1.a. Full Title: Validation and re-calibration of a Sudden Cardiac Death risk score

b. Abbreviated Title (Length 26 characters): SCD Risk Score Validation

2. Writing Group:
   Writing group members: Brittany Bogle, Nona Sotoodehnia, Wayne Rosamond, others welcome

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

First author: Brittany Bogle
Address: 137 East Franklin Street
         Suite 306
         Chapel Hill, NC 27514
         Phone: Fax:
         E-mail: bbogle@email.unc.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: Wayne Rosamond
   Address: 137 East Franklin Street
            Suite 306
            Department of Epidemiology
            Gillings School of Global Public Health
            University of North Carolina at Chapel Hill
            Chapel Hill, NC 27514
            Phone: (919)962-3230 Fax: (919) 966-9800
            E-mail: wayne_rosamond@unc.edu

3. Timeline: Analysis can begin immediately upon approval. Goal of a complete manuscript submitted by August 2016.

4. Rationale:
   A 10-year sudden cardiac death (SCD) risk score using traditional cardiovascular risk factors was previously developed by Bogle et al (under review [1]) in the public-use ARIC dataset and validated in the Framingham Heart Study. A Cox proportional hazards
model was used; SCD was defined as an out-of-hospital death fatal coronary heart disease event within one hour of the onset of acute symptoms (“one-hour SCD”). This analysis only considered SCD events within 10 years of the baseline measurements taken at Visit 1.

Two measures, discrimination and calibration, were used to assess the strength of this risk score. In a Cox model, discrimination is the ability of a model to distinguish between those with higher risk compared to those with lower risk. Calibration is the agreement between observed and predicted outcomes. Excellent discrimination and calibration within ARIC public use dataset for one-hour SCD was reported. While discrimination was still excellent in a validation cohort (Framingham), calibration was diminished.

An updated SCD definition has been applied to ARIC and is now available to ARIC investigators. This definition is the result of an ARIC ancillary study where SCD was defined as a “sudden pulseless condition from a cardiac origin in a previously stable individual.” [2]

We propose to assess the validity of this risk score in several ways. Visit 4 will be used to obtain the baseline variables needed to compute the estimated 10-year risk of SCD in the remaining ARIC population. These estimated risks will be compared to observed SCD events using the one-hour SCD definition and updated definition of SCD. We will also reexamine 10-year risk estimation using the updated SCD definition using Visit 1 as baseline. We hypothesize that the discrimination and calibration will be diminished when compared to the original analysis. As such, we will explore strength of recalibrating the score.

The lead author is a CVD postdoctoral trainee. One objective of the CVD program is to develop study validation and replication skills for the trainees in the UNC CVD epidemiology group. This study complements this goal.

5. Main Hypothesis/Study Questions:

Aim 1: Assess the validity of the SCD 10-year risk score in the aging ARIC population by using Visit 4 baseline variables.
   Aim 1.1: Determine the calibration and discrimination of the risk score with respect to the updated SCD definition
   Aim 1.2: Determine the calibration and discrimination of the risk score with respect to the one-hour SCD definition

Aim 2: Assess the validity of the SCD 10-year risk score using Visit 1 as the baseline with respect to the updated SCD definition

Aim 3: Recalibrate the risk score using the updated SCD definition

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).
Aim 1:

10-year SCD risk will be calculated for individuals in ARIC using risk factors obtained at Visit 4.

**Discrimination and Calibration**

Discrimination will be measured by Harrell’s c-statistic. Calibration will be computed through dividing participants into deciles by increasing average predicted probability of SCD according to the score’s predicted risk. Within each decile, the average observed SCD incidence rate will be compared to the predicted SCD incidence rate. Differences will be quantified using the Hosmer-Lemeshow statistic. For **Aim 1.1**, the observed SCD rate will be computed using the updated SCD definition. For **Aim 1.2**, the observed SCD rate will be computed using the one-hour SCD definition.

**Variables**: Follow-up time from Visit 1 until any death or SCD death within 10 years of follow-up, age, sex, traditional cardiovascular risk factors (e.g. blood pressure values, smoking status, lipid lowering medication use)

**Outcomes**: **Aim 1.1**: updated SCD outcomes; **Aim 1.2**: one-hour SCD outcomes

Aim 2:

10-year SCD risk will be calculated for individuals in ARIC using risk factors obtained at Visit 1.

**Discrimination and Calibration**

Discrimination and calibration will be measured as described in Aim 1. The observed updated SCD definition will be used in assessing discrimination and calibration.

Table 1: Example table of metrics that will be evaluated during Aim 1 and Aim 2

<table>
<thead>
<tr>
<th>Baseline</th>
<th>SCD outcomes through</th>
<th>Performance with respect to one-hour SCD</th>
<th>Performance with respect to updated SCD definition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Discrimination</td>
<td>Calibration</td>
</tr>
<tr>
<td>Visit 1</td>
<td>10 years from Visit 1 (approx. Visit 4)</td>
<td>N/A – used to construct the original risk score and previously reported</td>
<td>e-statistic</td>
</tr>
</tbody>
</table>
Aim 3: Recalibrate and explore the capability of developing a new predictive model using the updated SCD definition as the outcome

Recalibration is the process of using the same variables in a given model and optimizing the maximum likelihood equation to produce the coefficients used in the risk score for a different outcome or population. Recalibration will be performed in three phases.

Phase 1: Visit 1 will be considered baseline. SCD events (using the updated definition) within 10 years of Visit 1 will be the outcome of interest. Instead of a risk computed using the coefficients from previous work as in Aim 1 and Aim 2, coefficients will be fit to the new definition. Discrimination and calibration will be assessed. This score will be validated on all living ARIC participants using baseline values from Visit 4 to estimate the 10-year risk of SCD; discrimination and calibration will be reported for the validation set.

Phase 2: Visit 4 will be considered baseline. SCD events (using the updated definition) within 10 years of Visit 4 will be the outcome of interest. Coefficients will be fit to the new definition using the same variables as in the original score. Discrimination and calibration will be assessed. This score will be validated on all living ARIC participants using baseline values from Visit 4 to estimate the 10-year risk of SCD; discrimination and calibration will be reported for the validation set.

Phase 3: Visit 1 baseline and those living at Visit 4’s baseline values will be considered as from populations with respect to death; that is, those that experienced SCD will be removed from Visit 4 at baseline. Combining these two populations into a single baseline dataset allows us a more diverse range of ages and risk factors for which to recalibrate the SCD risk score.

70% of the combined dataset will be randomly selected to recalibrate the risk score to; this is referred to as that training set. The risk score will then be applied to the remaining 30% of the population to assess discrimination and calibration performance.

Table 2: Phases in Aim 3

<table>
<thead>
<tr>
<th>Phase</th>
<th>Baseline SCD outcomes through</th>
<th>Performance Using updated SCD defn</th>
<th>Baseline SCD outcomes through</th>
<th>Performance Using updated SCD defn</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Discrimination</td>
<td>Calibration</td>
<td>Discrimation</td>
</tr>
</tbody>
</table>
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 list under the Study Members Area of the web site at: http://www.cscce.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)?

MS #1888: Assessment of Conventional Cardiovascular Risk Factors and Multiple Biomarkers for the Prediction of Sudden Cardiac Death Writing group members: Rajat Deo, Suma Konety, Selcuk Adabag, Alvaro Alonso, Ronit Katz, Nona Sotoodehnia,
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  
  ___  A. primarily the result of an ancillary study (list number* __________)  
  _x_  B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* 2013.12)  

*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] A community-based risk score for sudden cardiac death. Bogle, Brittany; Mehrotra, Sanjay; Goldberger, Jeff; Ning, Hongyan; Lloyd-Jones, Donald M. under review at JACC.