



WG(s): N/A

New Delhi, 13-15 November 2019

DOCUMENT**Source:** TG-Cardio**Title:** Call for AI for Health Topic Group Participation: AI for Health Topic Group: Cardiovascular Disease (CVD) Risk Prediction**Purpose:** Engagement

Contact: Benjamin R.H. Muthambi, DrPH, MPH Email: brm5@caa.columbia.edu
IEPH, Inc.
6224 Spring Knoll Drive
Harrisburg, PA 17111, USA

Abstract: Calling on members of the health and artificial intelligence communities with a vested interest in application of AI in cardiovascular disease (CVD) risk prediction! Become engaged in the topic group dedicated to establishing a standardized benchmarking platform for application of AI in CVD risk prediction within the International Telecommunication Union (ITU)/World Health Organization (WHO) Focus Group on “Artificial Intelligence for Health” (FG-AI4H). This document is the same as seen in meeting E, reproduced for meeting G for easier reference. It does not include content relevant to the sub-topic group on artificial intelligence in coronary computed tomography angiography.

Call for Topic Group Participation:**AI for Health Topic Group: Cardiovascular Disease (CVD) Risk Prediction**

The International Telecommunication Union (ITU)/World Health Organization (WHO) Focus Group on “Artificial Intelligence for Health” (FG-AI4H; <https://www.itu.int/go/fgai4h>) seeks engagement from members of the health and artificial intelligence (AI) communities (including clinicians & public health practitioners/researchers, technologists, innovation entrepreneurs, potential benchmarking data providers, machine learning experts, software developers, researchers, regulators, policy-makers, companies, institutions including academic and research institutions, and field experts) with a vested interest in shaping the benchmarking process of applications of AI in cardiovascular disease (CVD) risk prediction.

1. About FG-AI4H

Over the past decade, considerable resources have been allocated to exploring the use of AI for health, which has revealed an immense potential. Yet, due to the complexity of AI models, it is difficult to understand their strengths, weaknesses, and limitations. If the technology is poorly designed or the underlying training data are biased or incomplete, errors or problematic results can occur. AI technology can only be used with complete confidence if it has been quality controlled through a rigorous evaluation in a standardized way. Towards developing this standard assessment framework of AI for health, the ITU has established FG-AI4H in partnership with the WHO.

Thus far, FG-AI4H has established thirteen topic groups. These are concerned with: AI and cardiovascular disease (CVD) risk prediction, child growth monitoring, dermatology, falls among the elderly, histopathology, neuro-cognitive diseases, ophthalmology (retinal imaging diagnostics), psychiatry, radiotherapy, snakebite and snake identification, symptom checkers, tuberculosis, and volumetric chest computed tomography.

Each topic group agrees upon representative benchmarking tasks in a pragmatic, best-practice approach, which can later be scaled and expanded to similar tasks. Every benchmarking task should address a health problem of relevance (e.g. impacting a large and diverse part of the global population or challenging to treat) and for which AI technology would provide a tangible improvement relative to the current practice (e.g. better care, results, and/or cost/time effectiveness).

For a rigorous and sound evaluation, undisclosed test data sets must be available (or have to be collected) for each task. All data must be of high quality and compliant with ethical and legal standards. In addition, the data must originate from a variety of sources so that it can be determined whether an AI algorithm can generalize across different conditions, locations, or settings (e.g. across different people, hospitals, and/or measurement devices). The format/properties of the data serving as input to the AI and of the output expected from the AI, as well as the benchmarking metrics are agreed upon and specified by the topic group.

Finally, the AI-to-be-evaluated will be benchmarked with the undisclosed test data on FG-AI4H computing infrastructure. Here, the AI will process single samples of the undisclosed test data set and predict output variables, which will be compared with the "ground truth." The results of the benchmarking will be provided to the AI developers and will appear on a (potentially anonymized) leaderboard.

2. About the AI for Health Topic Group: Cardiovascular Disease (CVD) Risk Prediction

A topic group is a community of stakeholders from the medical and AI communities with a shared interest in a topic. The topic group/community of stakeholders will be differentiated into 3 subgroups, and calls for participants will be issued for the 2nd & 3rd subgroups indicated below:

- Topic Driver Team: Dr Benjamin Muthambi (IEPH), the corresponding/ topic driver and proposer of CVD Risk Prediction topic in collaboration with his WatifHealth partners: Dr Nao Sipula, and Mr Jason Paul.
- Topic Peer-Reviewer Forum - Call for Participants Opening June 17, 2019 (A link to information on steps for joining topic group and scheduled activities will be provided under updates on the ITU Topic Group webpage: <https://www.itu.int/en/ITU-T/focusgroups/ai4h/Pages/cardio.aspx>)
- Sub-topics & Technical Subgroups/Forums (Project Phase-Specific) - Calls for Participants at this more specific level are still to be announced: also through the ITU Topic Group web Page at: <https://www.itu.int/en/ITU-T/focusgroups/ai4h/Pages/cardio.aspx>

i) Topic Driver Team: Responsibilities of the topic driver may evolve over time but an interim list of the responsibilities includes:

- Creating the initial draft version(s) of the topic description document.
- Reviewing the input documents for the topic and moderating the integration in a dedicated session at each Focus Group meeting.
- Organizing regular phone calls to coordinate work on the topic description document between meetings.

ii) Topic Group Peer-Review & Advisory Forum: The preliminary objectives of the Topic Stakeholder Community are manifold:

- to provide a forum for open communication among various stakeholders,
- to agree upon the benchmarking tasks of this topic and scoring metrics,
- to facilitate the collection of high quality labelled test data from different sources,
- to clarify the input and output format of the test data,
- to define and set-up the technical benchmarking infrastructure, and
- to coordinate the benchmarking process in collaboration with the Focus Group management and working groups.

The primary output of this topic group forum is one document that describes all aspects of how to perform the benchmarking for this topic. (The document will be developed in a cooperative way by suggesting changes as input documents for the next FG-AI4H meeting that will then be discussed and integrated into an official output document of this meeting. The process will continue over several meetings until the topic description document is ready for performing the first benchmarking.)

iii) Sub-topics & Technical Subgroups/Forums (Project Phase-Specific): The preliminary sub-topics & technical subgroups/forums subject to further calls for participants to be announced include:

- Conceptualization of Sub-Topics on CVD-related Predictive Analytics in Clinical & Public Health, & Applications thereof
- External Peer Review for Ethics (Pre-IRB) & Quality Assurance of Project Proceedings
- Methods: Epidemiology/Evaluation Study Design, Metrics & Statistical Methods
- Computer & Data Science: Machine Learning Algorithm Developer, Server Admins & Distributed Systems Engineering
- R Programming: Data Management, Predictive Analytics, Statistical Analyses, Shiny Web App
- Project Planning & Reproducible Reporting Tools & Writing

3. Introductory Subject Information: AI for Health Topic Group: Cardiovascular Disease (CVD) Risk Prediction.

This section (Section 3) introduces the subject matter of the topic including objectives and intended benchmarking task, relevance and data availability, followed by the next section (Section 5) on how to get involved.

3.1. Project Objectives/Problem to be addressed: Diabetics have higher CVD risk, hence improved CVD risk prediction is critical for better diabetes management and reducing mortality. The proposed project aims to:

- a) To assess: i) CVD risk prediction accuracy of various machine learning (ML) methods benchmarked against CVD risk based on actually observed occurrence of first CVD event (truth) documented in diverse cohorts/populations data, and ii) replicability/reproducibility of ML prediction of CVD risk using 'external data' from diverse populations meeting prescribed criteria but 'not previously accessed' (undisclosed) data to the ML algorithms under evaluation;
- b) Compare CVD risk prediction accuracy of ML algorithms referenced above in section 3.1. (a) to: i) several routine clinical-use CVD risk scoring tools/calculators, and ii) traditional multivariate statistical methods (in collaboration with other co-investigators who recently undertook similar risk prediction accuracy studies);

c) Determine which methods, if any, consistently show better predictive accuracy across diverse populations. Using the above-referenced methods, benchmarking, anticipated findings and peer-review thereof, the project expects to establish an evidence-based standards-setting blueprint.

3.2. Relevance/Background, significance and rationale: Cardiovascular disease (CVD) is the global leading cause of morbidity and mortality (WHO, 2014). CVD accounts for > 2/3 of mortality among type 2 diabetes patients (ADA, 2019). Widely used clinical CVD risk scoring tools/calculators incorporate several factors with well-established etiological associations with CVD such as age, sex, BMI, systolic blood pressure, smoking, A1C, lipid levels, age at diagnosis &/or onset of diabetes, diabetes duration, and antihypertensive and lipid-reducing drugs, but do not necessarily include a comparable set of predictors. In addition, these methods often fail to identify many people who would benefit from preventive treatment, while others receive unnecessary interventions. For example, approx. ~50% of myocardial infarctions (MIs) and strokes occur among persons predicted to be at risk of CVD by routinely-used risk calculators (Ridker et al, 2008). Highlighting the need for standardization, prior systematic CVD risk prediction accuracy studies often used: disparate study populations; either traditional multivariate statistical methods or disparate ML algorithms; incomparable sets of predictors often not considering the broader range of potential predictor data made possible by mining electronic health records, big data aggregation, or accounting for complex interactions that ML can handle more easily; and also used different measures of predictive accuracy. This study hypothesizes that ML algorithms can improve CVD predictive accuracy over CVD risk scoring tools/calculators used in the standard of practice across diverse populations. If demonstrated, ML-assisted DSS should be considered as the underlying approach for standard of practice in CVD risk prediction.

3.3. Data availability:

Public Data: Anonymized data already acquired for preliminary work for this project are publicly-accessible academic training-use data extracted from the NIH-funded Framingham Heart Study which were in turn sourced from the BioLINCC data repository under NIH data sharing terms.

Undisclosed Data: 'External data' meeting prescribed criteria but 'not previously accessed' (undisclosed data) to train the ML algorithms under evaluation, will be sourced for replication studies from various repositories identified as suitable potential data sources including multiple researcher-use data obtainable under well-established NIH-funded research data sharing terms applicable to the NHLBI/National Heart, Lung, and Blood Institute's Open BioLINCC Biologic Specimen and Data Repository, i.e. NIH/NHLBI BioLINCC; UK NHS/CPRD data repository; and a diverse range of other data sources with suitable data identified in the literature, and other data sources still to be identified through a planned call for data-contributing project participants (per above-referenced AI for cardiovascular disease risk prediction topic group's Project Phase-Specific Technical Contributor Subgroups/Forums. The identified potential sources of 'not previously accessed/undisclosed data' contain clinical and other patient data used in routine clinical care by CVD risk scoring tools/calculators, and in research using traditional multivariate statistical methods or ML algorithms.

4. Details about the activities of the topic group: More information is in documents [C-017-R1](#) & [D-031](#), which can be accessed with a free ITU account (cf. "Get involved").

As indicated earlier, the AI for Health topic group on cardiovascular disease (CVD) risk prediction is lead by Dr. Benjamin Muthambi who encourages interested parties to join as topic group members. As indicated under Section 2, the topic group is keenly interested in engaging expertise from the health and AI communities and from additional data-contributing partners.

5. Getting involved:

To join this topic group, please send an e-mail to the focus group secretariat (tsbfgai4h@itu.int) and the topic driver (brm5@caa.columbia.edu). Please use a descriptive e-mail subject (e.g.

"Participation topic group AI for cardiovascular disease risk prediction"), briefly introduce yourself and your organization, concisely describe your relevant experience and expertise, and explain your interest in the topic group.

Participation in FG-AI4H is free of charge and open to all. To attend the workshops and meetings, please visit the Focus Group website (<https://www.itu.int/go/fgai4h>), where you can also find the whitepaper, get access to the documentation, and sign up to the mailing list.
