Predictors contributing to the estimation of pulmonary tuberculosis among adults in a resource-limited setting: A systematic review of diagnostic predictions

Gebremedhin Berhe Gebregergs[1] iD**, Gebretsadik Berhe[1], Kibrom Gebreslasie Gebrehiwot[2] and Afework Mulugeta[1]**

Abstract

Background: Although tuberculosis is highly prevalent in low- and middle-income countries, millions of cases remain undetected using current diagnostic methods. To address this problem, researchers have proposed prediction rules.
Objective: We analyzed existing prediction rules for the diagnosis of pulmonary tuberculosis and identified factors with a moderate to high strength of association with the disease.
Methods: We conducted a comprehensive search of relevant databases (MEDLINE/PubMed, Cochrane Library, Science Direct, Global Health for Reports, and Google Scholar) up to 14 November 2022. Studies that developed diagnostic algorithms for pulmonary tuberculosis in adults from low and middle-income countries were included. Two reviewers performed study screening, data extraction, and quality assessment. The study quality was assessed using the Quality Assessment of Diagnostic Accuracy Studies-2. We performed a narrative synthesis.
Results: Of the 26 articles selected, only half included human immune deficiency virus-positive patients. In symptomatic human immune deficiency virus patients, radiographic findings and body mass index were strong predictors of pulmonary tuberculosis, with an odds ratio of >4. However, in human immune deficiency virus-negative individuals, the biomarkers showed a moderate association with the disease. In symptomatic human immune deficiency virus patients, a C-reactive protein level ⩾10 mg/L had a sensitivity and specificity of 93% and 40%, respectively, whereas a trial of antibiotics had a specificity of 86% and a sensitivity of 43%. In smear-negative patients, anti-tuberculosis treatment showed a sensitivity of 52% and a specificity of 63%.
Conclusions: The performance of predictors and diagnostic algorithms differs among patient subgroups, such as in human immune deficiency virus-positive patients, radiographic findings, and body mass index were strong predictors of pulmonary tuberculosis. However, in human immune deficiency virus-negative individuals, the biomarkers showed a moderate association with the disease. A few models have reached the World Health Organization's recommendation. Therefore, more work should be done to strengthen the predictive models for tuberculosis screening in the future, and they should be developed rigorously, considering the heterogeneity of the population in clinical work.

Keywords

Predictors, pulmonary tuberculosis, diagnostic algorithm, performance

Introduction

Tuberculosis (TB) continues to be highly prevalent in low- and middle-income countries (LMICs) where resources for TB screening and diagnosis are limited.[1] However, approximately 3.1 million people are missed with the current approach to TB case detection.[1] Therefore, reaching missing TB cases is critically important to save millions of lives and end TB.[2]

[1]School of Public Health, College of Health Sciences, Mekelle University, Mekelle, Ethiopia
[2]Department of Internal Medicine, College of Health Sciences, Mekelle University, Mekelle, Ethiopia

Corresponding author:
Gebremedhin Berhe Gebregergs, School of Public Health, College of Health Sciences, Mekelle University, P.Box: 1871, Mekelle, Ethiopia.
Email: bgebrish04@gmail.com

Currently, a molecular test for TB, namely, the GeneXpert Dx System (Cepheid, CA, USA), was co-developed by Cepheid, Inc. (Sunnyvale, CA, USA), and the Foundation for Innovative New Diagnostics.[3] This test has improved TB diagnosis; however, its access is limited in primary healthcare settings.[4] Truenat, a new molecular diagnostic tool for TB, has not been widely implemented in many developing countries.[5] Smear microscopy has low sensitivity and fails to detect nearly half of all TB cases.[6] The availability of radiography and specialist staff is limited. In such cases, a risk prediction rule helps identify patients at high risk of developing TB and thus requires additional diagnostic testing or consideration of anti-TB treatment.[7,8] In 2021, prediction rules, such as the chest radiograph abnormality score, were proposed by the World Health Organization (WHO) to increase the TB case detection rate.[9]

Several risk factors, such as HIV infection, silicosis, diabetes mellitus, kidney disease, and low body weight, affect the development of TB.[10] Statistically combining their effects can produce a more powerful risk prediction than considering one risk factor at a time.[11,12]

Several clinical prediction rules are available for the diagnosis of TB.[13–15] These rules include factors related to the patient's clinical findings, history, and investigation results. Pinto et al.[16] conducted a systematic review to assess the diagnostic accuracy and reproducibility of chest radiograph scoring systems for the diagnosis of pulmonary TB and reported that the scoring system appears useful for ruling out pulmonary TB in hospital settings. However, such a scoring system cannot be applied to peripheral health facilities, where radiography is scarce. The quality and characteristics of the prediction models for pulmonary TB have been reviewed elsewhere.[17] However, studies evaluating the accuracy of the symptoms or combinations of symptoms were excluded. Such algorithms are important to identify easily applicable and readily available predictors for future model development and validation. A recent systematic review investigated the utility of a clinical scoring system in improving TB case findings.[18] It showed variable sensitivity and specificity but did not demonstrate the power of the predictors and the effectiveness of clinical scoring.

A standardized scoring system for pulmonary TB using readily available, measurable, and widely generalizable predictors is useful for improving TB case detection. Such tools should be robust and simple to implement in low-resource settings.[19] Synthesizing appropriate algorithms for different clinical and epidemiological settings is important for TB screening and diagnosis. In this study, we examined the performance of prediction rules for diagnosing pulmonary TB in adult patients attending health facilities. We also identified common predictors of pulmonary TB with moderate-to-strong predictive ability.

Methods

Study design and population

We conducted a systematic review of observational studies that developed clinical algorithms for diagnosing pulmonary TB. We classified the algorithms into model based and criteria based. Model-based algorithms were developed using two or more predictors independently associated with pulmonary TB in multivariable logistic regression, classification trees, neural networks, or vector machine learning.[20] Criteria-based algorithms were developed empirically (without the use of a mathematical model) using signs and symptoms of TB, with some including laboratory and chest X-ray (CXR) findings.

This systematic review was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[21] This study was registered in the Open Science Framework (OSF, https://osf.io/4y7sz/).

Data sources and search strategy

We conducted a comprehensive literature search in MEDLINE/PubMed, Cochrane Library, Science Direct, Global Health for reports, and Google Scholar for studies published from 1 January 1990 to 14 November 2022. The literature search was conducted from 15 April to 14 November 2022. We used optimized search filters developed to identify diagnostic prediction studies in PubMed.[22] The search filter had a sensitivity of 0.95, ranging from 0.94 to 0.97. To address our review question, we added search strings related to the diagnosis and pulmonary TB. We used the following search terms: ("Stratification" OR "ROC Curve"[Mesh] OR "Discrimination" OR "Discriminate" OR "c-statistic" OR "c statistic" OR "Area under the curve" OR "AUC" OR "Calibration" OR "Indices" OR "Algorithm" OR "Multivariable") AND ("Diagnosis" OR "diagnosis" [MeSH]) AND ("Pulmonary TB" OR "PTB" OR "TB" [Mesh] OR "TB").

The results were limited to the English language based on the study team's ability. We conducted forward and backward citation searches in Google Scholar using the references of the included articles to find other potentially relevant studies, including gray literature. We searched clinical trial registers, conference proceedings, and gray literature databases for unpublished studies. All search results were managed using EndNote, version 5.

Inclusion and exclusion criteria

We used the population, index test, reference test, and diagnosis of interest framework for diagnostic accuracy studies[23] to guide the comparison of algorithms. Any original study that developed and/or validated a diagnostic score/algorithm for pulmonary TB in adults (aged ⩾ 15 years) seeking health

care was included in this review. We only included studies conducted in LMICs as defined by the World Bank in 2022.[24] High-income countries were excluded because of the exceptionally low incidence of TB and the remarkably high detection rate of TB cases. Studies on extrapulmonary TB and drug-resistant TB were excluded. We also excluded surveys because the presumed TB cases in the community were different from those presented to the health facility.

Study selection

Using the inclusion and exclusion criteria, two authors independently screened the titles and abstracts of the retrieved studies. Disagreements were resolved through a consensus discussion. Duplicate articles were excluded from the analysis. Eligible citations were selected for full-text review. Full-text articles that met all criteria for this review were retained for quality assessment, data extraction, and subsequent analyses.

Assessment of study quality

GB Gebregergs and G Berhe independently assessed study quality using the Quality Assessment of Diagnostic Accuracy Studies-2, a revised tool comprising four domains: patient selection, index test, reference standard, and flow and timing.[25] As recommended, each study was graded as "high," "low," or "unclear" for risk of bias or concerns about applicability for each domain.

If all signaling questions for a domain are answered "yes," the risk of bias is considered low. If any of the signaling questions are answered "no," the risk of bias is considered high.

Disagreements between the two reviewers were resolved by discussion.

Data extraction and analysis

A data extraction format was created using an Excel spreadsheet. We adopted the CHARMS (Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies) checklist[26] to develop the data extraction format. In case we were unable to obtain all the required information from the available reports, we contacted the corresponding authors.

Our primary interest was to identify parameters that significantly predicted active pulmonary TB along with their measure of effect. We classified the strength of association into three categories: weak (OR $\leqslant$ 2.0, moderate (OR: 2.1–3.0), and strong (OR > 3.0). The secondary outcome of interest was the performance of the algorithm. The most available algorithms were symptom screening, serial screening (symptom screening followed by chest radiography), and biomarker screening. A combination of these algorithms has also been studied. We also calculated the number needed to

screen (NNS) to detect one person with TB disease as the reciprocal of the prevalence of newly diagnosed TB.[27] Case–control studies were excluded from NNS calculations. We presented the effect size (odds ratio) of predictors and the diagnostic accuracy of different algorithms for symptomatic individuals, regardless of symptoms, and for HIV-positive and HIV-negative individuals. In the presence of two or more studies, we used a median estimate to avoid small-study effects.[28]

We only performed a narrative synthesis. We did not pool the estimates because the scoring systems were heterogeneous in terms of predictors and derivation methods.

Results

Study selection and characteristics

Of the 4873 citations identified, 4860 unique articles remained after removing duplicates. After screening the titles and abstracts, 82 articles met the inclusion criteria, and their full texts were retrieved. After a full-text review, 56 articles were excluded for various reasons, and 26 articles[7,29–53] were included in the systematic review (Figure 1).

In all, 16 (61.5%) studies[7,30,31,33,35,37,38,40–42,44,46,47,50–52] were from middle-income countries, 8 studies[29,32,34,39,43,48,49,53] were from low-income countries, and the remaining 2 studies were multi-country studies. In total, 24 (92.2%) articles[7,29–35,37–52] were from high-burden countries (HBCs), and 2 studies[36,53] were from both HBCs and non-HBCs.

All included studies were observational. Ascertainment of pulmonary TB was performed using culture in 20 studies,[30,35–53] and GeneXpert in six studies according to the WHO standard.[54] In all, 13 studies[29,30,34,35,37,38,40–43,45,46,49] focused solely on HIV-positive patients, while 12 studies[7,31,32,36,39,44,47,48,50–53] included participants, regardless of their HIV status.

Six studies[43,44,48,51–53] exclusively enrolled patients with smear-negative results. The median sample size was 575 (IQR: 345–1048). The details of the included studies are presented in Table 1.

Quality assessment

The QUADAS-2 results for risk of bias and applicability concerns are summarized in Figure 2. Only seven studies (27%) had a considerable risk of bias for patient selection and two (7.7%) for flow and timing, indicating an overall good quality in these areas. All studies were considered to have low applicability concerns for the index test and reference standard.

The algorithm was derived using a mathematical model in fifteen studies. Among the model-based prediction studies, seven[39,41,46,48,49,51,53] focused on model development, one employed independent external validation,[38] and seven[7,30,33,35,37,41,52] combined developments with internal validation. Among the validation methods, two
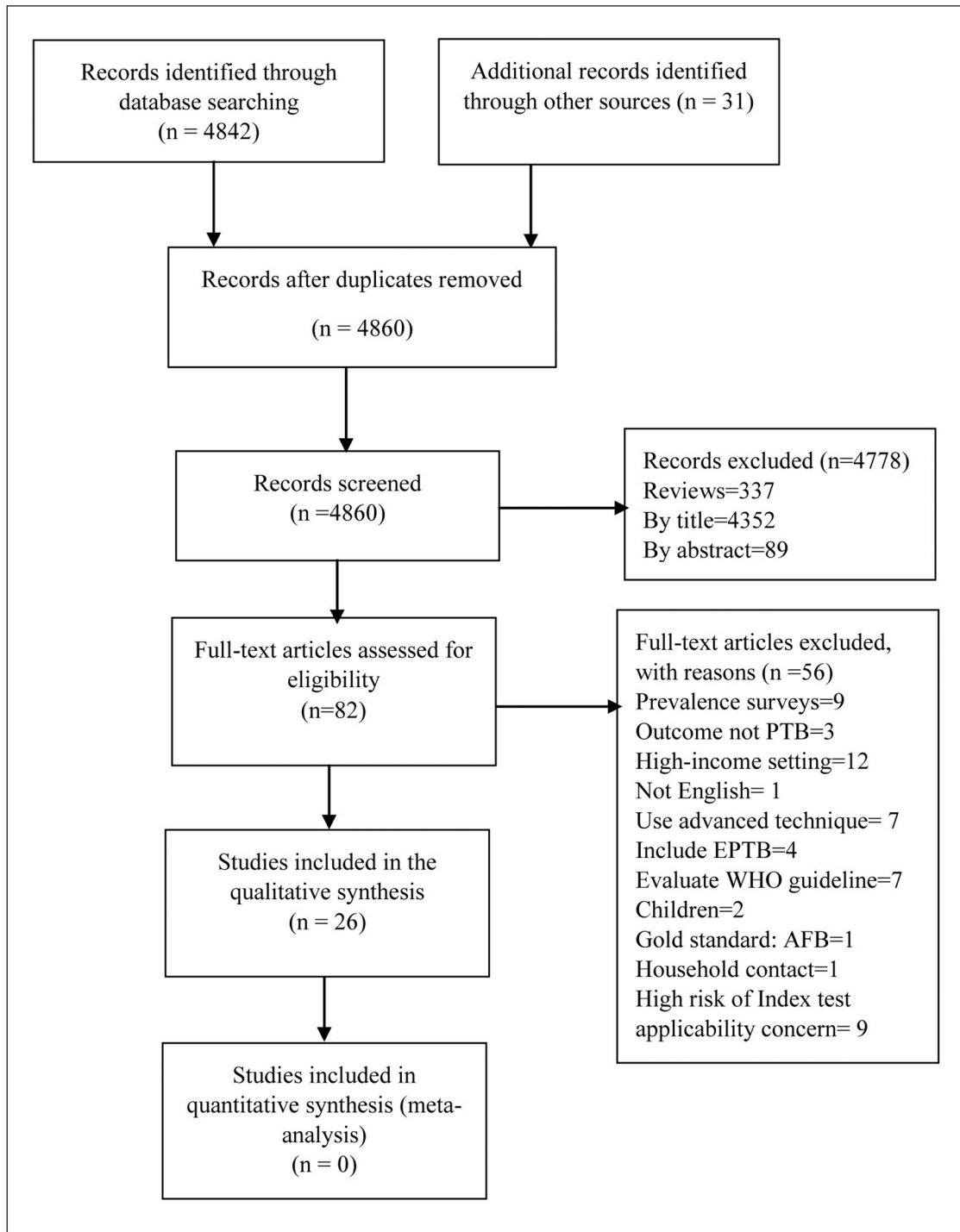
Figure 1. PRISMA flow diagram for included and excluded studies.
PRISMA: preferred reporting items for systematic reviews and meta-analyses.

were temporal[30,37]; three were data splits[31,41,52]; and two were geographical.[7,35] None of the prediction models considered expert opinions when selecting predictors. Five studies (33.3%) with multiple variables had an event-per-variable ratio <10. Six articles used the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) checklist.

Seven criteria-based algorithms reported accuracy tests and 12 model-based algorithms reported model classification (Table 2). Among the 15 model-based studies, 13 reported areas under the receiver operating characteristic curve (AUC), and the median AUC was 0.82 (IQR: 0.78–0.84). The maximum (0.92) AUC was reported by Yu et al.,[33] and the lowest (0.69) AUC was noted by Mello et al.[52]

Table 1. Characteristics of the included studies.

First author and year	Country	Study design	Study setting	Study population	Smear status	HIV status	No. with outcome/ sample size (%)	Outcome ascertainment
Acharya et al.[29]	India	Cross-sectional	ART center	Symptomatic and asymptomatic HIV patients	Not known	Positive	31/138 (22%)	GeneXpert positive
Chen et al.[30]	China	Cross-sectional	Hospital	Symptomatic HIV patients	Not known	Positive	146/766 (19%)	Culture positive
Hidayat et al.[31]	Indonesia	Cross-sectional	Hospital	Elderly patients	Not known	Not known	78/302 (26%)	GeneXpert positive
Jerene et al.[32]	Ethiopia	Cross-sectional	Hospital and health center	Diabetes mellitus patients	Not known	Not known	13/172 (8%)	GeneXpert positive
Yu et al.[33]	China	Case–control	Hospital	Active vs inactive pulmonary TB	Not known	Negative	NA	GeneXpert positive
Saripalli et al.[34]	India	Cross-sectional	Hospital	Symptomatic HIV patients	Not known	Positive	77/150 (51%)	GeneXpert positive
Auld et al.[35]	Botswana	Clinical trial	Hospital and clinic	Symptomatic and asymptomatic HIV patients	Not known	Positive	318/5418 (6%)	Culture positive
Samuels et al.[36]	MC	Case–control	Hospital	Patients with respiratory symptom	Not known	Not known	NA	Culture positive
Ouyang et al.[37]	China	Case–control	Hospital	Symptomatic HIV patients	Not known	Positive	NA	Culture positive
Boyles, et al.[38]	South Africa	Cross-sectional	Community clinic	Symptomatic HIV patients	Not known	Positive	75/217 (35%)	Culture positive
Baik et al.[7]	South Africa	Case–control	Clinic	Patients assessed for TB	Not known	Not known	NA	GeneXpert positive
George et al.[39]	Ethiopia	Cross-sectional	Hospital	Symptomatic pulmonary TB	Not known	Not known	43/425 (10%)	Culture positive
Shapiro et al.[40]	South Africa	Cross-sectional	Clinic	Symptomatic and asymptomatic HIV patients	Not known	Positive	42/425 (10%)	Culture positive
Hanifa et al.[41]	South Africa	Cohort	Hospital and health center	Symptomatic HIV patients	Not known	Positive	110/1048 (10%)	Culture positive
Drain et al.[42]	South Africa	Cross-sectional	Local hospital	Symptomatic and asymptomatic HIV patients	Not known	Positive	123/675 (18%)	Culture-positive
Padmapriyadarsini et al.[43]	India	Cross-sectional	Hospital and clinic	Symptomatic HIV patients	Negative	Positive	68/504 (13%)	Culture positive
Soto et al.[44]	Peru	Cross-sectional	Hospital	Smear-negative TB patients	Negative	Not known	182/663 (27%)	Culture positive
Swindell et al.[45]	MC	Cross-sectional	Health center	Symptomatic and asymptomatic HIV patients	Not known	Positive	31/801 (4%)	Culture positive
Rangaka et al.[46]	South African	Cross-sectional	Clinic	Symptomatic and asymptomatic HIV patients	Not known	Positive	126/1429 (9%)	Culture positive
Castro et al.[47]	Brazil	Cross-sectional	Hospital	Patients vising lung clinic	Not known	Not known	243/1365 18%	Culture positive
Swai et al.[48]	Tanzania	Cross-sectional	Hospital	Smear-negative TB patients	Negative	Not known	127/467 (27%)	Culture positive
Nguyen et al.[49]	Viet Nam	Cross-sectional	Clinic	Symptomatic and asymptomatic HIV patients	Not known	Positive	28/397 (7%)	Culture positive
Solari et al.[50]	Peru	Cross-sectional	Hospital	Patients with respiratory symptoms	Not known	Not known	109/345 (32%)	Culture positive
Soto et al.[51]	Peru	Cross-sectional	Hospital	Smear-negative TB patients	Negative	Not known	27/262 (10%)	Culture positive
Mello et al.[52]	Brazil	Cohort	Hospital	Smear-negative TB patients	Negative	Not known	270/551 (49%)	Culture positive
Samb et al.[53]	Burundi and Tanzania	Cohort	University Hospital	Patients with respiratory symptoms	Negative	Not known	41/141 (29%)	Culture positive

MC: multi-country; NA: not applicable; TB: tuberculosis; HIV: human immune deficiency virus.
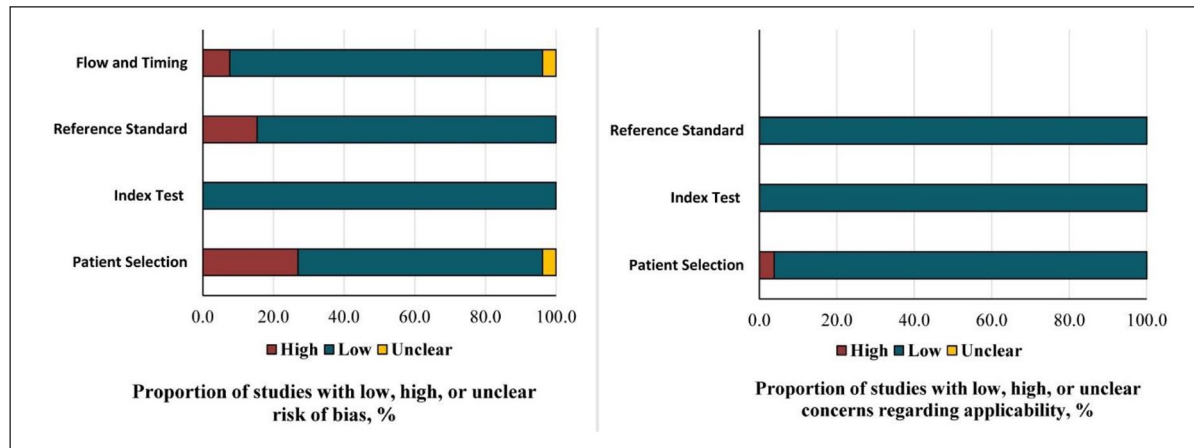
Figure 2. Summary of risk of bias and applicability concern with QUADAS-2.
QUADAS: quality assessment of diagnostic accuracy studies.

Predictors included in model-based predictions

To identify the most reliable and consistent predictors of pulmonary TB across different models, we focused on studies with at least 10 events per variable to ensure sufficient data for accurate analysis (Table 3).

In symptomatic HIV patients, cavity, hilar lymphadenopathy, and pleural effusion were strong predictors (OR > 3.0), while fever and cervical lymphadenopathy[37] also showed significant predictive power. For HIV patients, regardless of symptoms, the number of WHO, TB symptoms, body temperature, hemoglobin,[35] CD4 count < 200 cell/µL,[46] and body mass index[35,46] were moderate to strong predictors.

Biochemical indicators predicted pulmonary TB in HIV-negative people.[33] These included mean corpuscular volume, erythrocyte sedimentation rate, albumin level, adenosine deaminase level, monocyte/high-density lipoprotein ratio, and high-sensitivity CRP/lymphocyte ratio, with moderate to high strength of association. Among smear-negative cases, the absence of sputum and the presence of a cavity estimated pulmonary TB with moderate (OR = 2.0, 95% CI: 1.1–3.3) to strong (OR = 6.3, 95% CI: 2.6–10.0) predictive power, respectively.

Lung cavities or a miliary pattern on chest radiography were particularly strong predictors of pulmonary TB in patients with unknown HIV status, with odds ratios of 8.0 and 5.6, respectively.[50] The number and duration of WHO, TB symptoms, HIV infection, and diabetes mellitus were additional predictors of pulmonary TB.[7]

Diagnostic accuracy of clinical algorithms

Table 4 describes the performance and NNS of a selected population, using culture as the gold standard.

For symptomatic HIV patients, a combined approach using a trial of antibiotics and a C-reactive protein (CRP) level ≥ 10 mg/L demonstrated high sensitivity (94.0%, 95%

CI: 86.0%–98%) in identifying pulmonary TB, but with lower specificity (37.0%, 95% CI: 28%–46%). However, a trial of antibiotics individually displayed lower sensitivities (43%) and higher specificity (86%). For people living with HIV, a CD4 count below 200 cells/mm³ yielded an NNS of 6 (0–26). Similarly, a trial of antibiotics had a low NNS of 2 (0–8). A risk score combining TBSSR and hemoglobin levels demonstrated high sensitivity (88%) and low specificity (55%) for identifying TB in both symptomatic and asymptomatic HIV patients.[35] Urine LAM positivity showed high specificity (82%, 95% CI: 89%–98%) and low sensitivity (31%; 95% CI: 23–39). Furthermore, the duration of cough for at least 2 weeks had an NNS of 19, ranging from 11 to 26.[45,46]

In HIV-negative individuals, a risk score combining multiple biomarkers demonstrated significantly higher diagnostic accuracy than that relying on CRP levels alone. While a CRP level of 10 mg/L offered moderate sensitivity (72.6%) and specificity (71.1%),[36] the multi-biomarker risk score increased sensitivity to 84.1% and specificity to 86.4%.[33]

Discussion

This systematic review shows the strength of the predictors included in diagnostic models for pulmonary TB in a variety of patients. Different algorithms/risk scores have been developed for different clinical and epidemiological settings. Diagnostic accuracies and NNS are shown.

Literature indicates that TB symptoms vary between HIV-infected and non-infected individuals.[27,55] For example, many HIV-infected and culture-positive TB patients do not cough, and some may even be asymptomatic. In this review, typical CXR findings and cervical lymphadenopathy were powerful covariates (OR > 3) among symptomatic HIV patients. However, those on early ART (< 3 months), high body temperature (> 37.5°C), low BMI (less than 18.5 kg/m²), and

Table 2. Methods reported in prediction studies for diagnosis of pulmonary TB.

Characteristics	Studies reporting characteristics	Categories	N (%) or median (IQR)
Type of prediction	26	Criteria based	11 (42.3)
		Model based	15 (57.7)
Missing data	14	No missing	1 (7.1)
		Complete case analysis	4 (28.6)
		Simple imputation	2 (14.3
		Handling is not mentioned	7 (50.0)
Event per variable (EPV)	15	Median	10 (5–21)
Fulfill 10 EPV (≥10)	15	Yes	10 (66.7)
		No	5 (33.3)
Candidate predictor selection	14	Expert's opinion	0 (0.0)
		Less than 10% missing and easy of available variables	1 (7.1)
		Prior selected variables	1 (7.1)
		p-value from Univariate analysis only	11 (78.7)
		p-value from Univariate analysis and AUC	1 (7.1)
Model building procedures	15	Full model approach	4 (26.7)
		Backward elimination	7 (46.6)
		Stepwise selection	4 (26.7)
Purpose of criteria prediction	11	Accuracy test	10 (90.9)
		Updating	1 (9.1)
Purpose of model prediction	15	Development	7 (46.7)
		Development and Validation	7 (46.7)
		Validation	1 (6.6)
Model validation	8	Split sample	3 (37.5)
		Temporal	2 (25.0)
		Geographic	2 (25.0)
		Independent	1 (12.5)
Discrimination (criteria based)	5	AUC	0.77 (0.73–0.78)
Discrimination (model based)	13	AUC	0.82 (0.78–0.84)
Criteria performance	7	Sensitivity	83.0 (48.0–88.0)
		Specificity	66.0 (38.0–83.0)
Model classification	12	Sensitivity	85.0 (67.0–91.0)
		Specificity	65.4 (51.0–77.0)
Model calibration	15	Calibration plot only	1 (6.7)
		Calibration slope and intercept	1 (6.7)
		Calibration plot and Hosmer–Lemeshow	4 (26.6)
		Not reported	9 (60.0)
Model presentation	15	Risk score	7 (46.7)
		Nomo gram	2 (13.3)
		OR	6 (40.0)

AUC: area under the receiver operating characteristic curve; EPV: event per variable; OR: odds ratio; IQR: interquartile range.

low CD4 count; (<200 cells/mm^3) had higher odds of pulmonary TB in the presence of asymptomatic HIV patients in screening practice.

The WHO-recommended four-symptom screening for patients with HIV includes current cough, fever, weight loss, or night sweats.[56] However, these symptoms may not be sufficiently sensitive to detect all TB cases. In line with this, one study indicated that symptom-based screening missed approximately 25% of active TB cases among HIV-infected adults.[57]

Smear-negative pulmonary TB is more common in HIV-infected patients, leading to a delayed diagnosis.[58] Among smear-negative presumptive cases, the presence of a cavity was a powerful diagnostic predictor. This might be the reason for the routine use of CXR as a screening method for smear-negative individuals before GeneXpert and culture, due to resource constraints.[56] In this population, age, respiratory rate, absence of sputum production, and eosinophilia were moderately strong and should be considered candidate predictors for future predictions.

Table 3. Strength of the predictors of pulmonary TB in diagnostic models.

Predictor (reference)	OR (95% CI)
Symptomatic HIV patients	
Body mass index $<18.5\,kg/m^2$, Hanifa et al.[41]	6.8 (2.6–17.6)
Pulmonary cavity[30,37]	4.9 (2.4–7.3)
Hilary and/or mediastinal lymphadenopathy[37]	4.0
Pleural effusion[37]	4.0
Fever [37]	3.7
Cervical lymphadenopathy[37]	3.6
CD4 count $<200\,cells/\mu L$[30,41]	2.6
High body temperature[37]	2.5
Number of WHO symptoms[30,41]	2.4 (1.2–3.6)
Duration of ART $<3\,months$[41]	2.3 (1.2–4.5)
Smoking[30]	2.2
History of TB[30]	1.8
Erythrocyte sedimentation rate[37]	1.6
Symptomatic and asymptomatic HIV patients	
Number of WHO TB symptoms[35]	7.0 (4.7–10.5)
High body temperature[35]	5.5 (3.5–8.7)
Hemoglobin[35]	2.5 (1.3–4.9)
CD4 count $<200\,cells/\mu L$[46]	2.0 (0.9–4.5)
Body mass index[35,46]	2.0 (1.7–2.3)
ART $<3\,months$ [46]	1.9 (0.9–4.0)
Sex-male[35]	1.4 (0.9–2.1)
Smoking[35]	1.3 (1.0–2.1)
HIV-negative patients	
High-sensitivity CRP-to-lymphocyte ratio[33]	8.0 (5.1–12.9)
Adenosine deaminase level[33]	2.9 (1.8–4.8)
Monocyte-to-high-density lipoprotein ratio[33]	2.8 (1.8–4.4)
Erythrocyte sedimentation rate[33]	2.2 (1.4–3.4)
Mean corpuscular volume[33]	0.3 (0.7–0.4)
Albumin level[33]	0.2 (0.1–0.3)
Smear negative and HIV status unknown patients	
Presence of cavities[48,52]	6.3 (2.6-10.0)
Age (26–60) years[52]	2.7 (1.2–6.0)
Respiratory rate $>20/min$[48]	2.5 (1.3–4.8)
RBC morphology-mixed [48]	2.0 (1.1–3.5)
Absence of sputum[52]	2.0 (1.1–3.3)
Eosinophilia <0.4[48]	2.0 (1.3–5.0)
Weight loss[52]	1.6 (0.9–2.7)
Smear and HIV status unknown patients	
Cavity/miliary pattern[50]	8.0 (2.8–23.2)
Upper lobe infiltrate[50]	5.6 (3.2–10.0)
Number WHO TB symptoms[7]	5.4 (3.7–7.8)
HIV infected[7]	3.6 (2.6–4.8)
Sex (male)[7]	2.9 (2.1–4.1)
Weight loss	2.8 (1.5–5.2)
Age (25–34)[7]	2.7 (1.7–4.1)
Duration of symptoms[7]	2.4 (1.8–3.2)
Diabetes miletus[7]	2.0 (0.9–4.4)
History of TB[50]	1.7 (1.3–2.6)

OR: odds ratio; ART: anti-retroviral therapy; WHO: World Health Organization.

In this review, diabetes mellitus and HIV had moderate power in patients whose HIV status was unknown. Bacterial burden and silicosis influence the estimation of pulmonary TB.[10] However, none of the reviewed articles included them in the prediction models. Therefore, these risk factors should be considered in future studies.

In all the populations mentioned, the odds ratio from a series of studies could be different because of differences in the standard errors or sample sizes of the studies.[59] For example, the closer the odds ratio is to unity, the more likely it is to be explained by methodological difficulties such as confounding, misclassification, or other sources of bias.[60] Therefore, prediction rules must be rigorously developed to contribute to clinical practice and decision-making.[61] However, in the present review, more than half of the included studies did not report missing data and approximately 33% of the model-based studies had fewer than 10 events per variable ratio. Therefore, their results had a higher risk of selection bias and overfitting. Adherence to methodological standards such as TRIPOD is essential to avoid such problems.[20]

According to Porta M, an odds ratio alone is not sufficient for discrimination.[62] Even risk factors strongly associated with a disease may have a low ability to discriminate between cases and non-cases. Sensitivity and 1-specificity should be combined using AUC to assess the discriminatory accuracy of a variable.[62] In this review, the AUC for model-based predictions/algorithms showed a 5 percentage point increment over criteria-based prediction. This may be due to the synergistic value of the predictors in the estimation of pulmonary TB.

According to WHO recommendations, any screening algorithm for TB should have a sensitivity of at least 90% and a specificity of 70%.[19] However, the urgency to improve case detection has led to flexibility in the performance criteria for new TB tests. For example, some stakeholders prefer high sensitivity as a prerequisite, whereas others require modest improvements if they are to be accessible to more people.[63]

The target population for TB screening, that is, symptomatic, asymptomatic, smear negative, HIV positive, or HIV negative, should be identified. In symptomatic HIV patients, a CRP level $\geq 10\,mg/L$ increased sensitivity by 50 percentage points and decreased specificity by 46 percentage points compared to the antibiotic trial. However, the composite algorithms of the two did not improve. In addition, in terms of performance for diagnosing pulmonary TB, CXR findings and CD4 count had lower sensitivity than CRP, and the opposite is true when it comes to specificity.

In symptomatic and asymptomatic individuals with HIV, a risk score derived from TB signs and symptoms, risk factors, and hemoglobin levels was more sensitive and less specific. Conversely, urine LAM positivity showed the lowest sensitivity and specificity.

In this population, CRP level $\geq 10\,mg/L$ had a sensitivity and specificity of approximately 80% and 70%, respectively. The specificity of the test was similar in HIV-negative individuals, but the sensitivity was 6 percentage points lower in this group. However, a previous systematic review of the diagnostic accuracy of CRP found a high pooled sensitivity (93%) and moderate specificity (60%) for the detection of pulmonary TB, with no differences according to HIV status.[64]

Table 4. Diagnostic accuracy of clinical algorithms for pulmonary TB among a variety of patients using culture as a reference test.

Algorithm (references)	Sensitivity (%; 95% CI)	Specificity (%; 95% CI)	PPV (%; 95% CI)	NPV (%; 95% CI)	NNS (95% CI)
Symptomatic HIV patients					
Trial of antibiotics and CRP ⩾ 10 mg/L[38]	94.0 (86–98)	37.0 (28–46)	45.0 (37–53)	92.0 (81–98)	2(1–10)
CRP ⩾ 10 mg/L[36,38]	93.0 (91–95)	39.5 (26–52.9)	—	—	3 (1–13)
TBSSR and CXR finding and CD4 count[30,§]	93.0	35.0	—	—	5 ((1–25)
TBSSR and CD4 count and ART[41,§]	91.8 (85–96.2)	34.3 (31.3–37.5)	—	97.3 (94.9–98.7)	7(1–30)
TBSS and ESR and CXR finding[37,§]	76.4 (70.9–81.1)	87.7 (83.4–91.1)	90.7	—	NA
CXR suggestive of TB[43]	72.0(60–82)	57.0 (55–59)	21.0 (17–24)	93.0 (90–95)	5 (1–21)
CD4 count < 200cells/mm3, [43]	69.0(57–80)	46.0 (45–48)	17.0 (14–19)	91.0 (87–94)	6 (1–26)
Trial of antibiotics[38]	43.0 (32–55)	86.0 (78–91)	63.0 (48–76)	76.0 (65–80)	2 (1–8)
Symptomatic and asymptomatic HIV patients					
TBSSR and hemoglobin[35,§]	88.0	55.0	—	98.7	17(15–19)
CRP ⩾ 10 mg/L[40]	78.6 (63.2–89.7)	72.3 (67.6–76.7)	23.7(16.9–31.7)	96.9 (94.1–98.6)	10 (1–19)
Urine LAM positivity[42]	31.0 (23–39)	92.0 (89–94)	46 (35–58)	86.0 (83–88)	2 (1–10)
Cough ⩾ 2 weeks[45,46]	40.8 (20.6–61)	78.0 (61–95)	24.9 (18–31.7)	92.3 (92–92.6)	19 (11–26)
Any of WHO TB symptoms and CXR[45]	50 (36–64)	78.0 (74–82)	24 (16–33)	92.0 (88–95)	4 (1–18)
Symptomatic HIV-negative patients					
Multiple biomarker[33,§]	84.1 (80.4–87.3)	86.4 (81.8–90.1)	—	—	NA
CRP > 10 mg/L[36]	72.6 (67.1–77.8)	71.1 (65.3–76.4)	—	—	NA
Smear-negative and HIV-status unknown patients					
TBSS and hematocrit[53,§]	85.0	67.0	43.0	94.0	3 (3–5)
TBSSR and CXR finding[44,51,52,§]	70 (63.9–88)	82 (63.8–96)	—	—	4 (2–10)
Anti TB Treatment[48]	52 (43–60.9)	62.6 (56.7–68.2)	38.2(31–45.9)	74.6 (68.5–79.9)	4 (3–4)
HIV and smear-status unknown patients					
TBSSR and CXR[50,§]	93.0	42.0	43.0	93.0	3 (3–4)
TBSSR[47,§]	83.1 (77.8–87.6)	51.8 (48.5–55.1)	91.6 (83.2–90.0)	32.9 (30.1–35.7)	6 (5–6)

CI: confidence interval; §: risk score; NA: not applicable because it is a case-control study; TBSS: tuberculosis sign and symptom; TBSSR: tuberculosis sign, symptom and risk factors; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein; LAM: lipoarabinomannan assay; CXR: chest radiograph; PPV: positive predictive value; NPV: negative predictive value; NNS: numbers needed to screen.

This may be due to differences in sample size. In smear-negative patients, anti-TB treatment had the lowest sensitivity and specificity, while a risk score derived from TB symptoms, risk factors, and typical CXR findings had moderate sensitivity (70%) and high specificity (82%). This suggests a need to move beyond sputum smear tests for diagnosing smear-negative patients. A risk score that includes CXR findings offers a potentially better alternative, although further research might be needed to improve its specificity.

The rationale for when to use a test requires judgment about whom to assess, the cost of the test, and the errors that occur if it does not accurately classify patients.[59] Currently available TB diagnostics do not have an "ideal algorithm." However, an algorithm that achieves the lowest NNS and highest PPV, and whose validity is least affected by site-specific variation, is preferable.[65] In symptomatic HIV patients, lower and higher NNS were documented for antibiotic trials and risk scores from TBSSR, CD4 count, and ART, respectively. In symptomatic and asymptomatic HIV patients, urine LAM positivity had an NNS of 2. However, if prolonged cough is used as the initial screening rule for this population, 19 people should be evaluated to detect one person with active TB. This increases the number of healthy individuals undergoing diagnostic testing and exposes them to unplanned out-of-pocket costs.

This review provides a deeper insight into diagnostic prediction and demonstrates the wide range of predictors that influence the diagnosis of pulmonary TB in a resource-limited setting. However, this study has some limitations. The included studies focused on adults who sought healthcare; therefore, the results may not apply to people who do not seek healthcare. Finally, we restricted the inclusion of studies written in English. Therefore, some studies may have been overlooked.

Conclusions

The performance of predictors and diagnostic algorithms varies among patient subgroups. For example, in HIV patients, radiographic findings and body mass index were strong predictors of pulmonary TB. However, in HIV-negative individuals, the biomarkers showed a moderate association with the disease. The results showed that few models had reached the WHO's recommendation, as any screening algorithm for TB should have at least a sensitivity of 90% and a specificity of 70%. Therefore, more work should be done to strengthen the predictive models for TB screening in the future, and they should be developed rigorously, considering the heterogeneity of the population in clinical work.

Author contributions

GBG and GB designed and conceptualized the study. GBG, GB, and KGG searched for, screened, and extracted data. GBG analyzed the data and wrote the first draft of the manuscript. GBG, GB, KGG, and AM reviewed and edited the manuscript. All authors have approved the final manuscript for publication.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval

This study did not involve human participants and ethical approval was not required.

Informed consent

Not required.

ORCID iD

Gebremedhin Berhe Gebregergs 🔟 https://orcid.org/0000-0002-2561-1057

Supplemental material

Supplemental material for this article is available online.

References

1. WHO. *Global tuberculosis report 2023*. Geneva: World Health Organization, 2023.
2. The Global Fund. *Make a global priority of finding missing cases of tuberculosis*. The Global Fund to Fight AIDS, Tuberculosis and Malaria, https://www.theglobalfund.org/en/blog/2017-10-10-make-a-global-priority-of-finding-missing-cases-of-tuberculosis/ (2017, accessed 20 September 2022).
3. Lawn SD and Nicol MP. Xpert MTB/RIF assay: development, evaluation, and implementation of a new rapid molecular diagnostic method for tuberculosis and rifampicin resistance. *Future Microbiol* 2011; 6: 1067–1082.
4. Abdurrahman ST, Mbanaso O, Lawson L, et al. Testing pooled sputum with Xpert MTB/RIF for the diagnosis of pulmonary tuberculosis to increase affordability in low-income countries. *J Clin Microbiol* 2015; 53(8):2502–2508.
5. Meaza A, Tesfaye E, Mohamed Z, et al. Diagnostic accuracy of Truenat Tuberculosis and Rifampicin-Resistance assays in Addis Ababa, Ethiopia. *PLoS One* 2021; 16: e0261084.
6. Parry CM. Sputum smear negative pulmonary tuberculosis. *Trop Doc* 1993; 23: 145–146.
7. Baik Y, Rickman HM, Hanrahan CF, et al. A clinical score for identifying active tuberculosis while awaiting microbiological results: development and validation of a multivariable prediction model in sub-Saharan Africa. *PLoS Med* 2020; 17: e1003420.
8. Adams ST and Leveson SH. Clinical prediction rules. *BMJ* 2012; 344: d8312.
9. WHO. *WHO consolidated guidelines on tuberculosis. Module 2: screening—systematic screening for tuberculosis disease*. Geneva: World Health Organization, 2021.
10. CDC. TB risk factors, http://www.cdc.gov/tb/topic/basics/risk.htm (2022, accessed 09 December 2022).
11. Fletcher RH, Fletcher SW and Fletcher GS. *Clinical epidemiology, the essentials*. New York: Lippincott Williams & Wilkins, 2014.
12. WHO. *Global tuberculosis report*. Geneva, Switzerland: World Health Organization, 2021.
13. Htet KKK, Chongsuvivatwong V and Aung ST. Sensitivity and specificity of tuberculosis signs and symptoms screening and adjunct role of social pathology characteristics in predicting bacteriologically confirmed tuberculosis in Myanmar. *Trop Med Health* 2021; 49: 3.
14. Balcha TT, Skogmar S, Sturegård E, et al. A clinical scoring algorithm for determination of the risk of tuberculosis in HIV-infected adults: a cohort study performed at Ethiopian Health Centers. *Open Forum Infect Dis* 2014; 1(3): ofu095.
15. Gupta RK, Calderwood CJ, Yavlinsky A, et al. Discovery and validation of a personalized risk predictor for incident tuberculosis in low transmission settings. *Nat Med* 2020; 26(12): 1941–1949.
16. Pinto LM, Pai M, Dheda K, et al. Scoring systems using chest radiographic features for the diagnosis of pulmonary tuberculosis in adults: a systematic review. *Eur Respir J* 2013; 42: 480–494.

17. Wyk SSV, Lin HH and Claassens MM. A systematic review of prediction models for prevalent pulmonary tuberculosis in adults. *Int J Tuberc Lung Dis* 2017; 21: 405–411.

18. Jensen SB, Rudolf F and Wejse C. Utility of a clinical scoring system in prioritizing TB investigations—a systematic review. *Expert Rev Anti Infect Ther* 2019; 17: 475–488.

19. WHO. *High-priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting*. Geneva, Switzerland: World Health Organization, 2014.

20. Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Inter Med* 2015; 162(1): W1–W73.

21. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Inter Med* 2009; 151: 264–269.

22. Geersing G-J, Bouwmeester W, Zuithoff P, et al. Search filters for finding prognostic and diagnostic prediction studies in medline to enhance systematic reviews. *PLoS One* 2012; 7; e32844.

23. JBI. Joanna Briggs Institute Reviewers' Manual: 2014 Edition/ Supplement. *The systematic review of diagnostic test accuracy*. The Joanna Briggs Institute, Australia, 2014.

24. World Bank. *Data and statistics: country classification*. Washington, DC: World Bank, 2022.

25. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Inter Med* 2011; 55: 529–536.

26. Moons KGM, de Groot JAH, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: The CHARMS checklist. *PLoS Med* 2014; 11: e1001744.

27. Kranzer K, Houben RMGJ, Glynn JR, et al. Yield of HIV-associated tuberculosis during intensified case finding in resource-limited settings: a systematic review and meta-analysis. *Lancet Infect Dis* 2010; 10: 93–102.

28. Hartwig FP, Davey Smith G, Schmidt AF, et al. The median and the mode as robust meta-analysis estimators in the presence of small-study effects and outliers. *Res Synth Methods* 2020; 11: 397–412.

29. Acharya S, Deshpande P, Asirvatham ES, et al. Utility of the lateral flow urine lipoarabinomannan tuberculosis assay in patients with advanced HIV disease at antiretroviral therapy centers in Mumbai, India. *PLoS One* 2022; 17: e0273970.

30. Chen J, Li L, Chen T, et al. Predicting the risk of active pulmonary tuberculosis in people living with HIV: development and validation of a nomogram. *BMC Infect Dis* 2022; 22: 388.

31. Hidayat A, Murti B, Soedarsono S, et al. Simple tuberculosis screening tool using signs, symptoms, and risk factors to reduce the missed opportunity in the older population. *BMC Pulm Med* 2022; 22: 208.

32. Jerene D, Muleta C, Dressie S, et al. The yield of chest X-ray-based versus symptom-based screening among patients with diabetes mellitus in public health facilities in Addis Ababa, Ethiopia. *J Clin Tuberc Other Mycobact Dis* 2022; 29: 100333.

33. Yu Q, Yan J, Tian S, et al. A scoring system developed from a nomogram to differentiate active pulmonary tuberculosis from inactive pulmonary tuberculosis. *Front Cell Infect Microbiol* 2022; 12: 947954.

34. Saripalli A and Ramapuram J. C-reactive protein as a screening test for tuberculosis in people living with HIV in Southern India: a cross-sectional, observational study. *J Clin Med* 2022; 11: 3566.

35. Auld AF, Kerkhoff AD, Hanifa Y, et al. Derivation and external validation of a risk score for predicting HIV-associated tuberculosis to support case finding and preventive therapy scale-up: a cohort study. *PLoS Med* 2021; 18: e1003739.

36. Samuels THA, Wyss R, Ongarello S , et al. Evaluation of the diagnostic performance of laboratory-based c-reactive protein as a triage test for active pulmonary tuberculosis. *PLoS One* 2021; 16: e0254002.

37. Ouyang J, Yuan J, Chen Y, et al. The development and validation of a diagnostic scoring system to differentiate pulmonary tuberculosis from non-tuberculosis pulmonary infections in HIV-infected patients with severe immune suppression. *BMC Infect Dis* 2021; 21: 863.

38. Boyles TH, Nduna M, Pitsi T, et al. A clinical prediction score including trial of antibiotics and C-reactive protein to improve the diagnosis of tuberculosis in ambulatory people with HIV. *Open Forum Infect Dis* 2020; 7(2): ofz543.

39. George M, Dinant GJ, Kentiba E, et al. Evaluation of the performance of clinical predictors in estimating the probability of pulmonary tuberculosis among smear-negative cases in Northern Ethiopia: a cross-sectional study. *BMJ Open* 2020; 10: e037913.

40. Shapiro AE, Hong T, Govere S, et al. C-reactive protein as a screening test for HIV-associated pulmonary tuberculosis prior to antiretroviral therapy in South Africa. *AIDS* 2018; 32: 1811–1820.

41. Hanifa Y, Fielding KL, Chihota VN, et al. A clinical scoring system to prioritize investigation for tuberculosis among adults attending HIV clinics in South Africa. *PLoS One* 2017; 12: e0181519.

42. Drain PK, Losina E, Coleman SM, et al. Rapid urine lipoarabinomannan assay as a clinic-based screening test for active tuberculosis at HIV diagnosis. *BMC Pulm Med* 2016; 16: 147.

43. Padmapriyadarsini C, Tripathy S, Sekar L, et al. Evaluation of a diagnostic algorithm for sputum smear–negative pulmonary tuberculosis in HIV-infected adults. *J Acquir Immune Defic Syndr* 2013; 63: 331–338.

44. Soto A, Solari L, Agapito J, et al. Algorithm for the diagnosis of smear-negative pulmonary tuberculosis in high-incidence resource-constrained settings. *Trop Med Int Health* 2013; 18: 1222–1230.

45. Swindells S, Komarow L, Tripathy S, et al. Screening for pulmonary tuberculosis in HIV-infected individuals: AIDS Clinical Trials Group Protocol A5253. *Int J Tuberc Lung Dis* 2013; 17: 532–539.

46. Rangaka MX, Wilkinson RJ, Glynn JR, et al. Effect of antiretroviral therapy on the diagnostic accuracy of symptom screening for intensified tuberculosis case finding in a South African HIV clinic. *Clin Infect Dis* 2012; 55: 1698–1706.

47. Castro CB, Costa PA, Ruffino-Netto A, et al. Assessment of a clinical score for screening suspected pulmonary tuberculosis cases. *Rev Saúde Pública* 2011; 45: 1110–1116.

48. Swai HF, Mugusi FM and Mbwambo JK. Sputum smear-negative pulmonary tuberculosis: sensitivity and specificity of diagnostic algorithm. *BMC Res Notes* 2011; 4: 475.

49. Nguyen DTM, Hung NQ, Giang LT, et al. Improving the diagnosis of pulmonary tuberculosis in HIV-infected individuals

in Ho Chi Minh City, Vietnam. *Int J Tuberc Lung Dis* 2011; 15: 1528–1534.

50. Solari L, Acuna-Villaorduna C, Soto A, et al. A clinical prediction rule for pulmonary tuberculosis in emergency departments. *Int J Tuberc Lung Dis* 2008; 12: 619–624.

51. Soto A, Solari L, Agapito J, et al. Development of a clinical scoring system for the diagnosis of smear-negative pulmonary tuberculosis. *BJID* 2008; 12: 128–132.

52. Mello FC, Bastos LG, Soares SL, et al. Predicting smear-negative pulmonary tuberculosis with classification trees and logistic regression: a cross-sectional study. *BMC Public Health* 2006; 6: 43.

53. Samb B, Henzel D, Daley CL, et al. Methods for diagnosing tuberculosis among in-patients in Eastern Africa whose sputum smears are negative. *Int J Tuberc Lung Dis* 1997; 1: 25–30.

54. WHO. *Guidelines for treatment of tuberculosis: fourth edition*. Geneva: World Health Organization, 2010.

55. Reid MJ and Shah NS. Approaches to tuberculosis screening and diagnosis in people with HIV in resource-limited settings. *Lancet Infect Dis* 2009; 9: 173–184.

56. WHO. Guidelines for intensified tuberculosis case-finding and isoniazid preventive therapy for people living with HIV in resource-constrained settings. Geneva: World Health Organiation, 2011.

57. Getahun H, Kittikraisak W, Heilig CM, et al. Development of a standardized screening rule for tuberculosis in people living with HIV in resource-constrained settings: individual participant data meta-analysis of observational studies. *PLoS Med* 2011; 8: e1000391.

58. Saranchuk P, Boulle A, Hilderbrand K, et al. Evaluation of a diagnostic algorithm for smear-negative pulmonary tuberculosis in HIV-infected adults *South Afr Med J* 2007; 97: 517–523.

59. Rothman KJ, Greenland S and Lash TL. *Modern epidemiology*, 3rd edition. Philadelphia, PA: Lippincott Williams & Wilkins, 2008.

60. Brownson RC and Petitti DB. *Applied epidemiology theory to practice*. New York, Oxford: Oxford University Press, 1998.

61. Cowley LE, Farewell DM, Maguire S, et al. Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature. *Diagn Progn Res* 2019; 3: 16.

62. IEA. *A dictionary of epidemiology*. 6th ed. Chicago, IL: International Epidemiological Association, 2014.

63. McNerney R and Daley P. Towards a point-of-care test for active tuberculosis: obstacles and opportunities. *Nat Rev Microbiol* 2011; 9: 204–213.

64. Yoon C, Chaisson LH, Patel SM, et al. Diagnostic accuracy of C-reactive protein for active pulmonary tuberculosis: a meta-analysis. *Int J Tuberc Lung Dis* 2017; 21: 1013–1019.

65. Van't Hoog AH, Onozaki I and Lonnroth K. Choosing algorithms for TB screening: a modeling study to compare yield, predictive value, and diagnostic burden. *BMC Infect Dis* 2014, 14: 532.