



WG(s): Plen Helsinki, 20-22 September 2022

DOCUMENT

Source: TG-Outbreaks Topic Driver

Title: Att.2 – CfTGP (TG-Outbreaks) [same as Meeting I]

Purpose: Engagement

Contact: Auss Abbood, Stephane Ghazzi Email: ghozzis@rki.de,
Robert Koch Institute, Helmholtz stephane.ghozzi@helmholtz-hzi.de
HZI, Germany

Abstract: Calling on members of the medical and artificial intelligence communities with a vested interest in “Outbreak detection”. Become engaged in the group dedicated to establishing a standardized benchmarking platform for “AI for outbreak detection” within the International Telecommunication Union (ITU)/World Health Organization (WHO) Focus Group on “Artificial Intelligence for Health” (FG AI4H).

This version of the CfTGP is the same as seen in Meeting I (FGAI4H-I-018-A02), reproduced for easier reference as a Meeting N document.

Call for Topic Group Participation: “AI for Outbreak detection”

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

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 14 topic groups. The topic groups are: AI and cardiovascular disease 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 assessment, tuberculosis, volumetric chest computed tomography, and outbreak detection.

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 Topic group: Outbreak detection

A topic group is a community of stakeholders from the medical and AI communities with a shared interest in a topic. The objectives of the topic groups are manifold:

1. Provide a forum for open communication among various stakeholders,
2. Agree upon the benchmarking tasks of this topic and scoring metrics,
3. Facilitate the collection of high-quality labelled test data from different sources,
4. Clarify the input and output format of the test data,
5. Define and set-up the technical benchmarking infrastructure
6. Coordinate the benchmarking process in collaboration with the Focus Group management and working groups.

The primary output of a topic group is one document that describes all aspects of how to perform the benchmarking for this topic. The document will be developed cooperatively over several FGAI4H meetings. Iterative versions of the document will be submitted as input documents for each meeting and relevant changes will be discussed and integrated into an official output document. The process will continue over several meetings until the topic description document is ready for performing the first benchmarking.

This topic group is dedicated to outbreak detection algorithms.

Topic relevance

Infectious disease outbreaks pose a major risk to public health and are of global concern. Many established infectious diseases cause the death of millions of people every year and new infectious diseases are emerging. According to a WHO ranking, infectious diseases are ranked in the top 10 causes of death worldwide.

However, early detection of outbreaks can prompt fast interventions to limit spread of the disease or even prevent an outbreak altogether. Improved algorithms for outbreak detection can save lives, increase quality of life and will benefit the overall health of the world population.

Role of AI

AI algorithms can increase the timeliness and accuracy of outbreak detection, and further have the potential to improve an understanding of the warnings and the disease spread itself. AI algorithms are particularly powerful in incorporating multiple data sources with diverse properties. The integration of high-quality data sources, from e.g. mandatory reporting systems and laboratory tests, with data from real-time-surveillance systems is crucial to achieve earlier and more comprehensive detection of notifiable and non-notifiable pathogens. Different syndromic surveillance systems and valuable external data sources (google trends, health apps) can be incorporated. The gain of additional information on the underlying causes, by using explainable AI approaches, further enables for more specific actions to be taken for prevention.

Available Data

In Germany, the recording of infectious disease cases is regulated by the national infection protection law (the *IfSG*). Cases of notifiable pathogens are reported and collected via a mandatory reporting system at the national public health institute of Germany, the Robert-Koch-Institute. The available data set comprises 8 million reported infectious disease cases based on a collection since 2001. The data arrives pseudonymized at the RKI from about 400 local health agencies. The data holds expert labels relating cases to specific disease outbreaks. For each case, information is given on the pathogen, demographics (age, sex), location (NUTS-3 level, county) and on additional features such as hospitalization, fatality, and affiliation with care facilities and others. Some data is publicly available in an aggregated form, e.g. by counts for a specific disease, by week and county. However, details and single cases are not published. Most importantly, the expert outbreak labels have not been disclosed so far.

Benchmarking task

At present, outbreak detection algorithms are commonly parametrized and benchmarked on small sets of data or on simulations. These simulations are very simplistic outbreak representations, which capture only few aspects and often reduce benchmarking to the task of detecting elevated case count levels. By using real outbreak data from mandatory surveillance system, algorithms can be benchmarked on the actual task of detecting real world outbreak events.

The topic of outbreak detection is of national and international concern. The development of most detection algorithms is, however, naturally executed on national level. Thereby, each country relies on individual national disease surveillance systems.

To create a standardised benchmarking for output detection algorithms, the topic group aims to address all aspects, which are relevant and shared across countries.

Topic-specific questions to be addressed may include:

- Agreement on definitions of the term and event of outbreaks for detection algorithms
- How to deal with unlabelled outbreak cases for performance evaluation?
- Not all outbreaks are found or investigated by the local health agencies.
- Which score or metrics should be applied for benchmarking?
- Classic statistical metrics are accuracy, F1-score and prediction uncertainty. Further metrics to consider are timeliness, outbreak size or the number of false alarms.
- How do we optimally define a test set (undisclosed) to serve as a gold standard for benchmarking?
- How can additional data from molecular surveillance be used for quality assurance of given outbreak labels? Following the assumption that individual cases linked to an outbreak are reflected by closely related genome sequences.
- What is the benefit of AI approaches compared to established statistical models for outbreak detection?

All aspects will be discussed during regular focus group meetings and incorporated in the corresponding *topic description document* for this topic group.

Documents

More details about the activities of the topic group can be found in the following documents:

- [E-026](#): Proposal for new AI4H topic “Outbreak detection algorithms”
- [E-026-A01](#): Presentation on *Supervised Learning for Automated infectious disease outbreak detection*
- We wrote a review that collects published work that directly or indirectly describes how to benchmark outbreak detection algorithms named *How to benchmark disease outbreak detection algorithms: A review*

These can be accessed with a free ITU account (cf. “Get involved”).

Topic group members

The topic group on *outbreak detection algorithms* currently includes members from the

- Robert-Koch-Institute (National Public Health Institute Germany), involving members from different groups within the institute
 - Stéphane Ghozzi, Infectious disease epidemiology, Signale team (Robert Koch Institute, Unit 31)
 - Auss Abbood, Infectious disease epidemiology, Signale team (Robert Koch Institute, Unit 31)

3 Get involved

The creation of an international benchmarking framework for outbreak detection algorithms requires the involvement of diverse experts and institutes. The topic group would benefit from further expertise of the epidemiological, medical and AI communities.

We aim to extend the topic of outbreak detection beyond country borders, bringing together experts, algorithmic solutions and data outbreak collections from different countries. We are continuously working on diversity in our group. As of now, we especially are looking for members from the **Middle East, South America, and Australia**

Contributions to this topic group can be of different form:

- 1) By helping on the collection of labelled test data from different sources. Any data stream (case reporting systems, surveillance systems, etc.) directly linked to outbreak labels (expert/lab confirmed) is of high value. The ultimate goal is to cover outbreak data from different systems and countries.
- 2) By providing AI models and approaches for outbreak detection. Contributing to the development of a viable and accepted benchmarking framework
- 3) By supporting and advising the group on different aspects (data, methods, benchmarking, etc.) of this topic

To join this topic group, please send an e-mail to the focus group secretariat (tsbfgai4h@itu.int) and the topic driver (ghozzis@rki.de, abbooda@rki.de). Please use a descriptive e-mail subject (e.g. "Participation topic group AI for outbreak detection"), briefly introduce yourself and your

organization, concisely describe your relevant experience and expertise, and explain your interest in the topic group.

We highly encourage the involvement of experts from various disciplines, regarding relevant data sources, AI solutions or general expert knowledge on outbreak detection of infectious diseases.

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://itu.int/go/fgai4h>), where you can also find the whitepaper, get access to the documentation, and sign up to the mailing list.
