INTERNATIONAL TELECOMMUNICATION UNION

TELECOMMUNICATION STANDARDIZATION SECTOR

STUDY PERIOD 2022-2024

FG-AI4H-P-022-A01

ITU-T Focus Group on AI for Health

Original: English

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

DOCUMENT

Source:	TG-TB Topic Driver	
Title:	Att.1 – TDD update (TG-TB) [same as Meeting M]	
Purpose:	Discussion	

Contact:	Dr. Manjula Singh, ICMR, Under Department of Health Research, MOHFW, GOI, India	Tel: +91 9868245793 Fax: +91 26588896 Email: drmanjulasb@gmail.com
Contact:	DG, ICMR, Indian Council of Medical Research, (ICMR) India	Tel: Fax: +91 26588896 Email: dg@icmr.org.in
Contact:	Dr. Saurabh Gupta AIIMS, Delhi, India	Email: drsaurabhmd@gmail.com

Abstract: In 2017, 27% of the estimated 10 million global Tuberculosis (TB) cases developed in India. In the past few years, India has been actively implementing multiple strategies for reducing the burden of TB, including the web-based reporting system, the national TB prevalence survey, and the rollout of TB service delivery from all HIV clinics. An early adoption of Computer Assisted Diagnosis (CAD) systems based on artificial intelligence (AI) technologies for TB detection in India will synergize with the current endeavours to close the gap in TB control and will help global fight against TB and use of AI in the field of population health.

This TG works on the standardization of bench marking approach for development of AI tool for radiographic detection and screening of TB.

This version of the TDD is the same as seen in Meeting M (FGAI4H-M-022-A01), reproduced for easier reference as a Meeting N document.

Contents

1 Introduction

Machine learning (ML) techniques have been used world-wide for improving public health through development of better prognostic, diagnostic and predictive models. The application of ML technologies in healthcare has increased exponentially. In 2017, 27% of the estimated 10 million global TB cases were reported from India. There is a need to detect the missed cases by early adoption of CAD systems based on artificial intelligence (AI) technologies for TB detection. This will synergize with the global endeavours to close the gap in TB control. .

ICMR has access to large volume of high quality clinical data from various extramural and intramural programs. The data ranges from text based patient profiles to complex molecular sequences and structures and images. The proposal focuses on development of the AI tool for radiographic detection and screening of TB and will nucleate advance machine learning and analytics on data generated from various TB research activities of ICMR.

1.1 Document Structure and description

Description of topic: The topic deals with the development of an AI tool for radiographic detection of TB. In 2017, 27% of the estimated 10 million global TB cases developed in India.[1] In the past few years, India has been actively implementing multiple strategies for reducing the burden of TB, including the web-based reporting system, the national TB prevalence survey, and the rollout of TB service delivery from all HIV clinics. An early adoption of CAD systems based on artificial intelligence (AI) technologies for TB detection will synergize with the current endeavors to close the gap in TB control around the world and will help in the global fight against TB and use of AI in the field of population health.

Artificial intelligence technologies, including deep learning (DL) and natural language processing (NLP), draw attention of many public health professionals because of its potential to ease the shortfall of health workers.[4]In India, the quality of TB care varies widely depending on the geographical location and the socioeconomic status of a patient, resulting in delayed or missed diagnosis of active TB cases.[5]

There is approximately only one radiologist for every 100,000 population in India. In Zambia, also one of the high-TB/HIV-burden countries, there are only two radiologists in the country of over 11 million population.[3] The national goal to eliminate TB in India by 2025 and around the world by 2030 can be facilitated by early adoption of AI in TB control program, and successful use of AI in ending TB will help the world achieve the goals.

There are various AI tools in development in India and there is need to have a robust AI tool with a high sensitivity and specificity which can be used as a screening tool.This document is intended to propose a benchmark for AI in radiographic detection of Tuberculosis which include data format, desired data for AI training and testing as well as AI performance evaluation methodologies.

1.2 Relevance

The topic is of immense relevance in view of following:

a. AI technologies, including deep learning (DL) and natural language processing (NLP), draw attention of many public health professionals as it can overcome the shortage of healthcare professionals. Based on the study conducted by the Center for Internet and Society, India (CISI), it is evident that India has a well-established infrastructure for integration of AI into the Indian health technology infrastructure.

b. AI has gained a platform in India to enable the scientific community to deal with challenges related to cognitive disorders and social issues through the use of psychological tools & batteries, early diagnosis and better therapies, intervention technologies and rehabilitation programs.

c. India has conducted mobile TB diagnostic van intervention using X-ray diagnostic vans and sputum microscopy for diagnosis of TB intribal population which has resulted in increase in detection and reduction in out of pocket expenditure of patients . The AI technique can help in diagnosis of TB cases in difficult o reach areas in India and other low income countries.

d. Also National TB prevalence survey is being conducted covering 500000 population in entire coutry using mobile X-rays in field. The AI x-ray diagnosis would be of immense use in field diagnosis of TB and can be used across the globe in other countries as well where there are limited resources and expertise.

1.3 Current approaches and Gold Standard for detection

The gold standard for diagnosis of TB is the microbiological confirmation either by culture or CBNAAT. The sensitivity of the smear microscopy is low and it tends to miss many cases of TB and the availability of the CBNAAT is not there in peripheral areas because of the infrastructure that it requires and lack of expert manpower to run it and of course the high cost. X-rays are done to further confirm the findings. Therefore role of X-ray becomes more important specially as the triage test to detect the TB cases and moreover, the availability of the X-ray machines in the periphery make is more feasible. Currently the X-rays are read by the radiologists and co-related clinically. However in resource poor countries use of AI can help in radiographic detection of TB at a very low cost thus making a huge impact in saving lives.

1.4 Possible impact of AI on this Topic

Pulmonary TB being an infectious disease has the threat to spread in absence of its timely detection which is a major challenge. The current diagnostics available make it more challenging asmany millions across the world are missed by the conventional method. Use of AI for radiographic detection of TB would have greater public health impact around the world in view of its potential tobe used in remote and hard to reach areas for detection of TB.

1.5 Expected Impact of the benchmarking

The benchmark dataset for the X-ray detection of TB should be representative of not only of one region but the entire world to be robust enough to have >95% sensitivity and 100% specificity. The benchmarking for the AI tool would help in generating such a dataset which could help in validation of AI tool across the world.

The benchmark dataset for the X-ray detection of TB should be representative of not only of one region but the entire world to be robust enough to have >95% sensitivity and 100% specificity. The benchmarking for the AI tool would help in generating such a dataset which could help in validation of AI tool across the world.

1.6 Ethical considerations

- **Ethical consideration of benchmarking including its data acquisition**: Ethical considerations on collection of data and thereafter usage of AI for public health are very important. The major concern is about the data anonymization.The identifiers must be removed from the data and the data used for learning should be confirmed via gold standard tests. The data acquisition should be voluntary from the cases and their contacts.Consent of patients for use of the data for development of AI tool must be taken.
- **Ethical consideration for use of AI tool for public health use:**The radiographic detection of TB using AI must be 100% specific and >95% sensitive.Primarily the tool could be used for screening purposes in remote settings followed by the final diagnosis by other methods. The patients must be informed of the use of such tools for screening and detection of their disease and their implications.

2 Existing AI solutions

There are few AI tools available for radiographic detection or screening of TB. But they have been trained on a limited dataset and the entire range of variables have not been covered. Some of the systems have very low specificity and thus more false negative cases would be detected. These tools tend to fail on the undisclosed data sets. Therefore, there is a need for a standard benchmarking for the training and testing the data sets.

2.1 Existing work on benchmarking

There has been discussion on having a need for a benchmarking in area of AI for health but for the radiographic detection of TB, there is no benchmarking done till now. The tools available have been trained on the data set that was made available to them were mostly hospital based.

3 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 labelled 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 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.)

This Topic Group is for building AI based solution for radiographic screening/detection of Tuberculosis.

4 Method

The method for AI benchmarking includes input data format requirement and output data, testing data, labelling and testing and scoring matrixes.

4.1 AI Input data

Chest X-Ray images obtained from culture confirmed (Gold standard) cases would be included for AI benchmarking initially and subsequently the X-ray images from cases confirmed by molecular tests would also be included. Besides normal X-rays from various geographic locations, and confirmed Non-TB cases like pneumonia, Bronchitis etc would be included. Well annotated labelled images would be required.

4.2 AI output data structure

The AI output should include AI tool differentiating Normal and abnormal X-rays in 1st Phase and subsequently Tb and Non-TB abnormal cases and finally detecting various presentations of TB on Chest X-rays along with lesion position, area, classification

1) For clinical evaluation of AI algorithm for radiographic detection of Tuberculosis, images would be obtained from confirmed cases of TB and images labelled by Expert panel. The

panel would involve at least two experts with 3-5 years of practice in chest radiology. In case of any discrepancy among the two radiologists, the image would be referred to the third radiologist for final decision on disputed annotations from the other experts. All experts would receive prior specialized training regarding how to annotate TB lesions (including cavitary lesions) and lesion boundaries etc.

2) For clinical evaluation of cases for AI for radiographic detection of TB, gold standard is the microbiological confirmation of cases i.e either by sputum culture confirmation, CBNAAT confirmation or confirmation by clinical follow-up. The data set would also include cases from non-TB conditions which would include pneumonia, asthma, bronchitis, etc.

3) Confidentiality of gold standard testing data results would be maintained.

4.3 Model development

The model will be developed in following stages -

– **Stage 1:** Building an algorithm that interprets chest plain radiography and detects signs of abnormality and differentiates normal from abnormal X-ray and pulmonary tuberculosis among abnormal X-rays.

– **Stage 2:** Building a more comprehensive algorithm that combines imaging and other clinical information to provide more reliable prediction for diagnosis of tuberculosis and different presentations of TB on Chest radiography

– **Stage 3:** Expanding the model to be used in detection of other pulmonary diseases

Each stage will consist of three sub-phases;

– **Phase 1:** Retrospective data collection and model building

– **Phase 2:** Prospective validation and user feedback

– **Phase 3:** Full deployment of the system and continuous improvement

4.4 Scores and Metrices

Testing

– To evaluates AI tool’s performance Dataset comprising of a mix of each confounding variable case test data would be taken and tested against performance of AI. The initial diagnosis would be to differentiate normal with an abnormal ones and then TB and Non-TB. Further types of TB like cavitary, military, and lobes affected etc. would be differentiated.

– **Primary Benchmarking:** Primary testing would include detection of abnormal X-rays from normal X-rays from overall dataset. Abnormal X-rays would be further classified as TB and Non-TB based on detection of TB lesion. The data would also include normal X-rays and also X-rays from other non-TB cases. TB lesions detected by AI tool, would be compared with pre-labelled lesions to determine the true positive and false positive cases. Benchmarking metrics would include sensitivity of the tool to detect TB cases based on TB lesions and false positives. Initial detection would be between normal and abnormal cases and further classification would be for tuberculosis based on which sensitivity would be calculated. Specificity would be calculated based on ability of the AI tool to accurately detect and identify non-TB cases as Non-TB.

– **Secondary Benchmarking:** This would involve marking the lesion size, area, cavity, classification etc. Early cases of TB can be easily missed if not seen by experienced radiologists. Therefore, confirmed cases with early lesions in X-rays from confirmed TB cases clearly marked by radiologists would be included to train and test the performances.

4.5 Available data sets and undisclosed test data sets

1) In order to assess algorithm robustness, sufficient and diversified data from multiple heterogeneous sources (e.g., various types of digital images like, digital and Biochemical films converted into digital images, patient demographics like age, sex, socioeconomic status, geographical areas, smokers and other clinical conditions, etc.) would be used for testing to verify the generalization capacity of AI. Public and real-world undisclosed data (desensitized) would be collected.

2) Database currently would include59000 X-Ray images from a community base TB prevalence survey done in one of the states in South India which includes culture confirmed cases and normal X-rays. The Images Collected from various TB studies would also be included. The data set would also include the X-ray images from the ongoing National TB Prevalence survey which included 500,000 x-ray images to be collected from 625 clusters from all the states of the country over a period of one year. This would cover the geographic differences (rural and urban, plain and plateau etc.), gender differences, cases with different socioeconomic status and lifestyles. Data from various studies from multiple hospitals (Public and private) would also be collected. Besides data from cases enrolled in various clinical studies across various centers in India would also be included

3) The training data set would comprise of 80% of the entire data set and undisclosed test data would include 20 percent of the entire data set. The data sets would cover the entire variations as discussed above.

4) The entire data set would also have confirmation of cases by Gold standard test (microbiological or clinical).

5) The AI tool would be further validated in a prospective manner by the community based and hospital based study wherein the initial diagnosis of a new TB suspect would be given by the AI tool based on X-ray findings which would be further confirmed by physicians based on report from radiologists, microbiological tests and clinical follow-up.

6) A panel of Radiological and Pulmonary experts will examine labelled undisclosed test data to confirm data variance, quantity, heterogeneity, labelling and conformity to ethical and legal requirements.

4.6 AI tool development and status

ICMR, under DHR, Govt. of India, currently has large amount of data clinical and Pictorial generated through its 27 permanent Institutes and regional Research centres and also through ICMR funded studies. ICMR also has extensive clinical expertise in developing AI tool for diagnosis or screening of various communicable and non-communicable diseases and is in a position be to lead the study.

Collaboration with Institute of Plasma Research, DAE

We are currently collaborating with Institute of Plasma Research (IPR), Under Govt. of India for the development of the AI tool for radiographic detection of TB. The IPR has vast infrastructure and expertise in developing the tool for the Public health use. They have already developed the tool which can differentiate between TB, pulmonary TB foot prints and abnormal and normal Chest X-ray cases, however the tool needs to be trained on a larger data set including larger variant of tuberculosis as required.

Features of AI tool:

The automated tool can automatically detect foot print of Pulmonary Tuberculosis in Chest X ray at a rate of80 images per minute, can identify foot prints of Pulmonary tuberculosis and other chest ailments; can differentiate normal X ray from abnormal X ray using images of both biofilms as well as dicom version of digital X ray to some extent. The software also had an added advantage of being cloud independent and can be used in common desktops and laptops.

The process of training of the tool for detection of TB is in process. The tool would further be tested in a prospective study.

Salient features of the strategy to develop AI would be:

1. The tool must be trained to identify all possible variant of pulmonary tuberculosis.
2. Tool must have high sensitivity as well as specificity. Maximum tuberculosis cases should get screened out and no non tuberculosis case be identified as tuberculosis.
3. Tool must be able to screen tuberculosis considering all possible presentations of subject/chest X ray film as per geography, age, sex, occupation, stage of tuberculosis, quality of image, typeofimage.
4. Tool can be used in remote areas with minimal manpower.
5. Chest X ray films closely resembling tuberculosis but are non-tubercular in nature should ideally be screened as negative.

4.7 Reporting Methodology

– Reporting metrics would be two staged; first stage the AI tool would differentiate between Normal and abnormal. Normal cases would be true normal cases and the accurate detection would define specificity.

– Second would be abnormal but TB or Non-TB. This would define the sensitivity of the tool.

– Later on the tool can be trained to detect other non-TB chest lesions like pneumonia, Bronchitis, Cancer etc.

4.8 Results

Data is being collected at present. National TB Prevalence Survey has been initiated in the country.

Will cover 500,000 population in 625 clusters across the country

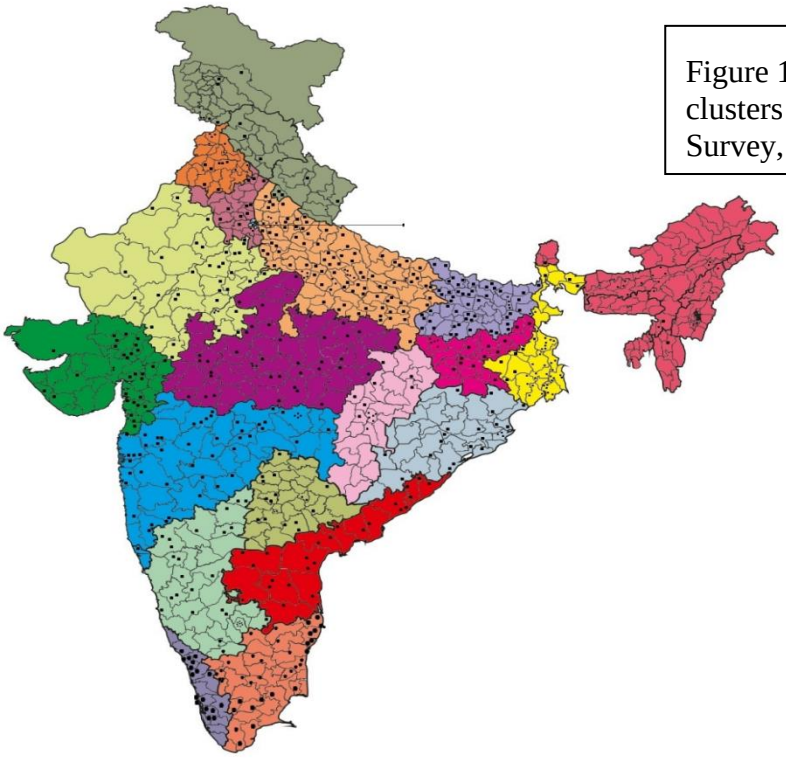

Figure 1: Map of India depicting survey clusters for National TB prevalence Survey, India

25 state of art buses fitted with Digital X-ray machines and CBNAAT for sputum examinations would go to 625 clusters for the survey. The X-rays being collected are dicom images. The reading would be done by 2 radiologists for each X-ray. IN case of discrepancy the image would be sent to 3rd radiologists. The final diagnosis would be confirmed with results from culture, CBNAAT, and clinical examination and subsequent follow up by local health authorities.

The field activities have been completed in August, 2021 and results of Gold standard tests are awaited. The survey restarted in October 2020; However again puton hold due to 2nd wave of pandemic and lockdown (April 21 to June 21). Field activities of survey have been completed and results of Gold standard tests are awaited. A total of 320864 population covered in 443 clusters. The annotated images of X-rays are being shared with IPR for development of AI tool.

4.9 Discussion

NA at present

4.10 Expected outcome

Development of a Cost-effective AI tool for radiographic screening/diagnosis for early detection of TB

4.11 Declaration of Conflict of Interest

None

5 Progress in terms of collaboration for development of AI tool

1. **Institute of Plasma Research, under DAE, Govt. of India** is developing the AI tool for radiographic detection of the TB using Chest X-ray. The MOU has been signed and the training of the tool has been initiated on the data of culture confirmed annotated X-rays from previous survey. The annotation of the X-rays is being done. Following the training, the tool would be placed in the survey vans/data from survey vans would be used for the training. However due to Covid-19 pandemic the survey has been put on hold.

   However the annotation of the X-rays is being done in consultation with Physician and radiologists.

2. **Collaboration with South Africa**, (Under BRICS collaboration): ICMR and WITS-CAD had MOU for the validation of the AI tool in National TB Prevalence Survey currently being undertaken by ICMR.The tool is well trained and is ready for use. It can work offline to give a diagnosis right in the survey buses. The tool was to be installed in Buses in last week of March 2020 for testing /validation in first phase. However due to Covid-19 Pandemic the team from South Africa could not visit India and also the survey is put on hold due to Covid-19 pandemic and would be planned soon. The survey resumed in October 2020 and validation restarted on the central X-rays data from collected from all parts of the country.

   Progress: A validation on 30000 X-rays collected from all parts of the country were validated.

   For the validation, 40 clusters were selected at random from the completed clusters of NTBPS which represents entire nation (20 Groups (i.e., State Group) 2 per State – Selected at random)

   There were discrepencies seen in data read by AI, MO and Teleradiologists. Line list of the discrepancy list was provided for the purpose ofEvaluation
   1. Medical Officer reading vs. AI
   2. Tele radiologist reading vs. AI
   3. Medical Officer/Tele radiologist reading vs. AI
   4. TB Disease Status vs. AI

Concerns: Need to understand the image quality variable

- What correction need to be done to improve the quality?
- If X-ray taken by same technician with same Machine. Then, what makes an impact on the quality?
- Is foreign bodies in images has any impact?
- Do running with multiple files (> 50 files, same time) make an impact on AI reading

The comparison with Gold standard is still ongoing and final results would be available once the microbiological results (Gold standard results are available

Screening of TB using Cough Sounds: Institute of Plasma Research (IPR) is also working on the screening of Signatures of Covid-19 using cough sounds and have developed a tool on limited data sets. IPR has developed a web application to detect COVID-19 by analysing voice recorded remotely. The developed web-based mechanism collects voice samples of COVID-18 patients (through doctors) and healthy individual volunteers and is compatible with PCs and mobile phones to record and upload voice samples, at https://covid19.ipr.res.in/app_Login/

IPR now plans to further extend it for TB for which the cough sound data is being collected for training.
