ARIC Manuscript Proposal # 3157

PC Reviewed: 5/8/2018 Status: _____ Priority: 2
SC Reviewed: _________ Status: _____ Priority: ____

1. a. Full Title: Retinal Vasculature Representation Learning for Cardiovascular Disease

   b. Abbreviated Title (Length 26 characters): Deep learning and CVS Disease.

2. Writing Group:
   Writing group members: Roomasa Channa, Alison G. Abraham, Sidra Zafar, Luca Giancardo.

I, first author, confirm that all the coauthors have given their approval for this manuscript proposal.

First author: Roomasa Channa
Address: 600 N Wolfe Street, Wilmer Eye Institute, Johns Hopkins Hospital, Baltimore, MD 21287 USA

   Phone: +14103504567       Fax:
   E-mail: Roomasa.channa@gmail.com

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: Alison G. Abraham
   Address: 600 North Wolfe Street, Woods 173, Baltimore, MD, 21287 USA

   Phone:            Fax:
   E-mail: Alison.abraham@jhu.edu

3. Timeline: Manuscript draft planned within 12 months from proposal approval.

4. Rationale:
Cardiovascular (CVS) disease remains the leading cause of mortality among adults aged 65 years or older in the United States. Elevated blood pressure, high cholesterol levels, low high-density lipoprotein (HDL) levels and diabetes have traditionally been used by clinicians for risk stratification of future CVS events. However, these are indirect markers of vascular health with no information regarding microvasculature status.\textsuperscript{1,2} Additionally, clinical algorithms such as the Framingham Risk Score may underestimate risk of future CVS events in certain sub-population groups, particularly women.\textsuperscript{3,4} Invasive tests such as coronary angiography, intracoronary Doppler wire, response to vasoactive agents or vasoconstrictor stimuli and measurement of myocardial blood flow using positron-emission tomography (PET) or cardiac magnetic resonance imaging (MRI) are used clinically as gold standard to assess coronary microvasculature.\textsuperscript{5,6} These are not readily available for population level screening purposes, especially in point-of-care applications.

There is increasing evidence that retinal vessels are believed to mirror organ (including heart) microvasculature in terms of anatomy and physiology.\textsuperscript{7} Visualization of retinal circulations, via fundus photos or recent optical coherence tomography angiography (OCT-A), offers a unique opportunity to non-invasively examine the systemic microcirculation. With widespread use of portable fundus cameras, and smartphone adaptors, fundus photos have become increasingly easy to acquire even at community level with minimal medical training. Estimation of candidate image based biomarkers such as vessel width, vessel tortuosity, artery to vein ratio or fractal analysis is possible through different software tools.\textsuperscript{8-11} These pre-determined vasculature features correlated with CVS condition.\textsuperscript{12} However, retinal vasculature is very complex and offers much more quantifiable data than what can be measured with pre-determined features. To address this gap, general purpose convolutional neural networks have been used for predicting CVS conditions without explicitly using vasculature features,\textsuperscript{13} but they are limited by low interpretability and the need for extensive training data.

We propose to use our new approach\textsuperscript{14} that leverages the power of convolutional neural networks to automatically learn image base features, specific to vasculature, thereby allowing for better interpretability. Easily deployable quantitative tools to predict future cardiovascular events are an unmet urgent medical need. Hence, the long-term objective is to develop and validate new quantitative “computational biomarkers” for screening of CVS and other microvasculature related conditions such as dementia, with tools that are easily deployable in existing infrastructure both at the patient’s homes and community clinics.

5. Main Hypothesis/Study Questions:

To apply machine learning:

1. To develop a quantitative model for cardiovascular disease risk screening using retina vasculature patterns automatically learned from images and, without using pre-defined vasculature features.

2. To develop retina image-based methods to predict existing and validated metrics for cardiovascular risk.
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:** A convolutional neural network-based learning model will be trained and tested for cardiovascular disease risk estimation using retinal vasculature features, derived from fundus photos and OCT-A images.

**Data sets:** The neural network will be trained using fundus photographs from both ARIC and EyeDOC and using OCT-A images from EyeDOC. For clinical validation, non-overlapping images derived from these datasets will be used.

**Inclusion criteria:** ARIC participants from visit 3

**Exclusion criteria:** ARIC participants with no retinal photographs or ungradable retinal photographs.

**Additional variables of interest:**
1. Age
2. Gender
3. Ethnicity/race
4. Education level
5. BMI
6. Smoking status
7. Diabetes
8. Systolic blood pressure
9. Diastolic blood pressure
10. Resting Heart Rate
11. Heart Rate Variability
12. Glycated hemoglobin (HbA1c) levels
13. Lipid panel
   a. Triglycerides
   b. Low-density lipoprotein
   c. High-density lipoprotein
   d. Total cholesterol
14. Treatment for Hypertension
15. Cognitive Status (normal, mild cognitive impairment or dementia with CVS diagnosis and mild cognitive impairment or dementia without CVS diagnosis)
16. Fundus retinal findings (microaneurysms, retinal hemorrhage, soft exudates, AV nicking, focal arteriolar narrowing, branch retinal arteriole-to-venular ratio (AVR), central retinal arteriolar equivalent, central retinal venular equivalent)

**Outcome variable:** Cardiovascular events

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

  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)?

ARIC MS#1110: ‘’Risk prediction of coronary heart disease based on retinal vessel caliber: The ARIC Study

ARIC MS#1424: ‘’Systematic review of the association of retinal microvascular signs and cardiovascular disease’’

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)* (Eye-DOC 2014.38)

*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

References:
