Accelerating AI for Lung Health: CAD for Silicosis Challenge

1. Background

Almost one in four people with tuberculosis (TB) globally are not diagnosed and reported according to the World Health Organization (WHO). Miners and other workers exposed to silica dust face occupational hazards, poor living and working conditions, and a high HIV burden, which coalesce to give a high risk for TB. Therefore, systematic TB screening of current and former miners using chest X-ray is strongly recommended by WHO. It is a crucial undertaking both for countries with established mining sectors (such as South Africa, India, China, and Brazil) and those where there are emerging mining industries and artisanal mining operations, including Mozambique, Tanzania, and Democratic Republic of Congo. Silicosis and TB are both responsible for substantial morbidity and mortality among mining communities and their co-occurrence (silico-tuberculosis) is associated with further elevated mortality.

In mining affected population, chronic dust exposure leads to silicosis which in turn is a risk factor for TB. Diagnosis of silicosis is important not only for providing care for silicosis but also for providing TB preventive treatment to prevent the development of active TB disease. In addition, people with silicosis are eligible for compensation in certain settings, such as in South Africa. The co-existence of TB and silicosis in the same person creates challenges in interpretation of X-ray chest for both diseases, Having demonstrated high accuracy in general populations, artificial intelligence (AI)-powered computer-aided detection (CAD) products are currently WHO-recommended for the screening and triage of TB disease in adults aged 15 years and older. However, many TB-CAD products are not yet independently validated for the detection of other, often co-occurring, lung diseases, including silicosis. Current literature suggests the high potential of several CAD products for silicosis detection, however both the literature and the breadth of commercially available CAD products capable of detecting both TB and silicosis are limited, despite the clear necessity of such tools.

Mining is a key sector in South Africa's economy, which has also shaped the public health of the southern African region through migration from neighboring countries to work in South African mines. Across the region, initiatives led by the Medical Bureau of Occupational Diseases (MBOD) of South Africa, mining companies, and international partners, such as the International Organization for Migration, World Bank, Global Fund and Stop TB Partnership, have been implemented to accelerate the detection of TB and silicosis in mining populations. Recognizing the significance of the intersection between occupational health and TB and its potential to improve the clinical management of TB and lung health outcomes more comprehensively for under-served populations, the Stop TB Partnership and MBOD, South Africa, (is collaboration with University of British Columbia and University of Cape Town) are launching a new initiative aimed at accelerating the development and validation of CAD to detect TB and silicosis disease in current and former miners.

This Expression of Interest (EOI) is issued to identify TB-CAD developers who are developing models to detect silicosis alongside TB. Specifically, through this EOI, we seek to identify potential partners interested in improving their CAD algorithms through access to the training set and validating such solutions.

2. Objectives

- To advance the development of CAD for silicosis detection through AI training.
- To contribute to global knowledge on the accuracy of AI for silicosis, TB, and silico-tuberculosis in mining populations.

3. Planned activities

Accelerating AI for Lung Health: CAD for Silicosis Challenge

Stage 0) Baseline validation (n = 501 CXR)

MBOD has made available 501 CXRs that include CXR containing silicosis only, TB only, silico-tuberculosis, and no disease.

For stage 0, Stop TB Partnership will evaluate the existing silicosis model for each interested vendor to document the baseline performance. Each CAD developer will install their software on the Stop TB Partnership server. The validation dataset will be analyzed to provide the numeric abnormality score output for both TB and silicosis. This baseline validation will include area under the receiver operating characteristic curve (AUROC), sensitivity and specificity across a range of threshold scores.

Stage 1) Training (n = 857 CXR)

MBOD has made the Stop TB Partnership the custodian of a training dataset of 857 high-quality CXRs taken during medical examinations of current and former miners. All CXR were independently validated by two experienced occupational health professionals, the dataset is enriched with most images containing silico-tuberculosis, silicosis and TB. Each image is labeled with a radiologist report, 99% also have lung function test results and 89% have the results of a certification panel providing ILO classification.

Stop TB Partnership will securely host this training dataset on its server and each developer will receive secure access to download the images. Training is intended to enrich an existing silicosis product and does not constitute the development of a new product alone.

Stage 2) Internal validation (n = 501 CXR)

After Stage 1, the same TB CAD developers will submit the enriched models for an internal validation in which the model will be rerun on the 501 CXR dataset used in Stage 0.

Dataset analysis by CAD will follow Stop TB Partnership's process, as described above. Each CAD developer will now install their new model, following training. The installed CAD will analyze the CXRs and provide only the numeric abnormality score output for further analysis, for both TB and silicosis. Output will be made available to Stop TB Partnership, the University of British Columbia, the University of Cape Town and the MBOD.

Stop TB Partnership, with support of University of British Columbia and University of Cape Town, will conduct an independent evaluation of submitted algorithms, including calculating the Area Under the Receiver Operating Characteristic curve (AUROC), sensitivity and specificity across the range of threshold scores, against the relevant reference standards of Xpert for TB and experienced occupational health doctors for silicosis.

Potential future activities

Developers of high-performing algorithms in internal validation may be invited to participate in external validations using separate set of CXRs sourced through ongoing initiatives in field settings targeting miners, ex-miners, and others exposed to silica dust.

4. Timeline

Baseline validation is planned for April and May 2025, with training taking place from May-June. Internal validation is planned to commence in June 2025.

5. Eligibility and terms

All AI vendors with a TB CAD product to detect silicosis are invited to train evaluate their algorithms using the MBOD dataset. Access to the training dataset will be contingent on agreement to:

- Provide the product free of charge for research purposes
- Participate in all stages of the challenge - training data will only be provided to those that consent to baseline and internal validation.
- Sign a data sharing agreement with Stop TB Partnership (as custodian of the data) to cover these activities.

6. Results and publication

The results of this activity will be presented in peer-reviewed manuscripts and shared in symposia. CAD developers will not be included in those publications but AI products will be named.

7. Next steps

Stop TB will convene a webinar on 26 March to outline the project. Please confirm if you are interested in attending by filling out this form.

Following this meeting, we will ask interested parties to formally confirm their participation by 2 April—further instructions and data sharing agreements will be shared post introductory webinar.