Open Forum Infectious Diseases

MAJOR ARTICLE

Improving Tuberculosis Diagnosis Through Artificial Intelligence (CAD4TB) and Stool Xpert MTB/RIF Testing: A Prospective Study From Oromia, Ethiopia

Giacomo Guido,[1] Berhanu Gulo,[2,a] Sergio Cotugno,[1,a] Worku Nigussa,[2] Kidist Bobosha,[3] Francesco Vladimiro Segala,[1,2,3,4,a] Beatrice Zauli,[5] Birhanu Kenate Sori,[6] Giovanni Putoto,[4] Francesco Cavallin,[7] Giordano Madeddu,[5] Flavio Antonio Bobbio,[2] Nicola Veronese,[8] Azmach Biset Asmare,[9] Abdi Reta,[9] Roberta Iatta,[10] Abata Surra,[9] Mulugeta Miressa,[2] Federico Gobbi,[11,12] Lorenzo Guglielmetti,[11] Melaku Tilahun,[3] Annalisa Saracino,[1] Abdissa Alemseged,[3] Fabio Manenti,[2] and Francesco Di Gennaro[1]

[1]Department of Precision and Regenerative Medicine and Ionian Area (DiMePRe-J), Clinic of Infectious Diseases, University of Bari "Aldo Moro," Bari, Italy, [2]Doctors with Africa CUAMM, Woljsso, Ethiopia, [3]Armauer Hansen Research Institute, Mycobacterial Diseases Research Directorate, Addis Ababa, Ethiopia, [4]Operational Research Unit, Doctors with Africa CUAMM, Padova, Italy, [5]Department of Medicine, Surgery and Pharmacy, Unit of Infectious Diseases, University of Sassari, Sassari, Italy, [6]Oromia Regional Health Bureau, Addis Ababa, Ethiopia, [7]Independent statistician, Solagna, Italy, [8]Saint Camillus International University of Health Sciences, Rome, Italy, [9]St Luke Hospital, Woljsso, Ethiopia, [10]Interdisciplinary Department of Medicine, University of Bari, Bari, Italy, [11]Department of Infectious, Tropical Diseases and Microbiology, IRCCS Sacro Cuore Don Calabria Hospital, Verona, Italy, and [12]Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy

Background. Tuberculosis remains the leading cause of death by a single infectious agent globally, with Ethiopia among the highest tuberculosis- and human immunodeficiency virus/tuberculosis–burden countries. Diagnostic gaps—particularly among household contacts (HHCs) unable to expectorate—hinder early case detection. Computer-aided detection software for chest radiography and nonrespiratory molecular assays, such as stool-based Xpert MTB/RIF testing, represent promising strategies for scalable screening.

Methods. We conducted a prospective diagnostic accuracy study at St Luke Catholic Hospital, Oromia, Ethiopia, enrolling 478 participants (152 tuberculosis index patients and 326 HHCs). All HHCs ≥4 years underwent digital chest radiographic screening, with or without CAD4TB (Delft Imaging) software assistance, and provided stool and sputum samples for Xpert MTB/RIF testing. The accuracy of CAD4TB and stool Xpert testing was evaluated against sputum Xpert testing as the reference.

Results. CAD4TB showed strong diagnostic performance, with a sensitivity of 0.77 (95% confidence interval, .70–.83) and specificity of 0.93 (.90–.96). Performance was higher among adults (sensitivity and specificity, 0.79 and 0.94) than in children (0.64 and 0.92). Stool and sputum Xpert testing demonstrated high concordance (Cohen's $\kappa = 0.76$), with a sensitivity of 0.77 (95% confidence interval, .70–.84) and specificity of 0.97 (.93–.99). During the study, 10.6% of HHCs (34 of 321) were newly diagnosed microbiologically with tuberculosis.

Conclusions. The combined use of CAD4TB and stool Xpert testing significantly improves tuberculosis detection, particularly among HHCs in high-burden, low-resource settings. This strategy is especially valuable in children and adults unable to produce sputum and where radiological expertise is limited.

Keywords. artificial intelligence; CAD4TB; household contacts; tuberculosis; Xpert MTB/RIF.

Tuberculosis has returned as the foremost cause of death due to a single infectious agent and is ranked in the top 10 among all causes of death [1]. The 2024 global tuberculosis report from the World Health Organization (WHO) [1] has underlined how, after the coronavirus disease 2019 pandemic, the global trend in tuberculosis incidence is rapidly increasing, with 8 countries accounting for more than two-thirds of the people with tuberculosis in 2023. Socioeconomic indicators, such as the gross domestic product and undernourishment, are correlated closely with tuberculosis incidence; in particular, undernourishment alone accounted for 1 million new cases in 2023, followed by human immunodeficiency virus (HIV) infection, alcohol use disorder, smoking, and diabetes [2, 3].

In high-income, low-tuberculosis-burden settings, chest computed tomography plays a central role in tuberculosis diagnosis—often following an initial screening modality—whereas in low-income, high-tuberculosis-burden countries, chest radiography remains the primary diagnostic tool [4]. However, chest radiographic sensitivity and specificity vary

Received 01 August 2025; accepted 21 November 2025; published online 10 December 2025

[a]B. G., S. C., and F. V. S. contributed equally to this work.

Correspondence: Giacomo Guido, MD, PhD, Department of Precision and Regenerative Medicine and Ionian Area (DiMePRe-J), Clinic of Infectious Diseases, University of Bari "Aldo Moro," Piazza G. Cesare, 11, Bari 70124, Italy (g.guido13@phd.uniba.it).

Open Forum Infectious Diseases®

© The Author(s) 2025. Published by Oxford University Press on behalf of Infectious Diseases Society of America. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

https://doi.org/10.1093/ofid/ofaf725

considerably by population, and interpretation is especially challenging in contexts where clinicians lack access to expert radiologists [5]. To address these limitations, artificial intelligence (AI)–based computer-aided detection (CAD) software has been integrated to provide clinicians with a quantitative tuberculosis likelihood score derived from chest radiographic images [6]. CAD4TB, like other AI-based tools, has demonstrated high sensitivity for tuberculosis detection, and newer versions have achieved progressively greater specificity [7] and have been endorsed by the WHO [8].

Concurrently, the WHO is advocating for the incorporation of nonrespiratory microbiological assays—including stool-based Xpert MTB/RIF testing for *Mycobacterium tuberculosis* complex—to facilitate the identification of asymptomatic and subclinical tuberculosis cases, which are particularly prevalent among household contacts (HHCs) [9]. The rationale behind the stool Xpert assay is that paucisymptomatic individuals may be unable to produce sufficient sputum for diagnosis, yet their bronchial secretions can be swallowed and subsequently excreted in stool. This test is particularly promising in the pediatric population, to avoid the need for more invasive procedures, such as gastric aspiration [5, 10].

Ethiopia is among the countries with the highest tuberculosis and HIV/tuberculosis burden and among the countries with the lowest Sustainable Development Goal indicators, with >27% of the population living below the international poverty threshold and >24% prevalence of undernutrition [11]. Moreover, 60% of the national budget for the fight against tuberculosis remains unfunded [12]. This situation has led to a high incidence of tuberculosis and poor coverage of community-based screening among HHCs [1]. Screening HHCs enables early case detection, interrupts transmission chains, identifies subclinical infections, and informs preventive therapy strategies, which is crucial for reducing the tuberculosis incidence in high-burden settings [13].

To address these knowledge gaps, our study evaluates the triage performance of CAD4TB and the diagnostic accuracy of stool Xpert testing in pediatric and adult HHCs in Ethiopia. These findings will guide scalable, automated chest radiographic and nonrespiratory molecular testing strategies to improve tuberculosis detection.

MATERIALS AND METHODS

A prospective diagnostic test accuracy study was conducted during tuberculosis active contact investigation at St Luke Catholic Hospital, Wolisso, in Ethiopia. This is the referral hospital in the Southwest Shoa zone, Oromia Region, serving a population from a catchment area of roughly 1.4 million people. As of 2024, the hospital bed capacity was 208, with a mean of 350 patients per day visiting the outpatient department. Routine tuberculosis screening at St Luke Catholic Hospital was based on a patient-initiated pathway with assessment of signs and symptoms suggestive of tuberculosis. In line with the Ethiopian national tuberculosis guidelines, this routine screening included assessment for persistent cough (lasting ≥2 weeks), fever, night sweats, weight loss, and hemoptysis as symptoms suggestive of tuberculosis. Tuberculosis cases were identified through clinical evaluation, which included chest radiographic assessment as the final step of the diagnostic process. Chest radiographs were evaluated in digital format.

All patients diagnosed with pulmonary tuberculosis (whether inpatients or outpatients) were eligible for inclusion as index patients if they had a positive Xpert MTB/RIF result for *M tuberculosis* on sputum or stool samples. Index patients were excluded if they had been on antituberculosis therapy for >72 hours at the time of eligibility screening, to avoid treatment-related reductions in bacillary load that could lower the sensitivity of Xpert testing at enrollment, consistently with previous studies [14]. In our setting, per Ethiopian national guidelines, acid-fast bacilli microscopy was used only for follow-up evaluation and not for diagnostic confirmation.

HHCs aged ≥5 years were eligible for inclusion if they lived or worked in the same dwelling as the index patient during the 2 months prior to tuberculosis diagnosis in the index patient [5]. Pregnant women were eligible for the inclusion as well but, to minimize radiation exposure risk to the fetus, chest radiography was not offered. HHCs already receiving treatment for active or latent tuberculosis were excluded from the study.

This study has been approved by the Oromia Health Bureau Research Ethics Committee (reference no. BFO/MBTFH/1-16/856). All information obtained was confidential. Eligible subjects received brief information about the study before being asked to participate; for children and adolescents aged 5–17 years, written consent was provided by the legal guardian and assent by the participant before any study procedure. The study was conducted according to the principles of the Helsinki Declaration. The protocol for this study was prospectively registered on ClinicalTrials.gov (NCT05818059) and has been published in an international peer-reviewed journal [15].

Close Contact Recruitment, Sequence Randomization, and Screening

Close contacts were identified via administration of a screening questionnaire to index patients. To minimize recruitment losses, close contacts were approached by research assistants and were offered transportation to St Luke Catholic Hospital and meal reimbursement to enter the screening program.

After enrollment, each study participant was assessed according to sequence AB (assessment by clinician 1 without CAD and by clinician 2 with CAD) or sequence BA (assessment by clinician 1 with CAD and by clinician 2 without CAD). Allocation to sequence AB or sequence BA was performed through randomization.

St Luke Catholic Hospital is provided with a digital radiographic system capable of delivering digital chest radiography and compatible with the CAD4TB cloud system. Once enrolled, all participants were screened for the presence of any symptom suggestive of tuberculosis, that is, cough of any duration, hemoptysis, fever, poor weight gain or weight loss, night sweats, chest pain, and shortness of breath.

Chest radiography was performed with posteroanterior imaging, following the procedures detailed by local and international guidelines to minimize radiation exposure [16]. Images were then uploaded on the CAD4TB cloud system. The output of CAD4TB is a score between 0 and 100 (where 0 represents normal and 100 very abnormal), indicating the likelihood that the person on the image has tuberculosis. Besides the abnormality score, a heat map indicating the position of the potential tuberculosis abnormalities is produced. This heat map is shown as a color overlay on top of the original chest radiograph.

A participant was labeled as a "possible tuberculosis" case according to the clinical judgment of the attending physician. Labeling was informed by all available data coming from symptom assessment and radiological abnormalities detected with or without CAD use. During assessment of possible tuberculosis cases, clinicians were asked to grade their level of confidence over the likelihood of tuberculosis on a scale from 0 to 100, following the same scale used in the CAD4Tb scoring system.

All contacts were referred for microbiological diagnosis of tuberculosis, regardless of clinical presentation and radiological findings. They were asked to provide sputum and stool samples for testing with Xpert MTB/RIF. Contacts who were not able to produce both samples were not excluded, and the inability to provide either sputum or stool was recorded. For the purposes of the present study, stool samples were processed according to the 1-step technique developed by de Haas et al [17]. Sputum samples were processed for Xpert testing according to routine local procedures and Ethiopian national guidelines [18].

Outcomes

The primary outcome was represented by the common accuracy metrics (sensitivity, specificity, positive predictive value [PPV], and negative predictive value [NPV]) when assessing the accuracy of CAD4TB compared with Xpert MTB/RIF assays. Secondary outcome measures were the following: (1) the concordance between stool and sputum Xpert testing; the common accuracy metrics (sensitivity, specificity, PPV, and NPV) when assessing the accuracy of stool Xpert compared with sputum Xpert testing and when assessing the accuracy of CAD4TB compared with Xpert MTB/RIF assays (both sputum and stool) within the strata; and the prevalence of previously unknown tuberculosis cases among participants.

Additional outcome measures including the following: (1) the additional true-positive cases when information from CAD4TB is available to the clinician; (2) the false-positive cases when information from CAD4TB is available to the clinician; (3) the number and type of discordant cases in CAD4TB versus clinician chest radiographic evaluation; (4) the association between the CAD4TB score and the clinician's confidence in his/her assessment when CAD4TB information is not available; and (5) the association between the CAD4TB score and the clinician's identification of a possible tuberculosis case when CAD4TB information is available.

Sample Size

The initial sample size was calculated assuming 90% confirmed tuberculosis diagnoses among possible tuberculosis cases; ≥34 subjects were needed to be 95% confident that such estimate would be within 10% of the true value in the population. Because we anticipated finding 15% possible tuberculosis cases among HHCs, the estimated minimum sample size included 231 HHCs [14]. However, the first interim analysis suggested a good enrollment pace but a lower proportion of confirmed tuberculosis diagnoses. Thus, to allow primary outcome evaluation, index patients were also included in the analysis. Assuming a 3:1 ratio for HHCs and index patients and an overall 60% confirmed tuberculosis diagnoses among possible tuberculosis cases in the updated sample, the new minimum sample size was 350. The study protocol was updated and approved by the local ethics committee (reference no. BFO/MBTFH/1-16/856).

Statistical Analysis

Categorical data were summarized as absolute and numerical data as median and interquartile range. The accuracy analysis included the standard measures (sensitivity, specificity, PPV, and NPV) with 95% confidence intervals (CIs). The concordance between stool and sputum Xpert testing was assessed using Cohen's κ and Gwet's AC1. CAD4TB scores were compared among clinician's identification categories using the Kruskal-Wallis test. Statistical significance was set at 5%. Statistical analysis was performed using R 4.4 software (R Foundation for Statistical Computing) [19].

RESULTS

Interim Analyses

At the 3-month interim analysis, 61 HHCs were enrolled (about 20 per month); 54 of them had valid data, 8 of 54 (15%) were judged as having possible tuberculosis, and 2 of 8 (25%) were confirmed with tuberculosis. At the 6-month interim analysis, 134 HHCs were enrolled (about 22 per month); 122 had valid data, 30 of 122 (25%) were judged as having possible tuberculosis, and 9 of 30 (30%) were confirmed with tuberculosis. At the 12-month interim analysis, 337 participants (236 HHCs and 101 index patients) were enrolled

Table 1. Characteristics of Study Participants

Characteristic	Participants, No./Total (%)[a] (N = 478)
Index patients	152/478 (31.8)
HHCs	326/478 (68.2)
Age at enrollment, median (IQR), y	28 (19–40)
Adults (aged≥18 y)	372/478 (77.8)
Sex	
Female	211/478 (44.1)
Male	267/478 (55.9)
Setting	
Rural	332/478 (69.5)
Urban	146/478 (30.5)
Educational level	
No education	90/478 (18.8)
Primary	319/478 (66.8)
Secondary	58/478 (12.1)
Tertiary	9/478 (1.9)
Not declared	2/478 (0.4)
Occupation	
No work or housewife	57/478 (11.9)
Daily laborer, merchant, or other	74/478 (15.5)
Farmer	206/478 (43.1)
Student	120/478 (25.1)
Not declared	21/478 (4.4)
Marital status	
Divorced	10/478 (2.1)
Married	274/478 (57.3)
Never married	190/478 (39.8)
Widowed	3/478 (0.6)
Not declared	1/478 (0.2)
Means of diagnosis for index case[b]	
Sputum Xpert assay	148/152 (97.4)
Stool Xpert assay	77/152 (50.7)
Microscopy	0/152 (0.0)
Tuberculosis LAM test	4/152 (2.6)
No. of HHCs reported by index patient	
1–3	56/152 (36.8)
4–6	39/152 (25.7)
≥7	16/152 (10.5)
Not declared	41/152 (27.0)

Abbreviations: HHCs, household contacts; IQR, interquartile range; LAM, lipoarabinomannan.

[a]Data represent no. (%) of study participants unless otherwise specified.

[b]The means of diagnosis were not mutually exclusive; some index cases were diagnosed using >1 method.

Table 2. Known Risk Factors and Symptoms of Tuberculosis

Risk Factors and Symptoms	Study Participants, No./Total (%)[a] (N = 478)
Known risk factors	
History of tuberculosis	18/477 (3.8)
Full course of antituberculosis treatment in those with a positive history	15/18 (83.3)
Pregnancy in women	5/189 (2.6)
Known HIV status	
Negative	106/454 (23.3)
Positive	9/454 (2.0)
Unknown	339/454 (74.7)
Results of HIV test	
Negative	348/355 (98.0)
Not valid	1/355 (0.3)
Positive	6/355 (1.7)
Diabetes mellitus	5/474 (1.0)
Alcohol use	7/475 (1.5)
Smoker status	5/474 (1.0)
Nutritional status	
Good	431/475 (90.7)
Malnourished/underweight	44/475 (9.3)
Silicosis	0/477 (0.0)
Cancer	0/477 (0.0)
Celiac disease	1/477 (0.2)
Renal disease	1/477 (0.2)
Previous gastric surgery	None
Cirrhosis	2/477 (0.4)
COPD	2/477 (0.4)
Symptoms	
Asymptomatic for tuberculosis	268/477 (56.2)
Symptomatic for tuberculosis	209/477 (43.8)
Cough of any duration	130/209 (62.2)
Cough for >2 wk	100/209 (47.8)
Hemoptysis	3/209 (1.4)
Fever	23/209 (11.0)
Poor weight gain or weight loss	29/209 (13.9)
Night sweats	116/209 (55.5)
Chest pain	16/209 (7.7)
Shortness of breath	28/194 (13.4)
Duration of symptoms at assessment	
1–7 d	29/209 (13.9)
8–21 d	70/209 (33.5)
3 wk to 3 mo	79/209 (37.8)
>3 mo	30/209 (14.3)
Unspecified	1/209 (0.5)

Abbreviations: COPD, chronic obstructive pulmonary disease; HIV, human immunodeficiency virus.

(about 28 participants per month); 318 had valid data, 111 of 318 (35%) were judged as having possible tuberculosis, and 90 of 111 (81%) were confirmed with tuberculosis.

Participant Information

The final analysis included 478 participants (152 index patients and 326 HHCs) who were enrolled between 27 April 2023 and 13 December 2024. Table 1 summarizes participant characteristics and specific information on index patients. Data on known risk factors for tuberculosis and symptoms are summarized in Table 2.

Primary Outcome Measure: Accuracy of CAD4TB Compared With Xpert Testing

Data on CAD4TB classification were recorded in 460 of 478 participants (96%). Valid data on both CAD4TB classification and Xpert assays were available for 455 of 460 participants (99%) (Table 3). Using Xpert testing as the reference standard, CAD4TB had a sensitivity of 0.77 (95% CI, .70–.83), a specificity of 0.93 (.90–.96), a PPV of 0.88 (.82–.93), and an NPV of 0.86 (.82–.90) (Table 4).

Secondary Outcome Measures

Comparison of Stool and Sputum Xpert Testing in Adults. Overall, 367 of 372 adults (99%) were able to produce a stool sample, and stool Xpert testing was properly performed in 362 of 367 samples (99%). Similarly, 369 of 372 adults (99%) were able to produce a sputum sample, and sputum Xpert testing as properly performed in 366 of 369 samples (99%). Data on both stool and sputum samples were available for 358 of 372 adults (96%), and valid Xpert results on both stool and sputum samples were available for 340 of 358 (95%) (Supplementary Table 1). Of 478 adults, 8 (1.7%) were unable to produce stool samples. Notably, among the 6 adults (1.2%) who could not produce sputum, 4 (67%) were nevertheless able to provide stool samples. The concordance between stool and sputum Xpert assays was good (Cohen's κ = 0.76; Gwet's AC1, 0.78). Using sputum Xpert testing as the reference, stool Xpert testing had a sensitivity of 0.77 (95% CI, .70–.84), a specificity of 0.97 (.93–.99), a PPV of 0.95 (.89–.98), and an NPV of 0.85 (.80–.90) (Table 4 and Supplementary Table 2).

Common Metrics (Sensitivity, Specificity, PPV, and NPV) for Assessing the Accuracy of CAD4TB Versus Xpert Testing Within Strata of Interest. Table 4 summarizes the common accuracy metrics for assessing the accuracy of CAD4TB compared with Xpert assays (both sputum and stool) within the strata of interest. The small number of participants who were smokers (n = 5) or positive for HIV (n = 7) prevented meaningful stratified analyses in those subgroups. In addition, the small number of malnourished/underweight individuals with negative Xpert results (n = 2) led to large CIs.

Table 3. Results According to CAD4TB and Xpert Testing (n = 455)

	Xpert Result, No.[a]	
CAD4TB Result	*M. tuberculosis* Detected	*M. tuberculosis* Not Detected
Unclear	6	4
Possible tuberculosis	131	18
Unlikely tuberculosis	40	256

Abbreviation: *M. tuberculosis, Mycobacterium tuberculosis.*

[a]Combined stool and sputum Xpert results.

New Tuberculosis Cases Among HHCs During the Study Period

During the study period, 34 of 321 HHCs with valid data (10.6%) were positive with stool or sputum Xpert assays.

Additional Outcome Measures

Additional Positive Cases When Information From CAD4TB Was Available to the Clinician. Among 298 cases categorized as negative or unclear according to clinician evaluation without CAD4TB, 22 (7.4%) were judged positive when information from CAD4TB was available (Supplementary Table 3).

Table 4. Common Accuracy Metrics (Sensitivity, Specificity, and Positive and Negative Predictive Values) According to Each Comparison and Strata of Interest

Comparison and Strata	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
CAD4TB vs Xpert testingt[a]				
All participants	.77 (.70–.83)	.93 (.90–.96)	.88 (.82–.93)	.86 (.82–.90)
Age group				
Children (<18 y)	.64 (.41–.83)	.92 (.83–.97)	.70 (.46–.88)	.89 (.80–.95)
Adults (>18 y)	.79 (.71–.85)	.94 (.90–.97)	.91 (.84–.95)	.86 (.80–.90)
Sex				
Female	.74 (.62–.84)	.92 (.85–.96)	.82 (.70–.90)	.88 (.81–.93)
Male	.78 (.69–.86)	.95 (.90–.98)	.92 (.84–.97)	.86 (.79–.91)
Nonsmoker status	.77 (.69–.83)	.93 (.90–.96)	.88 (.81–.93)	.87 (.82–.90)
HIV status				
Negative	.83 (.74–.90)	.91 (.59–.99)	.99 (.93–.99)	.40 (.21–.61)
Unknown	.72 (.60–.81)	.93 (.89–.96)	.78 (.66–.87)	.91 (.87–.94)
History of tuberculosis				
No	.76 (.69–.82)	.95 (.90–.96)	.88 (.81–.93)	.87 (.82–.90)
Yes	.89 (.52–.99)	.83 (.36–.99)	.89 (.52–.99)	.83 (.36–.99)
Nutrition status				
Good	.73 (.65–.81)	.94 (.90–.96)	.85 (.77–.91)	.88 (.83–.91)
Malnourished or underweight	.88 (.73–.96)	.50 (.01–.99)	.97 (.85–.99)	.17 (.00–.64)
Stool vs sputum Xpert testing				
Adults (aged ≥18 y)	.77 (.70–.84)	.97 (.93–.99)	.95 (.89–.98)	.85 (.80–.90

Abbreviations: CI, confidence interval; HIV, human immunodeficiency virus, NPV, negative predictive value; PPV, positive predictive value.

[a]Combination of stool and sputum Xpert testing. The small number of smoking participants (n = 5) and those positive for HIV (n = 7) prevented meaningful stratified analyses in those subgroups. The large CIs in the malnourished/underweight subgroup were due to the small number of participants with negative Xpert results (n = 2). Common accuracy metrics of stool versus sputum Xpert testing within strata of interest among adults are summarized in Supplementary Table 2.

False-Positive Cases When Information From CAD4TB Was Available to the Clinician. Among 149 cases categorized as positive for TB when information from CAD4TB was available to the clinician, 18 (12.1%) had negative Xpert results. In addition, 18 of 149 (12.1%) were judged negative when the clinician assessed the same cases without CAD4TB (Supplementary Table 3).

Discordant Cases With CAD4TB Versus Clinician Chest Radiographic Evaluation. Seventy-five of 462 assessments (16.2%) were discordant between CAD4TB and clinician chest radiographic evaluation (Supplementary Table 3). Among 20 cases assessed as unclear without CAD4TB, 4 were judged as possible tuberculosis and 12 as unlikely tuberculosis with CAD4TB. Among 164 cases assessed as possible tuberculosis without CAD4TB, 4 were judged as unclear and 33 as unlikely tuberculosis with CAD4TB. And among 280 cases assessed as unlikely tuberculosis without CAD4TB, 4 were judged as unclear and 18 as possible tuberculosis with CAD4TB.

Association Between CAD4TB Score and Clinician Confidence in Assessments Without CAD4TB Information Available. Lower CAD4TB scores were associated with lower clinician confidence on their assessments when CAD4TB information was not available (Spearman rank correlation coefficient $\rho = 0.61$; $P < .001$).

Association Between CAD4TB Score and Clinician Identification of Possible Tuberculosis Cases. When information from CAD4TB was available to the clinician, the median CAD4TB scores (interquartile range) were 78 (63–93), 33 (12–49), and 5 (3–13) in participants categorized as possible tuberculosis, unclear tuberculosis, and unlikely tuberculosis, respectively ($P < .001$) (Supplementary Table 4).

DISCUSSION

In a country with high burden of tuberculosis, our study demonstrated that CAD4TB achieved good sensitivity and excellent specificity in detecting active tuberculosis among HHCs, when compared with a reference standard of sputum Xpert MTB/RIF testing; in particular, we demonstrated that CAD4TB is a reliable tool to discriminate between individuals with and those without tuberculosis. The sensitivity observed in our study is consistent with that reported in other studies involving similar populations [20], and it appears higher than that reported in other subpopulations, such as pediatric cohorts [20]. We nonetheless observed a decline in both sensitivity and PPV when stratifying by age, with a sensitivity of 64% in the population <15 years of age. This finding can be explained by the radiological characteristics of pediatric tuberculosis, which more frequently presents with nodular and miliarylike abnormalities that are not always easily detectable on chest radiographs and, consequently, by AI [21, 22].

Similar misinterpretation may be expected in the HIV-positive population; however, the low representation of this subgroup did not allow any meaningful stratified analyses. Nevertheless, a multicenter study by Crowder et al [23] found no significant differences in sensitivity or specificity between HIV-positive and HIV-negative patients, although the median CD4 cell count was relatively high. In contrast, other studies have shown that the area under the curve with CAD systems was lower in HIV-positive than in HIV-negative patients [24]. Conversely, a history of prior tuberculosis was associated with strong CAD4TB sensitivity in detecting active disease, confirming that it may be an important risk factor for tuberculosis activation. However, in patients with a history of tuberculosis, using CAD4TB could lead to an increased risk of unnecessary retreatment, as the algorithm may misclassify residual lesions (eg, cavitations) from the previous infection as active disease [25], also proved by the lower specificity observed in this subgroup. This may pave the way for large-scale implementation of CAD4TB or other AI systems that can longitudinally compare predictive scores for the same individual, thereby distinguishing the increasingly recognized posttuberculosis lung disease changes from true reinfection or relapse.

Similarly, the malnourished subgroup also exhibited higher sensitivity compared to its counterpart, albeit with reduced specificity. We observed a low NPV with CAD4TB among malnourished individuals, but only 2 of them had negative Xpert results, which affected overall performance metrics. Therefore, from a preventive standpoint, it is essential to address both factors, particularly nutritional status, which more easily leads to increased extent of disease and cavitations in patients with tuberculosis secondary to impaired immune response [26].

Stool Xpert testing has already been demonstrated to be a useful and cost-effective tool for diagnosing tuberculosis in pediatric populations [27], to the extent that it is incorporated into WHO guidelines [5]. In our study, we observed strong concordance between stool and sputum Xpert results—both in index patients and during active case finding among HHCs. Furthermore, the stool Xpert assay is straightforward to perform, and obtaining a stool sample from the patient is generally uncomplicated. A few participants were unable to submit stool samples, mainly due to transportation challenges and long travel distances from the hospital, highlighting logistical barriers to sample collection that may affect diagnostic implementation in rural settings. Recent studies have demonstrated that stool testing in adults and adolescents achieves high diagnostic accuracy, with newer molecular assays, like in-house real-time quantitative polymerase chain reaction, outperforming the reference standard Xpert MTB/RIF Ultra [28, 29]. It is imperative to shift toward diagnostic strategies that use nonrespiratory specimens, as an increasing proportion of individuals no longer exhibit cough or sputum production as

primary symptoms [30]. It should nevertheless be noted that stool Xpert testing ought to augment or complement, rather than replace, sputum-based testing whenever feasible, as evidenced by the 33 false-negatives that we reported in our study. CAD4TB and stool Xpert testing represent 2 valuable options for tuberculosis diagnostic and screening campaigns.

In addition to demonstrating solid performance metrics, CAD4TB has been shown to increase the nonradiologist clinician's confidence in interpreting chest radiographs. Notably, among tuberculosis-negative cases, clinicians without access to CAD4TB information assigned higher likelihood scores—resulting in 7.4% of chest radiographs being misclassified as positive. This finding further illustrates that, in settings lacking expert personnel, chest radiographic interpretation is prone to frequent misinterpretations which, when combined with limited microbiological capacity, may lead to overtreatment, especially in certain subgroups [31]. Conversely, we observed that the median CAD4TB scores were lower among cases that clinicians labeled as "unlikely tuberculosis." This supports the use of CAD4TB as a first-line screening tool, even in the absence of trained clinicians.

It should also be acknowledged that multiple studies have demonstrated the cost of CAD systems to be substantially lower than that of using a radiologist in low-resource, high-tuberculosis-burden settings [32, 33]. While this might raise ethical questions about substituting AI for human expertise, it is more appropriately understood as a pragmatic investment to achieve a basic level of diagnostic capability in contexts where access to a radiologist is not feasible. In such environments, the implementation of CAD represents the most viable means to ensure that patients receive timely and accurate screening, rather than leaving them without any diagnostic service. The current study paves the way for the applicability of CAD4TB and stool Xpert in community-based active case finding, which in our cohort yielded >10% positivity among HHCs—consistent with anticipated rates in this high-risk population [15, 33].

Several limitations must be acknowledged. First, this was a single-center study conducted in a region where tuberculosis is highly endemic; it does not account for geographic variability in tuberculosis incidence or HIV prevalence, and therefore its generalizability may be limited, potentially affecting CAD4TB performance metrics. Second, we were unable to implement the latest Xpert MTB/RIF Ultra assay, which constrained our diagnostic capacity [34]. Third, although our protocol envisaged repeat screening of HHCs at 6 and 12 months for those with negative results at baseline, adherence to this rescreening schedule was poor, primarily owing to regional political instability [35] and participants' low perceived risk of tuberculosis infection.

Looking ahead, several areas warrant further investigation. In particular, recalibration of CAD4TB thresholds for specific subpopulations—such as pediatric patients and individuals living with HIV (especially those with AIDS)—is necessary to optimize sensitivity and specificity. Moreover, it would be valuable to compare CAD4TB against CAD systems applied to lung ultrasonographic imaging, given that chest ultrasonography may represent a more feasible modality for community-based screening in outreach programs [36, 37], where even portable digital chest radiography is not feasible, in the so-called hard-to-reach areas [38].

In conclusion, to summarize, the combined use of CAD4TB and stool Xpert testing presents a powerful strategy for improving tuberculosis detection in regions with high disease burden and limited resources. By integrating AI-assisted chest radiographic analysis with broader molecular testing that does not rely solely on sputum samples, health programs can identify cases that might otherwise be overlooked, thereby moving closer to the objective of comprehensive tuberculosis screening and timely intervention.

Supplementary Data

Supplementary materials are available at *Open Forum Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Author contributions. F. V. S., G. P., F. C., A. Sa., F. M., and F. D. G. conceived and developed the idea. G. G., S. C., W. N., F. V. S., B. Z., A. B. A., A. R., M. M., and A. Su carried out the study. B. K. S., F. A. B., M. T., A. A., and F. D. G. supervised the project. G. G., S. C., F. V. S., and F. D. G. wrote the original draft of the manuscript. F. C. performed the statistical analysis. G. M. and R. I. implemented the project. B. G., W. N., B. Z., B. K. S., G. P., F. C., F. A. B., A. B. A., A. R., R. I., A. Sa., M. M., F. G., L. G., M. T., and F. M. reviewed the manuscript. All authors accepted the final draft of the manuscript and agreed to submit the manuscript for publication.

Financial support. This work was supported by the 2022 call of the Italian Agency for Development Cooperation (Agenzia Italiana per la Cooperazione allo Sviluppo) and by the Italian Ministry of Health (funds to IRCCS Sacro Cuore Don Calabria Hospital—"Ricerca Corrente," Linea 3).

Potential conflicts of interest. All authors: No reported conflicts.

References

1. World Health Organization. Global tuberculosis report 2024. Geneva, Switzerland: World Health Organization, **2024**. Available at: https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2024. Accessed 2 June 2025.
2. Pizzol D, Di Gennaro F, Chhaganlal KD, et al. Tuberculosis and diabetes: current state and future perspectives. Trop Med Int Health **2016**; 21:694–702.
3. Di Gennaro F, Cotugno S, Guido G, et al. Disparities in tuberculosis diagnostic delays between native and migrant populations in Italy: a multicenter study. Int J Infect Dis **2025**; 150:107279.
4. Migliori GB, Wu SJ, Matteelli A, et al. Clinical standards for the diagnosis, treatment and prevention of TB infection. Int J Tuberc Lung Dis **2022**; 26:190–205.
5. World Health Organization. WHO Consolidated guidelines on tuberculosis: module 2: screening: systematic screening for tuberculosis disease. Geneva, Switzerland: World Health Organization, **2021**. Available at: https://www.who.int/publications-detail-redirect/9789240022676. Accessed 4 May 2025.

6. Philipsen RH, Sánchez CI, Maduskar P, et al. Automated chest-radiography as a triage for Xpert testing in resource-constrained settings: a prospective study of diagnostic accuracy and costs. Sci Rep **2015**; 5:12215.
7. Hua D, Nguyen K, Petrina N, 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**; 177:105159.
8. World Health Organization. WHO approves six software products for computer-aided detection of TB on chest X-ray. Geneva, Switzerland: World Health Organization, **2025**. Available at: https://www.who.int/news/item/11-06-2025-who-approves-six-software-products-for-computer-aided-detection-of-tb-on-chest-x-ray. Accessed 16 June 2025.
9. de Haas P, Yenew B, Diriba G, et al. The simple one-step stool processing method for detection of pulmonary tuberculosis: a study protocol to assess the robustness, stool storage conditions and sampling strategy for global implementation and scale-up. PLoS One **2022**; 17:e0264103.
10. Segala FV, Papagni R, Cotugno S, et al. Stool Xpert MTB/RIF as a possible diagnostic alternative to sputum in Africa: a systematic review and meta-analysis. Front Public Health **2023**; 11:1117709.
11. Sustainable Development Solutions Network. Ethiopia—SDG indicators [web dashboard]. Sustainable Development Report. 2024. Available at: https://dashboards.sdgindex.org/profiles/ethiopia/indicators. Accessed 13 June 2025.
12. Agizew TB, Dememew ZG, Leta T, et al. Prospects for tuberculosis elimination in Ethiopia: feasibility, challenges, and opportunities. Pan Afr Med J **2022**; 43:146.
13. Wolde HM, Zerihun B, Sinshaw W, Yewhalaw D, Abebe G. Comparison of the yield of two tuberculosis screening approaches among household contacts in a community setting of Silti Zone, Central Ethiopia: a prospective cohort study. BMC Pulm Med **2024**; 24:135.
14. Zar HJ, Workman LJ, Prins M, et al. Tuberculosis diagnosis in children using Xpert Ultra on different respiratory specimens. Am J Respir Crit Care Med **2019**; 200:1531–8.
15. Segala FV, Nigussa W, Guido G, et al. Active close contact investigation of tuberculosis through computer-aided detection and stool Xpert MTB/RIF among people living in Oromia Region, Ethiopia (CADOOL study): protocol for a prospective, cross-sectional study. BMJ open **2023**; 13:e074968.
16. World Health Organization. Chest radiography in tuberculosis detection: summary of current WHO recommendations and guidance on programmatic approaches. Available at: https://www.who.int/publications-detail-redirect/9789241511506. Accessed 13 June 2025.
17. de Haas P, Yenew B, Mengesha E, et al. The simple one-step (SOS) stool processing method for use with the Xpert MTB/RIF assay for a child-friendly diagnosis of tuberculosis closer to the point of care. J Clin Microbiol **2021**; 59:e0040621.
18. National guidelines for TB, drug resistant TB and leprosy in Ethiopia. 6th ed. Federal Ministry of Health of Ethiopia. Available at: https://www.afro.who.int/publications/national-guidelines-tb-drug-resistant-tb-and-leprosy-ethiopia-sixth-edition. Accessed 14 June 2025.
19. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, **2024**.
20. Abuzerr S, Zinszer K. Computer-aided diagnostic accuracy of pulmonary tuberculosis on chest radiography among lower respiratory tract symptoms patients. Front Public Health **2023**; 11:1254658.
21. Edem VF, Nkereuwem E, Agbla SC, et al. Accuracy of CAD4TB (computer-aided detection for tuberculosis) on paediatric chest radiographs. Eur Respir J **2024**; 64: 2400811.
22. Agustini D, Fediani Y, Liberty IA. Chest radiograph accuracy in diagnosing pediatric pulmonary tuberculosis (PTB): a systematic review and meta-analysis. JUMMEC **2024**; 27:21–30.
23. Crowder R, Thangakunam B, Andama A, et al. Diagnostic accuracy of tuberculosis screening tests in a prospective multinational cohort: chest-radiography with computer-aided detection, Xpert tuberculosis host response, and C-reactive protein. Clin Infect Dis doi:10.1093/cid/ciae549. Published 7 November 2024.
24. Gelaw SM, Kik SV, Ruhwald M, 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:e0000402.
25. 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**; 76:e894–901.
26. Hoyt KJ, Sarkar S, White L, et al. Effect of malnutrition on radiographic findings and mycobacterial burden in pulmonary tuberculosis. PLoS One **2019**; 14: e0214011.
27. Mafirakureva N, Klinkenberg E, Spruijt I, et al. Xpert Ultra stool testing to diagnose tuberculosis in children in Ethiopia and Indonesia: a model-based cost-effectiveness analysis. BMJ Open **2022**; 12:e058388.
28. Kay A, Vasiliu A, Carratala-Castro L, et al. Performance of a stool-based quantitative PCR assay for the diagnosis of tuberculosis in adolescents and adults: a multinational, prospective diagnostic accuracy study. Lancet Microbe **2024**; 5: e433–41.
29. Ness T, Maphalala N, Khumalo W, et al. Rapid diagnostic sequencing of stool DNA using targeted nanopore sequencing in patients with a pulmonary tuberculosis diagnosis. Open Forum Infect Dis **2025**; 12:ofaf121.
30. Trajman A, Campbell JR, Kunor T, et al. Tuberculosis. Lancet **2025**; 405:850–66.
31. Mateyo K, Kerkhoff AD, Dunn I, Nteeni MS, Muyoyeta M. Clinical and radiographic characteristics of presumptive tuberculosis patients previously treated for tuberculosis in Zambia. PLoS One **2022**; 17:e0263116.
32. Bashir S, Kik SV, Ruhwald M, et al. Economic analysis of different throughput scenarios and implementation strategies of computer-aided detection software as a screening and triage test for pulmonary TB. PLoS One **2022**; 17:e0277393.
33. Okada K, Yamada N, Takayanagi K, 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**; 52:2.
34. Christopher DJ, Coelho V, Ebby GS, Shankar D, Gupta R, Thangakunam B. Incremental yield of Xpert® MTB/RIF Ultra over Xpert® MTB/RIF in the diagnosis of extrapulmonary TB. Int J Tuberc Lung Dis **2021**; 25:939–44.
35. Gleixner-Hayat B. Ethiopia's fragile stability remains at risk. 1 November 2023. Carnegie Endowment for International Peace. Available at: https://carnegieendowment.org/posts/2023/11/ethiopias-fragile-stability-remains-at-risk?lang=en. Accessed 12 June 2025.
36. Di Gennaro F, Pisani L, Veronese N, et al. Potential diagnostic properties of chest ultrasound in thoracic tuberculosis—a systematic review. Int J Environ Res Public Health **2018**; 15:2235.
37. Guido G, Nigussa W, Cotugno S, et al. Comparative effectiveness of chest ultrasound, chest X-ray and computer-aided diagnostic (CAD) for tuberculosis diagnosis in low-resource setting: study protocol for a cross-sectional study from Ethiopia. Front Public Health **2024**; 12:1476866.
38. Useni S, Odume B, Tukur M, et al. Comparative analysis of the impact of portable digital X-ray on TB screening in hard-to-reach areas in Nigeria. Public Health Action **2025**; 15:17–20.