given the potential importance of embolic risk factors, the lack of inclusion of those embolic risk factors among existing stroke risk prediction tools such as the CHA2DS2-VASc score[13], and the PREVENT equation[14,15] and the aforementioned data that suggests that risk factors differ across age, we propose to determine how these tools might be best used to predict stroke in older adults.

To accomplish this work, we propose the following three aims:

5. Main Hypothesis/Study Aims:

Aim 1: To determine the difference in risk of cardioembolic stroke by age (≥70 versus <70 years), comparing differences by sex.

We hypothesize that among participants in ARIC who have their first (incident) ischemic stroke, that the risk of cardioembolic stroke will be higher among those at older ages compared to those of younger ages. We also hypothesize that there will be a difference by sex in that females at older ages will have a higher risk of cardioembolic stroke compared to females at younger ages, while the risk of cardioembolic stroke will remain consistent in males across age.

Aim 2: To determine the effect of embolic risk factors on the difference in risk of cardioembolic stroke by age (≥70 versus <70 years).

We hypothesize that among ARIC participants who have an incident ischemic stroke that the risk of cardioembolic stroke will be the same among those at older ages compared to those at younger ages after adjusting for embolic risk factors. We also hypothesize that specific embolic risk factors will mediate the difference in risk of cardioembolic stroke by age.

(Exploratory) Aim 3: To determine if the prediction of ischemic stroke among all participants in ARIC presenting at visit 5 can be improved by including embolic risk factors at older ages.

We hypothesize that traditional stroke risk prediction tools can be improved upon by adding embolic risk factors to these tools at older ages. This aim will be exploratory.

6. Design and analysis:

a) Inclusion/Exclusion criteria:

We will include ARIC participants without prevalent stroke at visit 5 who go on to have incident definite/probable ischemic stroke with complete covariates for analysis for Aims 1 and 2, and then will include all ARIC participants at visit 5 without prevalent stroke with complete covariates for Aim 3 (prediction).

b) Study design

This will be a prospective study to examine participants with available relevant stroke characteristics and other relevant covariates from visit 5 until the time of incident stroke.

c/d) Outcomes, and other variables of interest with specific reference to the time of their collection, and summary of data analysis

Aim 1: To determine the difference in risk of cardioembolic stroke by age (≥70 versus <70 years), comparing differences by sex.

Exposure: Age of stroke among ARIC participants who develop a stroke while on study and are present at visit 5 (dichotomized as older (≥70) versus younger (<70))

Outcome: Definite or probably incident ischemic stroke subtype [binned into embolic (cardioembolic) versus thrombotic (includes lacunar and carotid artery disease, potential for three-way analysis pending power)].

Other variables: We will consider all variables that could potentially serve as confounders or mediators in the analysis. This includes race-center (from visit 1) as well as stroke risk factors most proximal in time but preceding the ischemic stroke event as we have done in other work[16]. These include body mass index (BMI; kg/m2), smoking status (never; former; current), hypertension (defined as present if systolic blood pressure >140 mm Hg, diastolic blood pressure >90 mm Hg, or pharmacologic treatment for hypertension), diabetes (self-reported physician diagnosis of diabetes, nonfasting glucose ≥200 mg/dL, fasting glucose ≥126 mg/dL, or pharmacologic treatment for diabetes), low-density lipoprotein cholesterol (LDL; mg/dL), alcohol consumption (never; former; current). Consideration of specific embolic stroke risk factors is detailed below.

Statistical Analysis: Logistic regression models will be used to examine the odds of cardioembolic stroke comparing older ages (≥70) to younger ages (<70). Cox proportional hazard models were considered, but due to the fact that the date of stroke will be tied to participants' age, logistic regression models were instead chosen.

Covariates as described above will be included in a step-wise adjustment. We will test for a difference in the association between cardioembolic stroke and age by sex, with p-interaction set at ≤0.01 as evidence of possible statistical significance.

Aim 2: To determine the effect of embolic risk factors on the difference in risk of cardioembolic stroke by age (≥70 versus <70 years).

Exposure: Age of stroke among ARIC participants who develop a stroke while on study and are present at visit 5 (dichotomized as older (≥70) versus younger (<70)).

Outcome: Definite or probably incident ischemic stroke subtype [binned into embolic (cardioembolic) versus thrombotic (includes lacunar and carotid artery disease)].

Other variables: We will collect and consider all co-variates as described in aim 1 that could potentially serve as confounders in the analysis. For this aim, we have particular interest in cardiac history and covariates given our hypothesis that these are particularly relevant for stroke etiology at older ages and may actually mediate the difference in risk. These will include a history of coronary heart disease (defined as a myocardial infarction before or at the study visit or if the participant had coronary bypass surgery or an angioplasty of the coronary arteries), atrial fibrillation (visit based, and with ongoing surveillance), heart failure (visit based, and with ongoing surveillance/adjudication). We are also interested in looking at structural embolic risk factors including left ventricular hypertrophy, left ventricular ejection fraction and left atrial volume index from available echo data at visit 5 and p-wave terminal force in lead V1, prolonged p-wave duration, abnormal p-wave axis from available ECG data in ARIC.

Statistical Analysis: We will again use logistic regression models to examine the odds of cardioembolic stroke by age (as for Aim 1), adjusting for each embolic risk factors in separate models to see if and how the association changes. As power allows, we will then perform causal mediation analyses in consideration of total, direct and indirect effect of each embolic risk factors. Covariates as described above will be included in a step-wise adjustment model.

(Exploratory) Aim 3: To determine if the prediction of ischemic stroke among all participants in ARIC presenting at visit 5 can be improved by including embolic risk factors at older ages.

In this aim, we will use common stroke prediction tools to predict the risk of stroke among all participants in ARIC presenting to visit 5. We will then add our embolic stroke risk factors into the prediction model to see if we can improve our prediction of stroke at older ages where traditional risk prediction tools traditionally fall down in accuracy.

Variables for prediction model:
We will include the variables used in the CHA2DS2-VASc score[13] and the PREVENT equation[14,15]:

CHA2DS2-VASc score: age, sex, congestive heart failure, hypertension, diabetes, vascular disease (defined as having prior myocardial infarction, or peripheral artery disease, or aortic plaque at baseline visit 5). Prior stroke or TIA or thromboembolism included in the CHA2DS2-VASc score will be excluded for our prediction model as all ARIC participants included in our study were those without prevalent stroke at baseline visit 5.

PREVENT equation: age, sex, total cholesterol, HDL cholesterol, systolic blood pressure, body mass index, estimated GFR, diabetes, smoking status (former/never; current), anti-hypertensive medication (defined as medications prescribed for lowering blood pressure with ATC code C02, C03, C07, C08, C09), lipid-lowering medication (defined as statin use with ATC code C10AA and C10B).

Embolic risk factors described in aim 2 will be included. All variables for aim 3 will be collected from ARIC baseline visit 5.

Statistical Analysis:
We will start by first predicting the risk of stroke using two different risk prediction tools that are either used in clinical care (CHA2DS2-VASc excluding prior stroke or TIA or thromboembolism), or a more recently updated prediction tool (PREVENT) to predict the risk of stroke for each eligible participant starting at visit 5. Receiver operating characteristic (ROC) analysis will be conducted to compare observed versus predicted stroke risk. The area under the ROC curve will be calculated to obtain the C-statistic to assess the predictive accuracy. We will also separate the dataset by age into different validation sets and conduct prediction accuracy on each of these sets to obtain the ROC curve and its associated C-statistic for comparison across ages.

We will then explore the improvement of predicted stroke risk by introducing embolic risk factors as previously designated to see if this improves the prediction of stroke among those at older ages. We will include embolic stroke risk factors in a step-wise fashion to the model and see how well the model performs. ROC analysis and C-statistic will be used to determine the prediction as described above.

All statistical analyses will be conducted using Stata *v.18.0* (Stata Corp, College Station, Texas, USA).[17]

c) Any anticipated methodologic limitations or challenges if present

We anticipate that there may be some limitations to this work, and we have anticipated them in order to develop strategies to mitigate them prior to data collection and analysis.

We are interested in using cardiac values from the echocardiography evaluation as important embolic risk factors for aim 2. This is one reason, in addition to the age distribution of the ARIC cohort, that we have set our baseline as visit 5. It may be that not everyone completed an echo at visit 5. We will also consider performing our analysis only on those participants with complete echo data.

Risk factors at visits may not capture the actual status at the time of the stroke event. We will collect risk factors most proximal in time but preceding the incident stroke event in order to obtain the most relevant risk factor data. We recognize that this is a limitation of a longitudinal cohort study rather than a clinically based cohort.

We also understand that there are other prediction models that may be used to better predict stroke, but CHA2DS2-VASc score is commonly used clinically and the PREVENT equation is the one of the newest models with updated data/research included to improve the predicted risk of cardiovascular disease, including stroke.

d) Will the author need Limited data to complete the proposed manuscript? ☐ Yes, Limited data is needed (Provide a brief (2-3 sentences) justification for requesting PHI data) . ☒ No, De-identified data will be sufficient.

*Please note, Limited dataset access is strict and rarely provided. Limited data includes identifiable information such as dates (birthdays, visit dates, etc.). CMS, Genomic, Geocoded, Proteomics/Somalogic, and other -omic data all fall under the limited data category. De-identified data does not include dates. All dates are date adjusted to "Days since Visit 1".

7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? (Non-ARIC analysis means that the authors are not regarded as ARIC investigators and the "ARIC author" is essentially just a facilitator rather than an integral part of the writing group.) ☐ Yes ☒ 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 is distributed to ARIC PIs annually, 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

9. The lead author or the "sponsoring" ARIC 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 website at: https://aric.cscc.unc.edu/aric9/proposalsearch [ARIC Website□Publications□Proposal Search]

☒ 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)? x.

Associations between Low-Density Lipoprotein (LDL) Cholesterol Trajectories over the Adult Life Course and Risks of Cardiovascular Disease
Autumn M. Clemons

Age Differences in the Change in Cognition After Stroke: A Pooled Cohort study of ARIC, FOS, REGARDS
Mellanie Springer

The Association between midlife and late life risk factor and chronic disease clustering and stroke incidence and stroke severity
Marco Egle

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or does it use current [or ongoing] ancillary study data (this includes ACHIEVE)? ☐ Yes ☒ No → Skip to question 12

11.b. If yes to 11.a., is the proposal

☐ **A. primarily the result of an ancillary study**

☐ **B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables)**

11.c. If yes to 11.a., list number* __

*ancillary studies are listed by number
https://aric.cscc.unc.edu/aric9/researchers/ancillary_studies/approved_ancillary_studies [ARIC Website Ancillary Studies 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 https://aric.cscc.unc.edu/aric9/publications/policies_forms_and_guidelines [ARIC Website Publications Publication Policies, Forms, and Guidelines].
http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.

References: _

1. Martin SS, Aday AW, Almarzooq ZI, et al. 2024 heart disease and stroke statistics: A report of US and global data from the american heart association. *Circulation (New York, N.Y.)*. 2024;149(8):e347. https://www.ncbi.nlm.nih.gov/pubmed/38264914. doi: 10.1161/CIR.0000000000001209.

2. Scott CA, Li L, Rothwell PM. Diverging temporal trends in stroke incidence in younger vs older people: A systematic review and meta-analysis. *JAMA Neurol*. 2022;79(10):1036–1048. Accessed Jul 21, 2024. doi: 10.1001/jamaneurol.2022.1520.

3. Yousufuddin M, Young N. Aging and ischemic stroke. *Aging (Albany NY)*. 2019;11(9):2542–2544. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6535078/. Accessed Jul 21, 2024. doi: 10.18632/aging.101931.

4. Nacu A, Fromm A, Sand KM, Waje-Andreassen U, Thomassen L, Naess H. Age dependency of ischaemic stroke subtypes and vascular risk factors in western norway: The bergen norwegian stroke cooperation study. *Acta Neurol Scand*. 2016;133(3):202–207. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4744685/. Accessed Jul 25, 2024. doi: 10.1111/ane.12446.

5. Pierik R, Algra A, van Dijk E, et al. Distribution of cardioembolic stroke: A cohort study. *Cerebrovascular Diseases*. 2020;49(1):97–104. https://doi.org/10.1159/000505616. Accessed Aug 12, 2024. doi: 10.1159/000505616.

6. Howard G, Banach M, Kissela B, et al. Age-related differences in the role of risk factors for ischemic stroke. *Neurology*. 2023;100(14). doi: 10.1212/wnl.0000000000206837.

7. Soto-Cámara R, González-Bernal JJ, González-Santos J, Aguilar-Parra JM, Trigueros R, López-Liria R. Age-related risk factors at the first stroke event. *JCM*. 2020;9(7). doi: 10.3390/jcm9072233.

8. The ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. Major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs DiureticThe antihypertensive and lipid-lowering treatment to prevent heart attack trial (ALLHAT). *JAMA*. 2002;288(23):2981–2997. https://doi.org/10.1001/jama.288.23.2981. Accessed 7/18/2024. doi: 10.1001/jama.288.23.2981.

9. Koellhoffer EC, McCullough LD. The effects of estrogen in ischemic stroke. *Transl Stroke Res*. 2013;4(4):390–401. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4275797/. Accessed Aug 30, 2024. doi: 10.1007/s12975-012-0230-5.

10. Mosca L, Barrett-Connor E, Kass Wenger N. Sex/gender differences in cardiovascular disease prevention. *Circulation*. 2011;124(19):2145–2154. https://doi.org/10.1161/CIRCULATIONAHA.110.968792. doi: 10.1161/CIRCULATIONAHA.110.968792.

11. Rodgers JL, Jones J, Bolleddu SI, et al. Cardiovascular risks associated with gender and aging. *J Cardiovasc Dev Dis*. 2019;6(2):19. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6616540/. Accessed Aug 9, 2024. doi: 10.3390/jcdd6020019.

12. Morseth B, Geelhoed B, Linneberg A, et al. Age-specific atrial fibrillation incidence, attributable risk factors and risk of stroke and mortality: Results from the MORGAM

consortium. *Open Heart*. 2021;8(2):e001624. https://openheart.bmj.com/content/8/2/e001624. Accessed Aug 9, 2024. doi: 10.1136/openhrt-2021-001624.

13. Chen LY, Norby FL, Chamberlain AM, et al. CHA2DS2-VASc score and stroke prediction in atrial fibrillation in whites, blacks, and hispanics. *Stroke*. 2019;50(1):28–33. https://doi.org/10.1161/STROKEAHA.118.021453. doi: 10.1161/STROKEAHA.118.021453.

14. Razavi AC, Kohli P, McGuire DK, et al. PREVENT equations: A new era in cardiovascular disease risk assessment. *Circ Cardiovasc Qual Outcomes*. 2024;17(4):e010763. Accessed Jul 22, 2024. doi: 10.1161/CIRCOUTCOMES.123.010763.

15. Khan SS, Matsushita K, Sang Y, et al. Development and validation of the american heart association's PREVENT equations. *Circulation*. 2024;149(6):430–449. Accessed Jul 25, 2024. doi: 10.1161/CIRCULATIONAHA.123.067626.

16. Johansen MC, Chen J, Schneider ALC, et al. Association between ischemic stroke subtype and stroke severity. *Neurology*. 2023;101(9):e913–e921. https://www.neurology.org/doi/10.1212/WNL.0000000000207535. Accessed Jul 29, 2024. doi: 10.1212/WNL.0000000000207535.

17. StataCorp. 2023. stata statistical software: Release 18. college station, TX: StataCorp LLC. .