Chest X-ray-Based Telemedicine Platform for Pediatric Tuberculosis Diagnosis in Low Resource Settings: Development and validation

Juan J Gómez-Valverde, Ramón Sánchez-Jacob, José Luis Ribó, H Simon Schaaf, Lara García Delgado, Alicia Hernanz-Lobo, Daniel Capellán-Martín, Ángel Lancharro, Orvalho Augusto, Alberto L García-Basteiro, Begoña Santiago-García, Elisa López-Varela, María J Ledesma-Carbayo

Table of Contents

Chest X-ray-Based Telemedicine Platform for Pediatric Tuberculosis Diagnosis in Low Resource Settings: Development and validation

Juan J Gómez-Valverde[1, 2] PhD; Ramón Sánchez-Jacob[3] MD; José Luis Ribó[4] MD; H Simon Schaaf[5] MD, PhD; Lara García Delgado[1] MSc; Alicia Hernanz-Lobo[6, 7, 8, 9, 10] MD; Daniel Capellán-Martín[1] MSc; Ángel Lancharro[11, 12] MD; Orvalho Augusto[13, 14] MD, MPH; Alberto L García-Basteiro[8, 13, 15] PhD; Begoña Santiago-García[6, 7, 8] MD, PhD; Elisa López-Varela[13, 15] MD, PhD; María J Ledesma-Carbayo[1, 2] PhD

[1]Biomedical Image Technologies ETSI Telecomunicación Universidad Politécnica de Madrid Madrid ES
[2]Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) Madrid ES
[3]Department of Radiology Children's National Hospital & George Washington University School of Medicine Washington US
[4]Hospital Universitari General de Catalunya Barcelona ES
[5]Desmond Tutu TB Centre, Department of Paediatrics and Child Health Faculty of Medicine and Health Sciences Stellenbosch University Cape Town ZA
[6]Pediatric Infectious Diseases Department Gregorio Marañón University Hospital Madrid ES
[7]Gregorio Marañón Research Health Institute (IiSGM) Madrid ES
[8]Centro de Investigación Biomédica en Red de Enfermedades Infecciosas (CIBERINFEC) Instituto de Salud Carlos III Madrid ES
[9]RITIP Translational Research Network in Pediatric Infectious Diseases Madrid ES
[10]Centro de Salud Monóvar Madrid ES
[11]Radiología Pediátrica Hospital Materno Infantil Gregorio Marañón Madrid ES
[12]Radiología Pediátrica. HM Hospitales Madrid ES
[13]Centro de Investigação em Saúde de Manhiça (CISM) Maputo MZ
[14]Department of Global Health University of Washington Seattle US
[15]ISGlobal, Hospital Clínic Universitat de Barcelona Barcelona ES

Corresponding Author:
Juan J Gómez-Valverde PhD
Biomedical Image Technologies
ETSI Telecomunicación
Universidad Politécnica de Madrid
Avenida Complutense 30
Madrid
ES

Abstract

Background: Tuberculosis (TB) remains a major cause of morbidity and death worldwide, with a significant impact on children, especially those under 5 years old. The complex diagnosis of pediatric TB, compounded by limited access to sensitive diagnostic tests, underscores the need for improved tools to enhance diagnosis and care in resource-limited settings.

Objective: The objective of this paper is to present a telemedicine web platform, BITScreen PTB, aimed at improving the evaluation of pulmonary tuberculosis (TB) in children based on digital chest X-ray (CXR) imaging and clinical information in resource-limited settings.

Methods: The platform was evaluated by three independent expert readers through a retrospective assessment of a dataset with 218 imaging exams of children under 3 years of age, selected from a previous study performed in Mozambique. Key aspects assessed include the usability through a standardized questionnaire, the time needed to complete the assessment through the platform, the performance of the readers to identify TB cases based on the CXR, the association between the TB features identified in the CXRs and the initial diagnostic classification and the inter-reader agreement of the global assessment and the radiological findings.

Results: The platform usability and user satisfaction questionnaire evaluation received an average rating of 4.4 ± 0.59 out of 5. The average exam completion time ranged from 35 to 110 seconds. In addition, the study on CXR showed low sensitivity (16.3% to 28.2%) but high specificity (91.1% to 98.2%) in the assessment of the consensus case definition of pediatric TB using the

platform. The CXR finding with a stronger association with the initial diagnostic classification was air space opacification (?21 >20.38, P<.001). The study found varying levels of inter-reader agreement, with moderate/substantial agreement for air space opacification (kappa from 0.54 to 0.67) and pleural effusion (kappa from 0.43 to 0.72).

Conclusions: Our findings support the promising role of telemedicine platforms like BITScreen PTB in enhancing pediatric TB diagnosis access, particularly in resource-limited settings. Additionally, these platforms could facilitate the multireader and systematic assessment of CXR in pediatric TB clinical studies.

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✔ **Please make my preprint PDF available to anyone at any time (recommended).**
　Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
　Only make the preprint title and abstract visible.
　No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
　Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
　Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original Paper

Chest X-ray-Based Telemedicine Platform for Pediatric Tuberculosis Diagnosis in Low Resource Settings: Development and validation

Abstract

Background:

Tuberculosis (TB) remains a major cause of morbidity and death worldwide, with a significant impact on children, especially those under 5 years old. The complex diagnosis of pediatric TB, compounded by limited access to more accurate diagnostic tests, underscores the need for improved tools to enhance diagnosis and care in resource-limited settings.

Objective:

The objective of this paper is to present a telemedicine web platform, BITScreen PTB (Biomedical Image Technologies Screen for Pediatric Tuberculosis), aimed at improving the evaluation of pulmonary tuberculosis (TB) in children based on digital chest X-ray (CXR) imaging and clinical information in resource-limited settings.

Methods:

The platform was evaluated by three independent expert readers through a retrospective assessment of a dataset with 218 imaging exams of children under 3 years of age, selected from a previous study performed in Mozambique. Key aspects assessed include the usability through a standardized questionnaire, the time needed to complete the assessment through the platform, the performance of the readers to identify TB cases based on the CXR, the association between the TB features identified in the CXRs and the initial diagnostic classification and the inter-reader agreement of the global assessment and the radiological findings.

Results:

The platform usability and user satisfaction questionnaire evaluation received an average rating of 4.4 ± 0.59 out of 5. The average exam completion time ranged from 35 to 110 seconds. In addition, the study on CXR showed low sensitivity (16.3% to 28.2%) but high specificity (91.1% to 98.2%) in the assessment of the consensus case definition of pediatric TB using the platform. The CXR finding with a stronger association with the initial diagnostic classification was air space opacification (X^2_1 >20.38, P<.001). The study found varying levels of inter-reader agreement, with moderate/substantial agreement for air space opacification (kappa from 0.54 to 0.67) and pleural effusion (kappa from 0.43 to 0.72).

Conclusions:

Our findings support the promising role of telemedicine platforms like BITScreen PTB in enhancing pediatric TB diagnosis access, particularly in resource-limited settings. Additionally, these platforms could facilitate the multireader and systematic assessment of CXR in pediatric TB clinical studies.

Keywords: Telemedicine; telehealth; pediatric tuberculosis; tuberculosis; screening; chest radiograph; usability; low-resource settings

Introduction

Tuberculosis (TB) is a communicable disease caused by the *Mycobacterium tuberculosis*. According to the World Health Organization (WHO) TB remains one of the major causes of death worldwide from a single infectious agent with more than 1.6 million TB deaths in 2021[1]. Most children who die from TB are never diagnosed or treated[2]. The risk of death is particularly high (44%) in

children younger than 5 years with untreated TB, while less than 1% of children died receiving recommended treatment [3] .

The diagnosis of TB in children is complex, especially in infants and young children where the risk of rapid disease progression and mortality is higher than in any other age group[4,5]. The paucibacillary nature of TB in this age group and the absence of highly sensitive point-of-care diagnostic tests to microbiologically confirm pediatric TB makes diagnosis challenging[4]. Chest X-ray (CXR) remains a valuable diagnostic tool for TB in children, particularly in cases where laboratory testing is not available, not feasible, or yields negative results. Most children with pulmonary TB (PTB) show radiographic changes indicative of TB. For children under 5 years old, anteroposterior (AP) and lateral views are recommended, while posteroanterior (PA) CXRs are preferred for older children and adolescents[6]. The lateral radiograph is especially helpful in children under 5 years old for optimal evaluation of hilar or mediastinal lymphadenopathy[7]. CXR findings in children with PTB may lack specificity[8] and CXR alone is insufficient to determine the appropriate treatment for the child. Instead, CXR can support the clinical diagnosis of PTB when TB is presumed, and microbiological testing is negative.

Screening tests using symptoms or CXR may be useful in children who are TB contacts or living with HIV[2]. In these patients, and according to Vonasek et al.[2], any abnormality identified on CXR seems to be the most accurate screening test for PTB in children, although it could be influenced by CXR quality and inter-reader variability. A recent study in a cohort of HIV-negative children, [9] most of them (92%) under 5 years old, proposes a treatment decision algorithm in low-resource countries where CXR is reserved to confirm the diagnosis of patients without enough clinical evidence to initiate treatment. WHO guidelines stress the need of more research in relation to integrated treatment decision algorithms[6]. This emphasizes the importance of fostering research to enhance and validate these tools for the pediatric context, enabling informed recommendations in this regard [9,14].

Assessing disease severity in children is crucial for determining their eligibility for the recommended 4-month treatment regimen for non-severe TB in children and adolescents aged 3 months to 16 years. CXR is a valuable tool for this purpose. Additionally, CXR can also aid in evaluating treatment response and identifying alternative diagnoses in children who do not respond to TB treatment, as stated in recent WHO guidelines[6].

The limitations in accessibility and sensitivity of available diagnostic tests for childhood TB are likely reasons for the gap between the estimated 1.17 million annual incident child TB cases from which less than half are diagnosed or reported to the WHO[6] , and the gap is even larger for children under 5 years. Moreover, the COVID-19 pandemic reduced access to TB diagnosis and treatment, and has disproportionately affected children and young adolescents, with a relevant decrease in notifications for younger children. To close these gaps, the End TB Strategy defined by WHO highlighted the relevance of enhanced digital health tools for more efficient delivery, monitoring and evaluation of TB patient diagnosis, treatment and care[10,11] . Telemedicine tools could play an important role enabling accessibility for diagnosis or treatment. Prior works have demonstrated that telemedicine could be useful to optimize care of multidrug-resistant TB in resource-limited settings[12] and that providing specialist expertise directly through telemedicine tools in low-resource settings not only enhanced patient management but also delivered additional educational value to the local physicians, thereby benefiting other patients as well[13].

In this paper we present a new telemedicine web platform (BITScreen PTB) for the evaluation of pediatric TB based on digital CXR images and clinical information. The aim of the platform is to

enable remote reading, optimizing and standardizing the clinical evaluation of pediatric TB studies in resource-limited settings where availability of expert readers may be challenging. The platform was functionally evaluated in a pilot study by three independent expert readers through a retrospective assessment of a dataset with 218 exams of children under 3 years of age, selected from a previous study performed in Mozambique[5,14]. Additionally, using the results of the evaluations through the platform in the pilot study we present new insights of the performance, the agreement among evaluators and the challenges of the assessment of pediatric TB via CXR images considering different radiological findings.

Methods

BITScreen Platform

BITScreen is a store-and-forward telemedicine platform created using a Model-View-Controller (MVC) design pattern over open-source frameworks and tools. MVC provides a modular and scalable structure for organizing and developing software applications, enabling efficient development, maintenance, and expansion of the application. In an MVC application the "View" displays the information to the end user, the Controller processes the user's interaction using the information stored and organized in the "Model". The main functional requirement of the platform is to enable asynchronous medical evaluation of pediatric TB studies based on the assessment of clinical data and CXR images along with the corresponding clinical symptoms if desired. The global requirements identified in the design of the system are listed below:

1. Multi-study. Capacity to perform multiple clinical projects simultaneously.
2. Multi-center. The system must allow the participation of multiple medical centers and admit many-to-many relationships between medical centers and projects/studies.
3. Multi-device. Web-based access to the views of the platform allowing its use in different devices through an internet browser.
4. Security. The platform must warrant the security in terms of authentication, confidentiality, and integrity in compliance with European regulations.
5. Cloud storage. The system must enable the secure storage of images, tests and reports involved in the project in a remote environment.

Figure 1 describes the Unified Modelling Language (UML) use case diagram that shows the interaction between the users and the system. The user roles are the examiner, responsible for managing patients and creating new exams, the evaluator, in charge of the assessment of the studies through the identification of findings potentially present in the CXR images with TB, and the administrator user accountable of managing user and medical centers access as well as defining the examiners (persons who examine the patient) and evaluators (persons who evaluate the CXR) and following up the progress of the evaluations.

Figure 1. Use case diagram of BITScreen platform with the three roles considered (examiner, evaluator and administrator) and the operations associated to them. All the "Manage" operations include the sub-operations new, edit and delete.
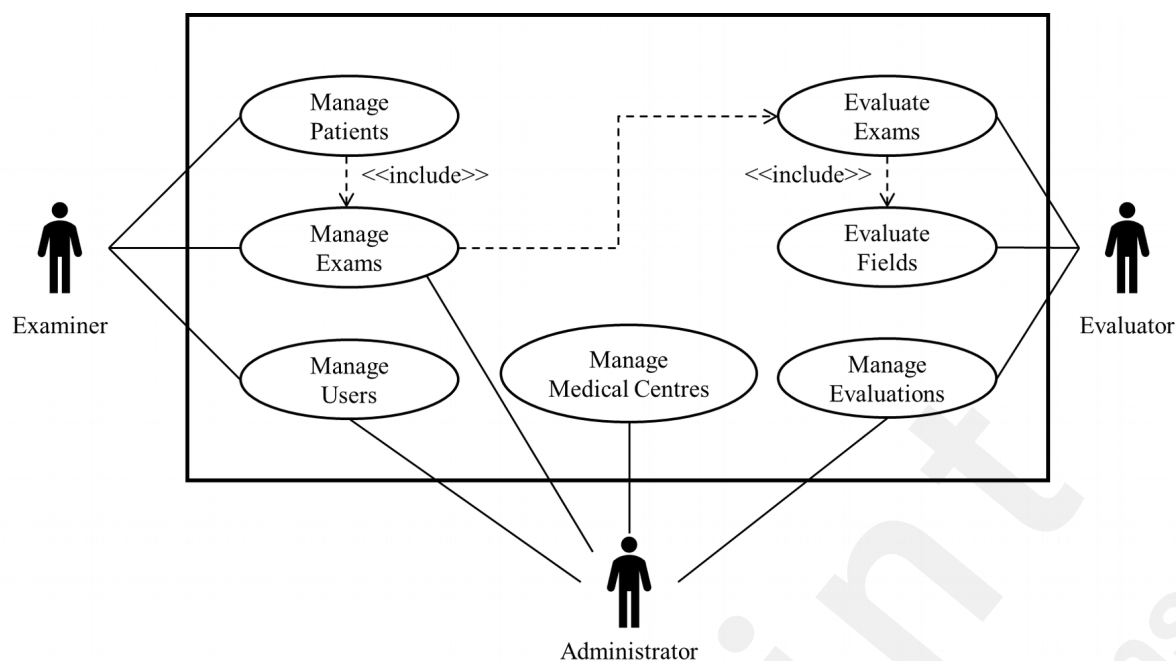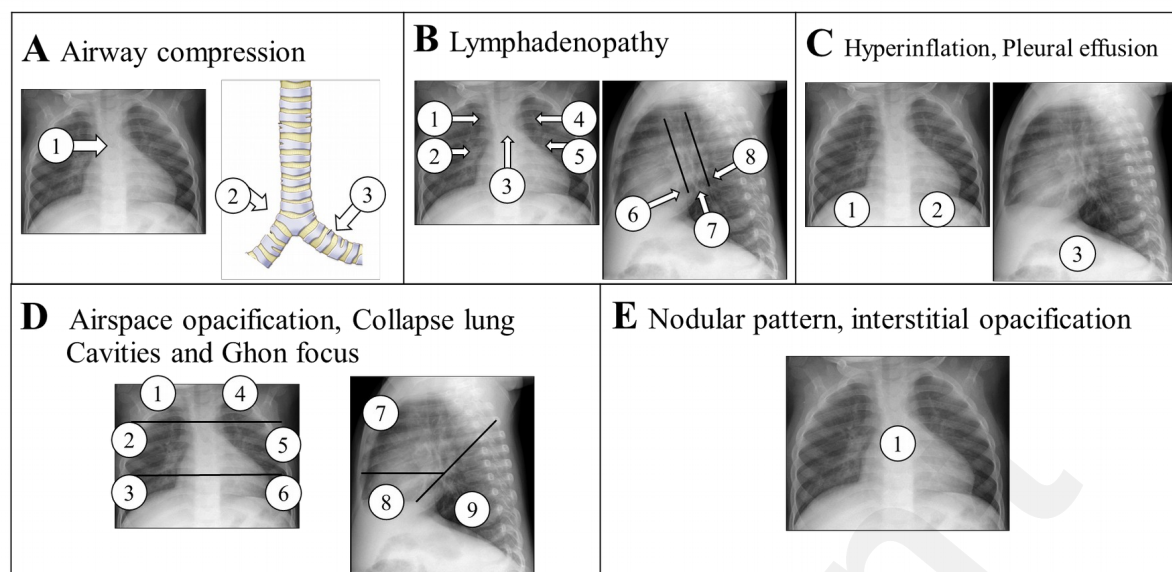
Figure 2 presents the activity diagrams used to design the functionality to upload a new exam to the platform by an examiner, including clinical information and the CXR images, and how the corresponding exam is sent to be evaluated by an evaluator user. The input fields included by the examiner to create a new exam were: month and year of birth, date of the exam, cough, fever, malnutrition, HIV (Human Immunodeficiency Virus) status, BCG (Bacillus Calmette-Guérin) vaccine scar, tuberculin skin test, TB diagnosis, TB contact, TB treatment, treatment starting date, and the CXR images (AP or PA and LAT views). In our pilot study, only the CXR images were presented to the evaluators. The age of the patient is calculated from the month and year of birth with respect to the acquisition date of the CXR. The examiner must upload at least one AP (anteroposterior) or PA (posteroanterior) view CXR image and if available the LAT (lateral) view also. In this pilot study the evaluators only were allowed to access the CXR images. In the evaluation process the platform was designed to include the assessment of the image quality of the CXR images, the presence of pulmonary TB radiological findings in different regions of the lungs considering different types of findings and a global evaluation of the CXR exam. In Figure 2, only one evaluation is included but the platform allows more than one (in the validation of the platform we included three evaluations for each exam). If more than one evaluation is configured, the process of evaluation of the exam will not end until all evaluations conclude their assessment in the platform.

Figure 2. Activity diagram of the process for creating and evaluating a new exam including the clinical information of the patient and the CXR images (anteroposterior or posteroanterior and lateral views).
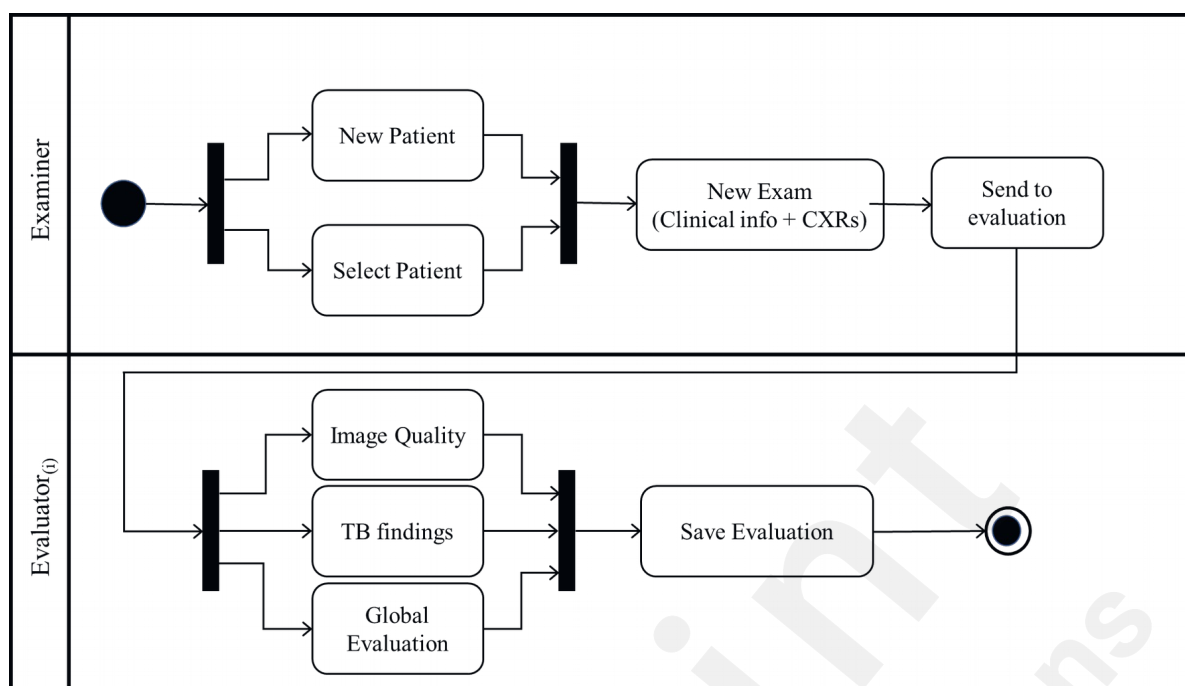
The evaluation of CXR images plays a critical role in the identification of presumed TB patients and is one of the primary focuses of the platform design. In order to provide an exhaustive and rigorous assessment of the CXR images the evaluators should mark "yes" or "no" to evaluate the presence or absence of the different radiological TB findings in different thoracic locations. For this purpose, we considered 10 sections corresponding with the different types of findings, leading to the total assessment of 55 independent observations, 36 in the AP/PA view and 19 from the lateral view. The 10 sections of pediatric CXR TB findings correspond to: airway compression and/or tracheal displacement, soft tissue density suggestive of lymphadenopathy, hyperinflation, pleural effusion, air space opacification, collapsed lobe or lung, cavities, calcified parenchymal lesions, nodular pattern and interstitial opacification. Figure 3 illustrates the templates provided to the evaluators, indicating the specific locations of the features to be assessed. The locations and types of findings have been considered following previous recommendations in the literature, including as main source "Diagnostic CXR Atlas for Tuberculosis in children"[15] and the CXR review tool developed by S. Andronikou and the South African Tuberculosis Vaccine Initiative (SATVI)[16].

Figure 3. Evaluation templates with the location of the specific findings that should be assessed by the evaluators with "yes" or "no" for each of the 10 sections. (A) Locations for the evaluation of possible airway compression and/or tracheal displacement. (B) Locations for the assessment of soft tissue density suggestive of lymphadenopathy. (C) Locations for the assessment of hyperinflation and pleural effusion. (D) Locations for the evaluation of air space opacification, collapsed lung, cavities and calcified parenchyma. (E) Location for the assessment of nodular pattern, either miliary or larger widespread and bilateral nodules, and interstitial opacification. Based on [15,16] and SATVI review tool by S. Andronikou.

For the implementation of the back end of the platform, we used the PHP's Laravel framework (version 6.2). Laravel includes a variety of built-in tools and features that were used in the project like routing, authentication, authorization, the management of the database connection and the blade templating engine. For the data storage we used MariaDB database (version 10.1.38), a fork of the MySQL database management system, because of its advantages in terms of efficiency, customization, portability, reliability, being open source, free and easy to use, and widely adopted by a large and active community. The front end was based on the public framework Bootstrap (v4.3.1), that provides several pre-designed components that can be easily incorporated into a website. Besides Bootstrap is also responsive, which facilitates the use of the application from a variety of devices and screen sizes. The server runs over a Debian 4.9 and was equipped with 2 virtual central processing units (CPUs) cores (Intel Xeon), 4GB of RAM and 100GB of hard disk space.

Dataset pilot study

The dataset to evaluate the platform in our pilot study was selected from a previous prospective descriptive study (ITACA) [17] of young children (< 3 years of age) presumed to have TB conducted at the Manhiça Health Research Center (CISM), located in Southern Mozambique[14,17]. ITACA study protocol was approved by the Mozambican National Bioethics Committee and the Hospital Clinic of Barcelona Ethics Review Committee. Written informed consent was obtained from the parent/legal guardians of all study participants. The sub-study for the digital processing of the CXR was further approved by the Mozambican National Bioethics Committee. We collected 218 exams selecting all the microbiologically confirmed and "probable" cases and randomly selecting 113 cases more from unlikely TB cases set. The cases were confirmed using Zielh Nielsen staining and rapid test as well as Xpert MTB/RIF and identified through mycobacterial molecular identification (HAIN GenoType® Mycobacterium CM/AS) [17]. Table 1 shows their demographic data. In order to improve the comparison between studies and promote the standardization of diagnostic procedures, we followed the classification of case definition for research reporting based on diagnostic evaluation studies of intrathoracic TB in children proposed by Graham et al.[18] In this update from the previous case definitions presented in 2012 and 2013 [16,19] the authors established three case definitions: confirmed, unconfirmed TB and unlikely TB. The collected cases were retrospectively classified [17]] following these definitions and using the

information collected from the previous study[14]. Table 2 shows the TB diagnosis categories identified together with the corresponding clinical data for each case. The symptoms definitions considered were[17]: cough for ≥14 days not responding to a course of antibiotics; fever greater than 38ºC ≥ 14 days; malnutrition defined as under 60% weight for height, failure to gain weight for more than 2 months or any loss of weight and not responded to nutritional interventional; TB contact in the last 12 months.

Table 1. Patient demographic characteristics of the dataset of the pilot study.

	Male	**Female**	**Total**
	n=122 (%)	n=96 (%)	n=218 (%)
Age range			
< 12 months	18 (14.8)	19 (18.8)	37 (17.0)
12-23 months	54 (44.3)	43 (44,8)	97 (44.5)
24-35 months	50 (41.0)	34 (35.4)	84 (38.5)

Table 2. Diagnostic categories and corresponding clinical characteristics considering the definitions from[17,18] .

TB Category	**Confirmed**	**Unconfirmed TB**	**Unlikely TB**
	n=10 (%)	**n=95 (%)**	**n=113 (%)**
Sex			
Female	6 (60.0)	44 (46.3)	46 (40.7)
Age range			
< 12 months	3 (30.0)	21 (22.1)	13 (11.5)
13-23 months	3 (30.0)	42 (44.2)	52 (46.0)
24-35 months	4 (40.0)	32 (33.7)	48 (42.5)
Cough			
Yes	5 (50.0)	17 (17.9)	14 (12.4)
Fever			
Yes	4 (40.0)	6 (6.3)	5 (4.4)
Malnutrition			
Yes	5 (50.0)	81 (85.3)	104 (92.0)
HIV status			
Positive	2 (20.0)	35 (36.8)	4 (3.5)
BCG scar			
Yes	10 (100.0)	94 (98.9)	113 (100.0)
TB contact			
Yes	2 (20.0)	11 (11.6)	4 (3.5)

Evaluation protocol

The 218 baseline exams, corresponding to the time of evaluation for presumptive TB, were uploaded by the administrator user through the automatic importing feature of the platform using a CSV file with the input fields described in Table 2 and the location of the CXR files with the AP (in all participants) and the LAT (in 207 of the participants) views. The platform assigned automatically all the cases to three pediatric CXR expert readers. All of them have extensive experience in the assessment of TB imaging in endemic settings of low-income, resource-limited countries[17,20]. The three evaluators performed a blind evaluation of the 218 exams using the platform and without any

other information but the CXR views and the reference templates (Figure 3). The evaluation included the assessment of the CXR images quality ("acceptable", "poor but readable" or "not acceptable not readable"), the 55 observations ("yes" or "no") over the 10 sections and a final global evaluation of the case ("suggestive of TB", "not suggestive of TB" or "not evaluable").

Performance metrics

To evaluate the performance of the evaluations we used the metrics: sensitivity, specificity, positive predictive value (PPV), F1-score and accuracy. We defined sensitivity or recall as the number of true positives cases with x ray findings suggestive of TB) divided by the sum of the number of true positives and false negatives. We defined specificity as the number of true negatives divided by the sum of true negatives and false negatives. The PPV is the proportion of true positive predictions out of all positive (true positive + false positive) predictions. It measures how many of the positive predictions are actually correct. The F1-score is a measure of a model's accuracy that combines the PPV and recall. It is used to evaluate the performance of a classification algorithm. The F1-score ranges between 0 and 1, where a score of 1 represents perfect PPV and recall, and a score of 0 represents the worst possible performance. We considered the accuracy as the sum of true positives and true negatives divided by the sum of true positives, true negatives, false positives and false negatives. A true positive case is defined when an evaluator selected as global evaluation of a case "suggestive of TB" and the exam was classified as "confirmed" or "unconfirmed TB". A true negative case is defined when the evaluator selected "not suggestive of TB" and the exam was classified as "unlikely TB". An exam is defined as false negative when the evaluator selected "not suggestive of TB" and the case was classified as "confirmed" or "unconfirmed TB". A case is defined as false positive when an evaluator selected "suggestive of TB" and the exam was classified as "unlikely TB". Additionally, we evaluated the association between the TB features identified in the CXRs and the global evaluation ("suggestive of TB" and "not suggestive of TB") with the initial diagnostic classification considering "confirmed" and "unconfirmed TB" together with chi-square P value <.05 representing statistical significance. Finally, we used Cohen kappa for the inter-reader agreement considering all the assessments performed by the evaluators (CXR image quality, TB feature evaluations and TB global evaluations). Kappa scores were classified as follows: ≤0 no agreement, 0.01–0.2 slight, 0.21–0.4 fair, 0.41–0.6 moderate, 0.61–0.8 substantial, and 0.81–1.00 almost perfect agreement.

Platform Usability Evaluation

We designed a comprehensive questionnaire with five sections and 15 items to effectively evaluate the usability of the platform. The questionnaire was adapted from the widely used Telehealth Usability Questionnaire (TUQ) proposed by Parmanto et al.[21], which has proven to be an effective tool for evaluating telemedicine services[22]. Our questionnaire covers several important usability perspectives, including usefulness (3 items), ease of use and learnability (2 items), interface quality (4 items), reliability (2 items), and global satisfaction (2 items). A detailed breakdown of the questionnaire components and associated items can be found in Table 3. We also conducted an analysis to determine the duration of the evaluation process for each exam. Specifically, we recorded the time elapsed from the point of request for a new exam to the moment when the evaluator submitted their final evaluation to the system. By calculating the difference between these two time points, we obtained an accurate and reliable estimate of the time required by the expert to complete a

thorough evaluation of an exam.

Results

The two primary views of the new BITScreen platform are depicted in Figures 4 and 5: the input form utilized by the examiner user and the evaluation form employed by evaluator users, respectively. In the top section of the input form (Figure 4), the examiner completes details such as cough, fever, last temperature, malnutrition, HIV, BCG scar, tuberculin skin test, TB category, contact with TB source patient, treatment, treatment starting date, and observations. In the bottom section, the examiner could upload CXR images for evaluation by the evaluators.

In the case of the evaluation form (Figure 5), the view used by the evaluators, presents the CXR images on the left side of the screen, enabling them to download or zoom in on each image. The view requires the assessment of the quality of each CXR image. On the right side of the screen, the 10 sections described previously are presented as separate tabs, where the evaluators must assess all the 55 different observations. The templates shown in Figure 3 are always present in the view to facilitate the evaluation's task. Readers had the option of marking all the locations without pathologic findings with "no", for all the criteria at once or for all the locations of a specific criterion to facilitate and expedite the evaluation. Finally, at the bottom of the view, the global evaluation field for the exam is displayed. All the fields are mandatory, except in the case where the CXR images are considered not evaluable.

Figure 4. Example of the BITScreen examiner user view of a new exam with the two different areas: Clinical data and Images.

Figure 5. Example of the BITScreen evaluator user view with three different areas: quality image assessment, identification of the presence of findings in the different locations presented in the templates and a global evaluation of the case.

The results of the usability questionnaire conducted on the telemedicine platform are presented in Table 3. The overall score for all questions was 4.4 ± 0.59 out of 5. The data indicates that users found the platform to be useful, with an average rating of 4.42 out of 5, easy to use and learn, with an average rating of 4.47 out of 5 and the interface quality received positive feedback, with an average rating of 4.13 out of 5. The platform was also perceived to be reliable, with an average rating of 4.26 out of 5 and a high level of variability (standard deviation of 0.82). All three evaluators reported a high level of satisfaction with the platform, with an average rating of 5.0 out of 5.

Some specific questions received lower ratings, particularly item 4 in the interface quality dimension ('The system is able to do everything I would want it to be able to do') and item 1 in the reliability dimension ('Whenever I made a mistake using the system, I could recover easily and quickly'). On the other hand, the items with higher feedback were item 1 in the ease of use and learnability section ('It was simple to use this system'), as well as the questions related to global satisfaction and future use, where the questions 'I would use the platform again' and 'Overall, I am satisfied with the platform' received the maximum feedback from all the evaluators.

Table 3. Results of the Usability Questionnaire (1=strongly disagree to 5=strongly agree)

Section	Mean ± Std
Usefulness	**4.42 ± 0.53**
1. It facilitates the assessment of CXRs in pediatric TB studies.	4.64 ± 0.58
2. It saves me time assessing CXRs in pediatric TB studies.	4.31 ± 0.58
3. It includes all the items I need to evaluate pediatric TB studies.	4.31 ± 0.58
Ease of Use and Learnability	**4.47 ± 0.52**
1. It was simple to use this system.	5.00 ± 0.00

2. It was easy to learn the system.	4.31 ± 0.58
3. The templates with the location of the findings facilitate the assessment of the cases.	4.31 ± 0.58
4. I believe I could become productive quickly using this system.	4.31 ± 0.58
Interface Quality	**4.13 ± 0.58**
1. The way I interact with this system is pleasant.	4.00 ± 0.00
2. I like using the system.	4.31 ± 0.58
3. The system is simple and easy to understand.	4.31 ± 0.58
4. The system is able to do everything I would want it to be able to do.	3.91 ± 1.00
Reliability	**4.26 ± 0.82**
1. Whenever I made a mistake using the system, I could recover easily and quickly.	3.91 ± 1.00
2. The system gave error messages that clearly told me how to fix the problems	4.64 ± 0.58
Satisfaction and Future Use	**5.0 ± 0.00**
1. I would use the platform again.	5.0 ± 0.00
2. Overall, I am satisfied with the platform	5.0 ± 0.00

Figure 6 shows the results of the evaluators' completion times. Evaluator 2 took the least time with an average of 35.3 ± 13.2 seconds, while evaluator 1 took an average of 37.8 ± 19.2 seconds and evaluator 3 took the longest with an average of 110.3 ± 63.2 seconds. As evaluator 3 identified more observations and performed better, the extra time taken is justified. Previous studies[23] have reported that radiologists take an average of 2 minutes and 9 seconds (129 seconds) to evaluate and report neonatal CXR images, which is longer than the time taken in our study. ItHowever, it's important to note that our reviewers only marked specific locations of findings, assessed image quality, and provided a global assessment without having to write or dictate a report. In any case, our results suggest that the platform could be a valuable tool for rapid case evaluation and marking of findings in CXR images.

Figure 6. Evaluation time in seconds by the three evaluators of the 218 exams
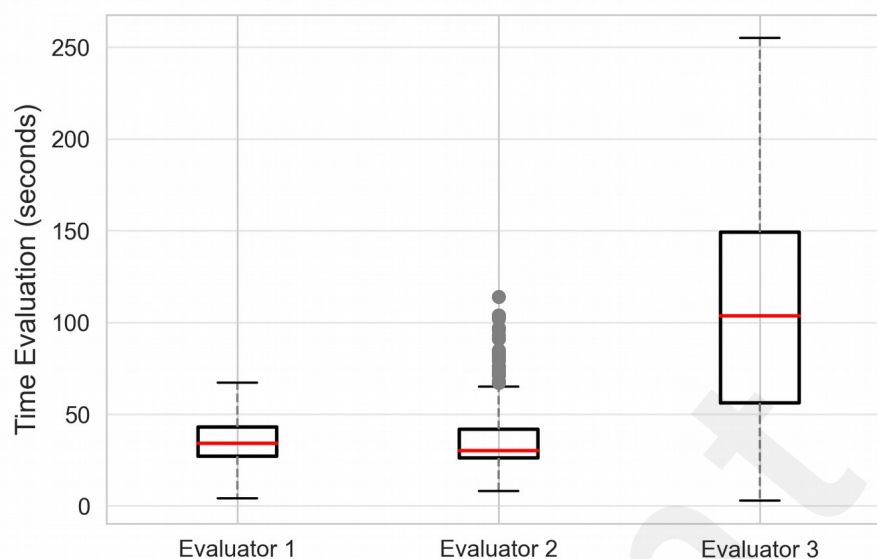
Next, we report the results from the assessment of 218 exams in this pilot study. The evaluation of the CXR AP views showed that 195 out of 219 (89.0%), 167 out of 193 (86.5%), and 150 out of 219 (68.5%) of the images were rated as "acceptable" by evaluator 1, 2, and 3, respectively, while 23 out of 219 (10.5%), 26 out of 193 (13.5%), and 65 out of 219 (29.7%) were rated as "poor but readable". However, in the case of lateral views, the image quality was lower, with 160 out of 209 (76.5%), 109 out of 161(67.7%), and 128 out of 208 (61.5%), of images rated as "acceptable" and 42 out of 209 (20.1%), 46 out of 161 (28.6%), and 59 out of 208 (28.4%), as "poor but readable" by the three evaluators, respectively. Additionally, 7 out of 209 (3.3%), 6 out of 161 (3.7%), and 21 out of 208 (10.1%) of lateral views were deemed "not acceptable, not readable". Notably, only evaluator 3 rated all views of the CXRs as "not acceptable, not readable" in two exams, and there was only one image that received this rating from all three evaluators. The number of images classified in each category by each expert is presented in Figure S1 in Multimedia Appendix 1, while Figure S2 in Multimedia Appendix 1 provides examples of images and their corresponding ratings.

Table 4 presents the performance metrics of the global evaluation. Among the three evaluators, evaluator 3 had the highest sensitivity (28.2%) and F1-score (40.8%) and accuracy (60.9%). However, evaluator 3 had the lowest specificity (91.1%), which suggests that it may have classified more unlikely TB cases as suggestive of TB compared to the other evaluators. Evaluator 2 had the highest specificity (98.2%) indicating that it was better at correctly identifying unlikely TB cases. However, it presented the lowest scores for sensitivity (12.4%) and F1-score (21.7%), implying that it struggled with correctly identifying confirmed and unconfirmed TB cases. Evaluator 1's scores were intermediate across all metrics, except for PPV where it had the lowest score (73.9%). This suggests that although evaluator 1 was not the best performer on any particular metric, it was consistently average across all metrics. To further illustrate the results, Figure 7 presents the confusion matrices with the corresponding number of cases identified as true negatives (top left), true positives (bottom right), false positives (top right), and false negatives (bottom left) and Table S1 in Multimedia Appendix 1 presents the evaluation for each TB diagnostic class.

Table 4. Performance metrics considering sensitivity, specificity, PPV, F1-score and accuracy. Total number of cases 218.

	Evaluator 1	**Evaluator 2**	**Evaluator 3**

Sensitivity %(95% CI)	16.3% (95% CI: 10.5 - 24.6)	12.4% (95% CI: 7.4 - 20.0)	28.2% (95% CI: 20.4 - 37.5)
Specificity %(95% CI)	94.6% (95% CI: 88.8 - 97.5)	98.2% (95% CI: 93.8 - 99.5)	91.1% (95% CI: 84.3 - 95.1)
PPV %(95% CI)	73.9% (95% CI: 50.3 - 63.4)	86.7% (95% CI: 62.1 - 96.3)	74.4% (95% CI: 58.9 - 85.4)
F1-score %(95% CI)	26.8% (95% CI: 19.8 - 35.1)	21.7% (95% CI: 15.2 - 29.9)	40.8% (95% CI: 33.1 - 49.1)
Accuracy %(95% CI)	56.9% (95% CI: 50.3 - 63.4)	56.9% (95% CI: 50.2 - 63.3)	60.9% (95% CI: 54.3 - 67.2)

Figure 7. Confusion matrices of the three evaluators.
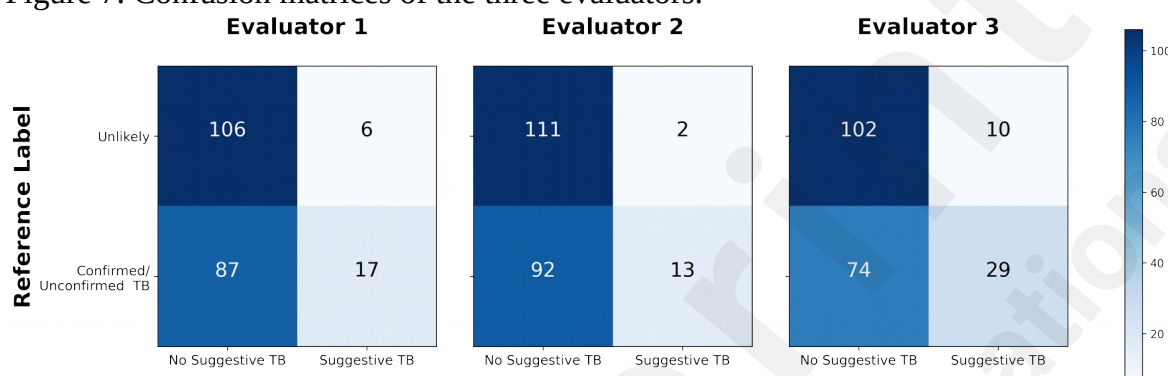
Table 5 shows the number of observations recorded by each of the three evaluators in the three diagnostic categories, namely confirmed, unconfirmed TB and unlikely TB, for the 10 examination fields. The total number of observations recorded by the three evaluators was 64, 59, and 150, indicating a significant difference between evaluator 3 and the other two evaluators. This difference was particularly pronounced in the unconfirmed TB and unlikely TB categories. Air space opacification was the category with the highest number of observations by all evaluators, especially in the unconfirmed TB category, where it ranged from 22 to 33 for the total of 95 cases. Lymphadenopathy was the second most frequently observed area, with evaluator 3 recording this finding in 34 examinations across all categories, 22 of which were in the unconfirmed TB category. In addition, a notable number of observations were recorded in the interstitial opacification field, with evaluator 3 being particularly active in marking this finding in 16 examinations. In contrast, the areas of cavities and calcified parenchyma were only identified by evaluator 3, who marked 4 and 6 exams, respectively. It is also worth noting that evaluator 3 recorded observations for all the exam areas, whereas evaluators 1 and 2 did not record any observations in the cavities and calcified parenchyma areas. Finally, Figure 8 shows examples of observations for four different patients with detailed marking of their findings.

Table 5. Results of the evaluation of the findings by the three experts considering the AP and Lateral CXRs without additional clinical information. Each data-point of the table represents the number of patients where the evaluators reported one or more times the presence of the finding. The last row includes all the patients with any of the previous abnormalities. The order of the data comes from the assessment of the findings by evaluator 1/ evaluator 2/ evaluator 3. Total number of cases 218.

	Overall	Confirmed	Unconfirmed TB	Unlikely TB
Number of patients	218	10	95	113

Airway compression and/or tracheal displacement	4/0/7	2/0/0	2/0/4	0/0/3
Lymphadenopathy	7/7/34	3/1/4	4/4/22	0/2/8
Hyperinflation	4/0/3	2/0/0	2/0/2	0/0/1
Air space opacification	31/42/52	7/7/7	22/30/33	2/5/12
Collapsed lung	5/1/9	0/0/2	4/1/6	1/0/1
Nodular picture	1/2/3	0/0/1	1/2/2	0/0/0
Interstitial opacification	7/1/28	1/0/2	5/1/16	1/0/10
Pleural effusion	5/6/4	0/0/0	5/6/4	0/0/0
Cavities	0/0/4	0/0/0	0/0/3	0/0/1
Calcified Parenchyma	0/0/6	0/0/0	0/0/4	0/0/2
Any abnormality	41/46/92	8/7/9	30/33/54	3/6/29

Figure 8. Example of evaluations of findings in different studies. The locations of the findings are defined in Figure 3. The color of the locations represents the number of evaluators that identified the presence of the finding in that location, being zero evaluators for the white color, yellow for one evaluator, orange for two evaluators and red for the three evaluators. (A) Presence of air space opacification in the AP and lateral CXR views of an exam of a female patient of 11 months classified as unconfirmed TB and as suggestive of TB by one out of the three evaluators. (B) Presence of lymphadenopathy in the AP and lateral CXR views of an exam of a female patient of 11 months classified as confirmed TB and as suggestive of TB by the three evaluators. (C) Presence of interstitial opacification on AP CXR views of two studies, the one on the left is from a male patient of one year and four months of age. Both studies were classified as unconfirmed TB and not suggestive of TB. The AP view on the right corresponds to a patient of female of 11 months of age. The exam was classified as unlikely TB and one out of three evaluators assessed it as confirmed TB. (D) Presence of pleural effusion in the AP and lateral view of an exam of a male of two years and two months of age classified as confirmed TB and evaluated as suggestive of TB by the three evaluators.

To better understand the impact of different evaluations on the final diagnosis of TB, we analyzed the association between the assessments made by each evaluator, including the final evaluation, and the initial diagnostic classification. Chi-square test results (Table S2 in Multimedia Appendix 1) revealed that the strongest association for the CXR features was between the identification of air space opacification with X^2_1 >20.38 and P<.001 for all evaluators. The second most noteworthy finding was lymphadenopathies significantly associated with the initial classification for evaluator 1 (X^2_1=5.79 P=.02), and evaluator 3 (X^2_1=11.88 P<.001). The final evaluation was also significantly associated with the initial classification for all three evaluators. These results align with those presented in Table 5, which demonstrated that these fields had the most observations among the rest.

Finally, we studied the agreement between evaluators using the Cohen kappa score for the inter-reader agreement for the image quality, the global evaluation and all the different findings (Figure S3 in Multimedia Appendix 1). Regarding the image quality we found substantial agreement between evaluator 1 and 2 (K=0.65) but only fair agreement between 1 and 3 (K=0.33) and 2 and 3 (K=0.31) because evaluator 3 assessed many more images as "poor quality". The agreement for the global evaluation was very similar with fair agreement (from 0.26 to 0.32). For the findings, we found that air space opacification had a moderate/substantial Cohen kappa index, ranging from 0.54 to 0.67. The number of observations identified by the evaluators (Table 5) and the association with the initial classification (Table S2 in Multimedia Appendix 1) showed that air space opacification was a crucial finding due to the large number of observations, strong association, and the homogeneity between different evaluators. Another field with significant agreement was pleural effusion, ranging from 0.43 to 0.72, although there were fewer observations and a weaker association with the initial classification. Lymphadenopathies also appeared to be an important finding in terms of observation and association, but the agreement was only slight, ranging between 0.13 to 0.21.

Discussion

Store-and-forward telemedicine has proven to be a valuable solution for enhancing the access to specialist and primary healthcare advice by leveraging technological advancements to overcome barriers in low-resource settings[12,13]. Our work demonstrates the potential use of this approach also in the assessment of TB in young children in underserved areas, where the lack of specialists and the difficulty of assessing TB in this population may have a greater impact. The positive

evaluation of the telemedicine system and the reduced time required for evaluation further supports the use of telemedicine for diagnosing PTB, ensuring timely intervention and efficient health care.

The low sensitivity of X rays to identify positive cases in our pilot study confirmed the difficulty of diagnosing TB in children reported in other studies[24–28]. Limited research has provided detailed insight into the global sensitivity and specificity of CXR in young children for the diagnosis of TB. Kaguthi et al.[25] reported sensitivities ranging from 50% and 75% and specificities between 72.9% and 85.2%. However, they noted the imprecise measurement of sensitivity due to the small number of definitive cases. Berteloot et al.[28] reported higher sensitivities (71.4%) and lower specificities (50.0%) although the evaluation process included a consensus, and an older age group of children. Other studies[26,27]have also examined the performance of TB diagnosis using CXR but focusing on the most relevant findings to support the diagnosis[26,27]. Similar to those, in our results the lymphadenopathies, opacifications and pleural effusions were the findings with the strongest association with positive evaluation (Table S2 in Multimedia Appendix 1). The integration of a treatment-decision algorithm that includes clinical evidence, CXR findings and Xpert MTB/RIF assay (or the current version Xpert MTB/RIF Ultra), as presented by several studies[2,9] could improve the performance of the diagnostic process and facilitate treatment decisions and could be considered in future developments.

In terms of inter-reader agreement, our findings align, to some extent, with other studies that have also reported slight to moderate agreement[25,26,29]. Kaguthi et al. [25] reported poor agreement on abnormalities consistent with TB (K=0.14) and moderate agreement (K=0.26) on lymphadenopathy. However, their lower agreement results compared to ours could be attributed to the variability in expertise among the readers. Our results are closer to other studies in terms of the reader profile[26,28,29]. Palmer et al.[26] presented moderate agreement (K>0.4) on specific features such as alveolar opacification, pleural effusion, expansile pneumonia and enlarged perihilar lymph nodes. Similarly, Berteloot et al.[28] reported a kappa value of 0.36 between radiologist and pediatric pulmonologist. Lastly, Andronikou et al.[29] presented a kappa value of 0.5 between trained pediatric radiologists, although their dataset included older children with a mean age of 9 years old.

Our pilot study has several limitations. The number of confirmed cases is small and the presence of some important features, relevant to diagnosis by CXR, such as airway compression and/or tracheal displacement, nodular pattern, cavities, or calcified parenchyma is also limited and may explain the lack of a stronger association with the TB classification highlighted in other studies[24]. The evaluators performance is compared to the case definition that includes abnormal CXR as one of the criteria for unconfirmed TB. As observed in analogous studies [28,29] our research was constrained by the limited number of studies and readers. Broader validation, including a wider range of studies and readers, may provide more robust insights into the agreement and performance of the evaluations. The expertise of our readers may not fully reflect the typical skill set available in resource-limited settings; however, this challenge can be overcome through the implementation of consensus classifications. Moreover, the approach of conducting double assessments by both non-experts and experts has been successfully tested in other projects[30,31], suggesting its potential effectiveness in enhancing diagnostic accuracy. By incorporating these methods into our telemedicine platform, we can overcome limitations related to reader expertise and improve the overall diagnostic process for pediatric TB in resource-limited settings.

Besides the current use of the platform as a diagnostic tool for remote evaluation of CXR exams, we have planned the future use for systematic assessment of datasets in clinical studies and as a labeling tool of TB findings present in CXR to facilitate the training of artificial intelligence segmentation and classification models. The inclusion of new datasets from multiple settings and the increase in the number of readers will allow for a thorough validation of the platform. It is also important to

consider the CXR as part of a larger diagnostic algorithm for pediatric TB, which involves assessing symptoms, signs of TB, exposure to a TB source patient, results from tests for *M. tuberculosis* infection (e.g., tuberculin skin tests or interferon-gamma release assays), as well as microbiological tests (e.g., Xpert MTB/RIF, microscopy, or culture for *M. tuberculosis*) and any other relevant supporting tests[15]. Following this approach, we consider leveraging the platform with a treatment-decision algorithm incorporating clinical evidence and artificial intelligence models to enabling automatic CXR scoring, that could significantly enhance the accuracy and efficiency of TB diagnosis in young children.

Conclusions

Tuberculosis continues to be a major global health problem, especially among children, and the COVID-19 pandemic has exacerbated the situation. CXR imaging plays a critical role in diagnosis, severity assessment, and evaluation of treatment response. In this work, we presented a new telemedicine web platform, BITScreen PTB, based on CXR images and clinical information, designed to facilitate remote reading and standardization of pediatric TB examinations in resource-limited settings.

Our platform was well received by users, and although further improvements may be needed to address concerns about reliability and interface quality, it has potential for future use. Our work highlights the potential of telemedicine platforms such as BITScreen PTB to improve access to TB diagnosis in children, particularly in low-resource settings. Finally, the platform has the potential to be used for labeling CXRs to create and integrate artificial intelligence models, which could improve the accuracy and speed of TB diagnosis in children, especially in resource-limited settings.

Acknowledgements

Conflicts of Interest

None declared.

Abbreviations

ACC: Accuracy
BCG: Bacillus Calmette-Guerin
CAD: Computer-aided diagnosis
CXR: Chest Radiographs
FN: False Negative
FP: False Positive
HIV: Human Immunodeficiency Virus
PPV: Positive Predictive Value
PTB: Pediatric Tuberculosis
TB: Tuberculosis
TN: True Negative
TP: True Positive

WHO: World Health Organization

References

1.   Global Tuberculosis Report 2022. Available from: https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022 [accessed Nov 11, 2022]

2.   Vonasek B, Ness T, Takwoingi Y, Kay AW, van Wyk SS, Ouellette L, Marais BJ, Steingart KR, Mandalakas AM. Screening tests for active pulmonary tuberculosis in children. Cochrane Database of Systematic Reviews 2021;2021(6). doi: 10.1002/14651858.CD013693.pub2

3.   Jenkins HE, Yuen CM, Rodriguez CA, Nathavitharana RR, McLaughlin MM, Donald P, Marais BJ, Becerra MC. Mortality in children diagnosed with tuberculosis: a systematic review and meta-analysis. Lancet Infect Dis 2017;17(3). doi: 10.1016/S1473-3099(16)30474-1

4.   Thomas TA. Tuberculosis in Children. Pediatr Clin North Am. 2017. doi: 10.1016/j.pcl.2017.03.010

5.   López-Varela E, Augusto OJ, Gondo K, García-Basteiro AL, Fraile O, Ira T, Ribó Aristizabal JL, Bulo H, Muñoz Gutierrez J, Aponte J, Macete E, Sacarlal J, Alonso PL. Incidence of tuberculosis among young children in rural mozambique. Pediatric Infectious Disease Journal 2015;34(7). doi: 10.1097/INF.0000000000000710

6.   World Health Organization. WHO consolidated guidelines on tuberculosis Module 5: Management of tuberculosis in children and adolescents. 2022;101. Available from: https://www.who.int/publications/i/item/9789240046764 [accessed Apr 11, 2023]

7.   Marais BJ. Childhood Tuberculosis: Epidemiology and Natural History of Disease. The Indian Journal of Pediatrics 2011;78(3):321–327. doi: 10.1007/s12098-010-0353-1

8.   De Villiers RVP, Andronikou S, Van De Westhuizen S. Specificity and sensitivity of chest radiographs in the diagnosis of paediatric pulmonary tuberculosis and the value of additional high-kilovolt radiographs. Australas Radiol John Wiley & Sons, Ltd; 2004 Jun 1;48(2):148–153. PMID:15230748

9.   Gunasekera KS, Walters E, Van Der Zalm MM, Palmer M, Warren JL, Hesseling AC, Cohen T, Seddon JA. Development of a Treatment-decision Algorithm for Human Immunodeficiency Virus-uninfected Children Evaluated for Pulmonary Tuberculosis. Clin Infect Dis Clin Infect Dis; 2021 Aug 15;73(4):E904–E912. PMID:33449999

10.   Uplekar M, Weil D, Lonnroth K, Jaramillo E, Lienhardt C, Dias HM, Falzon D, Floyd K, Gargioni G, Getahun H, Gilpin C, Glaziou P, Grzemska M, Mirzayev F, Nakatani H, Raviglione M. WHO's new end TB strategy. The Lancet. 2015. doi: 10.1016/S0140-6736(15)60570-0

11.   Falzon D, Timimi H, Kurosinski P, Migliori GB, Van Gemert W, Denkinger C, Isaacs C, Story A, Garfein RS, Bastos LGDV, Yassin MA, Rusovich V, Skrahina A, Van Hoi L, Broger T, Abubakar I, Hayward A, Thomas B V., Temesgen Z, Quraishi S, Von Delft D, Jaramillo E, Weyer K, Raviglione MC. Digital health for the End TB Strategy: developing priority products and making them work. Eur Respir J Eur Respir J; 2016 Jul 1;48(1):29–45. PMID:27230443

12.   Huang GKL, Pawape G, Taune M, Hiasihri S, Ustero P, O'Brien DP, du Cros P, Graham S, Wootton R, Majumdar SS. Telemedicine in Resource-Limited Settings to Optimize Care for Multidrug-Resistant Tuberculosis. Front Public Health Frontiers Media S.A.; 2019 Aug 13;7:222. doi: 10.3389/fpubh.2019.00222

13.   Delaigue S, Bonnardot L, Steichen O, Garcia DM, Venugopal R, Saint-Sauveur JF, Wootton R. Seven years of telemedicine in Médecins Sans Frontières demonstrate that offering direct specialist expertise in the frontline brings clinical and educational value. J Glob Health J Glob Health; 2018 Dec 1;8(2). PMID:30574293

14.   García-Basteiro AL, López-Varela E, Augusto OJ, Gondo K, Muñoz J, Sacarlal J, Marais B, Alonso PL, Ribó JL. Radiological Findings in Young Children Investigated for Tuberculosis in Mozambique. PLoS One Public Library of Science; 2015 May 28;10(5):e0127323. PMID:26020541

15.   Palmer M, Seddon JA, Goussard P, Schaaf HS. Diagnostic CXR Atlas for Tuberculosis in Children.

A Guide to Chest X-ray Interpretation, 2nd ed. Available from: https://theunion.org/technical-publications/diagnostic-cxr-atlas-for-tuberculosis-in-children [accessed Jul 25, 2023]

16. Graham SM, Ahmed T, Amanullah F, Browning R, Cardenas V, Casenghi M, Cuevas LE, Gale M, Gie RP, Grzemska M, Handelsman E, Hatherill M, Hesseling AC, Jean-Philippe P, Kampmann B, Kabra SK, Lienhardt C, Lighter-Fisher J, Madhi S, Makhene M, Marais BJ, McNeeley DF, Menzies H, Mitchell C, Modi S, Mofenson L, Musoke P, Nachman S, Powell C, Rigaud M, Rouzier V, Starke JR, Swaminathan S, Wingfield C. 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 Oxford University Press; 2012 May 5;205(Suppl 2):S199. PMID:22448023

17. López-Varela E, Augusto OJ, Gondo K, García-Basteiro AL, Fraile O, Ira T, Ribó Aristizabal JL, Bulo H, Muñoz Gutierrez J, Aponte J, Macete E, Sacarlal J, Alonso PL. Incidence of tuberculosis among young children in rural mozambique. Pediatric Infectious Disease Journal Lippincott Williams and Wilkins; 2015 Jul 4;34(7):686–692. PMID:26069945

18. Graham SM, Cuevas LE, Jean-Philippe P, Browning R, Casenghi M, Detjen AK, Gnanashanmugam D, Hesseling AC, Kampmann B, Mandalakas A, Marais BJ, Schito M, Spiegel HML, Starke JR, Worrell C, Zar HJ. Clinical Case Definitions for Classification of Intrathoracic Tuberculosis in Children: An Update. Clinical Infectious Diseases 2015;61. doi: 10.1093/cid/civ581

19. Cuevas LE, Browning R, Bossuyt P, Casenghi M, Cotton MF, Cruz AT, Dodd LE, Drobniewski F, Gale M, Graham SM, Grzemska M, Heinrich N, Hesseling AC, Huebner R, Jean-Philippe P, Kabra SK, Kampmann B, Lewinsohn D, Li M, Lienhardt C, Mandalakas AM, Marais BJ, Menzies HJ, Montepiedra G, Mwansambo C, Oberhelman R, Palumbo P, Russek-Cohen E, Shapiro DE, Smith B, Soto-Castellares G, Starke JR, Swaminathan S, Wingfield C, Worrell C. Evaluation of tuberculosis diagnostics in children: 2. Methodological issues for conducting and reporting research evaluations of tuberculosis diagnostics for intrathoracic tuberculosis in children. Consensus from an expert panel. Journal of Infectious Diseases 2012;205(SUPPL. 2). doi: 10.1093/infdis/jir879

20. du Preez K, Gabardo BMA, Kabra SK, Triasih R, Lestari T, Kal M, Tsogt B, Dorj G, Purev E, Nguyen TA, Naidoo L, Mvusi L, Schaaf HS, Hesseling AC, de Oliveira Rossoni AM, Carvalho ACC, Cardoso CAA, Sant'anna CC, Orti DGD, Costa FD, Vega LR, Sant'anna M de FP, Hoa NB, Phuc PH, Fiogbe AA, Affolabi D, Badoum G, Ouédraogo AR, Saouadogo T, Combary A, Kuate AK, Prudence BNA, Magassouba AS, Bangoura AM, Soumana A, Hermana G, Gando H, Fall N, Gning B, Dogo MF, Mbitikon O, Deffense M, Zimba K, Chabala C, Sekadde MP, Luzze H, Turyahabwe S, Dongo JP, Lopes C, Dos Santos M, Francis JR, Arango-Loboguerrero M, Perez-Velez CM, Koura KG, Graham SM. Priority Activities in Child and Adolescent Tuberculosis to Close the Policy-Practice Gap in Low- and Middle-Income Countries. Pathogens 2022, Vol 11, Page 196 Multidisciplinary Digital Publishing Institute; 2022 Feb 1;11(2):196. doi: 10.3390/PATHOGENS11020196

21. Parmanto B, Lewis, Jr. AN, Graham KM, Bertolet MH. Development of the Telehealth Usability Questionnaire (TUQ). Int J Telerehabil University Library Systems, University of Pittsburgh; 2016 Jul 1;8(1):3. PMID:27563386

22. Hajesmaeel-Gohari S, Bahaadinbeigy K. The most used questionnaires for evaluating telemedicine services. BMC Med Inform Decis Mak BioMed Central Ltd; 2021 Dec 1;21(1):1–11. PMID:33531013

23. Cowan IA, MacDonald SL, Floyd RA, Cowan MBChB I, MacDonald MBChB S, Floyd BAppSci R. Measuring and managing radiologist workload: Measuring radiologist reporting times using data from a Radiology Information System. J Med Imaging Radiat Oncol John Wiley & Sons, Ltd; 2013 Oct 1;57(5):558–566. PMID:24119269

24. Jain SK, Andronikou S, Goussard P, Antani S, Gomez-Pastrana D, Delacourt C, Starke JR, Ordonez AA, Jean-Philippe P, Browning RS, Perez-Velez CM. Advanced imaging tools for childhood tuberculosis: potential applications and research needs. Lancet Infect Dis Elsevier; 2020 Nov

1;20(11):e289–e297. PMID:32589869

25. Kaguthi G, Nduba V, Nyokabi J, Onchiri F, Gie R, Borgdorff M. Chest Radiographs for Pediatric TB Diagnosis: Interrater Agreement and Utility. Cinti S, editor. Interdiscip Perspect Infect Dis Hindawi Publishing Corporation; 2014;2014:291841. doi: 10.1155/2014/291841

26. Palmer M, Gunasekera KS, van der Zalm MM, Morrison J, Simon Schaaf H, Goussard P, Hesseling AC, Walters E, Seddon JA. The Diagnostic Accuracy of Chest Radiographic Features for Pediatric Intrathoracic Tuberculosis . Clinical Infectious Diseases 2022 Sep 15;75(6):1014–1021. doi: 10.1093/cid/ciac011

27. Swingler GH, du Toit G, Andronikou S, van der Merwe L, Zar HJ. Diagnostic accuracy of chest radiography in detecting mediastinal lymphadenopathy in suspected pulmonary tuberculosis. Arch Dis Child London: BMJ Publishing Group LTD; 2005 Nov;90(11):1153. PMID:1828270674

28. Berteloot L, Marcy O, Nguyen B, Ung V, Tejiokem M, Nacro B, Goyet S, Dim B, Blanche S, Borand L, Msellati P, Delacourt C. Value of chest X-ray in TB diagnosis in HIV-infected children living in resource-limited countries: the ANRS 12229-PAANTHER 01 study. International Journal of Tuberculosis and Lung Disease International Union Against Tuberculosis and Lung Disease; 2018 Aug 1;22(8):844–850. PMID:29991391

29. Andronikou S, Grier D, Minhas K. Reliability of chest radiograph interpretation for pulmonary tuberculosis in the screening of childhood TB contacts and migrant children in the UK. Clin Radiol W.B. Saunders; 2021 Feb 1;76(2):122–128. PMID:33010931

30. Daniel E, Pan W, Quinn GE, Smith E, Baumritter A, Ying GS, Quinn GE, Wade K, Baumritter A, Duros TB, Erbring L, Repka MX, Shepard JA, Emmert D, Herring CM, VanderVeen D, Johnston S, Wu C, Mantagos J, Ledoux D, Winter T, Weng F, Mansfield T, Bremer DL, McGregor M Lou, Jordan CO, Rogers DL, Fellows RR, Brandt S, Mann B, Wallace D, Freedman S, Jones SK, Tran-Viet D, Young R "Michelle," Barr CC, Bhola R, Douglas C, Fishman P, Bottorff M, Hubbuch B, Keith R, Bothun ED, DeBecker I, Anderson J, Holleschau AM, Miller NE, Nyquist DN, Siatkowski RM, Trigler L, Escobedo M, Corff K, Huynh M, Satnes K, Mills MD, Anninger W, Binenbaum G, Quinn G, Karp KA, Pearson D, Gong A, Stokes J, Harper CA, Weaver L, McHenry C, Conner K, Heemer R, Cokley E, Hoffman R, Dries D, Farnsworth KJ, Harrison D, Carlstrom B, Frye CA, Morrison D, Donahue S, Benegas N, Owings S, Phillips S, Ruark S, Ells A, Mitchell P, Ingram A, Sorbie R, Ying GS, Maguire M, Brightwell-Arnold M, Pistilli M, McWilliams K, Harris S, Whearry C, Daniel E, Martin ER, Parker Ostroff CR, Sepielli K, Smith E, Capone A, Hubbard GB, Lloyd P, Davis K, Gibson GC, Hansen R, Kemper AR, Prosser L, Musch DC, Christiansen SP, Hess DJ, Kymes SM, Sadda SVR, Spaulding R, Schron EB. Single grading vs double grading with adjudication in the telemedicine approaches to evaluating acute-phase retinopathy of prematurity (e-ROP) study. J AAPOS J AAPOS; 2018 Feb 1;22(1):32–37. PMID:29274372

31. Pow RE, Mello-Thoms C, Brennan P, Re P, Mbbs H, Mello-Thoms Bsc C, Brennan ; P, Correspondence P, Richard D, Pow E. Evaluation of the effect of double reporting on test accuracy in screening and diagnostic imaging studies: A review of the evidence. J Med Imaging Radiat Oncol John Wiley & Sons, Ltd; 2016 Jun 1;60(3):306–314. PMID:26996634
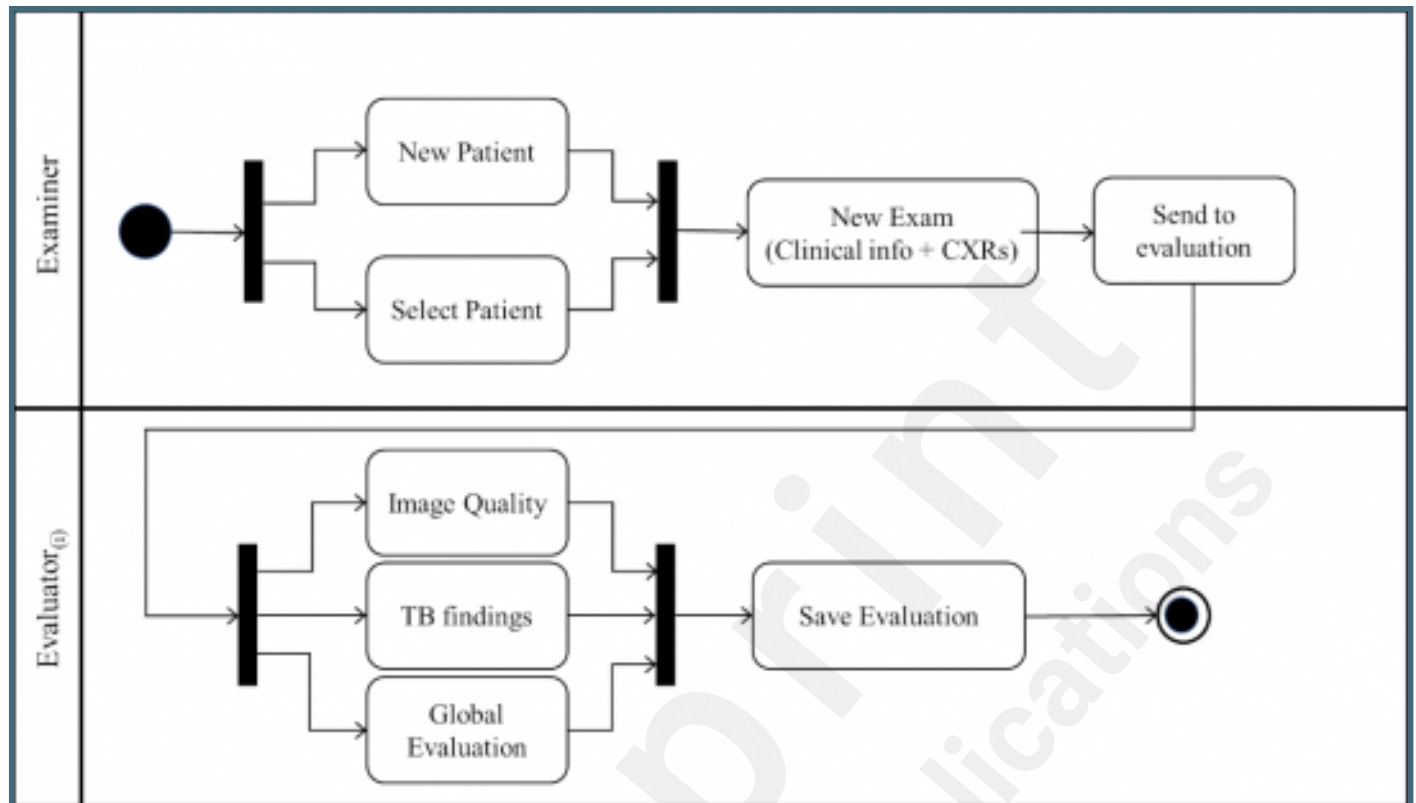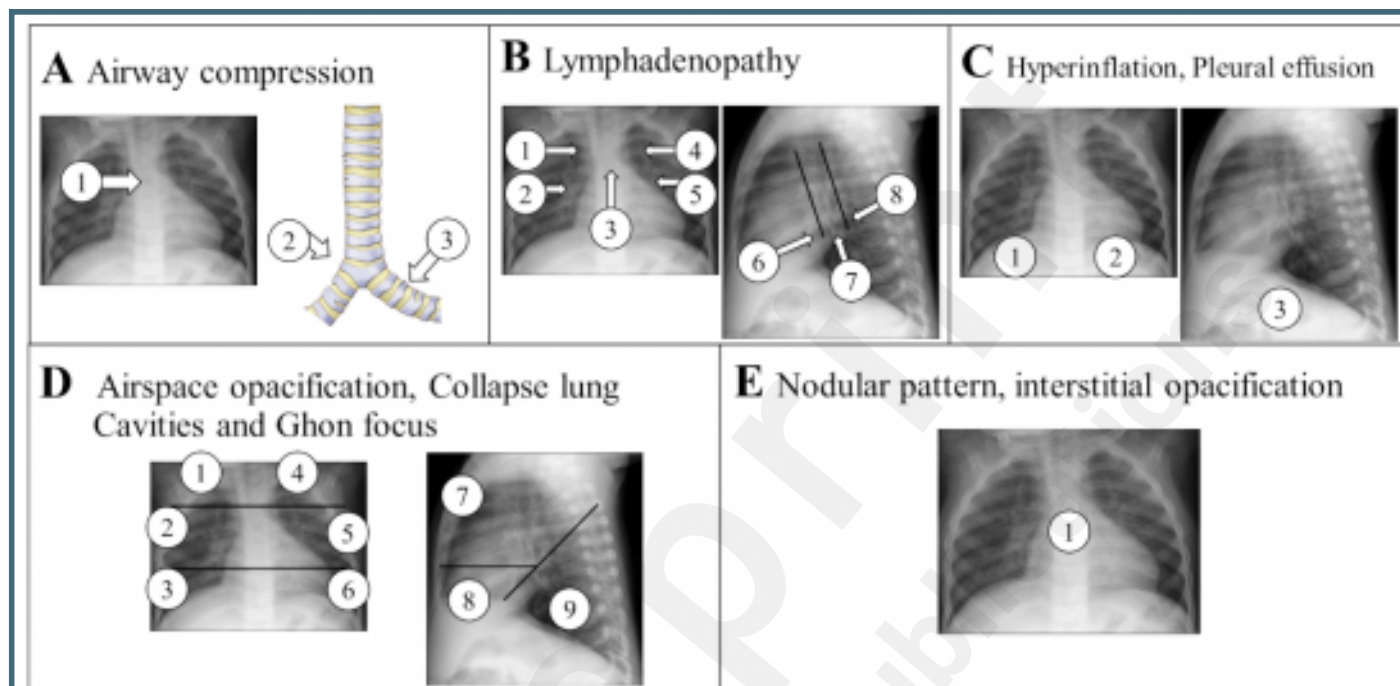
Supplementary Files

Figures

Use case diagram of BITScreen platform with the three roles considered (examiner, evaluator and administrator) and the operations associated to them. All the "Manage" operations include the sub-operations new, edit and delete.
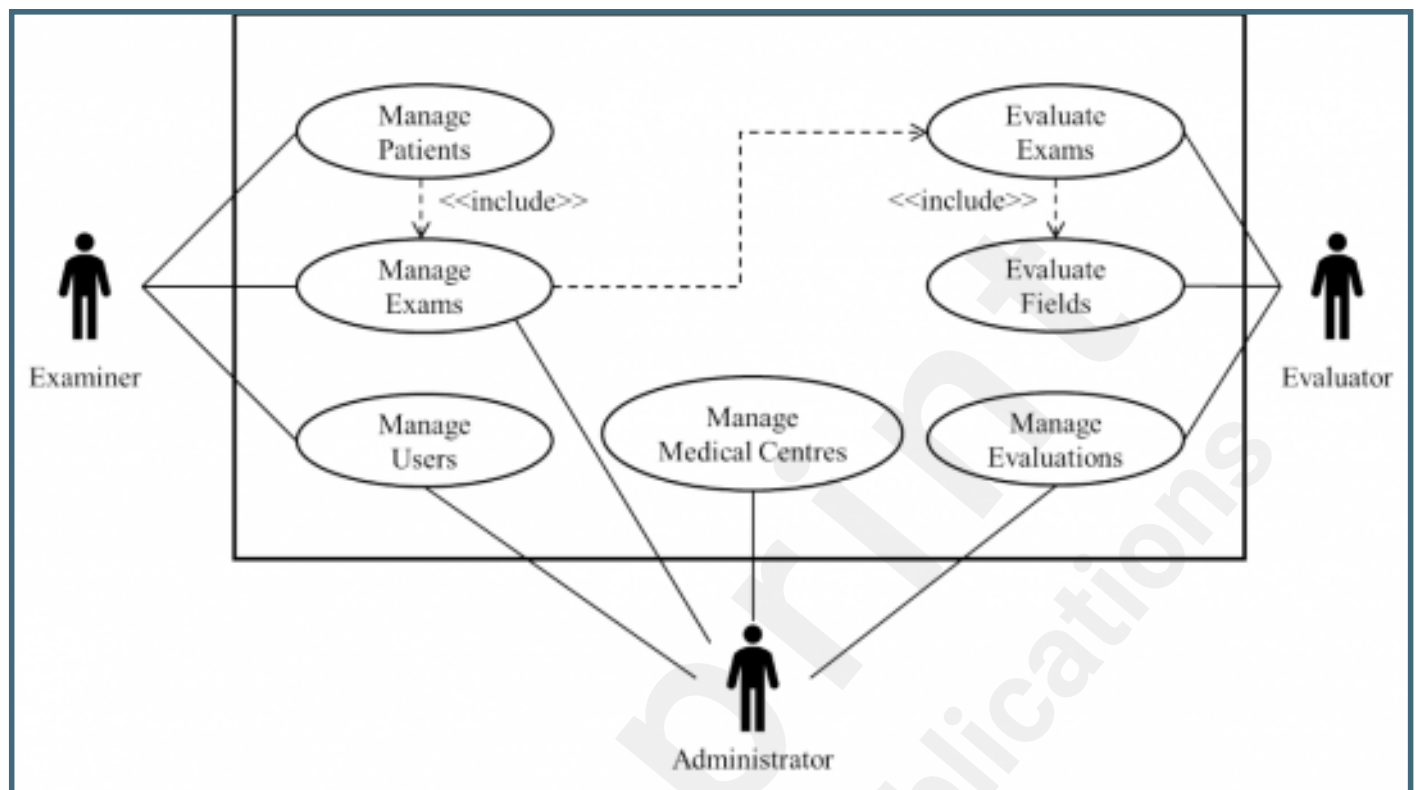
Activity diagram of the process for creating and evaluating a new exam including the clinical information of the patient and the CXR images (anteroposterior or posteroanterior and lateral views).

Evaluation templates with the location of the specific findings that should be assessed by the evaluators with "yes" or "no" for each of the 10 sections. (A) Locations for the evaluation of possible airway compression and/or tracheal displacement. (B) Locations for the assessment of soft tissue density suggestive of lymphadenopathy. (C) Locations for the assessment of hyperinflation and pleural effusion. (D) Locations for the evaluation of air space opacification, collapsed lung, cavities and calcified parenchyma. (E) Location for the assessment of nodular pattern, either miliary or larger widespread and bilateral nodules, and interstitial opacification. Based on [15,16] and SATVI review tool by S. Andronikou.

Example of the BITScreen examiner user view of a new exam with the two different areas: Clinical data and Images.

Example of the BITScreen evaluator user view with three different areas: quality image assessment, identification of the presence of findings in the different locations presented in the templates and a global evaluation of the case.
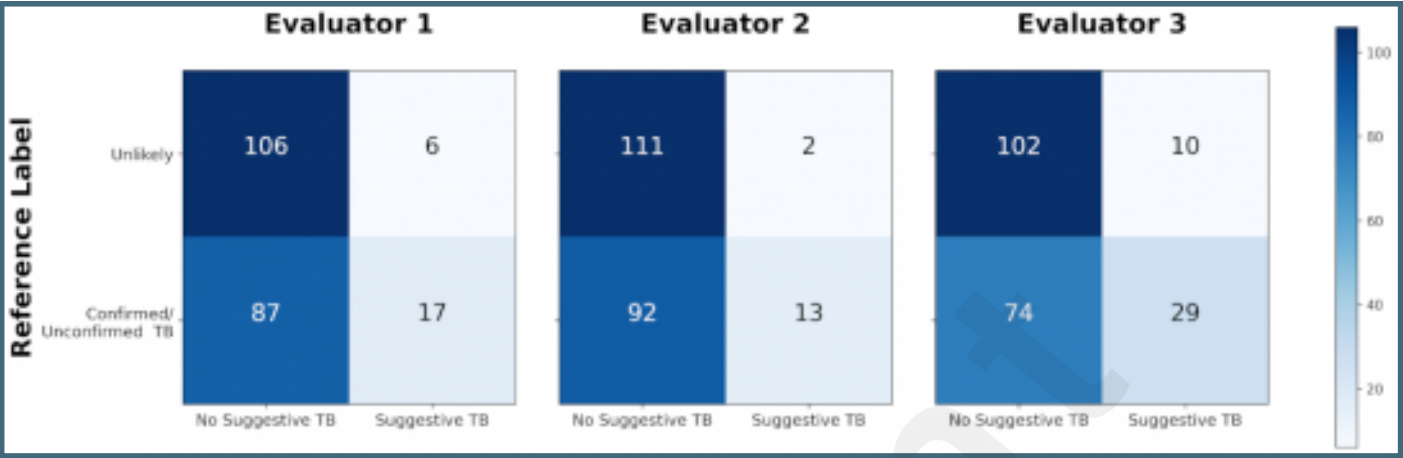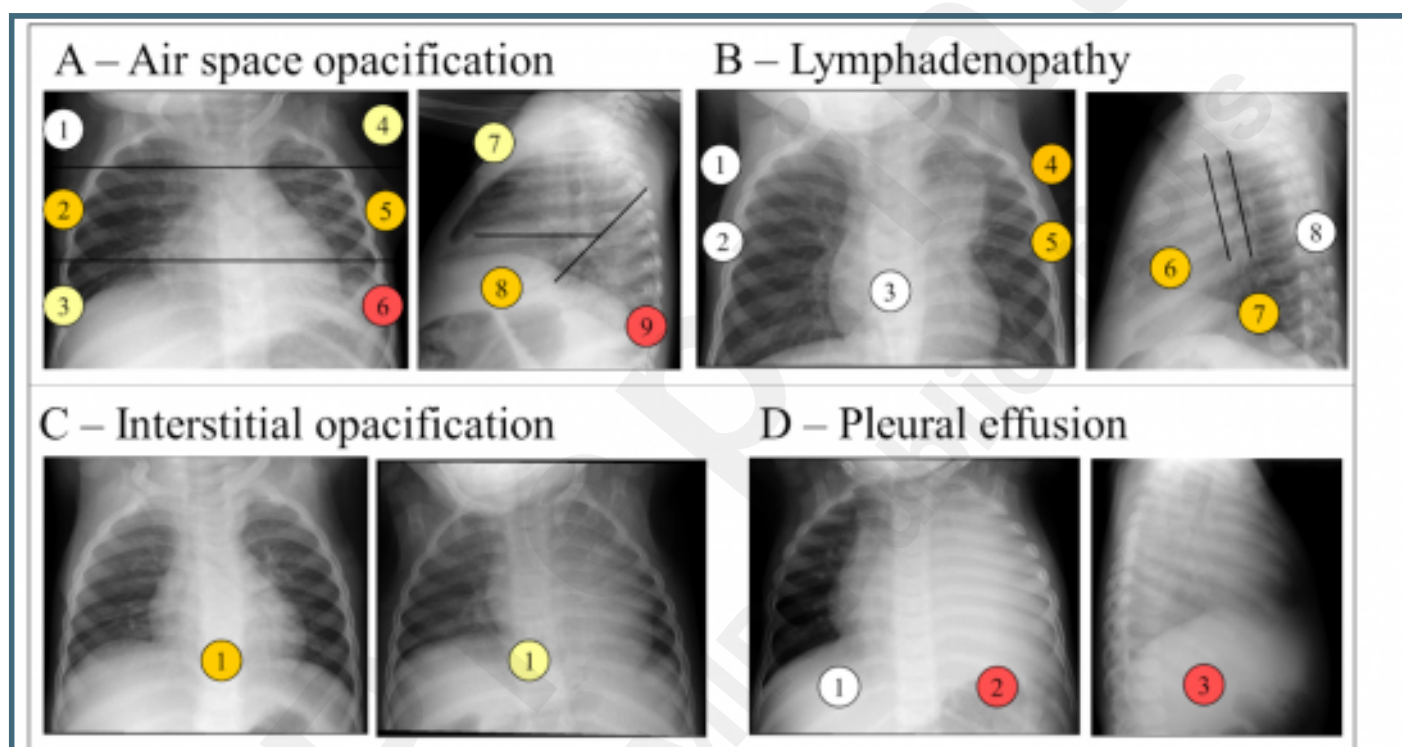
Evaluation time in seconds by the three evaluators of the 218 exams.

Confusion matrices of the three evaluators.

Example of evaluations of findings in different studies. The locations of the findings are defined in Figure 3. The color of the locations represents the number of evaluators that identified the presence of the finding in that location, being zero evaluators for the white color, yellow for one evaluator, orange for two evaluators and red for the three evaluators. (A) Presence of air space opacification in the AP and lateral CXR views of an exam of a female patient of 11 months classified as unconfirmed TB and as suggestive of TB by one out of the three evaluators. (B) Presence of lymphadenopathy in the AP and lateral CXR views of an exam of a female patient of 11 months classified as confirmed TB and as suggestive of TB by the three evaluators. (C) Presence of interstitial opacification on AP CXR views of two studies, the one on the left is from a male patient of one year and four months of age. Both studies were classified as unconfirmed TB and not suggestive of TB. The AP view on the right corresponds to a patient of female of 11 months of age. The exam was classified as unlikely TB and one out of three evaluators assessed it as confirmed TB. (D) Presence of pleural effusion in the AP and lateral view of an exam of a male of two years and two months of age classified as confirmed TB and evaluated as suggestive of TB by the three evaluators.

Multimedia Appendixes

Supplementary material.
URL: http://asset.jmir.pub/assets/1afcb66d227bfd04ceed8bc34dc49b64.docx