Artificial Intelligence in Paediatric Tuberculosis

Jaishree Naidoo[1] · **Susan Cheng Shelmerdine[2,3,4]** · **Carlos F. Ugas -Charcape[5]** · **Arhanjit Singh Sodhi[6]**

Abstract

Tuberculosis (TB) continues to be a leading cause of death in children despite global efforts focused on early diagnosis and interventions to limit the spread of the disease. This challenge has been made more complex in the context of the coronavirus pandemic, which has disrupted the "End TB Strategy" and framework set out by the World Health Organization (WHO). Since the inception of artificial intelligence (AI) more than 60 years ago, the interest in AI has risen and more recently we have seen the emergence of multiple real-world applications, many of which relate to medical imaging. Nonetheless, real-world AI applications and clinical studies are limited in the niche area of paediatric imaging. This review article will focus on how AI, or more specifically deep learning, can be applied to TB diagnosis and management in children. We describe how deep learning can be utilised in chest imaging to provide computer-assisted diagnosis to augment workflow and screening efforts. We also review examples of recent AI applications for TB screening in resource constrained environments and we explore some of the challenges and the future directions of AI in paediatric TB.

Keywords Artificial intelligence · Chest radiography · Children · Computer aided detection · Deep learning · Tuberculosis

Introduction: Background and Scope

In 2020, 1.1 million children fell ill with tuberculosis (TB) globally. It is a leading infectious cause of morbidity and mortality in children worldwide and a major health concern, particularly in Africa and Southeast Asia, where the highest number of children are infected [1]. Despite being both curable and preventable, management of TB has been adversely impacted first by the human immunodeficiency virus (HIV) and more recently by the coronavirus pandemic.

✉ Jaishree Naidoo
jaishreenaidoo@hotmail.com

1 Envisionit Deep AI LTD, Coveham House, Downside Bridge Road, Cobham KT11 3 EP, UK

2 Department of Clinical Radiology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK

3 Great Ormond Street Hospital for Children, UCL Great Ormond Street Institute of Child Health, London, UK

4 NIHR Great Ormond Street Hospital Biomedical Research Centre, London, UK

5 Department of Diagnostic Imaging, Instituto Nacional de Salud del Niño San Borja, Lima, Peru

6 Department of Computer Engineering, Thapar Institute of Engineering and Technology, Patiala, Punjab, India

The "2020 Global TB" report revealed that the identification of TB fell by 18% from 2019 to 2020 representing a drop from 7.1 million to 5.8 million estimated diagnoses. This significant reduction during the recent pandemic was due to the reduced capacity to provide adequate screening services because of various lockdowns, as well as restrictions in movement and the associated risks of visiting health care facilities, all of which impacted access to TB diagnosis, testing and treatment [2].

Important progress has been made in improving TB diagnosis in recent decades due to the availability of sensitive bacteriological tests for diagnosis and screening. "Xpert mycobacterium tuberculosis/resistance to rifampicin assay" is a rapid nucleic acid amplification (NAA) test that can detect *Mycobacterium tuberculosis* as well as drug resistance to rifampicin [3] Although it is a sensitive test, it is costly and requires additional resources including staff and time allocation, especially in paediatrics where obtaining sputum samples can be challenging. Imaging, therefore, plays a critical role in diagnosing pulmonary TB in children, due to their nonspecific clinical manifestations, low bacillary load and the difficulties in obtaining suitable samples [4].

In March 2021, the World Health Organisation (WHO) released consolidated guidelines for the management of tuberculosis (module 2) with the recommendation to use chest radiography. Radiography is a sensitive screening tool (pooled sensitivity 98%) [5] and, although lacking sufficient

specificity to confirm a TB diagnosis, has an important role in the early detection of the disease in children. Additionally, research has demonstrated that children are at a higher risk of TB (compared to adults), and that early detection has the potential to reduce the overall population burden of the disease when combined with TB preventative treatment [5].

Unfortunately, many countries impacted with a high caseload of TB lack sufficient numbers of paediatric radiologists to interpret the increased influx of radiographic images. This creates challenges when screening high volumes of cases across a large range of countries [6].

In recent years, there has been a renewed emphasis on developing and deploying innovative sustainable solutions to deal with some of the health care resource challenges related to TB screening.

Computer-aided detection and machine learning have the potential to provide a solution to the resource and diagnostic challenges mentioned above. In the consolidated guidelines module 2, WHO updated its recommendations for TB screening to include computer-aided detection software packages to automate the interpretation of digital chest radiography images in patients older than 15 years of age, and to produce a numerical score indicating the likelihood of TB [5].

The "Stop TB" Partnership and the "Foundation for Innovative New Diagnostics" (FIND) launched an online resource centre of computer-aided detection products for the diagnosis of TB. In addition, they undertook a landscaping analysis, gathering intervention-relevant information on different computer-aided detection products for TB acquired from companies known to have developed these products [7].

Kim Y et al. [8] reviewed peer-reviewed research articles on AI in the thorax between 2015 and 2021. The review focused on how AI (specifically, deep learning) could be applied to complement aspects of the current health care system. With advances in technology and appropriate preparation of physicians, AI can address clinical problems that have not been solved due to a lack of clinical resources or technological limitations [8]. Understanding how these AI products work, their real-world applications, key performance metrics and the value proposition is essential for radiologists to consider when evaluating and reviewing the array of AI products for paediatric applications in TB diagnosis.

Chest radiography and tuberculosis

Up to 80% of reported cases of paediatric tuberculosis develop thoracic disease [9]. Chest radiography is an inexpensive tool that provides a highly sensitive method for screening patients (pooled sensitivity 98%) [5]. It is also a widely available aid in diagnosing pulmonary TB — when it cannot be confirmed bacteriologically — and uses a minimal amount of radiation [10].

Interobserver variability of chest radiography signs

Studies have evaluated both the diagnostic accuracy of chest radiography for TB diagnosis via the detection of specific radiographic findings and the classification of a normal/abnormal study. Studies that used the classification of normal/abnormal chest radiography have reported variable performance, from low to moderate interobserver agreement [11], and therefore assessing for specific findings on chest radiographs alongside the clinical evaluation is recommended [12].

Classic radiographic findings that may suggest TB include cavitations, nodules, consolidation, pleural effusion and enlarged mediastinal lymph nodes, the latter of which is commonly the only manifestation of the disease [13]. Most studies focusing on the detection of specific radiographic findings using imaging reported a low interobserver agreement, with widely variable kappa values from -0.03 to 0.52 [14–16]. Specifically, the diagnostic accuracy of mediastinal lymphadenopathy was reported as low, related to the possibility of the superimposition of structures, which did not improve despite using lateral views [14].

Recently, WHO published an updated guideline of recommendations with important alterations on how to use and interpret chest radiography to aid TB treatment and management [17, 18]. The changes proposed were to resolve some limitations, especially related to the detection of lymph nodes, with recommendations to focus on improved detection and better inter-reader agreement signs. This allows classification of affected children into groups of severity, where they might benefit from shorter regimens and fewer drugs.

The challenging landscape of paediatric TB detection using chest radiographs highlights the critical roles played by image quality and interpretation [19].

Several studies report the limitations of non-radiologists interpreting chest radiographs in children [20]. Given the worldwide shortage of radiologists, especially pediatric radiologists [6], this poses a problem particularly in countries with a high burden of TB.

Effectiveness of chest radiography as a screening tool

Chest radiography as a screening tool was studied by Huang et al. [21] in a cohort of 4,468 children exposed to TB. They evaluated the protective efficacy of isoniazid preventive therapy in asymptomatic children with and without abnormal chest radiographs. They showed that exposed asymptomatic children with abnormal chest radiographs were

25.1-fold more likely to have co-prevalent TB and 26.7-fold more likely to be diagnosed with TB during follow-up than exposed asymptomatic children with normal chest radiographs [21]. This finding is important and highlights that there is a large portion of children with subclinical TB who do not show classic signs and are going undiagnosed [22]. Detecting subclinical TB provides an opportunity for health care professionals to deliver care early in the disease history and may limit the progression and risk of post-TB sequelae and extensive lung damage.

According to this evidence, WHO guidelines recommend chest radiography as a screening tool that includes symptom screening for children in all age groups as a part of systematic evaluation processes/protocols [17]. It is not recommended that chest radiography is used for follow-up/evaluating improvement; this must be done through monitoring weight and height as well as clinical evaluation [17].

Artificial intelligence for the diagnosis of tuberculosis from chest radiographs

In recent years, AI and computer-aided detection software have been developed to augment and automate the interpretation of digital chest radiography in TB screening [7]. A literature search for publications including studies on AI in TB imaging using search items related to AI, TB and thoracic imaging yielded 110 results of which 21 were published in the last 5 years (Table 1) [23–43]. It is important to note that all but 2 of the 21 articles [32, 37] included studies where the algorithms were validated on data sets with a population age group > 15 years old. This highlights the lack of clinical studies of AI in the niche area of paediatric imaging.

Artificial intelligence is the process of making a computer think and learn through programming, training and testing. As a subset of AI, machine learning uses statistical methods to enable machines to improve with and evolve as a result of that experience [44]. Deep learning is a subset of machine learning that is used in circumstances in which a very large amount of data must be processed. These deep learning networks rely on multlayered configurations called artificial neural networks to process data [45] (Fig. 1).

These artificial neural networks, often hundreds of layers deep, can train themselves from large data sets to make accurate predictions on newly input unknown data.

Machine learning and deep learning algorithms have been developed to improve workflows in radiology or to assist the radiologist by automating tasks such as lesion detection (Fig. 2) or medical imaging quantification.

Lopez Garnier et al. [46] in 2019 summarised data on computer-aided detection type, study design and diagnostic accuracy. The authors included 53 of the 4,712 articles reviewed: 40 focused on computer-aided detection design methods (development studies) and 13 focused on evaluation of computer-aided detection. Of the 40 development studies, 7 (17%) used deep learning methods while the remaining 33 (83%) used machine learning approaches. The authors concluded that AI-based computer-aided detection programs are promising, but more clinical studies are needed that minimise sources of potential bias to ensure validity of the findings outside of study settings [46].

Triage and worklist prioritisation are important applications of AI in thoracic radiology and may have clinical relevance for TB screening where a timely diagnosis is critical with such a highly transmissible infection [8]. Of course, this will be reliant on accurate AI image review if the triage tool is to prioritise abnormal radiographs for urgent reporting or a highly specific AI tool using Natural Language Processing (NLP) that can review the clinical indication for suspicion of TB and bring this to the attention of the radiologist.

In the last few years, several studies have been published on the use of AI in diagnosing pneumonia and TB on chest radiographs; however, the scientific literature that includes the performance measures of AI products in TB intended to be used in the paediatric population is restricted and, further, there are very few existing adult applications that could be applied to the paediatric population [47].

The study by Mouton, Pitcher and Douglas [48] was the first to examine AI detection of abnormalities in paediatric chest radiographs in a population suspected of a high incidence of TB. The system reached reasonable performance with an area under the curve (AUC) of 0.78 for correctly identifying abnormal regions in the image.

Schalekamp et al. [47] reviewed 40 CE (Conformité Européenne)-marked commercial software packages and discussed their performances in different chest findings and diseases in paediatrics, where it was found that none of these products was specifically designed for that population. The intended use of most of the current AI products in radiology with CE certification is more tailored to adult screening.

Hwang et al. [49] developed a deep learning-based automated detection (DLAD) algorithm and compared radiology and non-radiology physician performance in image interpretation for the detection of active TB with and without AI assistance using six independent external multicentre test data sets. They concluded that both non-radiology physicians and board-certified radiologists showed improvements in sensitivity with the assistance of DLAD thus highlighting the potential of AI as a second reader [8, 28, 49].

Table 1 Summary of included publications

Study no	Author	Journal	Source of data	Cohort size	AI algorithm type	Performance
1	Codlin AJ et al [23]	Scientific Reports (2021)	Chest radiographs (CXRs) were derived from a community-based, mobile CXR screening initiative in Vietnam	In the final test library, 1,032 participants were included	Developed a test library of CXRs that was blindly re-read by two TB clinicians with different levels of experience and then processed by 12 computer-aided detection (CAD) software solutions	Six computer-aided detection systems performed on par with the expert reader (Qure.ai, Deep Tek, Delft Imaging, JF Healthcare, OXIPIT and Lunit) and one additional software (Infervision) performed on par with the intermediate reader only. Qure.ai, Delft Imaging and Lunit were the only software to perform significantly better than the intermediate reader
2	Feng B et al [24]	European Radiology (2020)	This study comprised an evaluation of the database from two hospitals in China	CT images of 550 patients with solitary solid pulmonary nodules (SSPNs)	CT-based DLN. The deep learning signature (DLS) model was developed using the CNN method	The AUC in the training, internal validation and external validation cohorts were 0.889 (95% confidence interval [CI] 0.839–0.927), 0.879 (95% CI 0.813–0.928), and 0.809 (95% CI 0.746–0.862), respectively The CT-based Deep Learning Nomogram (DLN) can preoperatively distinguish between LAC (adenocarcinoma) and tuberculous granuloma (TBG) in patients presenting with solitary solid pulmonary nodules (SSPNs)
3	Huang T et al [25]	Journal of Healthcare Engineering (2021)	The Second Affiliated Hospital of Fujian Medical University, Quanzhou, China	100 patients	A new lung CT image segmentation algorithm (U-Net+deep convolution (DC)) was proposed based on U-Net network and compared with the CNN algorithm	The specificity (94.32%) and accuracy (97.22%) of CT image diagnosis based on U-Net+deep convolution algorithm was significantly higher than traditional diagnostic method (75.74% and 74.23%), and the differences were statistically significant ($P < 0.05$)
4	Khan FA et al [26]	Lancet Digital Health (2020)	Indus Hospital, Karachi, Pakistan	Authors included 2,198 (92.7%) of 2,370 enrolled participants: 2,187 (99.5%) of 2,198 were HIV-negative, and 272 (12.4%) had culture-confirmed pulmonary tuberculosis	Authors compared two software's, qXR version 2.0 (qXRv2) and CAD4TB version 6.0 (CAD4TBv6), with a reference of mycobacterial culture of two sputa. They tested for non-inferiority to preset WHO recommendations (0·90 for sensitivity, 0·70 for specificity) using a non-inferiority limit of 0·05	For both software's, accuracy was not inferior to WHO-recommended minimum values (qXRv2 sensitivity 0·93 [95% CI 0·89–0·95], non-inferiority $P=0.0002$; CAD4TBv6 sensitivity 0·93 [0·90–0·96], $P<0.0001$; qXRv2 specificity 0·75 [0·73–0·77], $P<0.0001$; CAD4TBv6 specificity 0·69 [0·67–0·71], $P=0.0003$)

Table 1 (continued)

Study no	Author	Journal	Cohort size	Source of data	AI algorithm type	Performance
5	Lakhani P et al. [27]	Radiology (2017)	Four deidentified HIPAA-compliant data sets were used in this study that was exempted from review by the institutional review board, which consisted of 1,007 posteroanterior CXRs	Four data sets. This includes two publicly available data sets maintained by the National Institutes of Health, which are from Montgomery County, Maryland, and Shenzhen, China. The other two data sets are from Thomas Jefferson University Hospital, Philadelphia, and the Belarus Tuberculosis Portal	Two different DCNNs, Alex Net and Google Net, were used to classify the images as having manifestations of pulmonary TB or as healthy	Deep learning with DCNNs can accurately classify TB at chest radiography with an AUC of 0.99
6	Lee JH et al [28]	European Radiology (2021)	20,135 radiographs in 19,686 individuals	Armed forces hospital, Seoul, South Korea	Deep learning–based automated detection algorithms	For the radiologically identifiable relevant abnormality, deep learning -based automated detection algorithms showed an AUC value of 0.967 (95% CI 0.938–0.996) with sensitivities of 0.821 and 0.679, specificities of 0.960 and 0.997, PPVs of 0.028 and 0.257, and NPVs of both 0.999 at high sensitivity and high specificity thresholds, respectively
7	Lee S et al [29]	Radiology (2021)	6,654 pre- and post-treatment radiographs from 3,327 patients with pulmonary tuberculosis and 3,182 normal radiographs from as many patients	Six Korean hospitals (hospitals A–F)	Efficient Net, which was adopted as a base feature extractor. The network was built and trained by using open-source software (TensorFlow, version 1.11.0; Keras, version 2.2.4)	In two test sets that included radiographs depicting active and healed tuberculosis (test set 1, $n=148$; test set 2 subset, $n=200$), a deep learning model ROCs, 0.83 and 0.84, respectively) differentiated active from healed tuberculosis on radiographs, with comparable performance to that of expert readers (AUCs, 0.69–0.80 [$P=0.001$ to $P=0.23$] and 0.71–0.80 [$P=0.001$ to $P=0.08$])
8	Ma L et al [30]	Journal of Xray Science and Technology (2020)	A CT image data set of 846 patients was retrospectively collected	Hospital of Hebei University of China	A U-Net deep learning algorithm was applied for automatic detection and segmentation of Active TB lesions	For an independent test, the AI tool yields an AUC value of 0.980. Accuracy, sensitivity, specificity, PPV and NPV are 0.968, 0.971, 0.971, and 0.964, respectively, which shows that the AI tool performs well detecting active TB and differential diagnosis of non-active TB

Table 1 (continued)

Study no	Author	Journal	Cohort size	Source of data	AI algorithm type	Performance
9	Nabulsi Z et al. [31]	Scientific Reports 2021	Data set of 248,445 patients	Five clusters of hospitals from five cities in India and six international datasets from India, China, and the United States	Developed a DLS that classifies CXRs as normal or abnormal	The ROCs were 0.87 (95% CI 0.87–0.88) in DS-1 and 0.94 (95% CI 0.93–0.96) in CXR-14. In a simulated workflow where the AI system prioritised abnormal cases, the turnaround time for abnormal cases reduced by 7–28% With the high-sensitivity operating point, the DLS predicted 29.9% of Dataset DS-1 and 24.0% of CXR-14 as normal, with negative predictive values (NPVs) of 0.98 and 0.85, respectively. With the high-specificity operating point, the DLS predicted 22.2% of Dataset DS-1 and 11.7% of CXR-14 as abnormal, with positive predictive values (PPVs) of 0.68 and 0.99, respectively
10	Nafisah SI, et al. [32][a]	Neural Computing and Applications (2022)	1,104 total cases in the data set	Three data sets used: Montgomery Shenzhen and Belarus. Dataset included children's CXRs	Used advanced deep learning models. Authors use different CNN models in experiments and compare their classification performance using three publicly available CXR data sets	EfficientNetB3, one of the CNN models, achieves the highest accuracy of 99.1%, with an ROC of 99.9%, and an average accuracy of 98.7%. Experiment results confirm that using segmented lung CXR images produces better performance than using raw lung CXR images
11	Nijiati M et al. [33]	Frontiers in Molecular Biosciences (2022)	Altogether, 5,000 patients with TB and 4,628 patients without TB were included in the study, totalling 9,628 CXRs in patients > 15 years old that were analysed	Kashgar, Xinjiang, China	Three different DCNN algorithms, including ResNet, VGG and AlexNet, were trained to classify the chest radiographs as images of pulmonary TB or without TB	Reaching an accuracy of 96.73% and marking the precise TB regions on the radiographs, ResNet algorithm-based AI outperformed the rest of the models and showed excellent diagnostic ability in different clinical subgroups in the stratification analysis
12	Nijiati M et al [34]	Journal of Xray Science and Technology (2021)	Total of 10,002 cases One training data set including 2,627 TB-positive cases and 7,375 TB-negative cases and one testing data set containing 276 TB-positive cases and 619 TB-negative cases	Cases are collected from a public data set (CHNCXR and MCUCXR) and several Chinese hospitals located in Shenzhen and Guangzhou, China	A segmentation model named TB-UNet 24 was built to detect diseased regions, which uses ResNeXt as the encoder of U-Net	AI system yields TB detection accuracy of 85%, which is much higher than the detection accuracy of radiologists (62%) without AI assistance. In addition, with AI assistance, the TB diagnostic sensitivity of local radiologists is improved by 11.8%

Table 1 (continued)

Study no	Author	Journal	Source of data	Cohort size	AI algorithm type	Performance
13	Qin ZZ et al. [35]	Lancet Digital Health (2021)	Three TB screening centres in Dhaka, Bangladesh	CXRs from 23,954 individuals were included in the analysis	Five commercial AI algorithms were evaluated: CAD4TB (version 7), Infer Read DR (version 2), Lunit INSIGHT CXR (version 4.9.0), JF CXR-1 (version 2), and qXR (version 3). Authors compared the performance of the AI algorithms with each other, with the radiologists and with the WHO's Target Product Profile (TPP) of triage tests (≥ 90% sensitivity and ≥ 70% specificity	All five AI algorithms significantly outperformed the radiologists. The areas under the receiver operating characteristic curve (ROCs) were 90.81% (95% CI 90.33–91.29) for qXR, 90.34% (89.81–90.87) for CAD4TB, 88.61% (88.03–89.20) for Lunit INSIGHT CXR, 84.90% (84.27–85.54) for Infer Read DR, and 84.89% (84.26–85.53) for JF CXR-1. Only qXR (74.3% specificity [95% CI 73.3–74.9]) and CAD4TB (72.9% specificity [72.3–73.5]) met the TPP at 90% sensitivity. All five AI algorithms reduced the number of Xpert tests required by 50% while maintaining a sensitivity above 90%
14	Qin ZZ et al [36]	Scientific Reports 2019	Outpatients in Nepal and Cameroon B.P. Koirala Institute of Health Sciences (BPKIHS) in Eastern Nepal and Tuberculosis Reference Laboratory Bamenda and the Bamenda Regional Hospital in Cameroon	1,196 cases	Evaluation of three DL systems (CAD4TB, Lunit INSIGHT and qXR) for detecting TB-associated abnormalities in CXRs	The AUC of the three systems was similar: Lunit (0.94, 95% CI: 0.93–0.96), qXR (0.94, 95% CI: 0.92–0.97) and CAD4TB (0.92, 95% CI: 0.90–0.95). When matching the sensitivity of the radiologists, the specificities of the DL systems were significantly higher except for one. Using DL systems to read CXRs could reduce the number of Xpert Mycobacterium tuberculosis/ Resistance to rifampacin (MTB/RIF) tests needed by 66% while maintaining sensitivity at 95% or better
15	Tang Y-X et al. [37][a]	NPJ Digital Media (2020)	NIH Indiana and Guangzhou	Used three different databases NIH Chest X-ray 14 data set: 112,120 frontal-view CXRs Indiana University: 3,813 de-identified frontal CXRs Guangzhou Women and Children's Medical Center in China: 5,856 pediatric CXRs	DCNNs: AlexNet, VGG, GoogLeNet, ResNet and DenseNet were trained and validated on the data sets	A DCNN-based model achieved an AUC of 0.9824 ± 0.0043 (with an accuracy of 94.64 ± 0.45%, a sensitivity of 96.50 ± 0.36% and a specificity of 92.86 ± 0.48%) for normal versus abnormal CXR classification. The CNN model obtained an AUC of 0.9804 ± 0.0032 (with an accuracy of 94.71 ± 0.32%, a sensitivity of 92.20 ± 0.34% and a specificity of 96.34 ± 0.31%) for normal versus lung opacity classification. The CNN model pretrained on cohorts of adults and fine-tuned on pediatric patients achieved an AUC of 0.9851 ± 0.0046 for normal versus pneumonia classification

Table 1 (continued)

Study no	Author	Journal	Cohort size	Source of data	AI algorithm type	Performance
16	Tavaziva G et al. [38]	Clinical Infectious Diseases (2022)	Authors included CXRs and individual patient data of 3,727/3,967 participants from 4/7 eligible studie	Eligible studies were identified through published systematic reviews	Authors analysed each CXR with 3 commercially available deep learning-based CAD programs: CAD4TB version 6 (Delft, Netherlands), Lunit INSIGHT version 3.1.0.0. (Lunit, South Korea) and qXR version 2 (qure.ai, India)	Despite using the same threshold score for classifying CXR in every study, sensitivity and specificity varied from study to study. Accuracy was similar to that of human readers. For computer-aided detection CXR analysis to be implemented as a high-sensitivity tuberculosis rule-out test, users will need threshold scores identified from their own patient populations
17	Wang L et al. [39]	European Journal of Nuclear Medicine and Molecular Imaging (2021)	A total of 1,105 patients were included in this study, comprising 301 patients with nontuberculous mycobacterium lung disease (NTM-LD) and 804 patients with *Mycobacterium tuberculosis* lung disease (MTB-LD) using computed tomography (CT) images	Tianjin Haihe Hospital (Tianjin University, China) and Xi'an Chest Hospital (Shanxi, China)	Three data sets for training, validating and testing of a deep learning model (3D-ResNet) at a ratio of 8:1:1	The AUCs of their model on training, validating and testing data sets were 0.90, 0.88, and 0.86, respectively, while the AUC on the external test set was 0.78. Additionally, the performance of the model was higher than that of the radiologist, and without manual labelling, the model automatically identified lung areas with abnormalities on CT > 1,000 times more effectively than the radiologists
18	Wong A et al [40]	Frontiers in Artificial intelligence (2022)	Data used in this study comprises 6,939 CXR images	Multinational patient cohort consists of patient cohorts curated by the Department of Health and Human Services in Montgomery County, Maryland, Shenzhen 3 People's Hospital in China, the National Institute of Allergy and Infectious Diseases in the U.S., as well as the Radiological Society of North America	TB-Net, a self-attention DCNN tailored for TB case screening	TB-Net is able to achieve accuracy/sensitivity/specificity of 99.86%/100%/99.71%, respectively. The proposed TB-Net not only achieves high tuberculosis case detection performance in terms of sensitivity and specificity, but also leverages clinically relevant critical factors in its decision-making process
19	Yan C et al [41]	European Radiology (2022)	Eight hundred ninety-two chest CT scans from pathogen-confirmed TB patients were retrospectively included	Nanfang Hospital of Southern Medical University	A deep learning–based cascading framework was connected to create a processing pipeline	The overall classification accuracy of six pulmonary critical imaging findings indicative of TB of the independent data sets was 81.08–91.05%. A moderate to strong correlation was demonstrated between the AI model–quantified TB score and the radiologist-estimated CT score

Table 1 (continued)

Study no	Author	Journal	Cohort size	Source of data	AI algorithm type	Performance
20	Yi PH et al. [42]	Clinical Imaging (2022)	Authors collected 10,951 CXRs from the NIH Chest X-ray 14 data set	NIH Chest X-ray14 data set	ResNet 50 Deep convolutional neural network (DCNN) training and validation done	The best-performing deep convolutional neural network (DCNN) had an area under the curve (AUC) of 0.88, which was trained on 10,951 images using the radiologist-annotated sets. DCNNs trained on chest X-rays (CXRs) labelled by a radiologist consistently outperformed those trained on the same CXRs labelled by Natural Language Processing (NLP), highlighting the benefit of radiologists in determining ground truth for machine-learning data set curation
21	Zhou W et al [43]	Quantitative Imaging in Medicine and Surgery (2022)	An internal data set with 7,025 images was used to develop the AI system, after which a 6-year dynamic cohort accumulation data set with 358,169 images was used to conduct an independent external validation of the trained AI system	Images were from five sources in the U.S. and China	A transfer learning approach was applied to train a pretrained ResNet model to build the final DCNN to identify and locate TB. A U-Net-based algorithm was trained to automatically segment the lung area	The AI system achieved an AUC of 0.99 and an accuracy of 0.948 on the internal data set, and an AUC of 0.95 and an accuracy of 0.931 on the external data set when it was used to detect TB from normal images

A literature search for publications including studies on artificial intelligence (AI) in tuberculosis (TB) imaging using search items related to AI, TB and thoracic imaging yielded 110 results of which 21 were published in the last 5 years

[a]The algorithms used in these 2 studies were validated on data sets which included population age groups < 15 years old

Fig. 1 A simplified view of how artificial neural networks function. Although 2 layers are shown, hundreds or thousands of layers are used
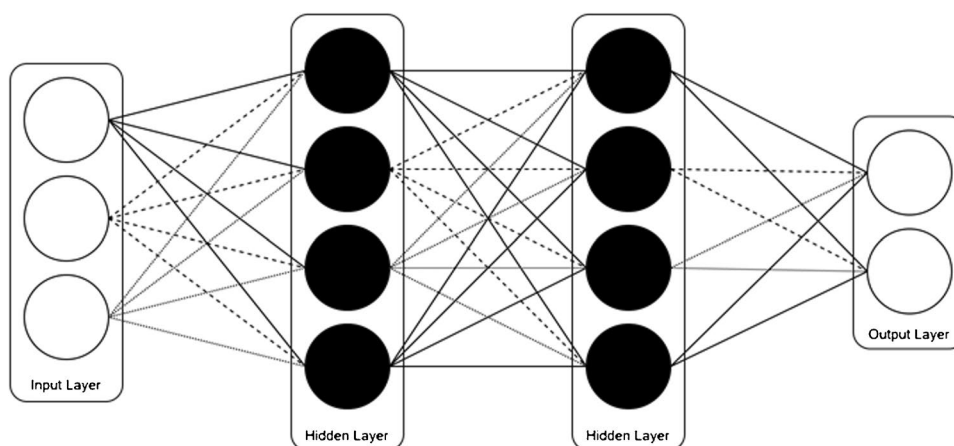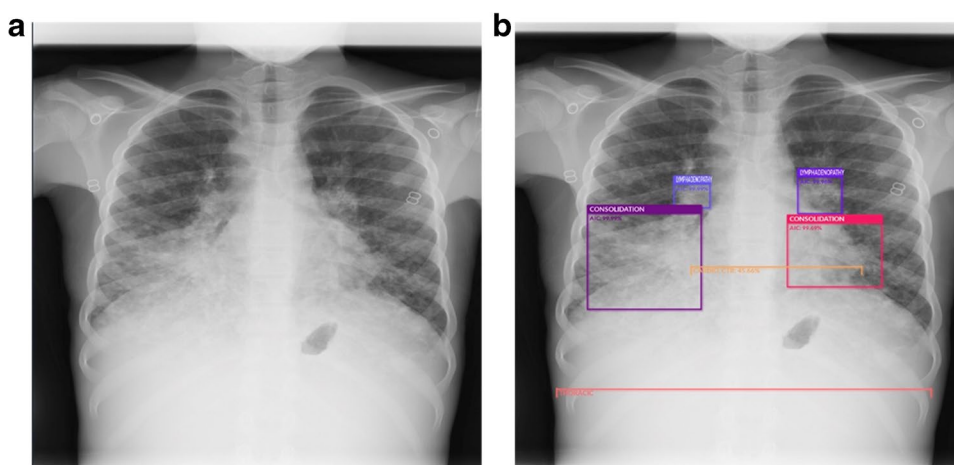

Fig. 2 a, b AP chest radiograph in a 2-year-old girl with pulmonary tuberculosis. **a** Before use of the artificial intelligence (AI) tool. **b** The AI tool identifies lymphadenopathy (*two smallest bounding boxes*) and consolidation (*two largest bounding boxes*) and measures the cardiothoracic ratio (*horizontal lines*)

AI as a double reader may replace the staffing needed for double radiologist interpretation and consensus and may increase the value radiologists are able to provide to their patients.

Computed tomography (CT) scans are not always accessible in high TB burden countries, but in countries where health care centres are better resourced with imaging equipment, these scans can aid radiologists in the diagnosis of suspected TB cases when chest radiographs are either inconclusive or in cases where there is clinical and imaging progression.

CT scans can provide additional information to guide diagnosis, monitor imaging changes and evaluate the severity of pulmonary TB.

Yan et al. [41] performed a retrospective, multicohort, diagnostic study where they developed an AI cascading model for fully automated diagnosis and triage of pulmonary TB based on the CT scans of 526 participants. Overall accuracy of 6 pulmonary critical imaging findings indicative of TB in the independent datasets was 81.1–91.1%. Spearman correlation analysis was used to assess the correlation between the radiologist-estimated CT score and the TB score determined by the algorithm, which was shown to be moderate to good (r = 0.453–0.761) [41].

Challenges and future directions of artificial intelligence in tuberculosis imaging

The challenges faced by using AI in TB imaging, particularly for the paediatric population, have many similarities to those of using AI in other diseases, namely a lack of diverse training data, lack of external validation of AI models, the possibility of bias, questionable reference standards (such as human reader opinion on radiographic diagnosis rather than correlation with microbiological reference) and real-world implementation data [50]. Future directions of AI for TB imaging, especially in children, will therefore need to focus on these aspects.

Training data for AI in TB imaging should include imaging from multiple centres and using multiple vendors' equipment and modalities from different manufacturers. This may also involve acquiring digital photographs of chest radiographs (either from radiographic film on a light box

or computer screens). This is important given that some AI tools are being developed as smartphone applications, bypassing any need for local information technology expertise in rural centres where picture archiving and communication systems (PACS) software may not be available, or where IT information technology specialists are lacking.

When considering the adaptability of AI for TB diagnosis to other settings and types of chest radiograph photographs, there are challenges. For example, Becker et al. [51] reported an AUC of 0.82 for their deep learning model in the localisation of pathological areas on chest radiograph photographs, although specific diagnostic labels were still challenging in their small data set of 138 adults. Adapting a deep learning tool to identify photographs of chest radiographs in children for TB would be a useful area for future research and can build on existing tools rather than starting from scratch. Researchers are attempting computational methods for recalibration of existing AI models for chest radiographic interpretation to allow for improved accuracy from smartphone acquired photographs [52].

Bias has been highlighted in some AI articles as contributing to poor performance [53]. This is of particular concern when AI models are applied to a wider population than they have been adequately trained to make diagnoses on (e.g., using a model trained on adults for use in children). As an example, Harris et al. [50] conducted a systematic review of diagnostic accuracy of AI-based computer programs to analyse chest radiographs for pulmonary tuberculosis. They found there was a risk of bias and higher mean AUC in development studies that used chest radiography databases compared to clinical studies [50].

Training without bias is critically important and requires that the initial data are free of any external influences that might cloud the information, such as demographic information or additional medical information that may suggest a certain diagnosis. Elimination/reduction of bias also requires that the training dataset is sufficiently large to encompass a variety of patients.

Although this article has outlined AI research for paediatric pneumonia [54, 55], only one commercial product (CAD4TB v6) [56] is licensed to interpret chest radiographs for use in children older than 4 years of age (others are for adults and adolescents). The future of this field is likely to see more emerging commercial solutions and products designed for the paediatric population.

In the future, randomised controlled trials implementing AI models across several centres and comparing outcomes with those that do not employ the AI models will be vital in understanding the patient benefit and efficiencies that AI might bring. This is not only important to determine cost-effectiveness across differing sites but also to understand whether there are certain conditions that would be more optimally suited for AI triage. There is also consideration to

be made on whether diagnostic thresholds require amending between sites with respect to different TB and HIV prevalence levels and patient ages. This would be helpful for raising funds to provide this standard of care and to enable greater access to emerging digital technologies. Eventually, as more robust AI models become developed across modalities other than radiography, it may be possible to use AI for point of care ultrasound in TB, for example [57].

Recent publications have reported key findings on chest ultrasound in TB [58–60], which could be highlighted by AI thus providing a point of care solution and diagnosis for less experienced users who may not know the significance of such appearances.

Artificial intelligence software could even become combined with a portable paediatric hands-free ultrasound device that remotely provides point of care diagnoses for other cardiothoracic conditions [61]. Such a device has been recognised as an innovative novel health technology for low resource settings by WHO [62] and a prototype is being developed.

The epidemiology of TB has been adversely impacted by HIV. Children with HIV often present without the classic imaging features of TB [63]. Understanding these features and training an algorithm to detect the pertinent signs and triage these patients can provide a solution in identifying the many patients who are missed in screening programmes.

Similarly, through AI analysis of historical data and the pathologies and signs on the chest radiograph and other imaging modalities, immune reconstitution inflammatory syndrome (IRIS), which is defined as clinical deterioration after initiation of antiretroviral therapy (radiologically manifested as worsening or new hilar or mediastinal lymphadenopathy with or without tracheobronchial compression, worsening or new reticular/nodular infiltrates, worsening or new air space consolidation or pleural effusion), can be identified and highlighted earlier.

Finally, without prospective, real-world testing of the AI models, we lack a true understanding of how these tools can aid patient outcomes and impact clinical workflow. This is not unique to TB imaging and relates to other avenues of AI research. One study by Philipsen et al. [64] performed a simulation test to determine the cost-effectiveness of using software (CAD4TB v3.07, Nijmegen, Netherlands) to pre-screen chest radiographs in adults with suspected TB. The software assigned each chest radiograph a score between 0 (normal) and 100 (highly abnormal) to determine which patients should undergo further, more expensive molecular testing for TB. They found that by adopting an optimal threshold score of 85, they would only require 40% of patients to undergo the molecular testing and overall costs per screened subject and costs per notified TB case were reduced by more than 50%. How generalisable this tool and such a workflow may be for paediatric cases is yet to be determined [45, 65].

Conclusion

Medical applications of AI are becoming increasingly important. The "End TB Strategy" was adopted by WHO with the aim of eliminating TB by reducing 90% of mortality and 80% TB incidence by 2030. This emphasises the need to ensure early and correct diagnosis and adequate treatment for people with TB.

An AI tool to assist with the diagnosis, triage and treatment of large numbers of patients with a highly transmissible infection will be transformative and will help us solve some of the challenges in combatting TB. However, more real-world AI applications and multicentre clinical studies are required in paediatric imaging to ensure that the strategy and goals set out to end TB in children are achievable.

Declarations

References

1. Global tuberculosis report 2019. Geneva: World Health Organization; 2019. Licence: CCBY-NC-SA3.0IGO. https://tbsouthafrica.org.za/resources/who-global-tuberculosis-report-2019 /. Accessed 1 September 2022
2. Global tuberculosis report 2020. Geneva, World Health Organization. https://www.who.int/publications/i/item/9789240013131 /. Accessed 1 September 2022
3. Wang XW, Pappoe F, Huang Y et al (2015) Xpert MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in children: a meta-analysis. Clin Lab 61:1775–1785
4. Pillay T, Andronikou S, Zar HJ (2020) Chest imaging in paediatric pulmonary TB. Paediatr Respir Rev 36:65–72
5. WHO consolidated guidelines on tuberculosis: Module 2: screening – systematic screening for tuberculosis disease [Internet] (2021) Geneva: World Health Organization https://www.who.int/publications/i/item/9789240022676 /.Accessed 1 September 2022
6. A global mapping of pediatric radiologists and pediatric radiology training: Source-WFPI. https://www.wfpiweb.org/Portals/7/Workrooms/Mapping-Framework-WFPI-EXCOM-circulation.pdf /. Accessed 1 September 2022
7. AI products for tuberculosis healthcare (2022) AI4HLTH https://www.ai4hlth.org/. Accessed 1 Sept 2022
8. Kim Y, Park JY, Hwang EJ et al (2021) Applications of artificial intelligence in the thorax: a narrative review focusing on thoracic radiology. J Thorac Dis 13:6943–6962
9. Cruz AT, Starke JR (2010) Pediatric tuberculosis. Pediatr Rev 31(13–25):126
10. Palmer M, Gunasekera KS, Van der Zalm MM et al (2022) The diagnostic accuracy of chest radiographic features for pediatric intrathoracic tuberculosis. Clin Infect Dis 75:1014–1021
11. Kaguthi G, Nduba V, Nyokabi J et al (2014) Chest radiographs for pediatric TB diagnosis: interrater agreement and utility. Interdisc Perspect Infect Dis 2014:291841
12. Frigati L, Maskew M, Workman L et al (2015) Clinical predictors of culture-confirmed pulmonary tuberculosis in children in a high tuberculosis and HIV prevalence area. Pediatr Infect Dis J 34:e206-210
13. Graham SM, Ahmed T, Amanullah F et al (2012) Evaluation of tuberculosis diagnostics in children: 1. Proposed clinical case definitions for classification of intrathoracic tuberculosis disease. Consensus from an expert panel. J Infect Dis 205 Suppl (Suppl 2):S199–208
14. Swingler GH, du Toit G, Andronikou S et al (2005) Diagnostic accuracy of chest radiography in detecting mediastinal lymphadenopathy in suspected pulmonary tuberculosis. Arch Dis Child 90:1153–1156
15. Triasih R, Robertson C, de Campo J et al (2015) An evaluation of chest X-ray in the context of community-based screening of child tuberculosis contacts. Int J Tuberc Lung Dis 12:1428–1434
16. Berteloot L, Marcy O, Nguyen B et al (2018) Value of chest X-ray in TB diagnosis in HIV-infected children living in resource-limited countries: the ANRS 12229-PAANTHER 01 study. Int J Tuberc Lung Dis 22:844–850
17. WHO consolidated guidelines on tuberculosis (Module 5: management of tuberculosis in children and adolescents) (2022) https://www.who.int/publications/i/item/9789240046764/. Accessed 1 Sept 2022
18. Andronikou S, Miranda-Schaeubinger M, Goussard P et al (2022) Changes in the role of chest radiographs for diagnosing and managing children with tuberculosis: the 2022 World Health Organization consolidated guidelines on tuberculosis. Pediatr Radiol https://doi.org/10.1007/s00247-022-05544-y
19. Sodhi KS, Bhalla AS, Mahomed N et al (2017) Imaging of thoracic tuberculosis in children: current and future directions. Pediatr Radiol 47:1260–1268
20. Fawole OA, Kelly MS, Steenhoff AP et al (2020) Interpretation of pediatric chest radiographs by non-radiologist clinicians in Botswana using World Health Organization criteria for endpoint pneumonia. Pediatr Radiol 50:913–922
21. Huang C-C, Tan Q, Becerra MC et al (2022) The contribution of chest radiography to the clinical management of children exposed to tuberculosis. Am J Respir Crit Care Med 206:892–900
22. Fritschi N, Wind A, Hammer J, Ritz N (2022) Subclinical tuberculosis in children: diagnostic strategies for identification reported in a 6-year national prospective surveillance study. Clin Infect Dis 74:678–684
23. Codlin AJ, Dao TP, Vo LNQ et al (2021) Independent evaluation of 12 artificial intelligence solutions for the detection of tuberculosis. Sci Rep 11:23895
24. Feng B, Chen X, Chen Y et al (2020) Solitary solid pulmonary nodules: a CT-based deep learning nomogram helps differentiate tuberculosis granulomas from lung adenocarcinomas. Eur Radiol 30:6497–6507
25. Huang T, Zheng X, He L, Chen Z (2021) Diagnostic value of deep learning-based CT feature for severe pulmonary infection. J Healthc Eng 2021:5359084
26. Khan FA, Majidulla A, Tavaziva G et al (2020) 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. Lancet Digit Health 2:e573–e581
27. Lakhani P, Sundaram B (2017) Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology 284:574–582

28. Lee JH, Park S, Hwang EJ et al (2021) Deep learning-based automated detection algorithm for active pulmonary tuberculosis on chest radiographs: diagnostic performance in systematic screening of asymptomatic individuals. Eur Radiol 31:1069–1080

29. Lee S, Yim J-J, Kwak N et al (2021) Deep learning to determine the activity of pulmonary tuberculosis on chest radiographs. Radiology 301:435–442

30. Ma L, Wang Y, Guo L et al (2020) Developing and verifying automatic detection of active pulmonary tuberculosis from multi-slice spiral CT images based on deep learning. J Xray Sci Technol 28:939–951

31. Nabulsi Z, Sellergren A, Jamshy S et al (2021) Deep learning for distinguishing normal versus abnormal chest radiographs and generalization to two unseen diseases tuberculosis and COVID-19. Sci Rep 11:15523

32. Nafisah SI, Muhammad G (2022) Tuberculosis detection in chest radiograph using convolutional neural network architecture and explainable artificial intelligence. Neural Comput Appl Apr 19:1–21

33. Nijiati M, Ma J, Hu C et al (2022) artificial intelligence assisting the early detection of active pulmonary tuberculosis from chest x-rays: a population-based study. Front Mol Biosci 9:874475

34. Nijiati M, Zhang Z, Abulizi A et al (2021) Deep learning assistance for tuberculosis diagnosis with chest radiography in low-resource settings. J Xray Sci Technol 29:785–796

35. Qin ZZ, Ahmed S, Sarker MS et al (2021) Tuberculosis detection from chest x-rays for triaging in a high tuberculosis-burden setting: an evaluation of five artificial intelligence algorithms. Lancet Digit Health 3:e543–e554

36. Qin ZZ, Sander MS, Rai B et al (2019) 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 9:15000

37. Tang Y-X, Tang Y-B, Peng Y et al (2020) Automated abnormality classification of chest radiographs using deep convolutional neural networks. NPJ Digit Med 3:70

38. Tavaziva G, Harris M, Abidi SK et al (2022) Chest x-ray analysis with deep learning-based software as a triage test for pulmonary tuberculosis: an individual patient data meta-analysis of diagnostic accuracy. Clin Infect Dis 74:1390–1400

39. Wang L, Ding W, Mo Y et al (2021) Distinguishing nontuberculous mycobacteria from Mycobacterium tuberculosis lung disease from CT images using a deep learning framework. Eur J Nucl Med Mol Imaging 48:4293–4306

40. Wong A, Lee JRH, Rahmat-Khah H et al (2022) TB-Net: A tailored, self-attention deep convolutional neural network design for detection of tuberculosis cases from chest x-ray images. Front Artif Intell 5:827299

41. Yan C, Wang L, Lin J et al (2022) A fully automatic artificial intelligence-based CT image analysis system for accurate detection, diagnosis, and quantitative severity evaluation of pulmonary tuberculosis. Eur Radiol 32:2188–2199

42. Yi PH, Kim TK, Lin CT (2022) Comparison of radiologist versus natural language processing-based image annotations for deep learning system for tuberculosis screening on chest radiographs. Clin Imaging 87:34–37

43. Zhou W, Cheng G, Zhang Z et al (2022) Deep learning-based pulmonary tuberculosis automated detection on chest radiography: large-scale independent testing. Quant Imaging Med Surg 12:2344–2355

44. Dunnmon JA, Yi D, Langlotz CP et al (2019) Assessment of convolutional neural networks for automated classification of chest radiographs. Radiology 290:537–544

45. He J, Baxter SL, Xu J et al (2019) The practical implementation of artificial intelligence technologies in medicine. Nat Med 25:30–36

46. Lopez-Garnier S, Sheen P, Zimic M (2019) Automatic diagnostics of tuberculosis using convolutional neural networks analysis of MODS digital images. PLoS ONE 14:e0212094

47. Schalekamp S, Klein WM, van Leeuwen KG (2022) Current and emerging artificial intelligence applications in chest imaging: a pediatric perspective. Pediatr Radiol 52:2120–2130

48. Mouton A, Pitcher RD, Douglas TS (2010) Computer-aided detection of pulmonary pathology in pediatric chest radiographs. Med Image Comput Comput Assist Interv 13:619–625

49. Hwang EJ, Park S, Jin K-N et al (2019) Development and validation of a deep learning-based automatic detection algorithm for active pulmonary tuberculosis on chest radiographs. Clin Infect Dis 69:739–747

50. Harris M, Qi A, Jeagal L et al (2019) A systematic review of the diagnostic accuracy of artificial intelligence-based computer programs to analyze chest x-rays for pulmonary tuberculosis. PLoS ONE 14:e0221339

51. Becker AS, Blüthgen C, van Phi VD et al (2018) Detection of tuberculosis patterns in digital photographs of chest X-ray images using Deep Learning: feasibility study. Int J Tuberc Lung Dis 22:328–335

52. Kuo P-C, Tsai CC, López DM et al (2021) Recalibration of deep learning models for abnormality detection in smartphone-captured chest radiograph. NPJ Digit Med 4:25

53. Goisauf M, Cano Abadía M (2022) Ethics of AI in radiology: A review of ethical and societal implications. Front Big Data 5:850383

54. Mahomed N, van Ginneken B, Philipsen RHHM et al (2020) Computer-aided diagnosis for World Health Organization-defined chest radiograph primary-endpoint pneumonia in children. Pediatr Radiol 50:482–491

55. Chen K-C, Yu H-R, Chen W-S et al (2020) Diagnosis of common pulmonary diseases in children by X-ray images and deep learning. Sci Rep 10:17374

56. CAD4TB (Computer-Aided Detection for Tuberculosis) software (2022) DELFT Imaging Solutions. https://www.delft.care/cad4tb/ Accessed 1 Sept 2022

57. Correa M, Zimic M, Barrientos F et al (2018) Automatic classification of pediatric pneumonia based on lung ultrasound pattern recognition. PLoS ONE 13:e0206410

58. Belard S, Heuvelings CC, Banderker E et al (2018) Utility of Point-of-care Ultrasound in Children with Pulmonary Tuberculosis. Pediatr Infect Dis J 37:637–642

59. Heuvelings CC, Belard S, Andronikou S et al (2019) Chest ultrasound findings in children with suspected pulmonary tuberculosis. Pediatr Pulmonol 54:463–470

60. Heuvelings CC, Belard S, Andronikou S et al (2019) Chest ultrasound compared to chest X-ray for pediatric pulmonary tuberculosis. Pediatr Pulmonol 54:1914–1920

61. Bloom Standard Automated Ultrasound Device (2022) https://www.bloomstandard.com/ Accessed 1 Sept 2022

62. WHO Compendium of Innovate Health Technologies for Low-Resource Settings 2021. https://www.who.int/publications/i/item/9789240032507 Accessed 1 Sept 2022

63. Naidoo J, Mahomed N, Moodley H (2017) A systemic review of tuberculosis with HIV coinfection in children. Pediatr Radiol 47:1269–1276

64. Philipsen RHHM, Sánchez CI, Maduskar P et al (2015) Automated chest-radiography as a triage for Xpert testing in resource-constrained settings: a prospective study of diagnostic accuracy and costs. Sci Rep 5:12215

65. Murphy K, Habib SS, Zaidi SMA et al (2020) Computer aided detection of tuberculosis on chest radiographs: An evaluation of the CAD4TB v6 system. Sci Rep 10:5492