Diagnostic Performance of a Computer-aided System for Tuberculosis Screening in Two Philippine Cities

Gabrielle P. Flores, MD,[1,2] Reiner Lorenzo J. Tamayo, MAHPS,[2,3] Robert Neil F. Leong, PhD,[2,4,5] Christian Sergio M. Biglaen,[2,6] Kathleen Nicole T. Uy, RN,[2,3,4] Renee Rose O. Maglente, RN, GDip (Epi),[2,7] Marlex Jorome M. Nuguid, MM[2] and Jason V. Alacapa, MD, MBA, MPH, MHM[2,4,8]

[1]*Royal Free London NHS Trust, London, United Kingdom*
[2]*Innovations for Community Health, Inc., Mandaluyong City, Philippines*
[3]*Department of Health Policy and Administration, College of Public Health, University of the Philippines Manila, Manila, Philippines*
[4]*metaHealth Insights and Innovations, Inc., Malabon City, Philippines*
[5]*Mathematics and Statistics Department, De La Salle University, Manila, Philippines*
[6]*Department of Epidemiology and Biostatistics, College of Public Health, University of the Philippines Manila, Manila, Philippines*
[7]*University of the Philippines Manila, Manila, Philippines*
[8]*Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA*

ABSTRACT

Background and Objectives. The Philippines faces challenges in the screening of tuberculosis (TB), one of them being the shortage in the health workforce who are skilled and allowed to screen TB. Deep learning neural networks (DLNNs) have shown potential in the TB screening process utilizing chest radiographs (CXRs). However, local studies on AI-based TB screening are limited. This study evaluated qXR3.0 technology's diagnostic performance for TB screening in Filipino adults aged 15 and older. Specifically, we evaluated the specificity and sensitivity of qXR3.0 compared to radiologists' impressions and determined whether it meets the World Health Organization (WHO) standards.

Methods. A prospective cohort design was used to perform a study on comparing screening and diagnostic accuracies of qXR3.0 and two radiologist gradings in accordance with the Standards for Reporting Diagnostic Accuracy (STARD). Subjects from two clinics in Metro Manila which had qXR 3.0 seeking consultation at the time of study were invited to participate to have CXRs and sputum collected. Radiologists' and qXR3.0 readings and impressions were compared with respect to the reference standard Xpert MTB/RiF assay. Diagnostic accuracy measures were calculated.

Results. With 82 participants, qXR3.0 demonstrated 100% sensitivity and 72.7% specificity with respect to the reference standard. There was a strong agreement between qXR3.0 and radiologists' readings as exhibited by the 0.7895 (between qXR 3.0 and CXRs read by at least one radiologist), 0.9362 (qXR 3.0 and CXRs read by both radiologists), and 0.9403 (qXR 3.0 and CXRs read as not suggestive of TB by at least one radiologist) concordance indices.

Conclusions. qXR3.0 demonstrated high sensitivity to identify presence of TB among patients, and meets the WHO standard of at least 70% specificity for detecting true TB infection. This shows an immense potential for the tool to supplement the shortage of radiologists for TB screening in the country. Future research directions may consider larger sample sizes to confirm these findings and explore the economic value of mainstream adoption of qXR 3.0 for TB screening.

Keywords: tuberculosis, diagnostic imaging, deep learning

eISSN 2094-9278 (Online)
Published: January 31, 2025
https://doi.org/10.47895/amp.vi0.8950

Corresponding author: Reiner Lorenzo J. Tamayo, MAHPS
Innovations for Community Health, Inc.
One Sulatan Tower, Sulatan Street, Mandaluyong, Philippines
Email: renztamayo@gmail.com
ORCiD: https://orcid.org/0000-0002-5130-4341

INTRODUCTION

The chest radiograph (CXR) is the primary screening tool for pulmonary tuberculosis (PTB) according to the National Tuberculosis Control Program's Manual of Procedures (NTP MOP).[1] It provides high sensitivity despite average specificity when interpreted by radiologists.[2] However, in low-resource, high-burden countries, the shortage of health-care professionals with skill sets to appropriately interpret CXRs for PTB limits the widespread implementation of CXR screening.[3]

In the Philippines, there is a significant increase in TB burden despite increased efforts and resources spent on TB care and prevention.[4] The Philippine Strategic TB Elimination Plan (PhilSTEP) cites the need for intensified CXR screening and highlights radiologists' role in achieving targets.[4] There is a severe disparity between the actual number of radiologists and the ideal coverage of one radiologist per 10,000 Filipinos, as evidenced by recent estimates being that only 1,500 radiologists are available to serve 100 million Filipinos. To address this, PhilSTEP cites digital CXRs and AI as critical components to meeting screening targets.[4]

Deep learning (DL) neural networks (NNs) have been extensively explored and used in many different settings for screening individuals with PTB.[5] DLNNs use computer-aided detection (CAD) to recognize abnormalities on CXRs. Among these CAD tools are C4DTB[6], Lunit INSIGHT[7], and qXR 3[8], all of which have been subjected to evaluation as a screening tool for PTB[5]. The results have been promising, with the diagnostic performances meeting the WHO-recommended minimal accuracies for PTB triage tests.[2,3,5]

Despite availability of the technologies, none of these have been widely used nor explored in the Philippines. Furthermore, the WHO has called for more evidence before endorsing CAD software in PTB pathways. Hence, this calls for the conduct of an assessment on the use of AI in PTB screening in the Philippines.

This study provides a context-specific assessment of a CAD software (qXR3.0) to screen for TB from CXRs of adult Filipinos. Specifically, this study (a) determined the specificity and sensitivity of qXR3.0 with respect to the reference standard of Xpert MTB/RiF assay, (b) compared these diagnostic performance to radiologists' readings and impressions on the same CXRs for screening for PTB, and (c) assessed the qXR3.0 with reference to the WHO-recommended minimal accuracies for PTB triage tests.

METHODS

This study was written in compliance with the Standards for Reporting of Diagnostic Accuracy Studies (STARD) statement.

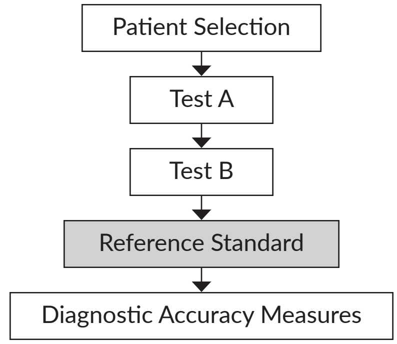

Figure 1. Randomized diagnostic accuracy comparative study.
Adapted from Chassé & Fergusson, 2019[9]

Study Design

A prospective cohort design was used to compare two tests, Test A: the deep-learning computer-aided technology qXR version 3.0; and Test B: the radiology report from licensed radiologists, as adapted from Chasse & Fergusson (Figure 1).[9] The reference standard was the Xpert MTB/RiF assay, a WHO-approved rapid diagnostic test for bacteriologically confirming pulmonary TB.[10]

Sampling Design

The study employed a convenience sample of adults who entered the clinics for a chest radiograph, and recruitment continued until the minimum number of positive cases (as detected by GeneXpert) had been met. The study had set a minimum number of positive cases as part of its recruitment process in order to assess sensitivity with better precision given that the tool assessed is a screening tool. All subjects were evaluated and screened for study eligibility by the investigators prior to the study entry and they were enrolled when one of the investigators were present in the clinics.

Sample size was calculated using the methods of Buderer et al.[11] following a 5% level of significance and assuming an estimated sensitivity of 93% and an estimated specificity of 75% based on Khan et al.[2] The prevalence used to compute the sample size is higher than that of the national average as it is based on the clinic's data which sees more TB patients as it functions as a presumptive TB referral center.

A desired precision of 0.07 required at least 278 subjects with at least 52 positive cases, as determined by Gene Xpert, to meet the sensitivity and specificity targets. However, due to resource constraints, the actual sample size used in the study did not meet these specified minimum numbers.

Participants

Participants included patients aged 15 and above with the following indications: presumptive TB based on the presence of chronic cough, weight loss, night sweats, or hemoptysis; screening of TB in individuals with risk factors or household contacts, or non-TB related indications such as general wellness and clearance to work assessments, as verified by the physician. Patients aged 15 and above were considered adults in line with the National Tuberculosis Control Program's Manual of Procedures.[1]

Individuals who were either in active TB treatment, had completed TB treatment in the past year, were diagnosed with extrapulmonary tuberculosis, or had referrals from HIV Treatment Hubs were excluded.

The study was conducted from October 2021 to February 2022 in Marikina and Mandaluyong cities, Philippines. The two primary care facilities provide consultation, treatment, and diagnostic services for a mix of patients, primarily presumptive TB cases.

Test Methods

Index Test: qXR

The qXR version 3.0 was evaluated in this study. The software was installed on local computers. The CXRs were uploaded to the software. qXR can detect abnormal findings on a CXR. qXR includes patented algorithms that can detect up to 29 different findings on the CXR. The CXR images were labeled using unique identifiers and interpreted by the AI software. QureAI's embedded software classifies CXR images as either normal or abnormal; furthermore, it identifies and highlights the location/s of potentially interesting findings on the CXR based on pre-specified cutoffs. The software did not take into account all other clinical information and microbiological test results.

Index Test: Radiology Reading

Two qualified radiologists with at least five years of professional experience interpreted CXRs that were labeled using unique identifiers and produced radiologic reports to officially communicate findings to the referring physicians and patients. The radiologists were blinded to all clinical information, qXR 3.0 results, interpretations from the other radiologist, and microbiological test results.

Reference Standard

The Xpert MTB/RIF is a rapid, automated nucleic acid amplification test. A rapid diagnostic test such as Xpert MTB/RIF is the primary diagnostic test for PTB and EPTB in adults and children according to the NTP MOP. It is the only WHO-recommended rapid diagnostic test with the added benefit of simultaneously assessing rifampicin resistance. A review showed that it achieves pooled sensitivity of 88% (95% CI 84-92%) and pooled specificity of 99% (95% CI, 98-99%) when used as an initial diagnostic tool for PTB.[9] As such, it is a suitable reference standard for the study due to its accuracy, availability, cost, and reduced waiting time. Results from the Xpert® MTB/RIF assay indicated whether or not *Mycobacterium tuberculosis* complex was detected. If detected, the results also stated whether rifampicin resistance was detected, not detected, or indeterminate. Samples were de-personalized and were given unique codes prior to testing to blind the assessors from the patients' clinical information while conducting the test.

Data Collection Procedures

Potential respondents who met the selection criteria were invited to participate in the study.

On the day of enrollment, a standardized demographic questionnaire was administered to the participants. Age, sex, presence of symptoms, comorbidities, as well as anthropometrics (height and weight) were collected.

A postero-anterior view digital CXR was performed. The x-ray images were labeled using their de-identified unique patient number and interpreted by the AI software and two qualified radiologists. CXRs were read within two days after receiving them.

A sputum sample was collected from all included patients on the same day of obtaining the CXR. Instructions on how to collect the sputum were provided to the patient. A specimen vial was provided before the patient was ushered to the data collection booth following standards on infection prevention and control detailed in the clinics' Standard Operations Manual. Sputum induction was conducted in an area with either natural ventilation or a negative pressure system installed, and only one patient was allowed into the sputum collection area at a time. For participants unable to expectorate, sputum conduction was performed through nebulization with saline solution (0.9% sodium chloride). Sputum specimens were analyzed within three days after collection. Molecular tests through the Xpert® MTB/RIF assay were used as the reference standard for pulmonary tuberculosis. The sputum was mixed with a reagent provided for the assay. A cartridge containing this mixture was placed in the GeneXpert® machine and was processed automatically. Results from the Xpert® MTB/RIF assay indicated whether or not *Mycobacterium tuberculosis* complex was detected in the sample.

Analysis

Patients were classified as either only having a positive or negative screen for PTB on imaging using both qXR and the radiologist's reading. These were then compared to the patient's disease status, categorized as PTB present or absent as determined bacteriologically through Xpert. Results were summarized and test performance measurements were calculated, e.g., sensitivity, specificity, and concordance indices. All 82 subjects had no missing test results or impressions by radiologists but those with indeterminate test results were removed from further analysis.

Data of participants who successfully completed a CXR and those who submitted a sputum sample for Xpert® after a CXR were included for statistical analysis. Diagnostic accuracy was calculated as sensitivity and specificity, together with their 95% confidence intervals. Positive and negative predictive values per 1000 patients being tested were estimated using the values of PTB prevalence and the derived sensitivities and specificities observed in the study. For inter-rater reliability, 15 samples of CXRs from bacteriologically-confirmed positive patients and 15 microbiologically-confirmed negative patients who were tested for PTB were used to assess the inter-rater reliability of both radiologists. Each radiologist analyzed the samples independently and was blinded to all clinical information and microbiological test results.

Ethical Considerations

The University of Santo Tomas (UST) College of Nursing (CON) Ethics Review Committee (ERC) approved the study (USTCON 2021-OR28). The protocol was not registered prospectively with a trials database. Despite the lack of this registration, the UST CON ERC authorized the continuation of the procedure. The full protocol is available upon request to the corresponding author.

The study objectives were explained to the respondents, and they were asked to sign an informed consent form to affirm approval to participate in the study without deception, coercion, undue influence, or inducement. Minors aged 15-17 were provided a separate assent form in a simpler language they could easily understand, to facilitate informed consent. In addition, a consent form was provided to the minor's parent or guardian. Both assent and consent forms were needed to be signed prior to enrolling a minor in the study. Each patient was assigned a unique identifier which was used for specimen labeling, encoding, and data analysis. Pursuant to Republic Act No. 10173 or the Data Privacy Act of 2012, the confidentiality of personal information was maintained and only authorized staff had access to these. Proper disposal of documents containing personal data was done in a secure manner as provided by law, and digital copies containing information were stored in a secure and encrypted database, with access provided only to the researchers.

RESULTS

Participant characteristics

From a total of 879 patients seen from Oct 2021 to Feb 2022 in the two clinic sites who met the inclusion criteria, 797 were excluded for any of the following reasons: (1) already being in active TB treatment, (2) had completed TB treatment in the past year, (3) were diagnosed with extrapulmonary tuberculosis, or (4) had referrals from HIV Treatment Hubs. Among the 82 patients recruited for the study, 36 (43.9%) were male, and 46 (56.1%) were female, with a mean age of 36.6 ± 15.8 years. Of the recruited patients, 14 had at least one comorbidity, with most commonly reported being hypertension (10.3%).

Test results

No adverse events were reported during and after the conduct of chest x-ray and sputum collection. Out of the 82 patients, 28 had their sputum samples collected for Xpert MTB, with M. tuberculosis detected in six (21.4%) and

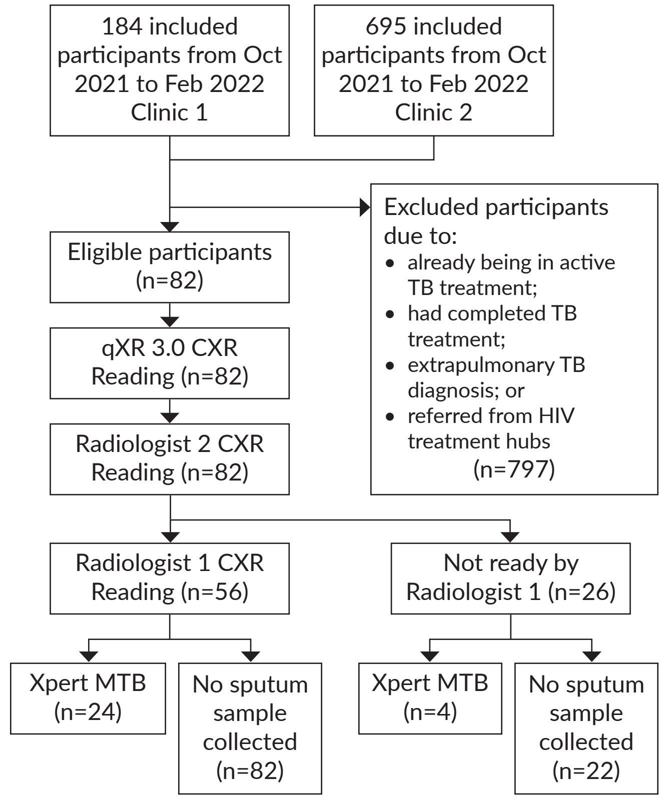

Figure 2. Participant flow.

not detected in 22 (78.6%). Radiologist 1 rated 12 out of 56 and radiologist 2 rated 23 out of 82 CXRs as abnormal, respectively. qXR 3 rated 26 out of 82 CXRs as abnormal, with 16 suggestive of tuberculosis (see Figure 2 for flow of participants with respect to TB screening methods).

Across tests, there was no observed significant difference in CXR gradings between sexes and presence of comorbidities, as shown in Table 1. Index and reference test results were summarized in Table 2.

Compared to the reference standard, the qXR3 (n=27) had a sensitivity of 100% (CI: 47.8 - 100%, α = 0.05) and specificity of 72.7% (CI: 49.8 - 89.3%). It had a higher specificity than the CXR gradings of two radiologists, as shown in Table 3.

The specificity of qXR3 meets the WHO requirement of 0.70, and is non-inferior to the standard at 5% significance level. The PPV of qXR3 was calculated as 454 per 1000 (CI: 167 to 766). The two radiologist gradings had PPVs of 400 (CI: 122 - 738) and 357 (128 - 659), respectively.

As for the agreement between qXR 3 and at least one of the two radiologist gradings (type 1), results (κ_1 = 0.7895; 95% CI [0.60,0.98]; $p<0.0001$) suggest a statistically very strong agreement, given that at least one radiologist evaluated the CXR impressions as suggestive of TB. In the four CXR readings with disagreeing interpretations, ambiguity in considering opacity, fibrosis, and calcifications as suggestive of TB was also observed, in addition to requesting an apicolordotic view of the chest x-ray for further evaluation, as in Figures 3 and 4.

In comparing radiologist-evaluated CXR with qXR3 gradings, where only CXRs read by both radiologists were considered (κ_2 = 0.9362; 95% CI [0.81,1]; p <0.0001), and where CXRs were read by at least one radiologist suggestive of TB as shown in Figure 4 (κ_3 = 0.9403; 95% CI [0.82,1]; p<0.0001), results also suggest a statistically very strong agreement in the CXR readings of the radiologists and qXR 3 in both cases. The single CXR where there were disagreements was interpreted by both radiologists as suggestive of TB. qXR 3 interpreted it as abnormal but not suggestive of TB.

Table 1. Participant Demographic and Clinical Characteristics

	Male	Female	Total/All	Percentage
Age (average in years)	33.7	38.9	36.6 ± 15.8	-
With comorbidities (one or more)				
Diabetes	2	6	8	9.2
Hypertension	1	8	9	10.3
Others	-	4	4	4.6
MTB/RiF Results (MTB) GX				
M. tuberculosis **detected**	5	1	6	21.4 (n=28)
M. tuberculosis **not detected**	8	14	22	78.6 (n=28)
Radiologist 1 CXR grading	(Heterogeneity: X^2: 0.2515, *p*-value: 0.112768)			
Findings **suggestive** of TB	8	4	12	21.4 (n=56)
Findings **not suggestive** of TB	3	11	14	25.0 (n=56)
Atherosclerotic aorta	-	1	1	
Bilateral apical pleural thickening	1	1	2	
Granuloma or Bone island	-	1	1	
Other	2	8	10	
No significant abnormality/findings	15	15	30	53.6 (n=56)
Radiologist 2 CXR grading	(Heterogeneity: X^2: 0.1358, *p*-value: 0.712454)			
Findings **suggestive** of TB	11	12	23	27.7 (n=82)
Findings **not suggestive** of TB	3	11	14	16.9 (n=82)
Calcified aortic knob	-	2	2	
Ill-defined opacities	1	-	1	
Pneumonia	-	1	1	
Other	2	8	10	
No significant abnormality/findings	22	23	45	55.4 (n=82)
qXR 3.0 CXR grading	(Heterogeneity: X^2: 1.0929, *p*-value: 0.295826)			
Findings **suggestive** of TB	9	7	16	19.3 (n=82)
Findings **not suggestive** of TB	1	9	10	12.0 (n=82)
No significant abnormality/findings	26	30	56	68.7 (n=82)

Table 2. Summary of Index and Reference Test Results

Test	Suggestive of TB (+)	Not Suggestive of TB (-)	No Significant Findings	Total
(MTB) GX	6	22	-	28
qXR 3.0	16	10	56	82
Radiologist 1	12	14	30	56
Radiologist 2	23	14	45	82
Radiologist 1 and Radiologist 2	10	2	44	56
Radiologist 1 or Radiologist 2	22	10	24	56

Table 3. Test Sensitivity and Specificity Confidence Intervals

Test	Sensitivity (CI: LI-UI)	Specificity (CI: LI-UI)	Accuracy
qXR 3 (n = 27)	100.0 (47.8 - 100.0)	72.7 (49.8 - 89.3)	77.8
Radiologist 1 (n = 22)	100.0 (39.8 - 100.0)	66.7 (41.0 - 86.7)	72.7
Radiologist 2 (n = 25)	100.0 (47.8 - 100.0)	55.0 (31.5-76.9)	64.0

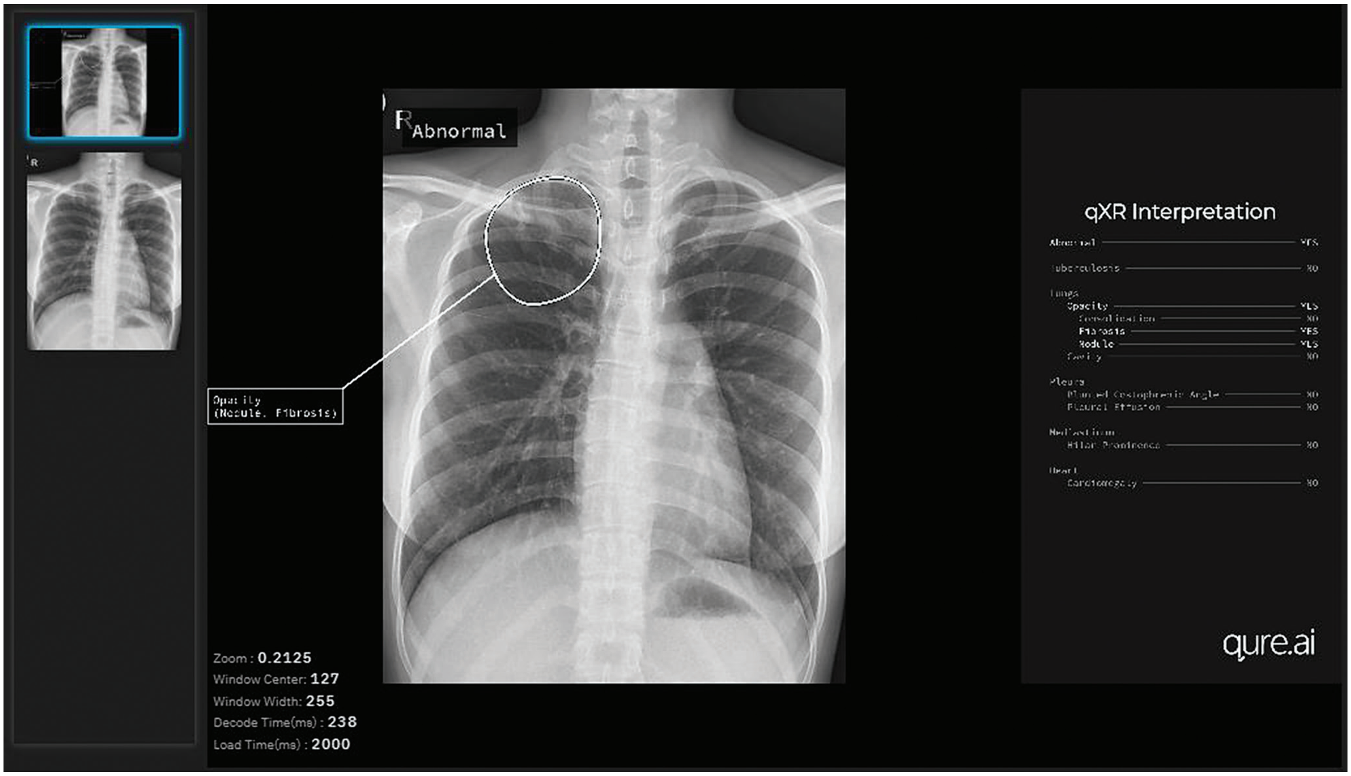

Figure 3. Abnormal qXR 3.0 reading of CXR interpreted as Normal by Radiologists 1 and 2.

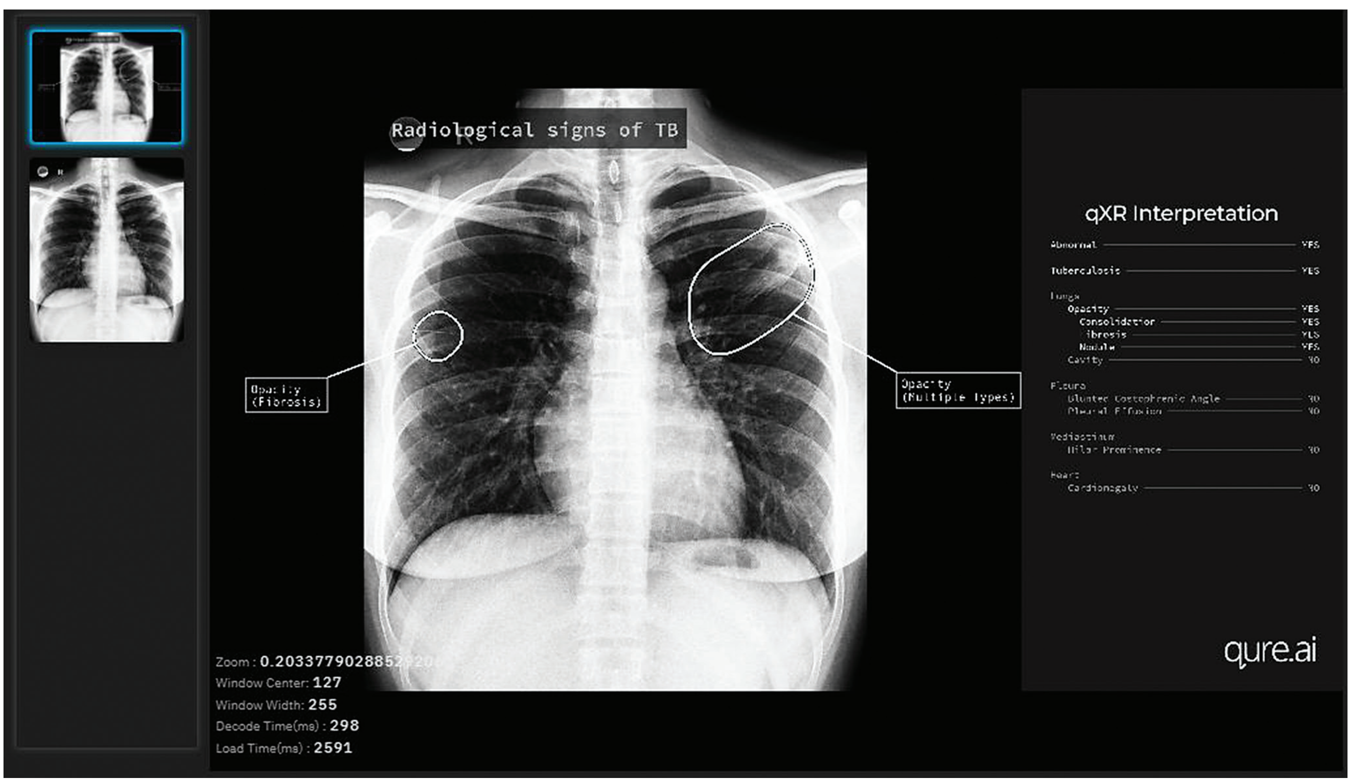

Figure 4. Abnormal qXR 3.0 reading of CXR interpreted as Normal by one of two radiologists.

DISCUSSION

The Philippines is a high TB burden country.[12] Concerningly, despite the country ranking fourth globally with the highest TB incidence[13], case detection has significantly dropped. This may be due to reduced health system capacity, less willingness and ability to seek care, concerns about the risk of going to healthcare facilities during a pandemic, and stigma associated with symptoms related to TB and COVID-19.[14] Looking into the reduced capacity to diagnose and attend to TB patients, several CADs are already and can be optimized for widespread adoption in the country to supplement the shortage as seen in high-burden settings.[15]

The calculated sensitivity and specificity of qXR 3, as well as those of the radiologists in the study, meet the WHO requirement of 0.90 and 0.70, and are non-inferior at 0.05 significance level, respectively. A higher specificity in the test translates to more patients being correctly screened as free from TB, potentially reducing the need for further Xpert testing or diagnostic tests. A study that included qXR 3 mentions a potential 50% reduction in the number of Xpert tests required, with screening test sensitivity kept above 90%.[14] These findings suggest that qXR 3 may supplement the shortage of radiologists, and augment the current screening procedure for TB.

Other studies have reported similar sensitivity (90.2% [95%CI 89.2–91.1]) and specificity (74.3% [95% CI 73.3–74.9]) values.[15] Other countries have reported the potential of using AI tools like qXR3 for the early detection of TB in augmenting screening procedures.[5,16] Despite this, the resulting sample size that was smaller than calculated limits the statistical power. The desired sampling requirements were unmet due to the limited foot traffic in the selected sites. Mobility restrictions have impacted the number of patients seen by clinics during the pandemic.[14]

Estimates of the Cohen's kappa coefficients suggest statistically strong agreements among and between gradings by qXR and radiologists. Nevertheless, a few agreements in some of the CXR gradings are expected.[17] Among the CXR gradings both by qXR3 and radiologists, disagreements in findings and impressions usually involved fibrosis and opacities. Aside from atelectasis or consolidation due to TB, opacity could be attributed to other conditions with a similar radiographic profile to TB.[18] The sole reliance on CXR grading in diagnosing TB could lead to over- or under-diagnosis that could cause more harm than good in the patient by subjecting them to unnecessary treatments at risk of developing drug-resistant TB.[18]

The study was conducted at clinics that typically receive referrals for presumptive TB cases. Careful considerations must be placed in interpreting these findings to be generalized in a larger population as the sampling design and locations may have introduced certain biases. Furthermore, the time frame used for study may have limited the representativeness of TB patients seeking care within that time frame which could have limited the representativeness of the study's sample with respect to different TB care seeking patterns throughout an entire year. Despite the study's limitations due to a small sample size, it is important to note that the population captured from the two clinics was diverse, potentially minimizing biases. This diversity within the sample, despite its size, lends credibility to the generalizability of the findings.

CONCLUSIONS

qXR 3 as a screening tool for TB met WHO requirements for sensitivity and specificity, with respect to the reference standard Xpert MTB/RiF. It further demonstrated high agreement with radiologist readings and impressions. Employing computer-aided detection of TB with acceptable sensitivity and specificity can help ease the burden on the limited number of trained medical professionals. This study shows an immense potential for the tool to supplement the shortage of radiologists, and provide more resource-efficient screening and diagnosis triage for TB, especially in low-resource, high-burden TB countries such as the Philippines. It is recommended that more research be done with bigger sample sizes to confirm these findings and explore the conclusions and determine their economic value of mainstream adoption of qXR 3.0 for TB screening.

Acknowledgments

The authors would like to express their sincere gratitude to Dr. Orven Jules Dumaoang for his assistance in the development of the protocol, and to the staff of Kalinga Health, Mandaluyong City Health Office, Marikina City Health Office, and the National TB Program for their valuable support in this endeavor.

Data Availability Statement

The protocol and data that support the findings of this study are available from the corresponding author, upon reasonable request.

Statement of Authorship

All authors certified fulfillment of ICMJE authorship criteria.

Author Disclosure

All authors declared no conflicts of interest.

Funding Source

Innovations for Community Health, the investigators' institution, shouldered all the expenses of this study. No external funding was received from either government or private entities.

REFERENCES

1. NTP Manual of Procedures 6th edition [Internet]. Gov.ph. 2020 [cited 2024 Feb 2]. Available from: https://ntp.doh.gov.ph/download/ntp-mop-6th-edition/
2. Khan FA, Majidulla A, Tavaziva G, Nazish A, Abidi SK, Benedetti A, et al. Chest x-ray analysis with deep learning-based software as a triage test for pulmonary tuberculosis: a prospective study of diagnostic accuracy for culture-confirmed disease. Lancet Digit Health. 2020 Nov;2(11):e573–81. doi:10.1016/s2589-7500(20)30221-1. PMID: 33328086.
3. Nash M, Kadavigere R, Andrade J, Sukumar CA, Chawla K, Shenoy VP, et al. Deep learning, computer-aided radiography reading for tuberculosis: a diagnostic accuracy study from a tertiary hospital in India. Sci Rep. 2020 Jan 14;10(1):210. doi:10.1038/s41598-019-56589-3. PMID: 31937802; PMCID: PMC6959311.
4. Updated Philippine strategic TB elimination plan 1: 2020-2023 [Internet]. Gov.ph. [cited 2023 Jul 18]. Available from: https://ntp.doh.gov.ph/download/updated-philippine-strategic-tb-elimination-plan-1-2020-2023/.
5. Qin ZZ, Sander MS, Rai B, Titahong CN, Sudrungrot S, Laah SN, et al. Using artificial intelligence to read chest radiographs for tuberculosis detection: A multi-site evaluation of the diagnostic accuracy of three deep learning systems. Sci Rep. 2019 Oct 18;9(1):15000. doi: 10.1038/s41598-019-51503-3. PMID: 31628424; PMCID: PMC6802077.
6. Thirona.eu. CAD4TB White paper [Internet]. 2018 [cited 2023 Jul 18]. Available from: https://thirona.eu/wp-content/uploads/2019/05/CAD4TB_6.0.0_WhitePaper.pdf.
7. Fujifilm.com. Lunit Insight CXR: Medical White Paper [Internet]. 2021 [cited 2023 Jul 18]. Available from: https://asset.fujifilm.com/www/in/files/2021-05/8cf5de8dc33b4e3316b4177fa47286c6/Lunit_INSIGHT_CXR_Medical_White_Paper.pdf.
8. Qure.ai. Qure qXR Brochure [Internet]. 2021 [cited 2023 Jul 18]. Available from: https://www.qure.ai/wp-content/uploads/2021/10/Qure_qXR_brochure-1.pdf.
9. Chassé M, Fergusson DA. Diagnostic accuracy studies. Semin Nucl Med. 2019 Mar;49(2):87–93. doi:10.1053/j.semnuclmed.2018.11.005. PMID: 30819399.
10. World Health Organization. Xpert MTB/RIF implementation manual: technical and operational 'how-to'; practical considerations [Internet]. 2014 [cited 2023 Jul 18]. Available from: https://www.who.int/publications-detail-redirect/9789241506700
11. Buderer NM. Statistical methodology: I. incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. Acad Emerg Med. 1996 Sep;3(9):895–900.doi: 10.1111/j.1553-2712.1996.tb03538.x. PMID: 8870764.
12. (PH) Department of Health. Joint program review 2019 report [Internet]. Gov.ph. 2020 [cited 2023 Jul 18]. Available from: https://ntp.doh.gov.ph/download/joint-program-review-2019-report/.
13. World Health Organization. Global tuberculosis report 2020 [Internet]. 2020 [cited 2023 Jul 18]. Available from: https://www.who.int/publications-detail-redirect/9789240013131.
14. World Health Organization. Global tuberculosis report 2021 [Internet]. 2021 [cited 2023 Jul 18]. Available from: https://www.who.int/publications-detail-redirect/9789240037021.
15. Qin ZZ, Ahmed S, Sarker MS, Paul K, Adel ASS, Naheyan T, et al. Tuberculosis detection from chest x-rays for triaging in a high tuberculosis-burden setting: an evaluation of five artificial intelligence algorithms. Lancet Digit Health. 2021 Sep;3(9):e543–54. doi:10.1016/s2589-7500(21)00116-3. PMID: 34446265.
16. Kazemzadeh S, Yu J, Jamshy S, Pilgrim R, Nabulsi Z, Chen C, et al. Deep learning detection of active pulmonary tuberculosis at chest radiography matched the clinical performance of radiologists. Radiology. 2023 Jan;306(1):124–37. doi:10.1148/radiol.212213. PMID: 36066366.
17. Vick A, Estrada CA, Rodriguez JM. Clinical reasoning for the infectious disease specialist: a primer to recognize cognitive biases. Clin Infect Dis. 2013 Aug;57(4):573–8. doi: 10.1093/cid/cit248. PMID: 23595833.
18. Nachiappan AC, Rahbar K, Shi X, Guy ES, Mortani Barbosa EJ Jr, Shroff GS, et al. Pulmonary tuberculosis: role of radiology in diagnosis and management. Radiographics. 2017 Jan-Feb;37(1):52–72. doi: 10.1148/rg.2017160032. PMID: 28076011.