1.a. Full Title: Exploration of Race Differences in Incidence of Venous Thromboembolism (VTE)

b. Abbreviated Title (Length 26 characters): Race and VTE

2. Writing Group:
   Writing group members: Aaron Folsom, Saonli Basu, Susan Heckbert, Pam Lutsey, Wayne Rosamond, Mary Cushman

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

First author: Aaron Folsom or TBN student
Address:

Phone: Fax:
E-mail: folso001@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:
Address:

Phone: Fax:
E-mail:

3. Timeline: Spring 2018

4. Rationale:

Racial differences in VTE incidence exist but are unexplained. In LITE, the age-adjusted VTE incidence rate was 1.6 (95% CI, 1.2-2.2) fold higher in African Americans, and, starting at any age after 45, the lifetime risk of VTE through age 85 was substantially higher for African Americans than for whites (Figure 1).
Most other US studies of VTE also suggest that African Americans have 30-60% higher VTE risk than whites do. The reasons for this are unclear, though African Americans have greater obesity, more medical conditions leading to provoked VTE, sickle trait, and generally higher levels of many plasma biomarkers associated with VTE (e.g., D-dimer, factor VIII, factor XI). In contrast, African Americans are less likely to have several of the strongest thrombophilic genetic variants related to VTE (e.g., Factor V Leiden). We recently reported that a 5-SNP genetic risk score was less strongly associated with VTE in African Americans than whites.

Our Longitudinal Investigation of Thrombembolism Etiology (LITE, involving ARIC and CHS) is one of the largest studies of VTE in African Americans and whites, so we will systematically examine (a) whether risk factor differences explain (or mediate) the excess VTE risks in African Americans and/or (b) whether race interacts with some risk factors. Of course, race is mainly a social construct in the US, but there are clear differences in frequencies of VTE-related gene variants among race groups, reflecting a contribution of biology, too.

(Citations not provided but available on request.)

5. **Main Hypothesis/Study Questions:**
   1. What are possible reasons for higher VTE incidence in African Americans than whites? 2. Are there clinically relevant race interactions with risk factors for VTE?

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).**

LITE is a prospective study of validated VTE in ARIC (through 2015) and CHS (through 2002). We exclude people with known baseline VTE. The outcome is VTE incidence, and variables of interest are described below.
Question 1: Explanation of race associations: The prospective cohort analyses will mainly be based on visit 1 risk factor measures, and 5,024 African Americans and 15,802 whites at risk, followed for approximately 26 years in ARIC and 12 years in CHS, and having an estimated 418 VTEs in African Americans and 847 VTEs in whites. Assuming no race interaction, VTE risk factors that we could test as potentially explaining the racial difference in VTE include BMI, physical activity, diabetes, smoking, hypertension, eGFR, socioeconomic status, hemostatic factors (fVIII, VWF, fXI, D-dimer, aPTT, protein C), CRP, cardiac biomarkers (NT-pro-BNP, troponin T), sickle cell trait and key genetic variants (F5, F2, ABO, F11, FGG). The majority of analyses can use all 20,826 baseline LITE participants, with an expected 1,265 VTE events, but some analyses will use ARIC alone with 15,385 participants (27% African American) and a total of 1,041 VTE events. The analysis will involve Cox proportional hazards regression, first putting age, sex, and race in the model and estimating the HR for race, and then entering singly or in combination possible explanatory risk factors to determine the degree to which they may “explain” the race HR. As previously reported, this approach to mediation analysis has some shortcomings, so we will explore alternative mediation models as well.

Power for this aim was estimated using the R package ‘powerSurvEpi’, which uses the power calculation formula derived by Hsieh and Lavori for Cox proportional hazards regression. Assuming that the age and sex adjusted HR for race (African American vs. white) is still 1.6, or even 1.4, our power should be >99% to show the HR for race is statistically significant in both all of LITE or in ARIC alone. To determine if other VTE risk factors explain (“mediate”) the race association with VTE, we used the function “powerEpi.default” in the ‘powerSurvEpi’ package, which is equivalent to testing for the mediation effect in Cox regression based on the method of Vittinghoff et al. At α=0.05, we will have 80% power to detect a change in the race HR after adding a potential mediating variable to the model, or equivalently detect a HR of 1.21 of a binary mediator at a 25% prevalence and a correlation of 0.30 with race.

Question 2: Interactions with race: The same factors will be studied for interactions with race, primarily using the Cox model; as most risk factors “operate” multiplicatively, the main test will be for interaction beyond multiplicative using cross-product terms and the likelihood ratio test. For the most part, variables will use their natural scale, but we will use cubic splines to verify the shapes of associations are modeled appropriately. In some cases (e.g., hemostatic factors), only high values affect risk of VTE and so these will be dichotomized. For a binary risk factor with 25% prevalence uncorrelated with race, we will have 80% power to detect a multiplicative interaction HR of 1.51 at α=0.05. The power will be lower if the risk factor is correlated with race. We also will explore additive interactions via additive models and the relative excess risk due to interaction (RERI).

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 exclude those with value RES_DNA = “No use/storage DNA”?  __x__ 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)?

   All related proposals are ours.

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* __2001.16__)  
   ___ 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 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.

13. Per Data Use Agreement Addendum, approved manuscripts using CMS data shall be submitted by the Coordinating Center to CMS for informational purposes prior to publication. Approved manuscripts should be sent to Pingping Wu at CC, at pingping_wu@unc.edu. I will be using CMS data in my manuscript ____ Yes _x___ No.