ARIC Manuscript Proposal #3456

1. a. Full Title: Improving risk prediction of atherosclerotic cardiovascular disease for people with chronic obstructive pulmonary disease: NHLBI pooled cohorts study

b. Abbreviated Title (Length 26 characters): ASCVD risk prediction in COPD

2. Writing Group:

Writing group members:
- Suurya Krishnan, MD
- Jean Bourbeau, MD MSc
- Benjamin Smith, MD MSc
- Elizabeth Oelsner MD PhD
- Pallavi P Balte, PhD
- Raquel Farias, MD PhD
- Andrea Benedetti, PhD
- Stephanie J London, MD DrPH
- David Couper, PhD
- Laura R Loehr, MD PhD
- Ravi Kalhan, MD MS
- Sachin Yende, MD MS
- Anne B Newman, MD MPH
- Jason Sanders, MD PhD
- Wendy White, PhD
- Ana Navas-Acien, MD MPH PhD

Affiliations:
- McGill University
- McGill University
- Columbia University
- Columbia University
- Columbia University
- McGill University
- McGill University
- McGill University
- NIEHS
- University of North Carolina
- University of North Carolina
- Northwestern Medicine
- University of Pittsburgh
- University of Pittsburgh
- Brigham and Women’s Hospital
- Tougaloo College
- Columbia University

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: Suurya Krishnan
Address: 3D.29 5252 Boulevard de Maisonneuve W
Montréal, QC H4A 3S9
Phone: (306)741-3066 E-mail: suurya.krishnan@mail.mcgill.ca

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: Stephanie J London
Address: A306 Rall Building 111 T W Alexander Dr
Research Triangle Park, NC 27709
Phone: 984-287-3688 E-mail: london2@niehs.nih.gov
3. **Timeline:** Proposal approval from all 9 cohorts in the NHLBI pooled cohorts study is expected by October 2019 and all analyses and preparation of the manuscript will be complete by January 2020. After approval of manuscript by all cohorts we plan to publish our work by July 2020, having completed everything within one year.

4. **Rationale:**

Chronic obstructive pulmonary disease (COPD) was the third leading cause of death globally in 2016 [1] and coexists with many other medical conditions, cardiac diseases being among the most important [2]. Epidemiological studies and large clinical trials have helped us understand the importance of comorbidities in COPD. Within the UK General Practice Research Database, patients with COPD were at increased risk for ischemic heart disease including myocardial infarction (MI) and angina compared to those without COPD [3]. In a large multisite randomized clinical trial, the Study to Understand Mortality and Morbidity (SUMMIT), which enrolled over 16,000 patients with moderate COPD and increased cardiovascular disease (CVD) risk, CVD was found to be the leading cause of mortality, accounting for 43% of deaths compared to 13% from pulmonary causes [4].

Undeniably, COPD and CVD share common risk factors such as sex, advanced age, smoking, and a sedentary lifestyle. However, large population-based studies have demonstrated that COPD increases the risk for developing atherosclerotic cardiovascular disease (ASCVD) even after accounting for sex, age, smoking and prior heart disease [5]. In the Lung Health Study, a decrease of 10% in forced expiratory volume in one second (FEV1) among COPD patients, indicating worsening airflow obstruction, was associated with a 30% increase in CVD risk [6]. FEV1 has also been associated with sudden death and stroke [6], and a more rapid rate of decline in FEV1 further increases CVD risk [7]. COPD patients also have elevated levels of inflammation and related biomarkers such as fibrinogen that is associated with an increased risk of ASCVD [8] and COPD exacerbations [9], which in turn further increases risk of an ASCVD event [10].

Being able to identify COPD patients at greatest risk to develop ASCVD is very important in clinical practice, as these patients will likely benefit the most from proven risk reduction strategies. The current guidelines report from the Global Initiative for Obstructive Lung Disease (GOLD) state, “ischaemic heart disease should be considered in all COPD patients depending on their risk factor profile. Cardiovascular risk may be assessed by the global risk calculator, which can be found on the US National Heart Lung Institute (NHBLI) website” [2]. However, the NHBLI ASCVD risk calculator or Pooled Cohorts Equation (PCE), does not include any markers specific to COPD.

We hypothesize that the standard PCE underestimates ASCVD risk in COPD patients. We furthermore hypothesize that incorporation of information on COPD status and severity will improve ASCVD risk prediction in COPD patients relative to the PCE risk prediction alone. We propose to test these hypotheses in a large, US general population-based sample.

**Hypotheses and specific objectives**

1. To estimate the extent to which the standard PCE for 10-year ASCVD risk underestimates risk in COPD patients compared to adults without COPD.
2. To determine if prediction of 10-year ASCVD risk is improved relative to the PCE estimate alone by including information on COPD status and COPD severity.
3. In exploratory analyses, to determine if prediction of 10-year ASCVD risk is further improved by including information on blood fibrinogen.
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).

Study design and population
This study will use data from the NHLBI Pooled Cohorts Study, which harmonized and pooled data from 9 US population-based adult cohorts [11, 12]:

1) Atherosclerosis in Communities (ARIC)
2) Coronary Artery Risk Development in young Adults (CARDIA)
3) Cardiovascular Health Study (CHS)
4) Framingham Heart Study – Offspring cohort (FHS-O)
5) Health Aging, and Body Composition (HABC)
6) Multi-Ethnic Study of Atherosclerosis (MESA)
7) Jackson Heart Study (JHS)
8) Hispanic Community Health Study/Study of Latinos (HCHS/SOL)
9) Strong Heart Study (SHS)

For this analysis, the following inclusion criteria will be applied at the baseline exams:

- Primary analysis: identical to the 2013 PCE derivation [11]
  - Age 40-79 at time of observation
  - No history of myocardial infarction (MI), stroke, heart failure, percutaneous coronary intervention, or atrial fibrillation, at time of observation
  - White or black race (note: majority/all of HCHS/SOL and SHS will be excluded)
  - No baseline statin use

- Secondary analyses:
  - Retaining all race/ethnicities (including all cohorts)
  - Excluding participants who started statin therapy over follow-up

Variables and measures

COPD status and severity (ARIC study visit 1)
COPD status will be defined on baseline spirometry as forced expiratory volume in one second/forced vital capacity (FEV1/FVC) < 0.70 measured by pre-bronchodilator (pre-BD) spirometry [13].
COPD severity will be defined by FEV1-percent predicted, continuous/categorized according to GOLD guidelines [2] using pre-BD spirometry measures. FEV1 and FEV1/FVC will be added as measures of COPD severity in secondary analyses.

2013 PCE 10-year ASCVD risk score (ARIC study visit 1)
The ASCVD risk score [11] will be calculated using self-reported age, sex, race, current smoking status (Y/N), treated or untreated systolic blood pressure (SBP), total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), and diabetes status (Y/N) (self-report of a physician diagnosis or diabetic medication use or HBA1C >7% or a fasting glucose >125mg/dl).

Other variables (ARIC study visit 1)
Height, weight, race, respiratory symptoms, physician diagnosis of COPD or asthma, inhaler medication use, physical activity.
Outcome: ASCVD Events (From ARIC visit 1 for the next 10 years: 1986 - 1999)
Hard ASCVD event will be defined as first occurrence of nonfatal MI, death from coronary heart
disease, fatal or nonfatal stroke. All cohorts were designed to prospectively ascertain and
adjudicate ASVCD events [11].

Exploratory analyses (ARIC study visit 1)
Fibrinogen levels were measured in blood in six cohorts (ARIC, CARDIA, CHS, FHS-O,
MESA, SHS) and will be considered as a potential additional COPD-specific ASCVD risk factor
in sub-group analyses of the six cohorts with this data.

Sample Size Calculation
The NHLBI Pooled Cohorts Study will provide adequate power for all planned analyses. A rule
of thumb in regression models is 10 outcomes per predictor variable and as the incident ASCVD
rate in our population is approximately 16%, in order to analyze up to 6 variables together we
need a sample size of 375 participants.

Statistical Analysis
All our risk models will be fit with random intercepts regression models to account for using
multiple cohorts in our study. Valid marginal risk predictions will be achieved by integrating
over the distribution of random effects in our final models [14]. Cox proportional regression
hazards will be constructed for all predictor variables together for time from baseline to incident
ASCVD event, censored at 10 years. If the proportional hazards assumption is violated, a logistic
regression model adjusted for censoring will be used as it produces more accurate coefficient
estimates than a Cox proportional hazards model [15, 16]. We will check for interactions
between all predictor variables and add a regression hazard for each significant interaction item
between two variables. To avoid overfitting, we will also use a penalized multiple regression
model with LASSO regularization as it is implemented in the R-package glmnet (the amount of
regularization is adjusted by cross validation). The bootstrapping technique will be applied for
internal validation of all models.

Performance of all risk models to predict ASCVD events will be evaluated by discrimination
using the C-statistic and calibration using the Brier Score/Greenwood–Nam–D'Agostino (GND)
test [15, 17]. All results will be stratified by age, sex, race, smoking status and by different
outcomes in sensitivity analyses. For results in the group with COPD status, we also perform
sensitivity analyses for COPD severity by GOLD status and stratified into groups by respiratory
symptoms and inhaler use, physical activity level and self-reported physician diagnosis of COPD
or asthma.

Objective 1. Baseline (time zero) will be the earliest study visit with all available measures
necessary to calculate the PCE risk score and spirometry performed. Ten-year ASCVD risk will
be calculated using the PCE risk score. The sample will be grouped by COPD status
(FEV1/FVC<0.07) (Y/N) and performance of the PCE risk model will be evaluated in each
group and compared.
Objective 2. Using the same baseline, we will estimate a regression risk model for the PCE risk
score, COPD status and COPD severity together. Model performance will be evaluated and
compared to the PCE risk score for improvement in discrimination and calibration.
Objective 3. Exploratory analyses. Baseline (time zero) will be the earliest study visit with all available measures necessary to calculate the PCE risk score and fibrinogen from six cohorts. We will estimate a regression risk model for the **PCE risk score**, **COPD status**, **COPD severity** and **fibrinogen** all together. Model performance will be evaluated and compared to our previous risk model and to the PCE risk score for improvement in discrimination and calibration, in this sub-group.

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/mantrack/maintain/search/dtSearch.html

   **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)?

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 https://www2.cscc.unc.edu/aric/approved-ancillary-studies