1. **Full Title:** Chronic inflammation and race-ethnic disparities in ischemic stroke: the ARIC study

b. **Abbreviated Title (Length 26 characters):** Inflammation, race and stroke

2. **Writing Group:**
   1. Writing group members: Cheryl Bushnell, MD, MHS, Lynne Wagenknecht, DrPH, ARIC Investigator, Beverly Snively, PhD, Wake Forest School of Medicine, Bradford Worrall, MD, MHS, University of Virginia, Carol Colton, PhD, Duke University Medical Center, Rebecca Gottesman, MD, PhD, Johns Hopkins School of Medicine, Ron Hoogeveen, PhD, Baylor College of Medicine, Tom Mosley, PhD, University of Mississippi, and Myriam Fornage, PhD, University of Texas, Houston

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

**First author:** Cheryl Bushnell, MD, MHS
Address: Wake Forest School of Medicine, Department of Neurology
Medical Center Boulevard
Winston Salem, NC 27157

Phone: 336-716-2357 Fax: 336-716-7790
E-mail: cbushnel@wakehealth.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:** Lynne Wagenknecht
Address: Wake Forest School of Medicine, Public Health Sciences
Medical Center Boulevard
Winston Salem, NC 27157

Phone: 336-716-7652 Fax: 336-716-6427
E-mail: lwgkncht@wakehealth.edu
3. **Timeline:** The initial pilot study results will be submitted in abstract form for the International Stroke Conference due August 13, 2013 (taking place in San Diego February 12-14, 2014). The manuscript presenting the pilot results will be prepared and submitted by November 2013. The ancillary study will be resubmitted to NIH November 5, 2013. The full manuscript resulting from the R01 (and the larger sample size) will be prepared once funding is received and experiments are completed (we anticipate 2016).

4. **Rationale:** Blacks have a higher risk of stroke than whites. Differences in risk factors explain some, but not all, of the disparity in stroke risk. Inflammatory and vascular remodeling biomarker levels may contribute to these differences.

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

   **Aim 1:** To determine whether visit 2 interleukin-1 receptor antagonist (IL-1ra) levels are associated with incident ischemic stroke and stroke severity in the ARIC cohort, and whether these levels explain the excess risk of stroke in African Americans relative to whites.
   - **Hypothesis 1a:** Elevated IL-1ra levels are associated with ischemic stroke.
   - **Hypothesis 1b:** Elevated IL-1ra levels partially explain the increased risk of ischemic stroke in African Americans relative to whites.
   - **Hypothesis 1c:** Among those with stroke, elevated IL-1ra levels are associated with a greater stroke severity.

   **Aim 2:** To determine the relationship between IL-1 gene family polymorphisms and the phenotype of circulating levels of IL-1ra and IL-1β, stratified by race-ethnicity and stroke status.
   - **Hypothesis:** The IL-1 gene family polymorphisms will be associated with variations of circulating IL-1ra and IL-1β levels, and these relationships will vary by race-ethnicity and stroke status.

   **Aim 3:** To determine whether a broad inflammatory profile is predictive of incident ischemic stroke beyond traditional risk factors in the ARIC cohort.
   - **Hypothesis:** A panel of inflammatory cytokines that have relevance for stroke or other cardiovascular risk factors, including IL-1ra, IL-1 β, IL-4, IL-6, IL-8, IL-10, IL-18, MCP-1, RANTES, MMPs 2, 3, 9, TIMP1, VCAM-1, and YKL40 will provide incremental prediction for incident ischemic stroke beyond traditional risk factors, including hypertension, diabetes, tobacco smoking, and hyperlipidemia.

   **Pilot study:** After the review of the original R01 submission in September 2011, we designed a pilot study to determine trends and signals associated with these biomarkers and race-specific stroke incidence. In addition, the feasibility of measuring biomarkers after many years of storage was addressed with these data. The hypotheses were the same, but we narrowed the scope to the biomarkers listed in the table due to budget constraints. We hypothesized that there are differences in inflammatory cytokine and vascular remodeling in blacks and white that might contribute to the risk of stroke and the severity.
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).

Experimental Design Main study:
A nested case control design will be used for this research question.
Study population: The cases will include African Americans and whites with incident ischemic stroke during follow-up after Visit 2. Controls will consist of participants randomly selected from those stroke-free during follow-up with plasma available from visit 2. Those with missing data will be excluded. There were 920 cases of ischemic stroke (388 African Americans and 532 Whites) occurring between ARIC baseline and end 2008. (We used this number in our sample size and budgetary considerations assuming that by the time our study is funded, several additional years of stroke cases will have accrued). The controls are selected at random from the cohort, resulting in an approximate distribution of 25% African Americans and 75% whites. In order to match the number of cases in African Americans (n=388), there will be 1600 controls, 400 of which will be African American. Detectable differences in IL-1ra are provided for the overall cohort (n=920) and for the race/ethnic subgroups (Tables 2 and 3).

Power calculation for Aim 1: We estimated the minimal detectable difference in IL-1ra for the available sample size. Samples sizes are based on available case samples in ARIC. An estimate of the within group standard deviation is available from our own data. With the sample of 920 stroke cases and 1600 controls, we have 80% power to detect a fairly small difference of 14.1 pg/ml in IL-1ra between cases and controls. Detectable differences for the African American group are larger (22.5 pg/ml) owing to the smaller sample size. With 80% power, we can detect an odds ratio of 1.12 for a one standard deviation increment in IL1ra. Based on our preliminary data, and a much larger observed difference in IL-1ra between stroke cases and controls, we should have sufficient power even in this smaller race group for Specific Aim 1. Natural logarithmic or an alternate transformation will be used, depending on our observed IL-1ra distribution relative to a Gaussian distribution.

The analysis of stroke severity will only include stroke cases. We estimate that 50% of stroke cases will be of mild severity and 50% will be of moderate/severe severity.24 With our estimated total of over 920 stroke cases, we will have power to detect a 27.0 pg/ml mean difference in IL-1ra levels by stroke severity (mild vs. moderate/severe).
Power calculation for Aim 2: Based on available allele frequency data for the SNP that tags IL-1RN VNTR allele 2, the following will be used to estimate the power for detecting differences in IL-1ra levels by genotype (Table 4).

Table. Detectable mean differences between genotype groups based on 80% power using a two-sample t-test for mean difference, assuming Hardy-Weinberg proportions [two-sided testing; alpha = 0.05; sd = 146 (stroke), sd = 61 (no stroke)]

<table>
<thead>
<tr>
<th>Model</th>
<th>MAF</th>
<th>White: Mean Differences</th>
<th>African American: Mean Differences</th>
<th>Whites and African Americans Combined: Mean Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Stroke</td>
<td>No stroke</td>
<td>Stroke</td>
</tr>
<tr>
<td>Dominant</td>
<td>0.10</td>
<td>45</td>
<td>13</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>0.20</td>
<td>37</td>
<td>11</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>0.30</td>
<td>36</td>
<td>10</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>0.40</td>
<td>37</td>
<td>11</td>
<td>44</td>
</tr>
<tr>
<td>Recessive</td>
<td>0.10</td>
<td>179</td>
<td>50</td>
<td>210</td>
</tr>
<tr>
<td></td>
<td>0.20</td>
<td>91</td>
<td>26</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>0.30</td>
<td>63</td>
<td>18</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>0.40</td>
<td>49</td>
<td>14</td>
<td>57</td>
</tr>
<tr>
<td>Trend</td>
<td>0.10</td>
<td>42</td>
<td>12</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>0.20</td>
<td>32</td>
<td>9</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>0.30</td>
<td>28</td>
<td>8</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>0.40</td>
<td>26</td>
<td>8</td>
<td>31</td>
</tr>
</tbody>
</table>

Methods Main Study:

1) Stroke case identification: We will use the ARIC study-wide identification and adjudication of incident stroke as the stroke cases and randomly selected stroke-free control subjects.

2) Covariates from ARIC (determined at visit 2 and during subsequent visits): Age, sex, race-ethnicity, field center, history of hypertension, diabetes, smoking history, alcohol consumption, body mass index, waist-hip ratio, total cholesterol, HDL, LDL, triglycerides, CRP levels, systolic blood pressure, diastolic blood pressure, history of CAD, history of PVD, atrial fibrillation, family history of stroke), history of asthma, fibrinogen level and socioeconomic status.

3) Stroke severity assignment. The ARIC source documents will be used for assigning stroke severity. Within the source documents for the stroke event, the neurologic examination by the neurology consultant, or the admission history and physical, will be used to assign the Wake Forest Stroke Severity Score (WFSSS). The WFSSS was initially developed by stroke physicians for the purpose of assigning appropriate clinical pathways for each category of severity (mild, moderate, and severe). The WFSSS scores represent a range of NIH Stroke Scale scores for each grade of severity, but since the variance in NIHSS increases dramatically with scores over 5 (full range of NIHSS is from 0-40), this will be essentially a qualitative score. The scores will be assigned by 4 stroke faculty at Wake Forest School of Medicine, who will form the stroke committee.
for the study. All of these neurologists have experience with adjudicating stroke events for other cohort studies and clinical trials. In addition, the PI has published studies focused on the value of hospital discharge summaries for estimating stroke severity, as well as prospective studies using stroke severity as the outcome. Each stroke case will be assigned WFSSS by two neurologists, and disagreements will be resolved by consensus from the stroke committee. The PI has reviewed 3 stroke cases from the Forsyth Field Center, and thereby determined that the neurologic exam was easy to find, scoring was straightforward, and it takes about 5 minutes to assign a score for each case. Those patients with imaging evidence of a new stroke, but no identifiable focal neurologic impairment will be scored mild.

Wake Forest Stroke Severity Scale Classification
Mild stroke: Minimal to mild weakness (4+ or better on MRC motor scale), no language deficit or mild dysphasia not interfering with communication needs, alert, or no evidence of inattention/neglect or isolated extinction to 1 modality.
Moderate stroke: Considerable weakness but with antigravity strength (>3+ MRC motor scale), considerable dysphasia that impairs communication ability, lethargic, considerable inattention or extinction to more than 1 sensory modality, unsafe gait (high fall risk, more than just unsteadiness)
Severe stroke: Severe weakness: at best, antigravity only (3+ or less), global or near-global aphasia, obtunded or comatose, severe hemi-inattention or anosognosia.

(3) Laboratory Biomarker assays. The stored plasma aliquots from visit 2 will be used for this study. These are currently stored at the ARIC Atherosclerosis Laboratory at Baylor. Dr. Hoogeveen will be supervising the assay analysis for each biomarker. Many of the biomarkers (IL-1ra, IL-1β, IL-4, IL-6, IL-8, IL-10, IL-18, MCP-1, VCAM-1, RANTES, MMP-2, MMP-3, MMP-9, TIMP-1, YKL-40) will be multiplexed using R&D commercially available assay kits. Because IL-1β is known to be present in low quantities, this will initially be run using high sensitivity assays separate from the multiplex arrays.

(4) Genotyping analysis. We will analyze the SNPs from the list in Table 5 in relationship to IL-1ra and IL-1β levels. These SNPs are already genotyped or imputed in the ARIC GWAS database (Table 6).

Analysis
Aim 1: To determine whether visit 2 interleukin-1 receptor antagonist (IL-1ra) levels are associated with incident ischemic stroke and stroke severity in the ARIC cohort, and whether these levels explain the excess risk of stroke in African Americans relative to whites.

Analysis of incident stroke: In this aim, the analysis will initially include a comparison of the mean IL-1ra levels in ischemic stroke cases vs. controls using a two-sample t-test. Similarly, two-sample t-tests will be used to compare IL-1ra levels between stroke cases and controls separately in African Americans and whites. A descriptive table will include IL-1ra means (SD) by stroke subtype (cardioembolic, large vessel, small vessel, other) and the mean (SD) IL-1ra levels for each type by race-
ethnicity. We will calculate the IL-1β:IL-1ra ratio based on non-transformed levels, and compare ratios in cases vs. controls, and in all of the subgroups mentioned above for IL-1ra levels. Distributions, including outliers will be evaluated, and transformations or non-parametric testing will be applied as needed.

Using cases vs controls as the outcome, logistic regression models will be constructed including the following possible covariates: IL-1ra levels, (or IL-1β:IL-1ra ratio), age, race-ethnicity, sex, vascular risk factors (diabetes, hypertension, dyslipidemia, smoking, CAD, PVD, atrial fibrillation, family history of stroke), history of asthma, fibrinogen and socioeconomic status. Our general approach will be to first describe the bivariate relationships among the covariates of main focus, IL-1ra and race-ethnic group, and the potential confounders, mediating factors, and other covariates. Particularly for the other covariates, this will be helpful to identify factors for which statistical adjustment may be important in the logistic regression modeling. We will examine the nature of these relationships carefully to ensure that we are able to identify outliers and other influential points using graphical methods. In the logistic regression modeling, we will first include IL-1ra level and select the set of additional covariates (excluding race-ethnic group) using forward variable selection with a nominal p<0.10. The resulting coefficient for IL-1ra will then be reported as an adjusted odds ratio estimate and approximate 95% confidence interval. Using this same forward selection approach except starting with race-ethnic group in the model (excluding IL-1ra initially), we will then generate a second model to evaluate whether any race-ethnic difference in risk of ischemic stroke is mediated by differences in IL-1ra levels. We will evaluate this by adding IL-1ra level to this second model, and by determining whether there is a statistically significant effect on the coefficient for race-ethnic group.31, 32

Analysis of stroke severity. The stroke severity analysis is restricted to stroke cases only. The descriptive analysis of severity will initially include a comparison of the IL-1ra levels mean (SD) in mild stroke severity vs. moderate/severe stroke severity using a two-sample t-test. Similarly, two-sample t-tests will be used to compare IL-1ra levels between mild and moderate/severe stroke groups separately in African Americans and whites. A descriptive table will include IL-1ra mean (SD) by stroke subtype (cardioembolic, large vessel, small vessel, other) and the mean (SD) IL-1ra for each type by race-ethnicity. We will calculate the IL-1β:IL-1ra ratio based on non-transformed levels, and compare ratios in mild vs. moderate/severe strokes, and in all of the subgroups mentioned above for IL-1ra levels.

A logistic regression for binary outcomes, or ordinal logistic regression for ordered categorical outcomes, will be used to evaluate whether IL-1a levels are associated with stroke severity, including the testing of proportional odds assumptions as needed. The model will be built using a forward variable selection approach similar to above, and accounting for our hypothesis that inflammatory biomarkers may predict severity of stroke. The model will include the following possible covariates: IL-1ra levels, (or IL-1β:IL-1ra ratio), age, race-ethnicity, sex, vascular risk factors (diabetes, hypertension history, total cholesterol, triglyceride level, HCL-C, LDL-C, systolic blood pressure, diastolic blood pressure, current smoker, history of CAD, history of PVD, atrial
fibrillation, family history of stroke), history of asthma, C-reactive protein, fibrinogen level and socioeconomic status.

If there are major differences in the levels of IL-1ra or the ratio with IL-1 β, we will build race-specific models for Blacks and whites to determine the association of these biomarkers with incident stroke by race.

**Aim 2:** To determine the relationship between IL-1 gene family polymorphisms and the phenotype of circulating levels of IL-1ra and IL-1β, stratified by race-ethnicity and stroke status.

This aim focuses on the specific polymorphisms of interest within the IL-1 gene family (Table 6) and their association with levels of IL-1ra and IL-1 β. The first step will be to calculate the mean (SD) for levels of IL-1ra and IL-1 β by genotype for each SNP. Associations between genotypes and quantitative IL-1ra and IL-1 β will be assessed using multiple linear regression under the additive, dominant, and recessive models, and under the general model for comparisons. Transformations will be performed if the biomarker levels are skewed. The linear regression models will be adjusted for age, sex, field center, principal components for genetic ancestry, BMI, smoking status, hypertension, diabetes, and hyperlipidemia. The regression coefficients (β) from each model will be used to estimate the covariate-adjusted, SNP-specific effect on the plasma biomarker level.

The SNP list for this analysis includes several that were imputed from the ARIC GWAS, in which case dosages will be used in the association analysis. The r-squared for these imputed SNPs was 0.899 or higher.

Haplotype- and gene-based analyses will be performed because of the strong association amongst the IL-1 gene cluster and their proximity on chromosome 2. Assuming ambiguity for many individual haplotypes based on GWAS genotype data, the corresponding statistical methods for unknown linkage phase will be performed. A design matrix for an additive haplotype model will be developed for the SNPs that are not in linkage disequilibrium and are independently associated with IL-1ra and IL-1 β levels in the linear regression model. An association model for each biomarker will be developed using haplotype analysis software. This model is based on a similar haplotype analysis of the IL-1 gene cluster in SLE, performed separately in African Americans vs. whites.

**Aim 3:** To determine whether a broad inflammatory profile will provide incremental prediction of incident ischemic stroke in the ARIC cohort and whether the models differ in African Americans vs. whites.

Potential prognostic biomarkers will be identified initially by testing each biomarker separately for association with the ischemic stroke outcome in a logistic regression model alone and with established clinical variables age, sex, race-ethnicity, hypertension, diabetes, tobacco smoking, and hyperlipidemia. Each biomarker with p<0.10 for
association will be considered further for risk prediction. Quartiles or tertiles of each biomarker level (IL-1ra, IL-1β, IL-4, IL-6, IL-8, IL-10, IL-18, CRP, MCP-1, RANTES, MMPs-2, 3, 9, TIMP1, VCAM-1, YKL40, Table 5) performed on visit 2 serum will be calculated and associations with the stroke outcome will be described using unadjusted odds ratios and 95% confidence intervals. A correlation analysis and other diagnostics for collinearity will be used among the biomarkers to identify collinearities. The biomarkers will then be grouped by function (cytokine, chemokine, or metalloproteinase) and by pro- vs. anti-inflammatory actions. Non-collinear markers within each group will be added to the clinical models to determine the incremental value of the biomarkers over and above these clinical variables and accounting for selection of controls from the cohort. Tests for confounding, e.g. by field center and genetic ancestry, will be performed with markers that significantly impact the model.

We will also model the stroke outcome in whites and African Americans separately for comparison.

**Pilot Study Methods:**

Cases: Participants with ischemic stroke, lacunar subtype, within 5 years of visit 2, including 40 African Americans and 40 whites enrolled at the Forsyth County Field Center. If there are too few stroke cases during the 5 year window from visit 2, we will extend the window. The justification for lacunar stroke cases is because our original preliminary data for the biomarkers are from a cohort of African Americans with lacunar stroke. Keeping the cohort homogeneous would be the best strategy for this small cohort. For the grant resubmission, we will add an exploratory aim that will include biomarkers measured in the other stroke types, as well.

Controls: Participants who are stroke-free with visit 2 blood and follow-up data available longitudinally up to the same time point as the stroke cases, enrolled at the Forsyth County Field Center, matched by age and gender (40 African Americans and 40 whites).

**Procedures**

The ARIC investigators at WFSM will select ID numbers based on the above criteria for cases and controls. WF ARIC investigators will work with the coordinating center and the ARIC lab to identify participants with the largest amount of sample available from visit 2, if possible.

For the cases, we will obtain the original stroke source documents used for adjudication from the Forsyth and Jackson Field Center data warehouse. Dr. Bushnell and one of her stroke section faculty will perform the stroke severity scoring (NIHSS and the Wake Forest Stroke Severity Score). Dr. Snively will assemble the new dataset. ID numbers will be sent to Dr. Hoogeveen who will coordinate the plasma analysis of the biomarkers, which will include IL-1ra, IL-1beta, IL-8, IL-10, MMP9, and TIMP1 using the multiplex assay platform. We anticipate using approximately 400 uL of previously unthawed plasma for the pilot study. We would perform a quick thaw and re-freeze of the samples (i.e. thaw, aliquot out 400 uL, refreeze the remaining sample in original vial and store at -80C). Use of citrated plasma samples will likely reduce the need to quick thaw and re-freeze.
Analysis
Statistical analysis of the pilot study data will include:

1. Comparison of cases and controls (frequency matched by age, gender, race-ethnicity) with respect to characteristics at baseline and follow-up through 5 years: vascular risk factors (diabetes, hypertension, dyslipidemia, smoking, coronary artery disease, atrial fibrillation, family history of stroke). Hypothesis tests will be performed using two-sample t, chi-square and Fisher’s exact tests.

2. Description of these baseline and follow-up characteristics in Blacks versus whites in case and control groups.

3. Comparison of biomarker levels between cases and controls using two-sample t-tests, accounting for multiple comparisons by Bonferroni correction.

4. Comparison of biomarker levels between cases and controls stratified by race-ethnicity.

5. Analysis of association between biomarker levels and severity of stroke in cases, using Spearman correlation.

6. Biomarker data will be transformed for analysis as needed.

Sample power calculations for mean differences in IL-1 ra based on stroke status with two-sided testing, alpha = 0.05, sd1 = 146, sd2 = 61, shown below.

<table>
<thead>
<tr>
<th>Power</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.70</td>
<td>64</td>
</tr>
<tr>
<td>0.75</td>
<td>68</td>
</tr>
<tr>
<td>0.80</td>
<td>72</td>
</tr>
<tr>
<td>0.85</td>
<td>77</td>
</tr>
<tr>
<td>0.90</td>
<td>83</td>
</tr>
</tbody>
</table>

7.a. Will the data be used for non-CVD analysis in this manuscript?  ____ Yes  ____ 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  ____ 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
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.cscuc.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)?


11a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data?  

___x___ Yes _______ No

11b. If yes, is the proposal

___x___ A. primarily the result of an ancillary study (list number* 2011.16)

___ 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
shows you which journals automatically upload articles to Pubmed central.