Real-World Performance of AI-Powered Chest X-Ray Screening for Pulmonary Tuberculosis across County and Township Healthcare Facilities in Yichang, China, 2022-2024

Wenjie Jiang, Hao Zhang, Zhili Li, Xinli Jiang, Jiamei Shao, Xuelin Yang, Jingjie Xiong, Hui Zhang, Hongsheng Wang, Jianxing Yu, Xiaoyou Su, Ye Wang, Jianhua Liu, Zhongjie Li

Submitted to: Journal of Medical Internet Research
on: August 27, 2025

Table of Contents

Real-World Performance of AI-Powered Chest X-Ray Screening for Pulmonary Tuberculosis across County and Township Healthcare Facilities in Yichang, China, 2022-2024

Wenjie Jiang[1*]; Hao Zhang[2*]; Zhili Li[3]; Xinli Jiang[1]; Jiamei Shao[1]; Xuelin Yang[1]; Jingjie Xiong[4]; Hui Zhang[5]; Hongsheng Wang[6]; Jianxing Yu[1]; Xiaoyou Su[1]; Ye Wang[1]; Jianhua Liu[2]; Zhongjie Li[1]

[1] School of Population Medicine and Public Health, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China Beijing CN
[2] Yichang Center for Diseases Control and Prevention, Yichang, China Yichang CN
[3] Beijing Chest Hospital, Capital Medical University / Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing, China School of Population Medicine and Public Health, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China Beijing CN
[4] The School of Public Health, Guilin Medical University, Guilin, Guangxi, China Yichang Center for Diseases Control and Prevention, Yichang, China Yichang CN
[5] National Center for Tuberculosis Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China Beijing CN
[6] Hospital for Skin Diseases, Institute of Dermatology, Chinese Academy of Medical Sciences & Peking Union Medical College, Nanjing, Jiangsu, China Nanjing CN
[*]these authors contributed equally

Corresponding Author:
Zhongjie Li

School of Population Medicine and Public Health, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China
31 Bei Ji Ge San Tiao, Beijing 100730, China
Beijing
CN

Abstract

Background: In resource-limited areas, severe shortages of radiologist contribute to high rates of missed pulmonary tuberculosis (PTB) cases when relying solely on conventional chest X-ray (CXR). Although artificial intelligence (AI)-powered computer-aided detection (CAD) has shown effective in PTB diagnosis, its real-world clinical utility and scalability in primary healthcare settings remained underexplored.

Objective: To evaluate the real-world performance of CAD technology for triaging of PTB patients in primary healthcare facilities in high-incidence areas, and to assess its potential for optimizing radiological resource allocation.

Methods: We conducted a retrospective paired-design diagnostic accuracy study using CXR images collected from 7 county- and 32 township-level healthcare facilities in Yichang city between 2022 and 2024. All images were retrospectively reprocessed with CAD software (JF CXR-1 v2), and the original radiology reports interpreted by radiologists were extracted. CAD and radiologist performances were compared using two primary evaluation indicators: diagnostic yield among diagnosed cases (DYD) and positive predictive value (PPV). Subgroup analysis (by region, age, sex, healthcare facility tier, and patients category) and sensitivity analysis were conducted to assess the robustness of the results.

Results: Among 93,319 enrolled study patients (including 273 bacteriologically confirmed PTB cases), CAD demonstrated a substantially higher DYD (83.88%, 229/273) than radiologists (25.64%, 70/273), though with much lower PPV (1.70% vs. 10.31%). This high-sensitivity performance achieved an 85.52% reduction (only 13,515 instead of 93,319 CXRs) in radiologist workload via selective review of CAD-positive images, without missing any radiologist-identified PTB cases. Furthermore, among CAD-positive images, probability scores >0.75 was a key threshold for identifying high-risk PTB patients, prioritizing for radiologist review. Subgroup analysis further revealed that CAD outperformed radiologists in identifying PTB cases across all scenarios, despite some heterogeneity. In township healthcare facilities, CAD demonstrated significantly better performance than in county-level facilities, with DYD of 86.72% and 62.50%, and PPV of 2.00% and 0.65%, respectively.

Conclusions: CAD technology demonstrates valuable PTB screening performance in primary healthcare facilities. Combined with a tiered "AI pre-screening with selective human review" strategy, this approach effectively alleviates workload of radiologist in resource-constrained regions, offering a scalable solution for tuberculosis prevention and control.

(JMIR Preprints 27/08/2025:83041)
DOI: https://doi.org/10.2196/preprints.83041

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
No. Please do not make my accepted manuscript PDF available to anyone. I understand that if I later pay to participate in <a href="https:/

Original Manuscript

Real-World Performance of AI-Powered Chest X-Ray Screening for Pulmonary Tuberculosis across County and Township Healthcare Facilities in Yichang, China, 2022-2024

Wenjie Jiang[1], BSc; Hao Zhang[2], MM; Zhili Li[1,3], MM; Xinli Jiang[1], MM; Jiamei Shao[1], BM; Xuelin Yang[1], BSc; Jingjie Xiong[2,4], BM; Hui Zhang[5], PhD; Hongsheng Wang[1,6], PhD; Jianxing Yu[1], PhD; Xiaoyou Su[1], PhD; Ye Wang[1], PhD; Jianhua Liu[2], PhD; Zhongjie Li[1], PhD

[1] School of Population Medicine and Public Health, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China
[2] Yichang Center for Diseases Control and Prevention, Yichang, China
[3] Beijing Chest Hospital, Capital Medical University / Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing, China
[4] The School of Public Health, Guilin Medical University, Guilin, Guangxi, China
[5] National Center for Tuberculosis Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China
[6] Hospital for Skin Diseases, Institute of Dermatology, Chinese Academy of Medical Sciences & Peking Union Medical College, Nanjing, Jiangsu, China
*Wenjie Jiang and Hao Zhang contributed equally to this work.

Corresponding Author:
Zhongjie Li, PhD
School of Population Medicine and Public Health,
Chinese Academy of Medical Sciences & Peking Union Medical College
31 Bei Ji Ge San Tiao, Beijing 100730, China
Phone: +86 13671194343
Email: lizhongjiecdc@163.com

Jianhua Liu, PhD
Yichang Center for Diseases Control and Prevention, Yichang, China
37 Lantai Road, Xiling District
Yichang, Hubei Province, China
Phone: +86 13872651886
Email: amour_1114@163.com

Ye Wang, PhD
School of Population Medicine and Public Health,
Chinese Academy of Medical Sciences & Peking Union Medical College
31 Bei Ji Ge San Tiao, Beijing 100730, China
Phone: +86 13552881677
Email: wangye_pumc@163.com

Abstract

Background: In resource-limited areas, severe shortages of radiologist contribute to high rates of missed pulmonary tuberculosis (PTB) cases when relying solely on conventional chest X-ray (CXR). Although artificial intelligence (AI)-powered computer-aided detection (CAD) has shown effective in PTB diagnosis, its real-world clinical utility and scalability in primary healthcare settings remained underexplored.

Objective: To evaluate the real-world performance of CAD technology for triaging of PTB patients in primary healthcare facilities in high-incidence areas, and to assess its potential for optimizing radiological resource allocation.

Methods: We conducted a retrospective paired-design diagnostic accuracy study using CXR images collected from 7 county- and 32 township-level healthcare facilities in Yichang city between 2022 and 2024. All images were retrospectively reprocessed with CAD software (JF CXR-1 v2), and the original radiology reports interpreted by radiologists were extracted. CAD and radiologist performances were compared using two primary evaluation indicators: diagnostic yield among diagnosed cases (DYD) and positive predictive value (PPV). Subgroup analysis (by region, age, sex, healthcare facility tier, and patients category) and sensitivity analysis were conducted to assess the robustness of the results.

Results: Among 93,319 enrolled study patients (including 273 bacteriologically confirmed PTB cases), CAD demonstrated a substantially higher DYD (83.88%, 229/273) than radiologists (25.64%, 70/273), though with much lower PPV (1.70% vs. 10.31%). This high-sensitivity performance achieved an 85.52% reduction (only 13,515 instead of 93,319 CXRs) in radiologist workload via selective review of CAD-positive images, without missing any radiologist-identified PTB cases. Furthermore, among CAD-positive images, probability scores >0.75 was a key threshold for identifying high-risk PTB patients, prioritizing for radiologist review. Subgroup analysis further revealed that CAD outperformed radiologists in identifying PTB cases across all scenarios, despite some heterogeneity. In township healthcare facilities, CAD demonstrated significantly better performance than in county-level facilities, with DYD of 86.72% and 62.50%, and PPV of 2.00% and 0.65%, respectively.

Conclusions: CAD technology demonstrates valuable PTB screening performance in primary healthcare facilities. Combined with a tiered "AI pre-screening with selective human review" strategy, this approach effectively alleviates workload of radiologist in resource-constrained regions, offering a scalable solution for tuberculosis prevention and control.

Keywords: tuberculosis detection; chest X-ray; computer-aided detection; resource-limited settings; real-world

Introduction

In 2023, an estimated 10.8 million new cases of tuberculosis (TB) were reported globally, among which 2.6 million went undiagnosed or unreported, contributing to ongoing transmission and high mortality rates [1-3]. Early screening and diagnosis are crucial for interrupting *Mycobacterium tuberculosis* (MTB) transmission, but the current passive case finding strategy, which relies on symptom-dependent access to healthcare, remains inefficient [4, 5].

Chest X-ray (CXR) remains the common screening tool in primary healthcare facilities due to its ease of implementation, rapid, and high acceptance [6]. However, the accuracy of CXR screening is influenced by radiologists' ability to identify pulmonary tuberculosis (PTB) cases [7]. The limited availability of well-trained radiologists in primary healthcare results in high rates of missed diagnoses for PTB, leading to delays in diagnosis and increased risks of disease transmission [8].

Artificial intelligence (AI)-assisted computer-aided detection (CAD) tools for CXR represent a promising innovation to enhance the TB screening efficiency, diagnostic accuracy, and cost-effectiveness [9, 10]. Currently, the performance of CAD tools has been shown to match or even surpass that of human experts, demonstrating high sensitivity and reducing diagnostic delays [11, 12]. However, most of existing research were conducted in controlled environments, like those using standardized equipment and pre-selected patient cohorts [13, 14]. The effectiveness of these tools in complex real-world settings has not yet been validated. Furthermore, CAD's potential to automate initial screening offers a solution to address the shortage of healthcare personnel [15]. However, challenges remain, including controlling the false positive rate (e.g., misclassifying non-TB abnormalities) and uncertainties regarding the adaptability of these tools within tiered healthcare systems.

In this study, we aimed to evaluate the performance of CAD to enhance PTB detection in primary healthcare settings, including county- and township-level healthcare facilities, and explore performance variations across hierarchical healthcare institutions in real-world clinical practice.

Methods

Study Design

Our Study was performed in three counties (Yidu County, Zigui County, and Changyang County) of Yichang city, Hubei province, where were affected by high PTB burden, with an overall incidence of PTB nearly 100 per 100,000 population during 2022-2024 year. This study involved in patients (including outpatients and inpatients) over 15 years of age who underwent CXR examinations at healthcare facilities due to symptomatic presentation. We employed a paired diagnostic accuracy design in which each chest radiograph received two independent assessments from: (1) AI-powered computer-aided detection; and (2) Clinical radiologists who generated original diagnostic reports during routine patient care, thereby reflecting real-world interpretation conditions. To address the issue of multiple CXRs potentially occurring for the same patient during the study period, we established a stratified selection protocol: for non-PTB cases, the initial CXR examination was utilized; for bacteriologically confirmed PTB cases, the CXR temporally closest to the confirmation date was selected. To reduce temporal misclassification bias between CXR acquisition and tuberculosis confirmation, we set a 6-month window period as the validity criterion.

Details of the Retrospective Data

All CXR images were extracted in Digital Imaging and Communications in Medicine (DICOM) format from the Picture Archiving and Communication System-Radiology Information System (PACS-RIS) of county and township-level healthcare facilities in three study counties. The data collection period spanned from January 1, 2022, to December 31, 2024. This system also concurrently captured patient’s data on gender, age, and radiology report information, including examination time, report date, radiological findings, and diagnostic conclusions. Data on PTB cases were retrived from the TBIMS, established and maintained by the Chinese Center for Disease Control and Prevention. TBIMS is a population-based surveillance system that mandates reporting of all bacteriologically confirmed PTB cases or clinically diagnosed PTB cases [16]. This dataset covered the period from January 1, 2022, to June 30, 2025, and included patient diagnosis type (bacteriologically confirmed PTB or clinically diagnosed PTB) and confirmation date.

AI System for CXR Interpretation

For AI-based CXR interpretation, we used the JF CXR-1 v2 software. It is an advanced AI system built on deep learning technology, primarily used for tuberculosis detection in CXRs with external validation support [17-19]. In 2022, the software was approved by China's National Medical Products Administration (NMPA) as a Class III medical device. The system generates probability scores ranging from 0 to 1, where higher values indicate greater likelihood of active PTB. Based on the manufacturer's clinical protocol, scores exceeding 0.35 mandate additional diagnostic evaluation and this study used >0.35 as the criterion for radiographic abnormalities consistent with active PTB.

The following steps were performed on the software:

1. DICOM images were input into the system for preprocessing: conversion to UINT8 format, contrast enhancement via histogram equalization, cropping to 1024×1024 resolution, conversion from single-channel to three-channel format for model compatibility, and pixel normalization by subtracting the mean value of 128.
2. Feature maps are extracted through the ResNet34+fpn backbone network. These feature maps then trigger parallel dual-branch processing: Branch 1 generates Logitmap via a fully connected layer and outputs lesion-localizing heatmaps through a heatmap rendering module; Branch 2 compresses features via a global pooling layer, computes logits through a fully connected layer, and derives disease probability values via the Sigmoid function.

Primary Indicators

We evaluated diagnostic performance using two primary indicators: diagnostic yield among all those diagnosed (DYD) and positive predictive value (PPV). DYD is defined as the proportion of confirmed tuberculosis patients who have received CXR examination in whom tuberculosis was detected by a diagnostic tool (CXR identified by CAD or radiologists) [20]. This indicator uniquely quantifies detection efficacy without requiring universal bacteriological verification, relying solely on a confirmed patient subset. PPV is defined as the proportion of patients confirmed to have tuberculosis among all results identified as positive by a diagnostic tool (CXR identified by CAD or radiologists).

Statistical Analysis

Descriptive Statistics

We summarized baseline characteristics of patients, presenting categorical variables as frequencies (n) and proportions (%) and comparing them using the chi-square test or Fisher's exact test. Non-normally distributed continuous variables were described as median with interquartile range (IQR) and analyzed using the Wilcoxon rank-sum test.

Diagnostic Performance

The ability of CAD and radiologists to identify PTB cases was evaluated using bacteriologically confirmed PTB patients from the TBIMS. To explore collaborative workflows between CAD and radiologists, performance was assessed under five distinct strategies: Radiologist interpretation alone; CAD results alone; positive if eithermethod flagged abnormality; positive only if bothmethods flagged abnormality; and CAD screening followed by radiologist verification: positive CAD results reviewed by radiologists. For CAD-positive patients, the primary CAD scores were divided into four classes according to quartiles. Tests for linear trend between score quartiles and the distribution of PTB patients (suspected by radiologists or confirmed) were conducted using Cochran–Armitage (chi-square) tests. Subgroup analysis evaluated the DYD and PPV of CAD across diverse demographic characteristics and clinical settings, with multivariate logistic regression models explaining the statistical significance of differences in DYD and PPV between subgroups.

Sensitivity analysis

Sensitivity analysis was performed to assess the robustness of CAD and radiologist performance in identifying PTB case using CXR under different scenarios: (1) time window validation: assessment of different PTB confirmation intervals (3-month and 9-month vs. the primary 6-month window); (2) image selection validation: impact of image selection strategy (using the last CXR examination vs. the initial CXR examination for non-TB patients); (3) case definition expansion: inclusion of clinically diagnosed PTB cases, to re-evaluate the performance of CAD and radiologists.

Ethical Considerations

The study was ethically approved by the Ethics Committee of Chinese Academy of Medical Sciences & Peking Union Medical College (CAMS&PUMC-IEC-2025-04).

Results

From January 2022 to December 2024, a total of 132,159 posteroanterior (PA) chest radiographs were collected from 7 county-level hospitals and 32 township health centers across three counties (Yidu County, Zigui County, and Changyang County) in Yichang City, China. All images were accompanied by complete radiological reports. By linking PTB case data through the TBIMS, the final study population comprised 93,319 patients. The study outline is presented in Figure 1.

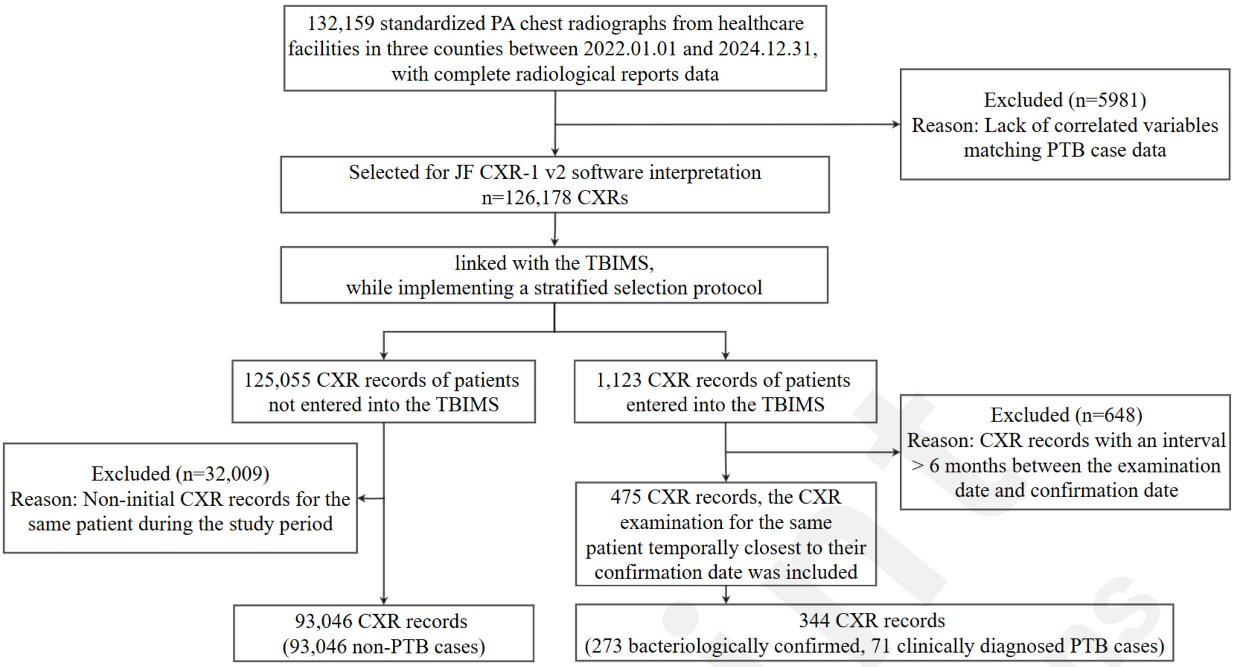


Figure 1. Flowchart depicting the study outline. PA: posteroanterior; CXR: chest x-ray; TBIMS: Tuberculosis Information Management System; We established a stratified selection protocol: for patients ultimately not classified as PTB cases, we utilized their initial CXR examination; for PTB cases, the CXR examination conducted ≤6 months preceding and temporally closest to their TBIMS-registered confirmation date.

Baseline characteristics of included patients

A total of 93,319 patients (refer to the Figure 1) were included in this study. Demographics and healthcare system classification variables are presented in Table 1, the median age of PTB patients was higher than that of non-PTB patients (69 vs. 60; P< .0001), and male gender preponderance (70.7% vs. 50.1%; P< .0001). A higher proportion of PTB patients came from township-level healthcare facilities compared to non-PTB patients (88.3% vs. 61.2%; P< .0001). The proportion of PTB patients from outpatient clinics were comparable with that of non-PTB patients (29.7% vs. 28.4%; P = .7021).

Table 1. Baseline characteristics of included patients (N=93,319).

Subgroup	Total patients (n=93,319)	Non-PTB patients (n=93,046)	Confirmed PTB patients (n=273)	*P*-value
Region, n (%)				.0013
Yidu County	35,020	34,941 (37.6%)	79 (28.9%)	
Zigui County	39,153	39,036 (42.0%)	117 (42.9%)	
Changyang County	19,146	19,069 (20.5%)	77 (28.2%)	
Age (years), Median (IQR)		60 (50-70)	69 (60-74)	< .0001[a]
Age (years), n (%)				< .0001
15-44	16,600	16,586 (17.8%)	14 (5.1%)	
45-64	40,571	40,489 (43.5%)	82 (30.0%)	
65-79	29,995	29,844 (32.1%)	151 (55.3%)	

>=80	6,153	6,127 (6.58%)	26 (9.5%)	
Sex, n (%)				< .0001
Male	46,827	46,634 (50.1%)	193 (70.7%)	
Female	46,492	46,412 (49.9%)	80 (29.3%)	
Healthcare Facility Tier, n (%)				< .0001
Township - level	57,182	56,941 (61.2%)	241 (88.3%)	
County - level	36,137	36,105 (38.8%)	32 (11.7%)	
Patients Category, n (%)				.7021
Outpatient	26,544	26,463 (28.4%)	81 (29.7%)	
Inpatient	66,775	66,583 (71.6%)	192 (70.3%)	

[a] *p* value was determined by Wilcoxon rank-sum test; others by Chi-square test for categorical variables.

Performance of PTB Detection: CAD versus Radiologists

Among 93,314 participants, 13,515 cases were identified as suspected PTB by CAD, with positive rate of 14.48% (13,515/93,319), which significantly exceeded the positive rate of suspected PTB (0.73%, 679/93,319) identified by radiologists. Among 273 confirmed PTB cases, CAD achieved DYD of 83.88% (229/273), which was substantially higher than that of radiologists (25.64%, 70/273). CAD captured all 70 PTB cases detected by radiologists, demonstrating low risk of missed detection. However, the PPV of CAD was only 1.70% (229/13,515), which was much lower than the radiologists' PPV of 10.31% (70/679) (Figure 2).

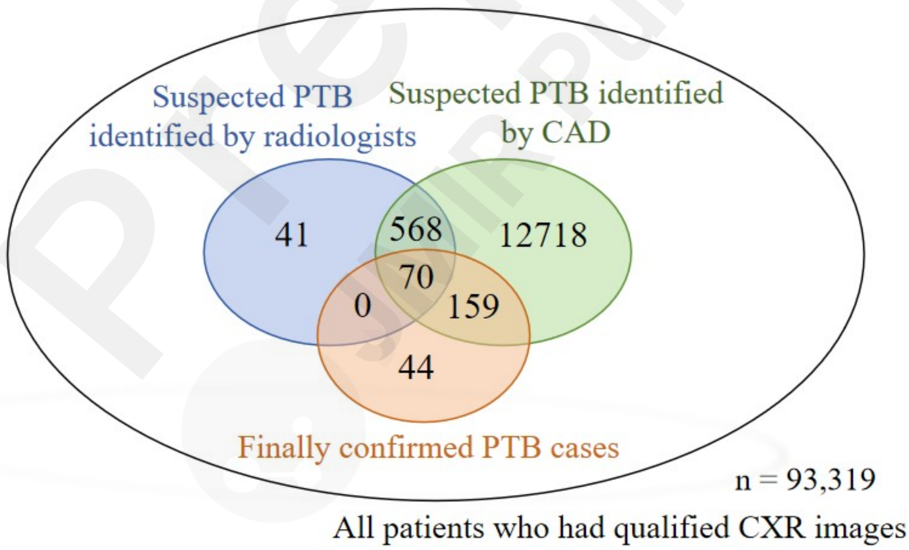


Figure 2. Distribution of PTB Cases Identified by CAD and Radiologists. CAD identified suspected PTB cases are represented in blue, radiologist identified suspected PTB cases are represented in green and finally confirmed PTB cases are represented in orange. The numbers inside intersections represent overlap cases.

Performance Evaluation of Combined Use of CAD and Radiologists

When combing the CAD and radiologists as PTB screening strategy listed in Table 2, PTB identified by radiologists or CAD (strategy 3) required radiologists to interpret all 93,319 CXR images. It achieved a PPV of 1.69% (229/13,556), reflecting a marginal decrease compared to CAD operating alone (strategy 2, PPV: 1.70%). The DYD of strategy 3 was 83.88% (229/273), consistent with CAD-alone (strategy 2) performance. PTB identified by both radiologists and CAD (strategy 4) similarly necessitated full interpretation of all CXRs. Its PPV increased to 10.97% (70/638), slightly exceeding radiologist-alone interpretation (strategy 1, PPV: 10.31%). The DYD of strategy 4 remained 25.64% (70/273), consistent with radiologist-alone (strategy 1) performance. PTB identified by radiologists among CAD abnormalities (strategy 5) demonstrated significant workflow efficiency: radiologists reviewed only CAD-positive images (13,515 CXRs; 14.48% of total volume). While this approach does not directly yield new PPV/DYD metrics, radiologists identified at least 70 PTB cases among CAD-positive images. Consequently, the final DYD of strategy 5 reached ≥25.64% (70/273), matching radiologist-alone (strategy 1) performance.

Table 2. Performance Comparison of Collaborative CAD-Radiologist Strategies for PTB Screening. CXR: chest x-ray; CAD: computer-aided detection; DYD: diagnostic yield among diagnosed. PPV: positive predictive value.

CXR-based PTB case identification strategies	Numbers of CXR to be read by radiologists n (%)	Numbers of reported suspected PTB n (%)	PPV (n/N)	DYD (n/N)
Strategy1: Identified by radiologists alone	93319 (100%)	679 (0.73%)	10.31% (70/679)	25.64% (70/273)
Strategy2: Identified by CAD alone	/	13515 (14.48%)	1.70% (229/13515)	83.88% (229/273)
Strategy3: Identified by radiologists or CAD	93319 (100%)	13556 (14.53%)	1.69% (229/13556)	83.88% (229/273)
Strategy4: Identified by radiologists and CAD	93319 (100%)	638 (0.68%)	10.97% (70/638)	25.64% (70/273)
Strategy5: Identified by radiologists among CAD abnormalities	13515 (14.48%)	≥ 638 (4.72%)	/	≥ 25.64% (70/273)

Distribution of CAD scores in suspected PTB cases

Among patients with abnormal CAD, each patient's CAD score was classified into four subgroups by quartiles. The distribution of CAD scores was found to be linearly correlated with the proportion of suspected PTB identified by radiologists and confirmed PTB cases (*P* for trend < .001). Most confirmed PTB patients (82.53%) had CAD scores exceeding the 75% quantiles (Q75) threshold, which accounting for 34.17% of all CAD-positive patients (Table 3).

Table 3. Distribution of abnormal CAD scores in suspected PTB cases identified by radiologists and confirmed PTB cases.

CAD score	CAD-positive patients n (%)	Suspected PTB cases identified by radiologists n (%)	Confirmed PTB cases n (%)

Total	**13515**	**638**	**229**
Classified by quartiles			
0.35 < score≤0.44	2828 (20.92%)	13 (2.04%)	6 (2.62%)
0.44 < score≤0.56	2995 (22.16%)	26 (4.08%)	11 (4.80%)
0.56 < score≤0.75	3074 (22.75%)	70 (10.97%)	23 (10.04%)
score > 0.75	4618 (34.17%)	529 (82.92%)	189 (82.53%)
P [a] for trend		**< .001**	**< .001**

[a] Tests for linear trend between score quartiles and the distribution of PTB patients (suspected by radiologists or confirmed) were conducted using Cochran–Armitage (chi-square) tests.

Subgroup Analysis of PTB Detection Performance by CAD and radiologists

CAD consistently demonstrated superior DYD compared to radiologists across all evaluated demographic and healthcare system classification variables (Figure 3). However, CAD showed significant variation in performance with respect to healthcare facility tier, performing poorer in the county-level group than in the township-level group ($P< .001$; Multimedia Appendix. Table 1). Multivariate logistic regression also demonstrated that advanced age (≥80 years) groups are less likely to achieve high PPV (OR=0.22; 95% CI: 0.09–0.53; $P< .001$). No significant differences were observed between the sexes and patients categories (Outpatient / Inpatient).

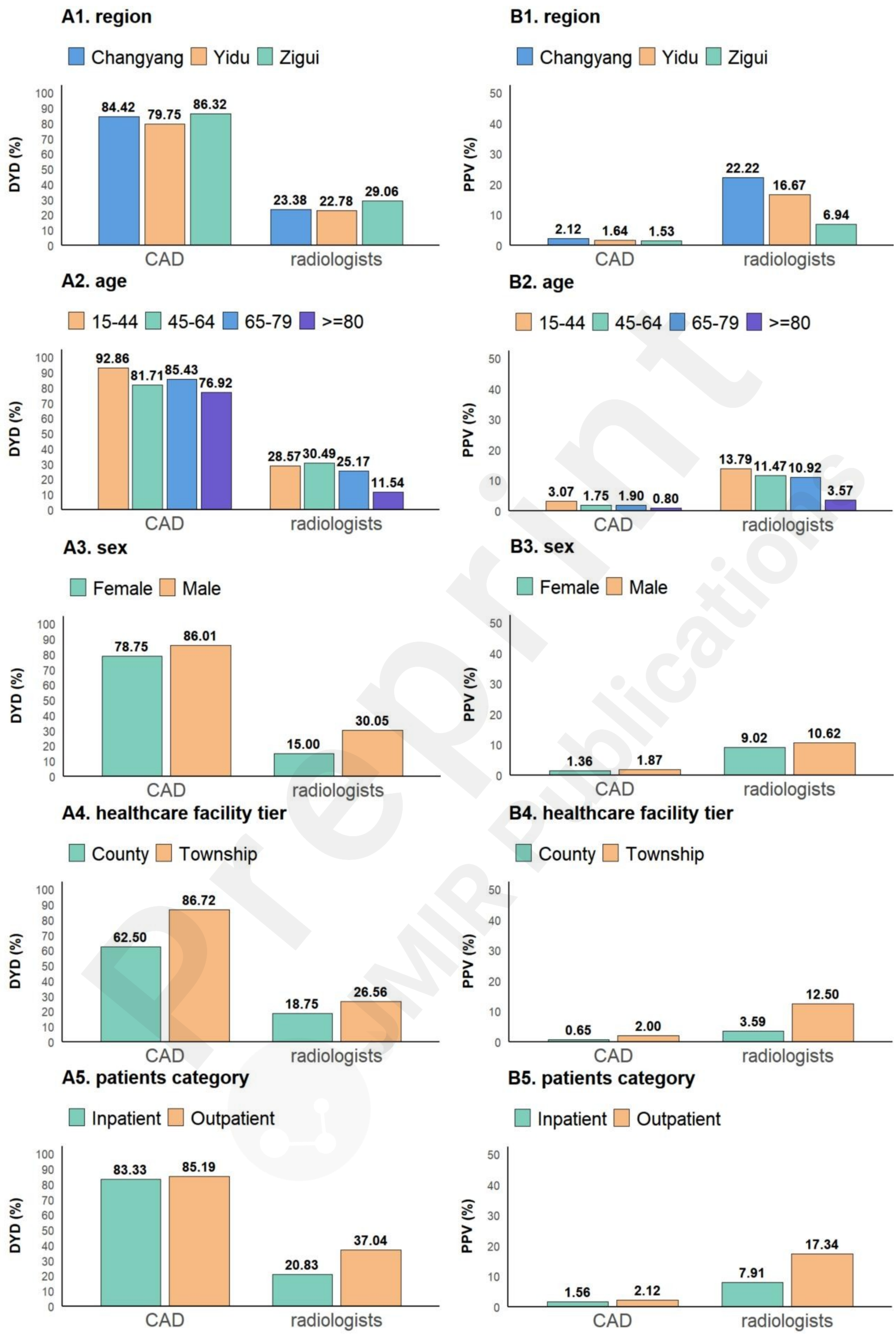


Figure 3. Comparion of DYD and PPV of CAD and radiologists by subgroups. Measures are disaggregated by A1 and B1: region; A2 and B2: age; A3 and B3: sex; A4 and B4: healthcare facility tier; A5 and B5: patients category. DYD=diagnostic yield among diagnosed. PPV=positive predictive value.

Sensitivity Analysis

During temporal window validation, shortening the PTB confirmation window to 3 months slightly increased CAD's DYD to 86.58% (increased to 86.58% from 83.88%; Multimedia Appendix. Table 2), whereas extending it to 9 months reduced DYD to 81.76% (decreased to 81.76% from 83.88%; Multimedia Appendix. Table 3), indicating a minor influence of temporal factors on performance. Image selection validation (using final CXRs vs. initial CXRs for non-PTB cases) revealed only subtle differences with unchanged core conclusions (Multimedia Appendix. Table 4), confirming the reliability of the original strategy. When clinically diagnosed PTB cases (n=71) were included (Multimedia Appendix. Table 5), the DYD of CAD decreased by only 2.78% (decreased to 81.10% from 83.88%), remaining significantly higher than that of radiologists (81.10% vs. 25.58%). The DYD of radiologists decreased slightly (decreased to 25.58% from 25.64%). The consistent trends in both groups confirm that the relative advantage of CAD was not substantially affected.

Discussion

Principal Results

This retrospective study evaluated the real-world performance of CAD in PTB detection using CXR interpretation. Among 93,319 patients analyzed from healthcare facilities, 273 were confirmed as bacteriologically confirmed PTB cases within three years via the TBMIS. The results demonstrated that CAD outperformed radiologists in identifying TB cases, detecting an additional 159 cases. Compared to "radiologist interpretation alone," the strategy of "CAD screening followed by radiologist verification" not only enhanced PTB case identification but also reduced the workload for radiologists. A CAD score exceeding 0.75 was a key threshold for identifying high-risk PTB patients, demonstrating high consistency with radiologists' judgments and supported by statistical trends confirming significantly higher PTB risk with elevated scores. Notably, CAD exhibited superior diagnostic performance in township medical institutions compared to county-level hospitals, effectively overcoming resource constraints inherent.

In the identification of PTB via CXRs, CAD demonstrate significant advantages over radiologists. This disparity primarily stems from CAD's highly sensitive algorithms, which effectively detect early-stage and subtle lesions [21]. In contrast, radiologists face limitations due to inherent constraints in human visual recognition, visual fatigue from prolonged image interpretation, and insufficient vigilance toward PTB under clinical resource constraints, resulting in considerably higher missed diagnosis rates. However, a persistent limitation of CAD products is the high false positive rate [22]. If relying solely on CAD, this increases unnecessary diagnostic follow-up burden, thereby causing waste of medical resources. In resource-constrained primary care settings, CAD systems can serve as a triage tool for rapid PTB diagnosis [22]. By adopting a strategy of "computer-aided screening followed by selective radiologist review", this approach optimizes the tuberculosis triage workflow. Such a stepwise process can effectively alleviate the workload of manual film reading while maintaining high sensitivity and reducing false-positive results [23]. Consistent with previous studies, CAD scores not only function as a diagnostic threshold but also possess predictive value for PTB risk based on continuous score variations [24]. CAD interpretation outcomes demonstrate high consistency with radiologists' chest X-ray readings, providing an objective reference for clinical decision-making [25-27].

CAD exhibit significant performance variations across medical institution tiers, necessitating tailored deployment strategies. In township medical facilities, CAD demonstrate higher DYD and PPV. Therefore, this tier is optimally suited for implementing a hybrid diagnostic model prioritizing "radiologist-assisted CAD interpretation". However, county hospitals require the integration of supplementary diagnostic methods, such as CT scans, alongside the hybrid model. This adaptation addresses the more significant false positive challenge encountered at this level. The difference may be related to the types of cases treated. Township institutions primarily handle common illnesses such as routine respiratory diseases, while county-level hospitals concentrate on more complex and difficult cases. The performance difference in CAD between the two tiers of hospitals might stem from limitations in recognizing complex cases, but this hypothesis requires further validation through more comprehensive studies.

Comparison With Prior Work

Previous research has centered on two PTB screening models: active screening in high-risk groups (e.g., high-burden populations [18], cross-border migrants [28], enclosed environment [29]); Passive screening in specialized facilities (e.g., tertiary hospitals [11], TB diagnostic centers [30]). This study pioneers PTB triage integration within China’s hierarchical healthcare system, offering novel evidence for AI deployment in primary care. While Kagujje et al. [31] also studied primary settings, their focus was on "prior TB history" as an AI diagnostic confounder. The observed DYD of JF CXR-1 v2 (83.88%) aligns with high sensitivities reported previously (e.g., Qin et al. [19]: 86.4%). Compared with the sensitivity shown in clinical validation (78.9% [12]), there is still a certain difference in performance in primary healthcare. Our proposed "AI pre-screening with selective human review" model, in collaboration with Munjal P et al. [32], reduces radiologist workload by 80%. Based on the TBIMS data, Xin et al. [24] confirmed the effectiveness of CAD in individuals aged 65 and above. This study further supplements evidence for the efficacy of CAD in other application scenarios.

Limitations

This study has several limitations. First, this study is subject to potential temporal misclassification bias due to the inherent lag between CXR acquisition and definitive PTB diagnosis. The analyzed CXRs were not necessarily obtained on the day of PTB confirmation but could precede it by up to 6 months. This temporal gap may have systematically underestimated CAD's sensitivity, as early-stage lesions might lack sufficient radiographic conspicuity for reliable CAD detection. The limitation empirically evidenced by our sensitivity analysis where shortening the confirmation window to 3 months increased CAD's DYD by 2.7% (86.58% vs. 83.88%). Second, the absence of follow-up verification for CAD-positive cases not clinically diagnosed as PTB (e.g., via additional microbiological testing or CT scans) precludes differentiation between true CAD false positives and occult TB cases missed by conventional diagnostics. This gap obscures CAD's actual impact on case detection enhancement. Finally, while we employed the clinically validated threshold (>0.35) for the JF CXR-1 v2 system, this does not constitute a study limitation but rather aligns with real-world clinical protocols. Future research should explore threshold optimization strategies without diminishing this study's focus on operational validity.

Conclusions

CAD dramatically improves PTB case identification in primary healthcare settings, while substantially reducing radiologist workload through a selective review strategy focused on CAD-positive images with reference to the probability score. Meanwhile, CAD demonstrates significantly superior performance in township medical facilities compared to county medical facilities, providing validated evidence for implementing tiered diagnosis and treatment models in resource-limited regions.

Acknowledgments

The authors would like to thank our funders China Medical Board, Peking Union Medical College Education Foundation, National Key Research and Development Program of China, and Jiangxi Zhongke Jiufeng Smart Medical Technology Co., Ltd. The authors were allowed to use the AI algorithms free of charge by Jiangxi Zhongke Jiufeng Smart Medical Technology Co., Ltd. for

research purposes, but the company had no influence over the research question, nor any other aspect of the work carried out, and had no impact on the transparency of the Article.

Funding

This study was supported by China Medical Board Open Competition Grants #24–564; Chinese Academy of Medical Sciences (CAMS) Innovation Fund for Medical Sciences (Grant numbers: 2021-I2M-1-044); National Key Research and Development Program of China (Grant number: 2023YFC2308701); Peking Union Medical College Education Foundation; Public Health Emergency Management Innovation Center; State Key Laboratory of Respiratory Health and Multimorbidity, Key Laboratory of Pathogen Infection Prevention and Control (Peking Union Medical College), Ministry of Education.

Conflicts of Interest

None declared.

Data Availability

The datasets generated or analyzed in this manuscript are not publicly available. However, any necessary clarifications or additional information is available from the corresponding author on reasonable request.

Authors' Contributions

WJJ and HZ contributed equally to this work and are co-first authors. WJJ and ZJL contributed to the conceptualization, formal analysis, investigation, and visualization of the study. HZ collaboratively implemented the study protocol, and managed multi-center data coordination. YW, JHL, and ZJL participated in the review and editing process, and supervised the overall project. WJJ, XLY, and JJX led fieldwork coordination and on-site implementation. The first draft of the manuscript was written by WJJ, and ZLL, XLJ, JMS, HZ, HSW, JXY, and XYS provided essential intellectual revisions and refinement suggestions for the manuscript. All authors read and approved the final manuscript.

Abbreviations

AI: artificial intelligence

CAD: computer-aided detection

CXR: chest X-ray

DYD: diagnostic yield among all those diagnosed

MTB: mycobacterium tuberculosis

MTBC：mycobacterium tuberculosis complex

NMPA: National Medical Products Administration

PPV: positive predictive value

PTB: pulmonary tuberculosis

TB: tuberculosis

TBIMS: Tuberculosis Information Management System

Multimedia Appendix

References

1. Motta I, Boeree M, Chesov D, Dheda K, Günther G, Horsburgh CR, Jr., et al. Recent advances in the treatment of tuberculosis. Clin Microbiol Infect. 2024 Sep;30(9):1107-14. PMID: 37482332. doi: 10.1016/j.cmi.2023.07.013.

2. Mao X, Wang J, Xu J, Xu P, Hu H, Li L, et al. Current diagnosing strategies for Mycobacterium tuberculosis and its drug resistance: a review. J Appl Microbiol. 2025 May 2;136(5). PMID: 40343775. doi: 10.1093/jambio/lxaf100.

3. Goletti D, Meintjes G, Andrade BB, Zumla A, Shan Lee S. Insights from the 2024 WHO Global Tuberculosis Report - More Comprehensive Action, Innovation, and Investments required for achieving WHO End TB goals. Int J Infect Dis. 2025 Jan;150:107325. PMID: 39631498. doi: 10.1016/j.ijid.2024.107325.

4. Ding C, Ji Z, Zheng L, Jin X, Ruan B, Zhang Y, et al. Population-based active screening strategy contributes to the prevention and control of tuberculosis. Zhejiang Da Xue Xue Bao Yi Xue Ban. 2022 Dec 25;51(6):669-78. PMID: 36915974. doi: 10.3724/zdxbyxb-2022-0426.

5. Yayan J, Franke KJ, Berger M, Windisch W, Rasche K. Early detection of tuberculosis: a systematic review. Pneumonia (Nathan). 2024 Jul 5;16(1):11. PMID: 38965640. doi: 10.1186/s41479-024-00133-z.

6. Hwang EJ, Park S, Jin KN, Kim JI, Choi SY, Lee JH, et al. Development and Validation of a Deep Learning-based Automatic Detection Algorithm for Active Pulmonary Tuberculosis on Chest Radiographs. Clin Infect Dis. 2019 Aug 16;69(5):739-47. PMID: 30418527. doi: 10.1093/cid/ciy967.

7. Pinto LM, Pai M, Dheda K, Schwartzman K, Menzies D, Steingart KR. Scoring systems using chest radiographic features for the diagnosis of pulmonary tuberculosis in adults: a systematic review. Eur Respir J. 2013 Aug;42(2):480-94. PMID: 23222871. doi: 10.1183/09031936.00107412.

8. Mollura DJ, Culp MP, Pollack E, Battino G, Scheel JR, Mango VL, et al. Artificial Intelligence in Low- and Middle-Income Countries: Innovating Global Health Radiology. Radiology. 2020 Dec;297(3):513-20. PMID: 33021895. doi: 10.1148/radiol.2020201434.

9. Du J, Su Y, Qiao J, Gao S, Dong E, Wang R, et al. Application of artificial intelligence in diagnosis of pulmonary tuberculosis. Chin Med J (Engl). 2024 Mar 5;137(5):559-61. PMID: 38369956. doi: 10.1097/cm9.0000000000003018.

10. Li LS, Yang L, Zhuang L, Ye ZY, Zhao WG, Gong WP. From immunology to artificial intelligence: revolutionizing latent tuberculosis infection diagnosis with machine learning. Mil Med Res. 2023 Nov 28;10(1):58. PMID: 38017571. doi: 10.1186/s40779-023-00490-8.

11. Tavaziva G, Harris M, Abidi SK, Geric C, Breuninger M, Dheda K, et al. Chest X-ray Analysis With Deep Learning-Based Software as a Triage Test for Pulmonary Tuberculosis: An Individual Patient Data Meta-Analysis of Diagnostic Accuracy. Clin Infect Dis. 2022 Apr 28;74(8):1390-400. PMID: 34286831. doi: 10.1093/cid/ciab639.

12. Yang Y, Xia L, Liu P, Yang F, Wu Y, Pan H, et al. A prospective multicenter clinical research study validating the effectiveness and safety of a chest X-ray-based pulmonary tuberculosis screening software JF CXR-1 built on a convolutional neural network algorithm. Front Med (Lausanne). 2023;10:1195451. PMID: 37649977. doi: 10.3389/fmed.2023.1195451.

13. Breuninger M, van Ginneken B, Philipsen RH, Mhimbira F, Hella JJ, Lwilla F, et al. Diagnostic accuracy of computer-aided detection of pulmonary tuberculosis in chest radiographs: a validation study from sub-Saharan Africa. PLoS One. 2014;9(9):e106381. PMID: 25192172. doi: 10.1371/journal.pone.0106381.

14. Tavaziva G, Majidulla A, Nazish A, Saeed S, Benedetti A, Khan AJ, et al. Diagnostic accuracy of a commercially available, deep learning-based chest X-ray interpretation software for detecting culture-confirmed pulmonary tuberculosis. Int J Infect Dis. 2022 Sep;122:15-20. PMID: 35597555. doi: 10.1016/j.ijid.2022.05.037.

15. Bhandari A. Revolutionizing Radiology With Artificial Intelligence. Cureus. 2024 Oct;16(10):e72646. PMID: 39474591. doi: 10.7759/cureus.72646.

16. Shao J, Zhang H, Wang Y, Su X, Xin H, Zhou P, et al. Diagnostic Delay Among Pulmonary Tuberculosis Patients Before, During and After COVID-19 Pandemic in Yichang City, China: A Longitudinal Study Based on Tuberculosis Surveillance Data. J Epidemiol Glob Health. 2025 May 26;15(1):74. PMID: 40418412. doi: 10.1007/s44197-025-00419-5.

17. 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-e54. PMID: 34446265. doi: 10.1016/s2589-7500(21)00116-3.
18. Codlin AJ, Dao TP, Vo LNQ, Forse RJ, Van Truong V, Dang HM, et al. Independent evaluation of 12 artificial intelligence solutions for the detection of tuberculosis. Sci Rep. 2021 Dec 13;11(1):23895. PMID: 34903808. doi: 10.1038/s41598-021-03265-0.
19. Qin ZZ, Van der Walt M, Moyo S, Ismail F, Maribe P, Denkinger CM, et al. Computer-aided detection of tuberculosis from chest radiographs in a tuberculosis prevalence survey in South Africa: external validation and modelled impacts of commercially available artificial intelligence software. Lancet Digit Health. 2024 Sep;6(9):e605-e13. PMID: 39033067. doi: 10.1016/s2589-7500(24)00118-3.
20. Broger T, Marx FM, Theron G, Marais BJ, Nicol MP, Kerkhoff AD, et al. Diagnostic yield as an important metric for the evaluation of novel tuberculosis tests: rationale and guidance for future research. Lancet Glob Health. 2024 Jul;12(7):e1184-e91. PMID: 38876764. doi: 10.1016/s2214-109x(24)00148-7.
21. Cao XF, Li Y, Xin HN, Zhang HR, Pai M, Gao L. Application of artificial intelligence in digital chest radiography reading for pulmonary tuberculosis screening. Chronic Dis Transl Med. 2021 Mar;7(1):35-40. PMID: 34013178. doi: 10.1016/j.cdtm.2021.02.001.
22. Hua D, Nguyen K, Petrina N, Young N, Cho JG, Yap A, et al. Benchmarking the diagnostic test accuracy of certified AI products for screening pulmonary tuberculosis in digital chest radiographs: Preliminary evidence from a rapid review and meta-analysis. Int J Med Inform. 2023 Sep;177:105159. PMID: 37549498. doi: 10.1016/j.ijmedinf.2023.105159.
23. de Vries G, Gainaru D, Keizer S, Mahler B, Radulescu I, Zamfirescu M, et al. Human reading versus computer-automated reading of chest radiographs in a tuberculosis screening programme in Romania. Eur Respir J. 2021 Jul;58(1). PMID: 33692121. doi: 10.1183/13993003.04628-2020.
24. Xin H, Wang W, Wang X, Huang J, Di Y, Du J, et al. The performance of computer-aided detection for chest radiography in tuberculosis screening: a population-based retrospective cohort study. Emerg Microbes Infect. 2025 Dec;14(1):2470998. PMID: 40260691. doi: 10.1080/22221751.2025.2470998.
25. WHO Guidelines Approved by the Guidelines Review Committee. WHO consolidated guidelines on tuberculosis: Module 2: screening – systematic screening for tuberculosis disease. Geneva: World Health Organization
© World Health Organization 2021.; 2021.
26. Okada K, Yamada N, Takayanagi K, Hiasa Y, Kitamura Y, Hoshino Y, et al. Applicability of artificial intelligence-based computer-aided detection (AI-CAD) for pulmonary tuberculosis to

community-based active case finding. Trop Med Health. 2024 Jan 2;52(1):2. PMID: 38163868. doi: 10.1186/s41182-023-00560-6.

27. Mahler B, de Vries G, van Hest R, Gainaru D, Menezes D, Popescu G, et al. Use of targeted mobile X-ray screening and computer-aided detection software to identify tuberculosis among high-risk groups in Romania: descriptive results of the E-DETECT TB active case-finding project. BMJ Open. 2021 Aug 24;11(8):e045289. PMID: 34429305. doi: 10.1136/bmjopen-2020-045289.

28. Gelaw SM, Kik SV, Ruhwald M, Ongarello S, Egzertegegne TS, Gorbacheva O, et al. Diagnostic accuracy of three computer-aided detection systems for detecting pulmonary tuberculosis on chest radiography when used for screening: Analysis of an international, multicenter migrants screening study. PLOS Glob Public Health. 2023;3(7):e0000402. PMID: 37450425. doi: 10.1371/journal.pgph.0000402.

29. Soares TR, Oliveira RD, Liu YE, Santos ADS, Santos P, Monte LRS, et al. Evaluation of chest X-ray with automated interpretation algorithms for mass tuberculosis screening in prisons: a cross-sectional study. Lancet Reg Health Am. 2023 Jan;17:100388. PMID: 36776567. doi: 10.1016/j.lana.2022.100388.

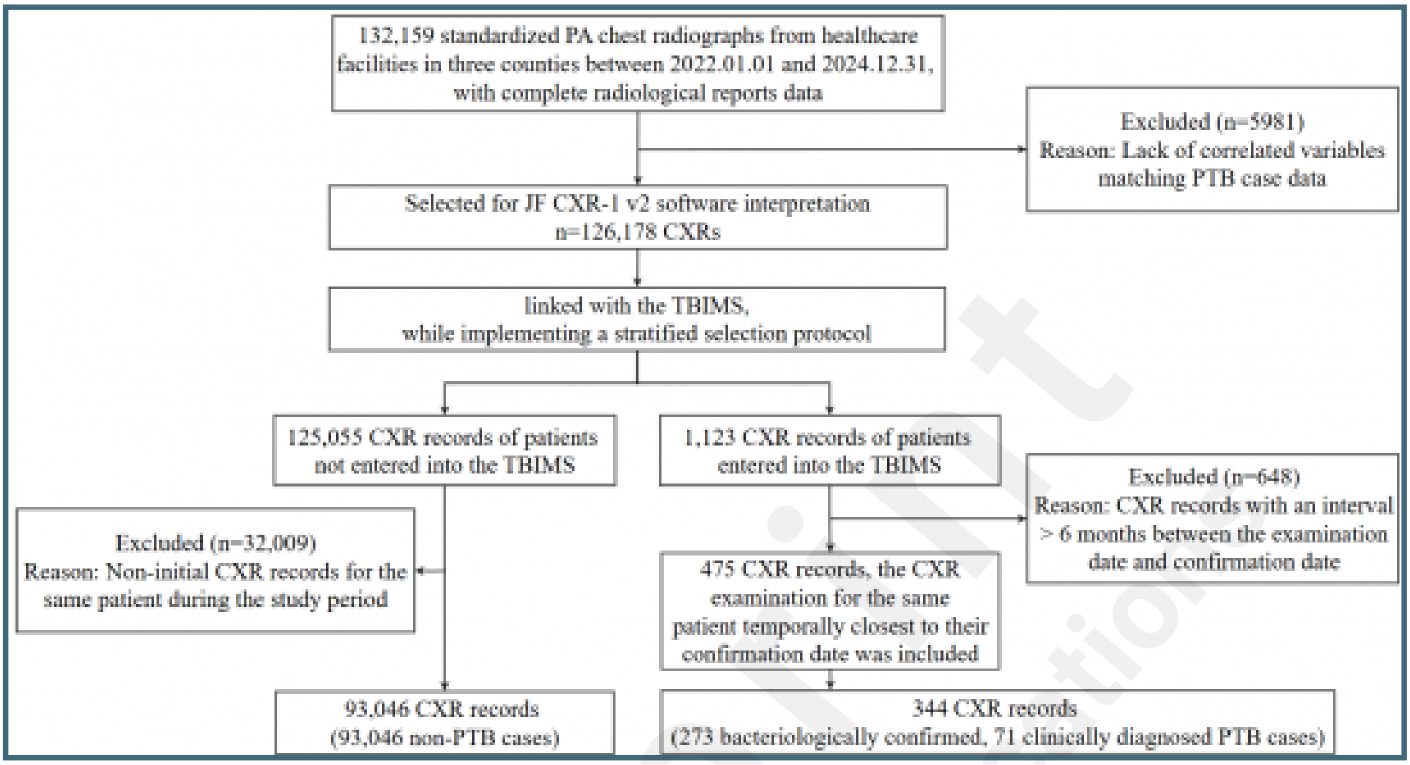
30. Murphy K, Habib SS, Zaidi SMA, Khowaja S, Khan A, Melendez J, et al. Computer aided detection of tuberculosis on chest radiographs: An evaluation of the CAD4TB v6 system. Sci Rep. 2020 Mar 26;10(1):5492. PMID: 32218458. doi: 10.1038/s41598-020-62148-y.

31. Kagujje M, Kerkhoff AD, Nteeni M, Dunn I, Mateyo K, Muyoyeta M. The Performance of Computer-Aided Detection Digital Chest X-ray Reading Technologies for Triage of Active Tuberculosis Among Persons With a History of Previous Tuberculosis. Clin Infect Dis. 2023 Feb 8;76(3):e894-e901. PMID: 36004409. doi: 10.1093/cid/ciac679.

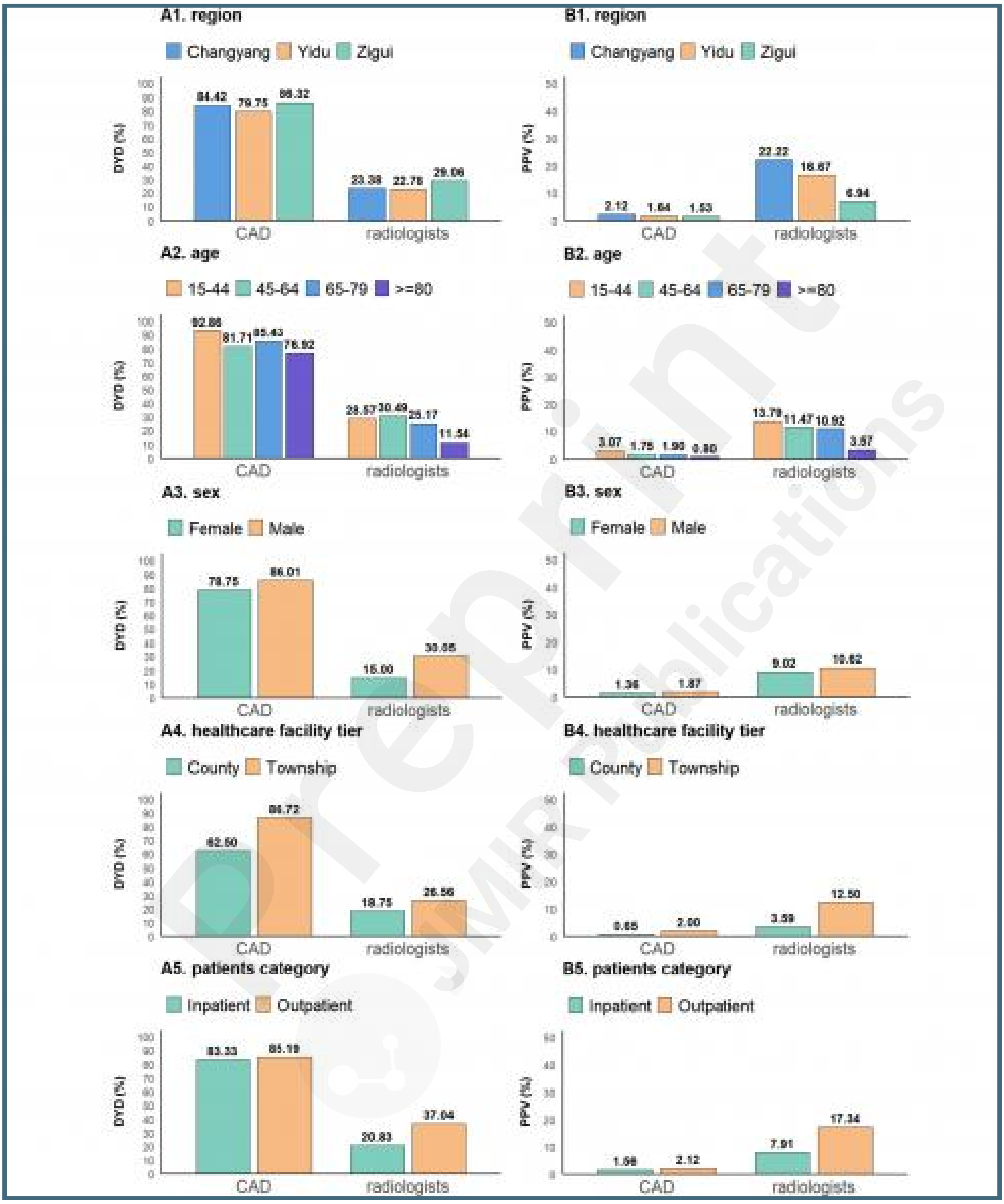
32. Munjal P, Mahrooqi AA, Rajan R, Jeremijenko A, Ahmad I, Akhtar MI, et al. Population-scale cross-sectional observational study for AI-powered TB screening on one million CXRs. NPJ Digit Med. 2025 Jul 9;8(1):418. PMID: 40634545. doi: 10.1038/s41746-025-01832-7.

Supplementary Files

Untitled.


132,159 standardized PA chest radiographs from healthcare facilities in three counties between 2022.01.01 and 2024.12.31, with complete radiological reports data
Excluded (n=5981) Reason: Lack of correlated variables matching PTB case data
Selected for JF CXR-1 v2 software interpretation n=126,178 CXRs
linked with the TBIMS, while implementing a stratified selection protocol
125,055 CXR records of patients not entered into the TBIMS
1,123 CXR records of patients entered into the TBIMS
Excluded (n=648) Reason: CXR records with an interval > 6 months between the examination date and confirmation date
Excluded (n=32,009) Reason: Non-initial CXR records for the same patient during the study period
475 CXR records, the CXR examination for the same patient temporally closest to their confirmation date was included
93,046 CXR records (93,046 non-PTB cases)
344 CXR records (273 bacteriologically confirmed, 71 clinically diagnosed PTB cases)

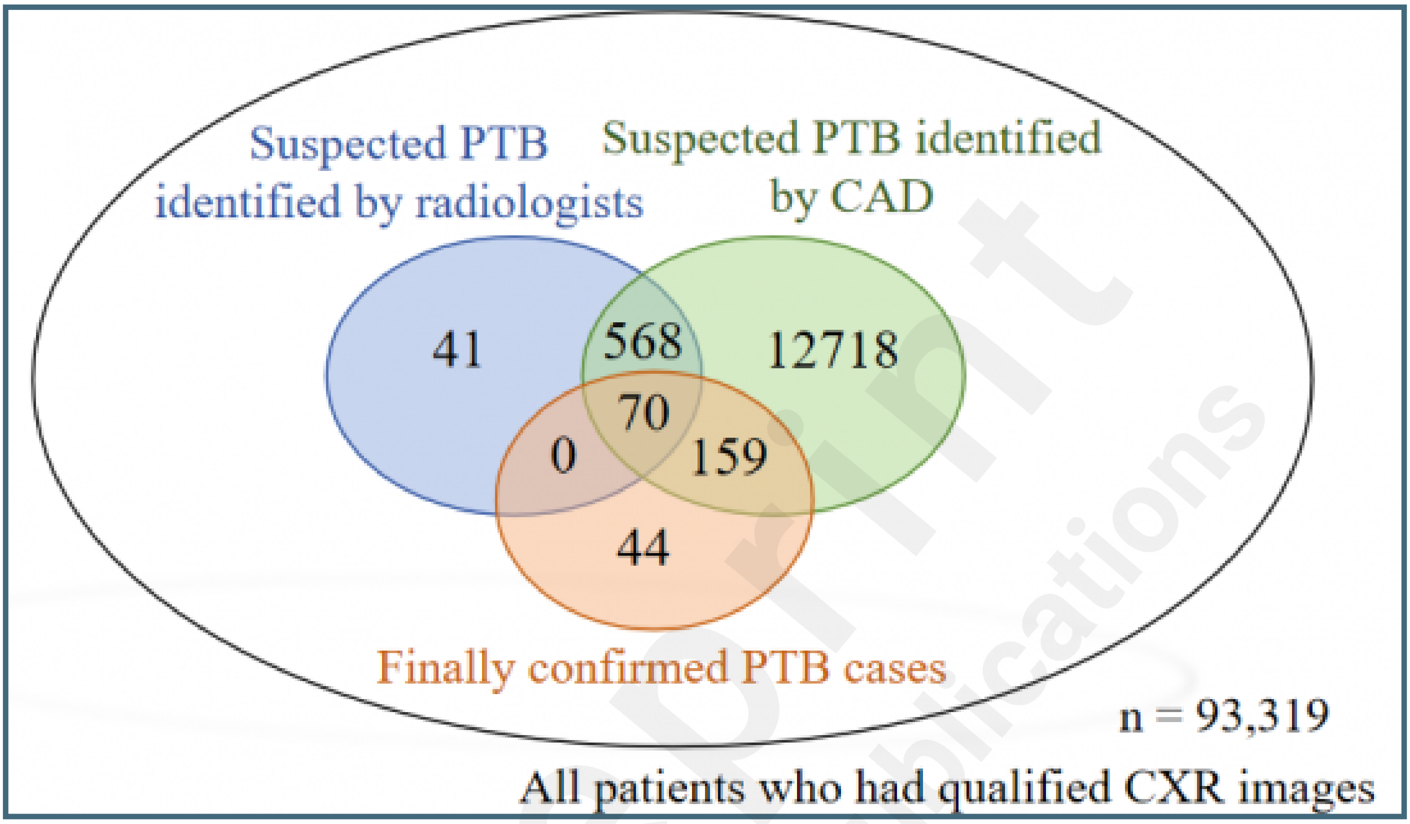
Untitled.



A1. region
Changyang Yidu Zigui
DYD (%)
84.42 79.75 86.32
23.38 22.78 29.06
CAD radiologists
B1. region
Changyang Yidu Zigui
PPV (%)
2.12 1.64 1.53
22.22 16.67 6.94
CAD radiologists
A2. age
15-44 45-64 65-79 >=80
DYD (%)
92.86 81.71 85.43 76.92
28.57 30.49 25.17 11.54
CAD radiologists
B2. age
15-44 45-64 65-79 >=80
PPV (%)
3.07 1.75 1.90 0.80
13.79 11.47 10.92 3.57
CAD radiologists
A3. sex
Female Male
DYD (%)
78.75 86.01
15.00 30.05
CAD radiologists
B3. sex
Female Male
PPV (%)
1.36 1.87
9.02 10.62
CAD radiologists
A4. healthcare facility tier
County Township
DYD (%)
62.50 86.72
18.75 26.56
CAD radiologists
B4. healthcare facility tier
County Township
PPV (%)
0.65 2.00
3.59 12.50
CAD radiologists
A5. patients category
Inpatient Outpatient
DYD (%)
83.33 85.19
20.83 37.04
CAD radiologists
B5. patients category
Inpatient Outpatient
PPV (%)
1.56 2.12
7.91 17.34
CAD radiologists

Figures

Distribution of PTB Cases Identified by CAD and Radiologists. CAD identified suspected PTB cases are represented in blue, radiologist identified suspected PTB cases are represented in green and finally confirmed PTB cases are represented in orange. The numbers inside intersections represent overlap cases.



Multimedia Appendixes

Untitled.
URL: http://asset.jmir.pub/assets/7ee9b6925c49c65279a91f149b71cb2e.docx