



Perspective

# Application of artificial intelligence in digital chest radiography reading for pulmonary tuberculosis screening

Xue-Fang Cao <sup>a</sup>, Yuan Li <sup>b</sup>, He-Nan Xin <sup>a</sup>, Hao-Ran Zhang <sup>a</sup>, Madhukar Pai <sup>c</sup>,  
Lei Gao <sup>a,\*</sup>

<sup>a</sup> NHC Key Laboratory of Systems Biology of Pathogens, Institute of Pathogen Biology, And Center for Tuberculosis Research, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China

<sup>b</sup> JF Healthcare, Nanchang, Jiangxi 330072, China

<sup>c</sup> McGill International TB Centre, McGill University, Montreal, Canada

Received 11 December 2020

Available online 3 March 2021

## Abstract

Currently, the diagnosis of tuberculosis (TB) is mainly based on the comprehensive consideration of the patient's symptoms and signs, laboratory examinations and chest radiography (CXR). CXR plays a pivotal role to support the early diagnosis of TB, especially when used for TB screening and differential diagnosis. However, high cost of CXR hardware and shortage of certified radiologists poses a major challenge for CXR application in TB screening in resource limited settings. The latest development of artificial intelligence (AI) combined with the accumulation of a large number of medical images provides new opportunities for the establishment of computer-aided detection (CAD) systems in the medical applications, especially in the era of deep learning (DL) technology. Several CAD solutions are now commercially available and there is growing evidence demonstrate their value in imaging diagnosis. Recently, WHO published a rapid communication which stated that CAD may be used as an alternative to human reader interpretation of plain digital CXRs for screening and triage of TB.

Copyright © 2021 Chinese Medical Association. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Keywords:** Tuberculosis; Artificial intelligence; Digital chest radiography; Diagnosis; Triage

## Introduction

Tuberculosis (TB) remains one of the major infectious diseases that seriously endanger human health. Nearly 1.4 million people died from TB in 2019. Of the estimated 10 million people who developed TB that year, some 3 million were either not diagnosed, or were not officially reported to national authorities.<sup>1</sup> The early diagnosis of pulmonary TB is of great significance for disease control, which is mainly based on comprehensive consideration of clinical suspected

\* Corresponding author. NHC Key Laboratory of Systems Biology of Pathogens, Institute of Pathogen Biology, and Center for Tuberculosis Research, Chinese Academy of Medical Sciences and Peking Union Medical College, No. 9 Dong Dan San Tiao, Dongcheng District, Beijing 100730, China.

E-mail address: [gaolei@ipbcams.ac.cn](mailto:gaolei@ipbcams.ac.cn) (L. Gao).

Peer review under responsibility of Chinese Medical Association.



symptoms, laboratory examinations, and chest imaging.

In China, TB patients are mainly identified by the routine surveillance system through presenting themselves to health-care facilities with suspected TB symptoms.<sup>2</sup> However, pulmonary TB patients without typical symptoms and those who have symptoms but do not visit health-care facilities might be missed. Symptom screening for TB was a key component for combating TB, and the health care provider-initiated active TB screening in high-risk populations had become an important strategy recommended by WHO.<sup>3</sup> In China, the active case finding had been implemented in close contacts of sputum smear positive TB patients, mainly via suspected symptoms screening. Those who present suspected TB symptoms would accept the subsequent chest radiography (CXR) and laboratory examinations in routine practice. However, such active case finding based on suspected symptoms screening showed lower sensitivity and specificity (especially among people living with HIV), and even smear-positive TB cases might be missed.<sup>4</sup>

WHO has recommended that CXR screening can be used for TB case finding. It not only can immediately check the presence of lesions suggestive of TB, but also can determine the location, nature and severity of the lesions. According to the previous studies, CXR examination based case finding showed higher accuracy than symptoms screening.<sup>4–6</sup> However, despite the high sensitivity when interpreted by experienced radiologists, CXR has its challenges, mostly due to its modest specificity, high inter- and intra-reader variability.<sup>5–8</sup> Besides, it has been reported that there was a shortage of qualified certificated-radiologists in many TB prevalent settings, which may impair screening efficacy.<sup>9,10</sup> High cost of CXR hardware also means that these tests are rarely available at the primary care level.<sup>11</sup>

Along with the application of computer-aided detection (CAD) system in reading medical images, several CAD software products have been developed to interpret digital CXRs for abnormalities suggestive of TB or other diseases.<sup>1</sup>

### **Application of CXR in TB detection**

Once hardware has been acquired, digital CXR has the advantages of low running cost and easy operation. Especially in low-income areas, digital radiography (DR) may be the only available medical imaging modality. CXR is used for a variety of conditions as a good tool for strengthening health systems.

For pulmonary TB, the radiographic presentation is wide-ranging, making it a challenging diagnosis. Active pulmonary TB can be subdivided into different forms: primary TB, secondary TB, and miliary TB. Primary pulmonary TB always develops shortly after infection, and CXR images mainly manifest as lymphadenopathy, consolidation, pleural effusion, and miliary nodules.<sup>12</sup> Secondary TB develops after a long period of latent infection, with CXR images that usually manifest as cavities, consolidations, and centrilobular nodules. Miliary TB is caused by hematogenous dissemination, which shows miliary lung nodules in a random distribution on CXR images.<sup>12</sup> In addition, prior old TB lesions often manifests as fibronodular opacities in the apical and upper lung zones, and fibronodular change is associated with a considerably higher risk of TB reactivation.<sup>12</sup>

As many diseases may present with similar radiologic patterns, differential diagnosis of diseases must be considered when reading a CXR, including different forms of pneumonia, occupational lung diseases, neoplastic disease, pleural effusion, and cardiac disease.<sup>13</sup> Hence, experienced radiologists are needed to interpret the images. However, even among trained human readers, inter- and intra-reader variability is common because they are susceptible to fatigue, perceptual biases, and cognitive biases, which may lead to errors.<sup>14</sup> Moreover, it is sometimes difficult to distinguish similar lesions or find very obscure nodules, particularly when the contrast between the lesion and the surrounding tissue is very low or the lesion overlaps the tissue structures or larger pulmonary blood vessels.<sup>15</sup> Therefore, the examination of TB using CXR will cause a certain degree of over- and under-diagnosis of TB. This is why an abnormal CXR should be followed up with a microbiological test for TB (such as smear microscopy, molecular test or liquid cultures).

### **CAD with machine learning (ML) technology in TB detection**

With the development of artificial intelligence (AI), CAD systems have become a hot research topic, which can analyze the data related to the patient under investigation and suggest a diagnosis based on the analysis.

ML and Deep Learning (DL) are two frequently used AI methods for creating CAD programs. ML is the scientific discipline, which relies less on human specification and instead allows algorithms to decide what variables are important.<sup>16</sup> As the CAD system in

CXR can help doctors detect suspicious lesions that are easy to miss, thereby improving the accuracy of detection, many CAD algorithms of CXRs have been developed for routine diagnostic procedures. Previous studies reviewed the development of CAD technology between 2001 and 2009, and found the CAD systems mainly focused on a single aspect, such as detecting lung cancer nodules.<sup>17,18</sup>

In terms of TB detection, many studies about ML-based CAD systems on CXRs have been published in the past two decades and two studies have reviewed them.<sup>19,20</sup> A majority of the studies primarily focused on how to create a CAD program for TB, with the AUC ranged from 0.78 to 0.99.<sup>20</sup> However, these studies usually used the same databases for training and testing as well as used human readers as the reference standard, which might overestimate the diagnostic accuracy of CAD. Thus, studies focusing on the assessment of the accuracy outside of the study setting are needed to evaluate the value of the systems in TB detection. Besides, CAD has also been used in conjunction with clinical information; achieving better performance than using CAD scores or clinical information alone for TB diagnosis.<sup>21</sup> These previous works suggested the potential of using CAD systems in detecting TB.

### CAD with DL technology in TB detection

DL is a subset of ML, which attempts to model brain architecture and has great application prospects in the field of AI in the past few years.<sup>16</sup> It is expected to revolutionize CAD systems and to play an important role in precision medical imaging. Especially in the field of computer vision, deep convolutional neural networks (DCNNs) have been proven to be very promising algorithms for various visual tasks. [Table 1](#) shows the DL CAD solutions that are commercially available.

Since DCNNs can perform end-to-end training from images to classification, there is no need for target-specific manual feature engineering. Several studies primarily focusing on reporting methods for creating CAD programs for TB have published since 2016.<sup>22–25</sup>

In 2017, two DCNNs (AlexNet and GoogLeNet), including pretrained and untrained models were used to evaluate the efficacy for detecting TB on CXR, which found that DL with DCNNs could accurately classify TB from healthy controls with an area under the receiver operating characteristic (ROC) curve (AUC) of 0.99.<sup>22</sup> Later, a deep DCNN model was trained by a TB-specific CXR dataset of one population (National Library of Medicine Shenzhen No.3 Hospital, China) and was tested by non-TB-specific CXR dataset (ChestX-ray8) of another population (National Institute of Health Clinical Centers, USA). This DCNN model exhibited an AUC of 0.98 and 0.85 for detecting TB in the training and intramural test sets, respectively, but the AUC dropped dramatically to 0.70 on the ChestX-ray8 dataset.<sup>24</sup> About one third (36.51%) of abnormal radiographs in the ChestX-ray8 dataset were estimated to be TB-related. This suggested weaknesses of DL, such as poor generalizability and overdiagnosis. Therefore, technical specification of CXR images, disease severity distribution, dataset distribution shift, and overdiagnosis should be evaluated before applying the model in practice; especially in settings that differ from those where training dataset were from.

CAD4TB (Delft Imaging Systems, Netherlands) is one of the earliest commercially available CAD for TB. However, DL was not employed until the development of version 6, which was released in 2018.<sup>26</sup> More CAD software products with DL technology were released in the past three years in worldwide, including a few from China, thereby more clinical studies primarily focusing on the assessment of the accuracy of the already-developed CAD software were conducted. A prospective study was conducted in a tertiary hospital from Pakistan to evaluate the performance of qXR(v2) and CAD4TB(v6) with mycobacterial culture being a standard reference, which collected a dataset of 2187 CXRs and reported that both software achieved non-inferior accuracy (qXR: sensitivity = 0.93; specificity = 0.75; CAD4TB: sensitivity = 0.93; specificity = 0.69) to WHO-recommended minimum values, while sensitivity might

Table 1  
Characteristics of AI-based solutions for interpretation of digital CXRs.

AI Solution	Developer	Category	Certification	Commercially available
CAD4TB version 6	DIAG, Radbound University, The Netherlands	DL	CE	Yes
qXR	Qure.AI, India	DL	CE	Yes
Lunit INSIGHT CXR	Lunit, South Korea	DL	CE	Yes

CXR: chest radiography; CE: Conformite Europeenne; DL: deep learning.

be lower when smear-negative pulmonary TB was more prevalent.<sup>27</sup> A retrospective case–control study was conducted in a tertiary hospital from India using microbiologically-confirmed TB as the reference standard, which reported that qXR (v2) could detect TB with an AUC of 0.81, a sensitivity of 71% and a specificity of 80%.<sup>28</sup> As many patients who present at tertiary hospitals have symptoms suggestive of TB, accurate and rapid triage tests that can rule out the disease are needed. Thus, the CAD software as a triage test for TB might be more suitable for the primary care level in settings where access to radiologists is limited than the tertiary care level.

In 2019, the Stop TB Partnership published a retrospective evaluation of three DL systems (CAD4TB (v6), Lunit INSIGHT (v4.7.2), and qXR(v2)) for detecting TB-associated abnormalities in CXRs from outpatients in Nepal and Cameroon. All 1196 individuals received a Xpert MTB/RIF assay and a CXR read by two groups of radiologists and the DL systems, with all three systems performing high AUCs (Lunit: 0.94, qXR: 0.94 and CAD4TB: 0.92) and significantly better than human radiologists.<sup>29</sup> A larger dataset of 23,566 CXRs from Bangladeshi was used to evaluate five AI software platforms specific to TB-associated abnormalities, including CAD4TB (v6), InferRead®DR (v2), Lunit INSIGHT CXR (v4.9.0), JF CXR-1 (v2) and qXR (v3) by the same study group one year later. All AI systems significantly outperformed the human readers, who were certified radiologists with

different years of working experience, and the AUCs from high to low are qXR (0.91), Lunit INSIGHT CXR (0.89), InferReadDR (0.85), JF CXR-1 (0.85), and CAD4TB (0.82).<sup>30</sup> Fig. 1 shows a screenshot of JF CXR-1, which automatically generates probability score for TB and a heat map localizing the most indicative region of the image, providing reference for radiologists to identify TB-related abnormalities on CXRs.

## WHO endorsement of CAD for CXR

In December 2020, WHO published a rapid communication on systematic screening for TB.<sup>31</sup> Evidence on the use of CXR as a screening tool for detection of TB disease in several populations were reviewed by WHO in 2020. Across all populations considered, “CXR was found to be a sensitive screening tool that, while lacking sufficient specificity to confirm a TB diagnosis, has an important role in the early detection of TB in children and adults at higher risk of TB, as well as to reduce the population burden of TB disease when combined with early treatment and other public health action.”

Three independent evaluations of CAD were reviewed by WHO for both the screening and triage use cases. Each evaluation considered all three CAD technologies that were CE marked at the start of 2020 and available on the market. Evaluations showed substantial variation in CAD performance (diagnostic

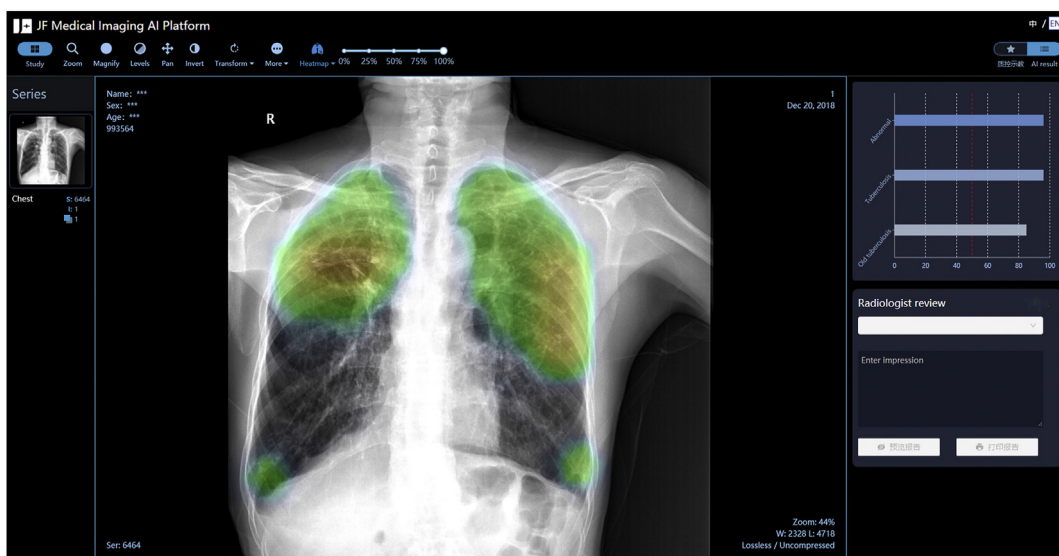


Fig. 1. A screenshot of JF CXR-1, a TB screening and triaging tool. This CAD system applies DL technology, which automatically generates a heat map and probability score for TB.

accuracy) across different contexts, implying that the use of CAD may require adjustment for the purpose and setting in which CAD will be implemented. Nonetheless, the diagnostic accuracy and the overall performance of CAD software were similar to the interpretation of digital CXR by a human reader, in both the screening and triage contexts.

Based on the review, WHO now recommends that CAD may be used as an alternative to human reader interpretation of plain digital CXR for TB screening and triage among adults.

## Conclusion and prospects

Accurate, accessible, and affordable tools are needed for precise detection and diagnosis of TB, the first and a very challenging step in ending TB. CAD system can automatically detect and localize abnormalities in CXRs at a level comparable to human radiologists, saving time for radiologists and more importantly addressing the challenge of human resource shortage. Therefore, CAD systems are more needed in low- and middle-income countries where TB burden are high and lack of trained radiologists.

As DL algorithm is more accurate and efficient than ML algorithm in the effective training of medical image big data,<sup>19</sup> the number of commercially available CAD systems with DL technology have increased substantially in recent years. They have been proved with highly sensitive and could save large proportions of the subsequent confirmatory tests. However, DL system performance at similar thresholds might vary greatly with the prevalence of TB. Therefore, while the WHO guidance for CAD software for TB emphasizes the need of predefined threshold scores, the implementers need to conduct their own pilot study to tailor specific thresholds according to the characteristics of the screened populations.<sup>29,32</sup>

With the recent WHO endorsement, CAD solutions can now be implemented in TB programs, guided by implementation research and operational studies to optimize impact. As for the development of future CAD systems, the correlation of demographics, clinical data and AI abnormal scores should be explored to achieve individualized risk scores, thereby optimizing performance. Besides, future research should also pay attention to improve the sensitivity of detecting smear-negative TB and to improve the generalizability of such novel technologies by eliminating diagnostic heterogeneity. In addition, as computed tomography (CT) scans allow practitioners to examine both the pulmonary parenchyma and the lymph nodes in greater

detail than can be done with plain CXR alone, the development and application of CAD for CT image reading should be attached importance as well along with the popularization of CT in TB diagnosis.

## Funding

This study was supported by a grant from the National Science and Technology Major Project of China [2017ZX10201302-008].

## Conflicts of interest

None.

## References

1. World Health Organization. *Global Tuberculosis Report 2020*. Geneva: World Health Organization; 2020.
2. Cheng J, Wang L, Zhang H, Xia Y. Diagnostic value of symptom screening for pulmonary tuberculosis in China. *PLoS One*. 2015;10, e0127725. <https://doi.org/10.1371/journal.pone.0127725>.
3. World Health Organization. *Interim Policy on Collaborative TB/HIV Activities*. Geneva: World Health Organization; 2004.
4. van't Hoog AH, Meme HK, Laserson KF, et al. Screening strategies for tuberculosis prevalence surveys: the value of chest radiography and symptom. *PLoS One*. 2012;7, e38691. <https://doi.org/10.1371/journal.pone.0038691>.
5. van Cleeff MR, Kivihya-Ndugga L, Githui W, Nganga L, Odhiambo J, Klatser PR. A comprehensive study of the efficiency of the routine pulmonary tuberculosis diagnostic process in Nairobi. *Int J Tubercul Lung Dis*. 2003;7:186–189.
6. Day JH, Charalambous S, Fielding KL, Hayes RJ, Churchyard GJ, Grant AD. Screening for tuberculosis prior to isoniazid preventive therapy among HIV-infected gold miners in South Africa. *Int J Tubercul Lung Dis*. 2006;10:523–529.
7. van Cleeff MR, Kivihya-Ndugga LE, Meme H, Odhiambo JA, Klatser PR. The role and performance of chest X-ray for the diagnosis of tuberculosis: a cost-effectiveness analysis in Nairobi, Kenya. *BMC Infect Dis*. 2005;5:111. <https://doi.org/10.1186/1471-2334-5-111>.
8. den Boon S, White NW, van Lill SW, et al. An evaluation of symptom and chest radiographic screening in tuberculosis prevalence surveys. *Int J Tubercul Lung Dis*. 2006;10:876–882.
9. Hoog AH, Meme HK, van Deutekom H, et al. High sensitivity of chest radiograph reading by clinical officers in a tuberculosis prevalence survey. *Int J Tubercul Lung Dis*. 2011;15:1308–1314. <https://doi.org/10.5588/ijtld.11.0004>.
10. Qin C, Yao D, Shi Y, Song Z. Computer-aided detection in chest radiography based on artificial intelligence: a survey. *Biomed Eng Online*. 2018;17:113. <https://doi.org/10.1186/s12938-018-0544-y>.
11. Pande T, Pai M, Khan FA, Denkinger CM. Use of chest radiography in the 22 highest tuberculosis burden countries. *Eur Respir J*. 2015;46:1816–1819. <https://doi.org/10.1183/13993003.01064-2015>.

12. Nachiappan AC, Rahbar K, Shi X, et al. Pulmonary tuberculosis: role of radiology in diagnosis and management. *Radiographics*. 2017;37:52–72. <https://doi.org/10.1148/rg.2017160032>.
13. World Health Organization. *Chest Radiography in Tuberculosis Detection: Summary of Current WHO Recommendations and Guidance on Programmatic Approaches*. World Health Organization; 2016.
14. Donovan T, Litchfield D. Looking for cancer: expertise related differences in searching and decision making. *Appl Cogn Psychol*. 2013;27:43–49.
15. Kulkarni S, Jha S. Artificial intelligence, radiology, and tuberculosis: a review. *Acad Radiol*. 2020;27:71–75. <https://doi.org/10.1016/j.acra.2019.10.003>.
16. Beam AL, Kohane IS. Big data and machine learning in health care. *J Am Med Assoc*. 2018;319:1317–1318. <https://doi.org/10.1001/jama.2017.18391>.
17. van Ginneken B, Hogeweg L, Prokop M. Computer-aided diagnosis in chest radiography: beyond nodules. *Eur J Radiol*. 2009;72:226–230. <https://doi.org/10.1016/j.ejrad.2009.05.061>.
18. van Ginneken B, ter Haar Romeny BM, Viergever MA. Computer-aided diagnosis in chest radiography: a survey. *IEEE Trans Med Imag*. 2001;20:1228–1241. <https://doi.org/10.1109/42.974918>.
19. Pande T, Cohen C, Pai M, Ahmad Khan F. Computer-aided detection of pulmonary tuberculosis on digital chest radiographs: a systematic review. *Int J Tubercul Lung Dis*. 2016;20:1226–1230. <https://doi.org/10.5588/ijtld.15.0926>.
20. Harris M, Qi A, Jeagal L, et al. A systematic review of the diagnostic accuracy of artificial intelligence-based computer programs to analyze chest x-rays for pulmonary tuberculosis. *PLoS One*. 2019;14, e0221339. <https://doi.org/10.1371/journal.pone.0221339>.
21. Melendez J, Sanchez CI, Philipsen RHHM, et al. An automated tuberculosis screening strategy combining X-ray-based computer-aided detection and clinical information. *Sci Rep*. 2016;6:25265. <https://doi.org/10.1038/srep25265>.
22. Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*. 2017;284:574–582. <https://doi.org/10.1148/radiol.2017162326>.
23. Hwang EJ, Park S, Jin KN, et al. Development and validation of a deep learning-based automatic detection algorithm for active pulmonary tuberculosis on chest radiographs. *Clin Infect Dis*. 2019;69:739–747. <https://doi.org/10.1093/cid/ciy967>.
24. Sathitratnacheewin S, Sunanta P, Pongpirul K. Deep learning for automated classification of tuberculosis-related chest X-Ray: dataset distribution shift limits diagnostic performance generalizability. *Heliyon*. 2020;6, e04614. <https://doi.org/10.1016/j.heliyon.2020.e04614>.
25. Heo SJ, Kim Y, Yun S, et al. Deep learning algorithms with demographic information help to detect tuberculosis in chest radiographs in annual workers' health examination data. *Int J Environ Res Publ Health*. 2019;16:250. <https://doi.org/10.3390/ijerph16020250>.
26. Murphy K, Habib SS, Zaidi SMA, et al. Computer aided detection of tuberculosis on chest radiographs: an evaluation of the CAD4TB v6 system. *Sci Rep*. 2020;10:5492. <https://doi.org/10.1038/s41598-020-62148-y>.
27. Ahmad Khan F, Majidulla A, Tavaziva G, et al. Chest x-ray analysis with deep learning-based software as a triage test for pulmonary tuberculosis: a prospective study of diagnostic accuracy for culture-confirmed disease. *The Lancet Digital Health*. 2020;2:e573–e581. [https://doi.org/10.1016/S2589-7500\(20\)30221-1](https://doi.org/10.1016/S2589-7500(20)30221-1).
28. Nash M, Kadavigere R, Andrade J, et al. Deep learning, computer-aided radiography reading for tuberculosis: a diagnostic accuracy study from a tertiary hospital in India. *Sci Rep*. 2020;10:210. <https://doi.org/10.1038/s41598-019-56589-3>.
29. Qin ZZ, Sander MS, Rai B, et al. Using artificial intelligence to read chest radiographs for tuberculosis detection: a multi-site evaluation of the diagnostic accuracy of three deep learning systems. *Sci Rep*. 2019;9:15000. <https://doi.org/10.1038/s41598-019-51503-3>.
30. Qin ZZ, Ahmed S, Sarker MS, et al. Can Artificial Intelligence (AI) Be Used to Accurately Detect Tuberculosis (TB) from Chest X-Ray? A Multiplatform Evaluation of Five AI Products Used for TB Screening in a High TB-Burden Setting. arXiv:2006.05509 [Preprint]. Available from <https://arxiv.org/abs/2006.05509?context=cs.LG>.
31. *Rapid Communication on Systematic Screening for Tuberculosis*. Geneva: World Health Organization; 2020.
32. World Health Organization. *Chest Radiography in Tuberculosis Detection: Summary of Current WHO Recommendations and Guidance on Programmatic Approaches*. World Health Organization; 2016.