1.a. **Full Title**: Accuracy of the Pooled Cohort Equation to Estimate Adverse Cardiovascular Events in Patients with Obesity

b. **Abbreviated Title (Length 26 characters)**: Accuracy of the PCE in Obesity

2. **Writing Group**:
   Writing group members: Ian J. Neeland, Rohan Khera, Vijay Nambi (ARIC), Colby Ayers (statistician), Chiadi Ndumele (ARIC)

I, Ian Jason Neeland, confirm that all the coauthors have given their approval for this manuscript proposal. IN

**First author**: Ian J. Neeland, MD
Address: 5323 Harry Hines Blvd., Dallas, TX 75390-8830

Phone: 214-645-1267   Fax: 214-645-2480
E-mail: ian.neeland@utsouthwestern.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: Vijay Nambi MD
Address: Alkek Tower, Room A679B, 6565 Fannin St., Houston, TX, 77030

Phone: 713-798-5800   Fax:
E-mail: vnambi@bcm.edu

3. **Timeline**: 3-6 months

4. **Rationale**:
The pooled cohort equation (PCE) was introduced in 2013 as a sex- and race-specific method of estimating expected rates of 10-year absolute risk for atherosclerotic cardiovascular disease (ASCVD) in a primary prevention population. The risk estimates were based on a combination of previously established cardiovascular risk factors that were examined prospectively in
selected cohorts. Patient factors currently included in the PCE are age, sex, race (Caucasian, African American or other), smoking status, as well as presence of hypertension, diabetes and recorded total and HDL-cholesterol levels.\textsuperscript{1, 2} \textit{A priori} defined cut-offs for risk-estimates are then used to define need for future preventative therapies, e.g. lifestyle modification and statins.\textsuperscript{3} Therefore, given the impact of PCE estimates on risk-stratification and future risk-modification strategies, it is critical to ensure that they perform adequately in populations in high-risk populations.\textsuperscript{4}

Individuals with obesity represent one such high-risk population that constitutes around 35\% of the US adult population.\textsuperscript{5-7} While obesity confers an elevated cardiovascular risk, some of this excess risk may not be explained by derangements in obesity-associated comorbidities present in the PCE.\textsuperscript{8} Hence, estimates of risk derived predominantly from risk factors may underperform in assessing the risk of future adverse cardiovascular events in the obese population. Furthermore, risk estimates may be enhanced by the degree (obesity class) or location (e.g. abdominal) of adiposity, not currently utilized in the PCE.

5. \textbf{Main Hypothesis/Study Questions:}

Therefore, using a database constructed from five large community-based epidemiologic cohorts, we aim to:

1. Assess the performance of the PCE in an obese population using metrics of discrimination and calibration.
2. Evaluate potential obesity-specific modifications to the PCE, which may improve the discrimination, calibration, and reclassification of the PCE model in this high-risk subset.
3. Secondary analysis: Compare findings in the obese population with the non-obese population of this dataset

\textbf{Hypotheses:} The PCE does not accurately predict 10-year ASCVD risk in an obese population. Addition of obesity-specific covariates will improve discrimination and calibration of the PCE in this population.

6. \textbf{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).}

\textbf{Data Sources:}
MESA – Multiethnic Study of Atherosclerosis
ARIC - Atherosclerosis Risk In Communities
CARDIA - Coronary Artery Risk Development in Young Adults
DHS – Dallas Heart Study
CHS - Cardiovascular Health Study

\textbf{Variables of interest:}
Visit 1 Data:
Body mass index, age, total cholesterol level, HDL cholesterol level, systolic blood pressure, diabetes status, smoking status, sex, race (white, African American, other), and antihypertensive medication use.
Waist circumference (not used to define exclusion criteria).

Visit 4 Data:
High sensitivity C-reactive protein (not used to define exclusion criteria).

**Study outcomes:** Primary: Risk of acute myocardial infarction, coronary heart disease death, or fatal or non-fatal ischemic stroke at 10 year follow up. Secondary outcome: inclusion of coronary or cerebral revascularization in the composite endpoint.

**Inclusion criteria:**
All adults 40-79 years of age with a BMI measurement.
Participants will subsequently be stratified by obesity status (< or ≥ 30 kg/m²)

**Exclusion criteria**
1. Patients with established atherosclerotic cardiovascular disease at study entry, manifested by prior history of coronary artery disease (stable angina, myocardial infarction), cerebrovascular disease (prior stroke or transient ischemic attack) or peripheral vascular disease.
2. Patients with missing information on follow up events included in primary outcome.
3. Patients with missing data on BMI
4. Patients with missing data on any of the variables necessary to calculate their risk estimate using the PCE during their enrollment visit.
5. Patients using a statin or other lipid altering medication (niacin, fibrates, bile acid binding resins) at the time of risk factor assessment
6. Patients with LDL cholesterol levels <70 or ≥190 mg/dl at initial assessment

**Analysis plan:**
- First, we will create a combined cohort of all individuals with a BMI of ≥ 30 kg/m² and calculate their 10 year estimated risk of ASCVD based on:
  1. The “out of the box” PCE calculator published online.
  2. PCE risk estimates recalibrated to the current dataset

- Second, we will categorize all individuals into risk groups based on their estimated 10-year risk of ASCVD: low (<5%), intermediate (5 to 7.5%), and high (>7.5%).

- Third, baseline characteristics will be compared across predicted ASCVD risk categories by using ANOVA for continuous variables and chi-square tests for categorical variables. Event rates per-1000 person years for the risk categories will be calculated.

- Fourth, we will examine the observed and predicted 10 year incidence of ASCVD among patients overall, and separately stratified according to sex (M/F), race (White vs. Non-White), diabetes status (Y/N), and obesity classification (class I/II-III). A bar graph of cumulative 10 year risk vs. Risk group (both observed and predicted) will be made. Calibration will be assessed using plots of observed vs. predicted risk within deciles of predicted risk using a previously suggested approach. Here, PCE calibration is assessed by calculating the ratio of predicted/observed event rates. A ratio of 1 would indicate
perfect calibration, and ratios above 1 and below 1 would indicate over- and under-
estimation of risk, respectively. Calibration will be tested using the Hosmer-Lemeshow
test. Discrimination of the model will be assessed using Harrell’s C-statistic.

- Fifth, we will explore inclusion of various measures of obesity (BMI-continuous, obesity
class - categorical, waist circumference-continuous, waist circumference in metabolic
syndrome cutoff - categorical) in a stepwise manner and compare model performance
measured using changes in C-statistic compared to the baseline model. We will also
cross-tabulate the PCE with and without each additional obesity-specific marker and
calculate the net reclassification improvement (NRI). Bootstrapping will be used to
calculate 95% Cis. The NRI for events and non-events will be calculated separately as
previously recommended. Analyses will be performed in the overall cohort and the
previously discussed stratification groups.

- Finally, we will compare the discrimination and calibration of the PCEs in the obese
population with the non-obese population in the dataset using both the standard PCE (out
of the box and recalibrated) and the obesity-specific PCE. Comparison will also be done
between groups with and without abdominal obesity, defined as >102 cm (40 in) in men
and >88 cm (35 in) in women. Analyses will be performed in the overall cohort and the
previously discussed stratification groups.

Table 1: Characteristics of included patients by data source

<table>
<thead>
<tr>
<th>Patient characteristics</th>
<th>Overall</th>
<th>MESA</th>
<th>ARIC</th>
<th>CARDIA</th>
<th>DHS</th>
<th>CHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient demographics (age, sex and race)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient comorbid conditions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean estimated risk (+/- SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event rates per 1000 person-years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean follow up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2: Characteristics of included patients by estimated risk cut-offs

<table>
<thead>
<tr>
<th>Patient characteristics</th>
<th>Low (&lt;5%)</th>
<th>Intermediate (5 to 7.5%)</th>
<th>High (&gt;7.5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient demographics (age, sex and race)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient comorbid conditions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean estimated risk (+/- SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event rates per 1000 person-years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean follow up</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Characteristics of included patients by obesity class

<table>
<thead>
<tr>
<th>Patient characteristics</th>
<th>Obesity Class I BMI 30-34.9</th>
<th>Obesity Class II BMI 35-39.9</th>
<th>Obesity class III BMI ≥ 40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient demographics (age, sex and race)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient comorbid conditions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean estimated risk (+/- SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event rates per 1000 person-years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean follow up</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Other tables:
1. NRI tables (cross-tabulation) of events and non-events with addition of obesity-specific risk markers
Figures:
1. Rates of outcome per 1000-person years by
   a. Cohort
   b. Estimated risk
   c. Obesity class
   d. Tertiles of waist circumference
2. Bar graph of cumulative 10-year risk (%) vs. Predicted risk groups (predicted and observed)
3. Calibration plots of observed vs. predicted with line of unity and deciles of predicted risk (dots)
4. ROC curves of baseline model and then with individual addition of obesity-specific risk markers

References:


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 lists under the Study Members Area of the web site at: [http://www.cscc.unc.edu/ARIC/search.php](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)? None

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? _____ Yes   _X_ No

11.b. If yes, is the proposal

   ___ A. primarily the result of an ancillary study (list number* _________)
   ___ 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.cscu.unc.edu/aric/forms/](http://www.cscu.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.csecc.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.