|  |  |  |
| --- | --- | --- |
| ITU Logo | INTERNATIONAL TELECOMMUNICATION UNION**TELECOMMUNICATIONSTANDARDIZATION SECTOR**STUDY PERIOD 2017-2020 | FG-AI4H-H-035 |
| **ITU-T Focus Group on AI for Health** |
| **Original: English** |
| **WG(s):** | Plenary | Brasilia, 22-24 January 2020 |
| **DOCUMENT** |
| **Source:** | Artificial Intelligence in Medical Epidemiology (AIME) and Sarawak State Health Department, Malaysia |
| **Title:** | TG-Malaria: Proposal for sub-topic: Malaria surveillance and predictive modelling |
| **Purpose:** | Discussion |
| **Contact:** | Helmi ZakariahArtificial Intelligence in Medical Epidemiology (AIME), Malaysia | Email: drhelmi@aime.life |
| **Contact:** | Yong Kui ChoonSarawak State Health DepartmentMalaysia |  |

|  |  |
| --- | --- |
| **Abstract:** | This document contains an outline of a sub-group topic description document on AI-based surveillance and predictive modelling of Malaria (TG-Malaria). |

**Table of Contents**

[1 Introduction 3](#_Toc24214579)

[1.1 Document Structure 3](#_Toc24214580)

[1.2 Topic Description 3](#_Toc24214581)

[1.3 Impact of the benchmarking AI Solution 4](#_Toc24214582)

[1.4 Ethical Considerations 4](#_Toc24214583)

[1.5 Existing AI Solutions 4](#_Toc24214584)

[1.6 Existing work on benchmarking 4](#_Toc24214585)

[2 AI4H Topic Group: Current topic group and its mandate 4](#_Toc24214586)

[2.1 Topic group structure 5](#_Toc24214587)

[2.1.1 Topic group participation 5](#_Toc24214588)

[2.1.2 Tools/process of TG cooperation 5](#_Toc24214589)

[2.1.3 TG interaction with WG and FG 5](#_Toc24214590)

[2.1.4 Next meetings 5](#_Toc24214591)

[2.1.5 Next steps 5](#_Toc24214592)

[3 Method 6](#_Toc24214593)

[3.1 AI Input Data Structure 6](#_Toc24214594)

[3.1.1 Available Data 6](#_Toc24214595)

[3.2 AI Output Data Structure 6](#_Toc24214596)

[3.3 Test Data Labels 6](#_Toc24214597)

[3.4 Score and Metrics 6](#_Toc24214598)

[3.5 Undisclosed Test Data Set Collection 6](#_Toc24214599)

[3.6 Benchmarking Methodology and Architecture 7](#_Toc24214600)

[3.7 Reporting Methodology 7](#_Toc24214601)

[4 Results 7](#_Toc24214602)

[5 Discussion 7](#_Toc24214603)

[6 Declaration of Conflict of Interest 7](#_Toc24214604)

[References 7](#_Toc24214605)

# Introduction

Malaria is one of the most endemic diseases in resource-poor settings such as Sub-Saharan Africa and South East Asia [1]. The continued public health significance of Malaria necessitates targeted action. In addressing the unique challenges of these settings, there is a need for reliable surveillance systems and disease control via proactive measures including predictive modelling. This explains the role of malaria outbreak detection and outbreak prediction using Artificial Intelligence (AI) [2]. The aim is to harness novel technologies including AI, predictive algorithms and real-time data analysis to automate the surveillance of Malaria and to empower public health efforts.

Recent AI and machine learning techniques have been successful in vector-borne disease surveillance and prediction; demonstrating a capacity to improve public health by optimizing resource allocation and providing valuable insights for decision-making.

The aim of this document is to present an approach for a multi-modal Malaria surveillance and prediction platform and to support standardised benchmarking of the aforementioned.

## Document Structure

This TDD is dedicated to support standardised benchmarking of AI-based surveillance and outbreak prediction of Malaria. The document will cover all core aspects relevant to topic including data collection and processing, selection of the AI model for prediction, algorithm performance and evaluation, as well as benchmarking methodologies.

## Topic Description

According to the World Health Organization (WHO) report of 2016, nearly half of the world population is at risk of malaria [3]. In 2015, 212 million cases were reported with Malaria accounting for over 480 000 deaths: 90% of which were from Africa, 7% from South-East Asia and 2% from Eastern Mediterranean region [4]. Despite a reduction in case load from 2010 to 2017, this progressed has stalled with no significant progress in reducing malaria incidence as per the World Malaria Report 2018 [5]. This highlights the need continued public health efforts in this area.

## Impact of the benchmarking AI Solution

The benchmark solution for AI-based surveillance and prediction of malaria should focus on developing tools that can enable the highly accurate identification of outbreaks; real-time analytics and outbreak visualisation; and predictive modelling in a fast, accurate, reliable, cost effective and resource-efficient manner.

The benchmark will be impactful if the selected solutions are robust and relevant in developing and highly Malaria-endemic countries. This is envisioned to improve public health and directly addresses Target 3.3 of Sustainable Development Goal 3 (SDG) which aims to end the Malaria endemic by 2030 [6].

## Ethical Considerations

**Ethical consideration of benchmarking including its data acquisition**:

* Acquisition of human ethical approvals from relevant medical authorities through working Institutional Research Boards (IRB),
* Data capture and processing must be done by qualified experts,
* Data anonymity must be respected by removing all references to personal identifiers,
* Integration of other datasets must also prevent personal identifiers.

**Ethical consideration on usage of AI**:

* A principled approach to scientific research to be followed,
* Field validation of the AI models must be aimed for.

## Existing AI Solutions

Existing AI solutions in Malaria outbreak surveillance and prediction are not forthcoming. However, there are strong parallels in similar mosquito-borne diseases such as Dengue. Artificial Intelligence in Medical Epidemiology (AIME) have employed support vector models and Multimodal Bayesian modelling to achieve near-instantaneous outbreak identification, highly sensitive and precise outbreak triangulation and predict future areas highly susceptible to outbreak. These predications are stratified by risk. By using mosquito biobehaviour, detailed ecological data and demographic data, these models have been deployed to much success in both medium- and high-disease burden areas.

Given the success of these models in Dengue and Zika, and the strong parallels in the above parameters, there is much optimism in replicating an AI-driven methodology for Malaria surveillance and prediction.

## Existing work on benchmarking

There has been discussion on the need for a benchmark in the area of Malaria outbreak surveillance and prediction for disease control. Currently, there is no benchmark. Comparable Malaria burden projections and their benchmarking are not reflective of an artificially intelligent platform.

# AI4H Topic Group: Current topic group and its mandate

The current topic group is specific and relevant to AI4H. The objectives are:

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

The primary output of a sub-topic group 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.)

## Topic group structure

TBC

Topic Groups summarize uses cases of a certain health topic or problem and similar AI benchmarking requirements. The Malaria Topic Group is currently focussed on Malaria detection and microbiological diagnosis. We propose a sub-Topic Group to explore Malaria surveillance, analytics and outbreak prediction.

### Sub-Topic group participation

Current participation in this Sub-Topic Group includes industry-experts Artificial Intelligence in Medical Epidemiology (AIME) and public health physicians of Sarawak State Health Department (Sawarak State, Malaysia).

The participation in both the focus and Topic Group is generally open and free of charge. To participate, one can follow the “call for participation” document outlining the process for joining the Focus Group and the Topic Group. For this topic, the corresponding call can be accessed through the official website (<https://itu.int/go/fgai4h>).

### Tools/process of STG cooperation

TBC

As will be regulated by the FG.

### TG interaction with WG and FG

TBC

As will be regulated by the FG.

### Next meetings

The Focus Groups meets about every two months at changing locations. The upcoming meetings are:

* H: Brasilia, Brazil; January 2020
* I: March 2020 (TBC)​
* J: Geneva, Switzerland, 4-8 May 2020 (TBC)

An up-to-date list can be found at the official [ITU FG AI4H website](https://www.itu.int/en/ITU-T/focusgroups/ai4h/Pages/default.aspx).

### Next steps

We aim to extend the topic of Malaria surveillance and outbreak prediction to all Malaria endemic Countries, while consolidating AI solutions and data from different countries. In order to do this, we will be expanding the pivotal work of the Malaria Topic Group to also encompass all areas of detection, surveillance and prediction.

 Next steps for this sub-topic group may includes:

1. Collection of annotated data from different sources. This includes data on Malaria cases and their incidence, the corresponding epidemiological investigations, surrounding geographical features, weather parameters, relevant ecology, vector and reservoir population.
2. Feature selection and implementation of AI models related to malaria detection **of cases and outbreaks.**
3. Suggestions on AI model performance metrics.
4. Contributing to the development of a viable and accepted benchmarking framework.
5. Support to the group on different aspects (data, methods, benchmarking, etc.) of this topic
6. **Discussion of the AI implementation to improve disease surveillance and prediction.**

# Method

TBC

## AI Input Data Structure

TBC

Deidentified Malaria case data is collected from relevant local health authorities. Important data points including onset of fever, onset of illness and locality of incidence are to be provided. Meteorological parameters include, among others – windspeed and direction, humidity, temperature, dewpoint, and others. Ecological parameters include the presence of plantations, fruit seasonality, body of water, and others – sourced from relevant authorities, or commercially-available GIS data.

### Available Data

Currently, we have access to about 4000 Malaria cases from the Sarawak state of Malaysia including granularity of date of fever onset, date of presentation to health system and geographical location of the health centre. This data is combined with industry-leading satellite imagery, location analytics and meteorological data. However, more data from other malaria endemic countries would be required to improve the accuracy of machine learning models as well as undisclosed test data for evaluation and validation of the tool.

## AI Output Data Structure

TBC

The output of the AI algorithm should include real time analytics including outbreak identification as well as future outbreak prediction with relevant ecological, meteorological and demographic variables displayed.

## Test Data Labels

TBC

Test data must include the same parameters as the training data with the annotated details of identified outbreak(s).

## Score and Metrics

TBC

To evaluate the AI tool’s performance, labelled dataset outbreaks will be compared against performance of AI. The algorithm evaluation mechanism should include metrics like ROC accuracy, precision, recall, specificity F1 scores, specificity, sensitivity and the choice will base on the algorithm used and purpose of the task.

## Undisclosed Test Data Set Collection

TBC

1. In order to assess algorithm robustness, sufficient undisclosed malaria case data would be collected. This is envisioned to come from different health facilities both public and private as collected by public health workers.
2. An agreeable number of test data for a benchmarking task will be specified.
3. Data processing standards should also be agreed upon.

## Benchmarking Methodology and Architecture

TBC

* mathematical validation using annotated test data
* field validation when compared to epidemiological outbreak detection
* protocol for performing the benchmarking

## Reporting Methodology

Disclosure of results through working group and peer reviewed publication.

# Results

TBC

# Discussion

TBC

* Discussion of the insights derived from executing the benchmarking including the areas
	+ external feedback on the whole sub-topic and its benchmarking
	+ data acquisition
	+ benchmarking process
	+ benchmarking results
	+ field implementation success stories
	+ comparability between deployment areas

# Declaration of Conflict of Interest

Dr Helmi Zakariah is the CEO of AIME (Malaysia).

References

[1] Gallup JL, Sachs JD. The Economic Burden of Malaria. In: Breman JG, Egan A, Keusch GT, editors. The Intolerable Burden of Malaria: A New Look at the Numbers: Supplement to Volume 64(1) of the American Journal of Tropical Medicine and Hygiene. Northbrook (IL): American Society of Tropical Medicine and Hygiene; 2001 Jan. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK2624/>

[2] Wong ZSY, et al., Artificial Intelligence for infectious disease Big Data Analytics, Infection, Disease & Health (2018), https://doi.org/10.1016/j.idh.2018.10.002

[3] WHO. World malaria report. Geneva, Switzerland, World Health Organization., 2016.

[4] WHO. World malaria report. Geneva, Switzerland, World Health Organization., 2015.

[5] WHO. World malaria report. Geneva, Switzerland, World Health Organization., 2017.

[6] WHO. Sustainable Development Goals. Geneva, Switzerland, World Health Organization.,

Available from: <https://www.who.int/sdg/targets/en/>

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_