Cochrane Library

Cochrane Database of Systematic Reviews

Symptom- and chest-radiography screening for active pulmonary tuberculosis in HIV-negative adults and adults with unknown HIV status (Review)

van't Hoog A, Viney K, Biermann O, Yang B, Leeflang MMG, Langendam MW

van't Hoog A, Viney K, Biermann O, Yang B, Leeflang MMG, Langendam MW.
Symptom- and chest-radiography screening for active pulmonary tuberculosis in HIV-negative adults and adults with unknown HIV status.
Cochrane Database of Systematic Reviews 2022, Issue 3. Art. No.: CD010890.
DOI: 10.1002/14651858.CD010890.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Review]

Symptom- and chest-radiography screening for active pulmonary tuberculosis in HIV-negative adults and adults with unknown HIV status

Anja van't Hoog[1], Kerri Viney[2,3,4], Olivia Biermann[2], Bada Yang[5], Mariska MG Leeflang[6], Miranda W Langendam[5]

[1]Anja van't Hoog, Health Research & Training Consultancy, Utrecht, Netherlands. [2]Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden. [3]School of Public Health, The University of Sydney, Sydney, Australia. [4]Global Tuberculosis Programme, World Health Organization, Geneva, Switzerland. [5]Epidemiology and Data Science, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands. [6]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands

Contact: Miranda W Langendam, m.w.langendam@amsterdamumc.nl.

Editorial group: Cochrane Infectious Diseases Group.
Publication status and date: Edited (no change to conclusions), published in Issue 5, 2022.

Citation: van't Hoog A, Viney K, Biermann O, Yang B, Leeflang MMG, Langendam MW.Symptom- and chest-radiography screening for active pulmonary tuberculosis in HIV-negative adults and adults with unknown HIV status. *Cochrane Database of Systematic Reviews* 2022, Issue 3. Art. No.: CD010890. DOI: 10.1002/14651858.CD010890.pub2.

ABSTRACT

Background

Systematic screening in high-burden settings is recommended as a strategy for early detection of pulmonary tuberculosis disease, reducing mortality, morbidity and transmission, and improving equity in access to care. Questioning for symptoms and chest radiography (CXR) have historically been the most widely available tools to screen for tuberculosis disease. Their accuracy is important for the design of tuberculosis screening programmes and determines, in combination with the accuracy of confirmatory diagnostic tests, the yield of a screening programme and the burden on individuals and the health service.

Objectives

To assess the sensitivity and specificity of questioning for the presence of one or more tuberculosis symptoms or symptom combinations, CXR, and combinations of these as screening tools for detecting bacteriologically confirmed pulmonary tuberculosis disease in HIV-negative adults and adults with unknown HIV status who are considered eligible for systematic screening for tuberculosis disease. Second, to investigate sources of heterogeneity, especially in relation to regional, epidemiological, and demographic characteristics of the study populations.

Search methods

We searched the MEDLINE, Embase, LILACS, and HTA (Health Technology Assessment) databases using pre-specified search terms and consulted experts for unpublished reports, for the period 1992 to 2018. The search date was 10 December 2018. This search was repeated on 2 July 2021.

Selection criteria

Studies were eligible if participants were screened for tuberculosis disease using symptom questions, or abnormalities on CXR, or both, and were offered confirmatory testing with a reference standard. We included studies if diagnostic two-by-two tables could be generated for one or more index tests, even if not all participants were subjected to a microbacteriological reference standard. We excluded studies evaluating self-reporting of symptoms.

Data collection and analysis

We categorized symptom and CXR index tests according to commonly used definitions. We assessed the methodological quality of included studies using the QUADAS-2 instrument. We examined the forest plots and receiver operating characteristic plots visually for heterogeneity. We estimated summary sensitivities and specificities (and 95% confidence intervals (CI)) for each index test using bivariate random-effects methods. We analyzed potential sources of heterogeneity in a hierarchical mixed-model.

Main results

The electronic database search identified 9473 titles and abstracts. Through expert consultation, we identified 31 reports on national tuberculosis prevalence surveys as eligible (of which eight were already captured in the search of the electronic databases), and we identified 957 potentially relevant articles through reference checking. After removal of duplicates, we assessed 10,415 titles and abstracts, of which we identified 430 (4%) for full text review, whereafter we excluded 364 articles. In total, 66 articles provided data on 59 studies. We assessed the 2 July 2021 search results; seven studies were potentially eligible but would make no material difference to the review findings or grading of the evidence, and were not added in this edition of the review.

We judged most studies at high risk of bias in one or more domains, most commonly because of incorporation bias and verification bias. We judged applicability concerns low in more than 80% of studies in all three domains.

The three most common symptom index tests, cough for two or more weeks (41 studies), any cough (21 studies), and any tuberculosis symptom (29 studies), showed a summary sensitivity of 42.1% (95% CI 36.6% to 47.7%), 51.3% (95% CI 42.8% to 59.7%), and 70.6% (95% CI 61.7% to 78.2%, all very low-certainty evidence), and a specificity of 94.4% (95% CI 92.6% to 95.8%, high-certainty evidence), 87.6% (95% CI 81.6% to 91.8%, low-certainty evidence), and 65.1% (95% CI 53.3% to 75.4%, low-certainty evidence), respectively. The data on symptom index tests were more heterogenous than those for CXR. The studies on any tuberculosis symptom were the most heterogeneous, but had the lowest number of variables explaining this variation. Symptom index tests also showed regional variation.

The summary sensitivity of any CXR abnormality (23 studies) was 94.7% (95% CI 92.2% to 96.4%, very low-certainty evidence) and 84.8% (95% CI 76.7% to 90.4%, low-certainty evidence) for CXR abnormalities suggestive of tuberculosis (19 studies), and specificity was 89.1% (95% CI 85.6% to 91.8%, low-certainty evidence) and 95.6% (95% CI 92.6% to 97.4%, high-certainty evidence), respectively. Sensitivity was more heterogenous than specificity, and could be explained by regional variation.

The addition of cough for two or more weeks, whether to any (pulmonary) CXR abnormality or to CXR abnormalities suggestive of tuberculosis, resulted in a summary sensitivity and specificity of 99.2% (95% CI 96.8% to 99.8%) and 84.9% (95% CI 81.2% to 88.1%) (15 studies; certainty of evidence not assessed).

Authors' conclusions

The summary estimates of the symptom and CXR index tests may inform the choice of screening and diagnostic algorithms in any given setting or country where screening for tuberculosis is being implemented. The high sensitivity of CXR index tests, with or without symptom questions in parallel, suggests a high yield of persons with tuberculosis disease. However, additional considerations will determine the design of screening and diagnostic algorithms, such as the availability and accessibility of CXR facilities or the resources to fund them, and the need for more or fewer diagnostic tests to confirm the diagnosis (depending on screening test specificity), which also has resource implications.

These review findings should be interpreted with caution due to methodological limitations in the included studies and regional variation in sensitivity and specificity. The sensitivity and specificity of an index test in a specific setting cannot be predicted with great precision due to heterogeneity. This should be borne in mind when planning for and implementing tuberculosis screening programmes.

PLAIN LANGUAGE SUMMARY

How accurate are asking about symptoms and doing a chest X-ray to screen for tuberculosis of the lungs among adults who are HIV-negative or with unknown HIV status?

Why is improving screening for tuberculosis of the lungs important?

Systematic screening in settings where tuberculosis is common is a recommended strategy for early detection of tuberculosis. Screening helps identify people who are more likely to have tuberculosis so they can have confirmatory testing. These are additional tests to confirm the presence of *Mycobacterium tuberculosis*, the bacterium that causes tuberculosis. Asking about tuberculosis symptoms (for example, cough, coughing up blood, fever, and fatigue) and doing a chest X-ray (CXR), which shows lung abnormalities, are commonly used screening methods. Tuberculosis is treatable with antibiotics, which means that early detection may result in lower mortality and morbidity, less transmission of tuberculosis, and more equitable access to care.

Not recognizing lung tuberculosis when it is present (a false-negative result) may result in delayed treatment and further transmission. Conversely, a screening result that is thought to be positive while it is not may result in unnecessary confirmatory tests, which burdens both the individual and the public health system.

Knowing how often screening tests lead to false-positive and false-negative results – this is called accuracy - may help in choosing a screening method.

What is the aim of this review?

To find out how accurate asking about symptoms and CXR are as screening tests for lung tuberculosis in adults with unknown or negative HIV status.

What was studied in the review?

We studied the accuracy of three types of symptom questions: (i) cough for two or more weeks, (ii) cough of any duration, and (iii) any tuberculosis symptom. For CXR, we studied two definitions for a positive result: (i) any CXR lung abnormality and (ii) CXR lung abnormalities suggestive of tuberculosis. The results are interpreted by staff trained in radiology.

What are the main results in this review?

The review included 59 studies, of which 48 reported on one or more symptom screening questions and 37 reported on CXR.

The results below indicate a situation in which five individuals (0.5%) have lung tuberculosis among a group of 1000 screened individuals.

Cough for two weeks or more: if 1000 individuals were screened, 58 would screen positive, meaning they report cough for two weeks or more and, of these, 56 (97%) would not have lung tuberculosis. Of 1000 individuals, 942 would screen negative, meaning they do not report cough for two weeks or more and, of these, three (0.3%) would have lung tuberculosis.

Cough of any duration: of 1000 individuals, 127 would screen positive and, of these, 124 (98%) would not have lung tuberculosis. Of 1000 individuals, 873 would screen negative and, of these, two (0.2%) would have lung tuberculosis.

Any tuberculosis symptom: of 1000 individuals, 351 would screen positive and, of these, 348 (99%) would not have lung tuberculosis. Of 1000 individuals, 649 would screen negative and, of these, one (0.2%) would have lung tuberculosis.

Any CXR lung abnormality: of 1000 individuals, 113 would show lung abnormalities on CXR and, of these, 108 (96%) would not have lung tuberculosis. Of 1000 individuals, 887 would not show lung abnormalities and, of these, no one (0%) would have lung tuberculosis.

CXR lung abnormalities suggestive of tuberculosis: of 1000 individuals, 48 would screen positive and, of these, 44 (92%) would not have lung tuberculosis. Of 1000 individuals, 952 would screen negative and, of these, one (0.1%) would have lung tuberculosis.

How reliable are the results of the studies in this review?

In the included studies, the diagnosis of tuberculosis was made by assessing the study participants with confirmatory tests (the reference standard). This is the best available method for deciding whether participants really had lung tuberculosis.

However, there were problems with how the studies were conducted. In many studies, those without symptoms or CXR abnormalities were not tested with a confirmatory test. Therefore, the numbers of those without symptoms or CXR abnormalities, but nevertheless having tuberculosis (people who tested falsely negative), may have been underestimated in these studies. Consequently, screening for symptoms or CXR abnormalities might appear more accurate than it really is.

In addition, results from individual studies included in the review varied, for example, because of regional variation. Therefore, we cannot be sure that screening for symptoms and CXR abnormalities will always have the same accuracy.

What are the implications of this review?

The results of the review suggest that screening for tuberculosis with symptom questions or CXR might result in a high yield of persons with tuberculosis disease. However, this screening might also result in a high proportion of persons without the disease screening positive. Additional considerations for the best design of screening programmes include the local epidemiological situation, availability and accessibility of CXR, and the need for confirmatory tests.

How up to date is this review?

The review authors searched for and included studies published from 1 January 1992 to 10 December 2018. A repeat of the search to 2 July 2021 revealed no further studies that would inform the results of the analysis.

SUMMARY OF FINDINGS

Summary of findings 1. Cough for two weeks or more

Review question: what is the accuracy of questioning for the presence of prolonged cough (2 weeks or more) as a screening test for detecting pulmonary tuberculosis disease in a general population of people with HIV-negative or unknown HIV status who are considered eligible for systematic screening for tuberculosis disease.

Role of index test: individuals with a positive screening test are offered further confirmatory testing to establish a tuberculosis diagnosis.

Reference standards: any one or combination of mycobacterial culture (on solid or liquid medium), sputum smear microscopy, Xpert MTB/RIF, or other nucleic acid amplification test.

Study design: cross-sectional studies.

Cough for two weeks or more **summary sensitivity** (95% CI): 42.1% (36.6 to 47.7); **summary specificity** (95% CI): 94.4% (92.6 to 95.8)

Index test result	Results per 1000 participants tested (95% CI)*			Number of participants (studies)	Certainty of evidence (GRADE)
	Prevalence of 0.5%*	Prevalence of 1%*	Prevalence of 2%*		
True positives (participants with tuberculosis disease)	2 (2 to 2)	4 (4 to 5)	8 (7 to 10)	7179 (41)	VERY LOW[a] ⊕◯◯◯
False negatives (participants incorrectly classified as not having tuberculosis disease)	3 (3 to 3)	6 (5 to 6)	12 (10 to 13)		
True negatives (participants without tuberculosis disease)	939 (921 to 953)	935 (917 to 948)	925 (907 to 939)	1,540,179 (41)	HIGH[b] ⊕⊕⊕⊕
False positives (participants incorrectly classified as having tuberculosis disease)	56 (42 to 74)	55 (42 to 73)	55 (41 to 73)		

Abbreviations: CI: confidence interval

*Tuberculosis disease prevalences of 0.5%, 1% and 2% were chosen based on the median prevalences among the studies for each index test, which ranged from 0.5% to 1.7% and to be consistent with the WHO guidelines (WHO 2021a).

[a]Downgraded by three, due to very serious risk of bias and serious inconsistency.

Very serious risk of bias is because in the QUADAS-2 Reference Standard domain more than three-quarters of the studies did not require all participants to undergo microbacteriological testing, but classified tuberculosis negative in those participants based on results of chest radiography (CXR) and symptoms (incorporation bias). Flow and Timing: more than half of the studies scored high risk of bias. Of all participants who required microbacteriological testing based on the protocol, less than 95% had a result. Sensitivity analysis showed that studies with low risk bias in these QUADAS-2 domains had considerably lower sensitivity (most extreme: studies with low risk for Reference Standard (8 studies): sensitivity 29.3% (95% confidence interval (CI) 19.4% to 41.7%).

Serious inconsistency is because of the very wide range in point estimates (10% to 100%), with some overlap of the CIs. In stratified analysis, population level variables that significantly ($P < 0.05$) modified the summary estimates were economic region and higher vs lower (< 0.5%) tuberculosis prevalence among the study participants. Study design variables that significantly modified the summary estimates were presence of incorporation bias and whether the reference standard included culture or not (but a combination of smear and Xpert MTB/RIF).

[b]Inconsistency was not judged as serious. Although the range in point estimates is wide (specificity 68% to 99%), the CIs considerably overlap. A few outlying values are of studies that share a quality concern in the patient selection domain. Variables that may explain heterogeneity in specificity were economic region and tuberculosis prevalence among the study participants.

Summary of findings 2. Cough of any duration (any cough)

Review question: what is the accuracy of questioning for the presence of cough of any duration (any cough) as a screening test for detecting pulmonary tuberculosis disease in a general population of people with HIV-negative or unknown HIV status who are considered eligible for systematic screening for tuberculosis disease.

Role of index test: individuals with a positive screening test are offered further confirmatory testing to establish a tuberculosis diagnosis.

Reference standards: any one or combination of mycobacterial culture (on solid or liquid medium), sputum smear microscopy, Xpert MTB/RIF, or other nucleic acid amplification test.

Study design: cross-sectional studies.

Cough of any duration **summary sensitivity** (95% CI): 51.3% (42.8 to 59.7); **summary specificity** (95% CI): 87.6% (81.6 to 91.8)

Index test result	Results per 1000 participants tested (95% CI)*			Number of participants (studies)	Certainty of evidence (GRADE)
	Prevalence of 0.5%*	Prevalence of 1%*	Prevalence of 2%*		
True positives (participants with tuberculosis disease)	3 (2 to 3)	5 (4 to 6)	10 (9 to 12)	2734 (21)	VERY LOW[a] ⊕◯◯◯
False negatives (participants incorrectly classified as not having tuberculosis disease)	2 (2 to 3)	5 (4 to 6)	10 (8 to 11)		
True negatives (participants without tuberculosis disease)	871 (812 to 913)	867 (808 to 908)	858 (800 to 899)	768,291 (21)	LOW[b] ⊕⊕◯◯
False positives (participants incorrectly classified as having tuberculosis disease)	124 (82 to 183)	123 (82 to 182)	122 (81 to 180)		

Abbreviations: CI: confidence interval
*Tuberculosis disease prevalences of 0.5%, 1% and 2% were chosen based on the median prevalences among the studies for each index test, which ranged from 0.5% to 1.7% and to be consistent with the WHO guidelines (WHO 2021a).
[a]Downgraded by three, due to very serious risk of bias and serious inconsistency.
Very serious risk of bias because in the QUADAS-2 Reference Standard domain more than half of the studies did not require all participants to undergo microbacteriological testing, but classified tuberculosis negative in those participants based on results of CXR and symptoms (incorporation bias). Flow and Timing: about one-third of the studies scored high risk of bias. Of all participants who required microbacteriological testing based on the protocol, less than 95% had a result. Sensitivity analysis showed that studies with low risk bias in these QUADAS-2 domains had considerably lower sensitivity (most extreme: studies with low risk for Reference Standard (8 studies): sensitivity 35.6% (95% CI 18.8% to 56.8%).

Serious inconsistency due to a very wide range in point estimates (0% to 100%), with some overlap of the CIs. Some of the heterogeneity could be explained by economic region. Studies in low-income countries showed higher sensitivity (64.8%, 95% CI 54.8% to 73.6%), in upper/middle/high income studies sensitivity was lower (34.4%, 95% CI 23.3% to 47.5%).

[b]Downgraded by two due to serious inconsistency and serious imprecision.

Serious inconsistency due to a wide range in point estimates (specificity 43% to 99%) without overlap of CIs. No statistical significant variables that could explain heterogeneity, however in low-income countries the sensitivity was somewhat lower (80.8%, 95% CI 69.1% to 88.9%) than in the upper/middle/high-income studies.

Serious imprecision since the CI around the false positives is as such that the proportion of the population requiring follow-up testing can vary by more than a factor two, which has serious resource implications.

Summary of findings 3. Any tuberculosis symptom

Review question: what is the accuracy of questioning for the presence any tuberculosis symtpom as a screening test for detecting pulmonary tuberculosis disease in a general population of people with HIV-negative or unknown HIV status who are considered eligible for systematic screening for tuberculosis disease.

Role of index test: individuals with a positive screening test are offered further confirmatory testing to establish a tuberculosis diagnosis.

Reference standards: any one or combination of mycobacterial culture (on solid or liquid medium), sputum smear microscopy, Xpert MTB/RIF, or other nucleic acid amplification test.

Study design: Cross-sectional studies.

Any tuberculosis symptom **summary sensitivity** (95% CI): 70.6% (61.7 to 78.2); **summary specificity** (95% CI): 65.1% (53.3 to 75.4)

Index Test result	Results per 1000 participants tested (95% CI)*			Number of participants (studies)	Certainty of evidence (GRADE)
	Prevalence of 0.5%*	Prevalence of 1%*	Prevalence of 2%*		
True positives (participants with tuberculosis disease)	4 (3 to 4)	7 (6 to 8)	14 (12 to 16)	4180 (29)	VERY LOW[a] ⊕◯◯◯
False negatives (participants incorrectly classified as not having tuberculosis disease)	1 (1 to 2)	3 (2 to 4)	6 (4 to 8)		
True negatives (participants without tuberculosis disease)	648 (530 to 750)	644 (528 to 746)	638 (522 to 739)	506,712 (29)	LOW[b] ⊕⊕◯◯
False positives (participants incorrectly classified as having tuberculosis disease)	347 (245 to 465)	342 (241 to 458)	355 (252 to 473)		

Abbreviations: CI: confidence interval

*Tuberculosis disease prevalences of 0.5%, 1% and 2% were chosen based on the median prevalences among the studies for each index test, which ranged from 0.5% to 1.7% and to be consistent with the WHO guidelines (WHO 2021a).

[a]Downgraded by three, due to very serious risk of bias and serious inconsistency.

Very serious risk of bias because in the QUADAS-2 Reference Standard domain more than half of the studies did not require all participants to undergo microbacteriological testing, but classified tuberculosis negative in those participants based on results of CXR and symptoms (incorporation bias). Flow and Timing: about one-third of the studies scored high risk of bias. Of all participants who required microbacteriological testing based on the protocol, less than 95% had a result. Sensitivity analysis showed that studies with low risk bias in these QUADAS-2 domains had considerably lower sensitivity (most extreme: studies with low risk for Reference Standard (12 studies): sensitivity 62.9% (95% CI 47.4% to 76.1%) and Flow and Timing (9 studies): sensitivity 62.9% (95% CI 43.5% to 78.9%).
Serious inconsistency. Very wide range in point estimates (18% to 100%), with overlap of the CIs. Some of the heterogeneity could be explained by economic region. Studies in low-income countries showed higher sensitivity (78.9%, 95% CI 69.3% to 86.2%), in upper/middle/high-income studies sensitivity was lower (56.3%, 95% CI 40.6% to 70.8%).

[b]Downgraded by two due to serious inconsistency and serious imprecision.
Serious inconsistency due to a wide range in point estimates (13% to 99%) without overlap of CI and no variables that statistically significantly explained heterogeneity.
Serious imprecision since the CI around the false positives is as such that the proportion of the population requiring follow-up testing can vary by almost a factor two, which has serious resource implications.

Summary of findings 4. Any CXR abnormality

Review question: what is the accuracy of any (pulmonary) CXR abnormality as a screening test for detecting pulmonary tuberculosis disease in a general population of people with HIV-negative or unknown HIV status who are considered eligible for systematic screening for tuberculosis disease.

Role of index test: individuals with a positive screening test are offered further confirmatory testing to establish a tuberculosis diagnosis.

Reference standards: any one or combination of mycobacterial culture (on solid or liquid medium), sputum smear microscopy, Xpert MTB/RIF, or other nucleic acid amplification test.

Study design: cross-sectional studies.

Any CXR abnormality **summary sensitivity** (95% CI): 94.7% (92.2 to 96.4); **summary specificity** (95% CI): 89.1% (85.6 to 91.8)

Index Test result	Results per 1000 participants tested (95% CI)*			Number of participants (studies)	Certainty of evidence (GRADE)
	Prevalence of 0.5%*	Prevalence of 1%*	Prevalence of 2%*		
True positives (participants with tuberculosis disease)	5 (5 to 5)	9 (9 to 10)	19 (18 to 19)	4532 (23)	VERY LOW[a] ⊕◯◯◯
False negatives (participants incorrectly classified as not having tuberculosis disease)	0 (0 to 0)	1 (0 to 1)	1 (1 to 2)		
True negatives (participants without tuberculosis disease)	887 (852 to 913)	882 (847 to 909)	873 (839 to 900)	1.034.525 (23)	LOW[b] ⊕⊕◯◯
False positives (participants incorrectly classified as having tuberculosis disease)	108 (82 to 143)	108 (81 to 143)	107 (80 to 141)		

Abbreviations: CI: confidence interval

*Tuberculosis disease prevalences of 0.5%, 1% and 2% were chosen based on the median prevalences among the studies for each index test, which ranged from 0.5% to 1.7% and to be consistent with the WHO guidelines (WHO 2021a).

[a]Downgraded by three, due to very serious risk of bias and serious inconsistency.

Very serious risk of bias since only 2 studies had low risk of bias in the Reference Standard domain. Less than half of the studies had low risk in the Flow and Timing domain.

Serious inconsistency because of a moderate range in sensitivity (70% to 100%) with some overlap in CIs. Variables that may explain observed variation are WHO region (Africa versus Asia/Pacific/other), prevalence of tuberculosis in the study population, and prevalence of smoking in the population (10% or more versus lower).

[b]Downgraded by two due to serious inconsistency and serious imprecision.

Serious inconsistency due to a moderate range in specificity (71% to 99%). The variable that may explain observed variation is whether the CXR was read of any abnormality including other visible organs (82.4%, 95% CI 73.8% to 88.6%) versus pulmonary abnormalities (91.1%, 95% CI 87.8% to 93.5%).

Serious imprecision because the CI around the false positives is as such that the proportion of the population requiring follow-up testing can vary by almost a factor two, which has serious resource implications.

Summary of findings 5. CXR abnormalities suggestive of tuberculosis

Review question: what is the accuracy of CXR abnormalities suggestive of tuberculosis as a screening test for detecting pulmonary tuberculosis disease in a general population of people with HIV-negative or unknown HIV status who are considered eligible for systematic screening for tuberculosis disease.

Role of index test: individuals with a positive screening test are offered further confirmatory testing to establish a tuberculosis diagnosis.

Reference standards: any one or combination of mycobacterial culture (on solid or liquid medium), sputum smear microscopy, Xpert MTB/RIF, or other nucleic acid amplification test.

Study design: cross-sectional studies.

CXR abnormalities suggestive of tuberculosis **summary sensitivity** (95% CI): 84.8% (76.7 to 90.4); **summary specificity** (95% CI): 95.6% (92.6 to 97.4)

Index Test result	Results per 1000 participants tested (95% CI)*			Number of participants (studies)	Certainty of evidence (GRADE)
	Prevalence of 0.5%*	Prevalence of 1%*	Prevalence of 2%*		
True positives (participants with tuberculosis disease)	4 (4 to 5)	8 (8 to 9)	17 (15 to 18)	2152 (19)	LOW[a] ⨁⨁◯◯
False negatives (participants incorrectly classified as not having tuberculosis disease)	1 (0 to 1)	2 (1 to 2)	3 (2 to 5)		
True negatives (participants without tuberculosis disease)	951 (922 to 969)	946 (917 to 964)	937 (908 to 954)	464,818 (19)	HIGH ⨁⨁⨁⨁
False positives (participants incorrectly classified as having tuberculosis disease)	44 (26 to 73)	44 (26 to 73)	43 (26 to 72)		

Abbreviations: CI: confidence interval

*Tuberculosis disease prevalences of 0.5%, 1% and 2% were chosen based on the median prevalences among the studies for each index test, which ranged from 0.5% to 1.7% and to be consistent with the WHO guidelines (WHO 2021a).

[a]Downgraded by two, due to serious risk of bias and serious inconsistency.

Serious risk of bias since only 3 of the 19 studies had low risk of bias in the Reference Standard domain and only 3 of 19 the studies had low risk in the Flow and Timing domain. The sensitivity in studies with low risk in domain 3 or domain 4 is lower compared to studies with high or unknown risk.

Serious inconsistency due to a wide range in sensitivity (37% to 100%) with some overlap in CIs. Variables that may explain observed variation are WHO region (Africa versus Asia/Pacific/other), and HIV prevalence, although the latter was not statistically significant (P = 0.074).

BACKGROUND

Target condition being diagnosed

Tuberculosis is an important infectious cause of morbidity and mortality among adults worldwide. In 2019, there were an estimated 10 million new cases of tuberculosis disease with 1.2 million tuberculosis deaths among HIV-negative people and an additional 208,000 deaths among people living with HIV infection (WHO 2020). An estimated one-quarter of the world's population is infected with *Mycobacterium tuberculosis*, the micro-organism that causes tuberculosis (Houben 2016). In humans, *M tuberculosis* infection typically affects the lungs (although it can affect other sites) and spreads by airborne transmission (Lawn 2011). Patients with infectious tuberculosis spread bacilli, most commonly through coughing. After initial infection, approximately 5% to 10% of infected people develop tuberculosis disease, also called active tuberculosis. Between 90% to 95% of infected people develop a tuberculosis infection, which may reactivate at a later stage, especially in the presence of conditions that affect immunity (including HIV infection, undernutrition, and old age) (Rieder 1999). It can take months to years for people to develop symptomatic and bacteriologically detectable tuberculosis. Tuberculosis infection and tuberculosis disease are increasingly seen as two ends of a continuous spectrum. In between these are early disease states that may be described as incipient tuberculosis and subclinical tuberculosis (Achkar 2011; Pai 2016). In the absence of diagnosis and treatment, people with tuberculosis disease may be infectious for prolonged time periods. Among HIV-negative people with tuberculosis disease, the average duration until self-cure or death is three years, and case fatality with no treatment is approximately 70% for people with sputum smear-positive tuberculosis (that is, tuberculosis detectable using sputum smear microscopy) and 20% for smear-negative tuberculosis (Tiemersma 2011).

Estimated incidence of tuberculosis disease is declining globally, albeit slowly. The cumulative decline in estimated global tuberculosis incidence has been 9% between the years 2015 and 2019, however this is below the cumulative decline of 20% required to reach the 2020 milestone target of the World Health Organization's (WHO) End TB Strategy, the goals set at the United Nations High Level Meeting on Tuberculosis and the Sustainable Development Goals (WHO 2020). In 2019, an estimated 2.9 million people who developed tuberculosis remained undiagnosed or unreported (WHO 2020). National tuberculosis prevalence surveys have revealed a considerable burden of undiagnosed culture-positive (that is, detectable with mycobacterial sputum culture), smear-negative tuberculosis, and only a minority of those cases report classical symptoms of tuberculosis (Law 2020; Onozaki 2015). Most tuberculosis cases are detected passively, among symptomatic people seeking health care (Golub 2005). Passive tuberculosis case detection results in considerable delays in tuberculosis detection (Sreeramareddy 2009), and at the time of diagnosis tuberculosis patients identified through passive tuberculosis case detection have more signs of illness compared to patients found through active tuberculosis case detection (den Boon 2008; van't Hoog 2013). Thus, a large proportion of patients with infectious tuberculosis will go undiagnosed if only passive case detection is used. Improving tuberculosis case detection to ensure early detection and treatment of those with undiagnosed tuberculosis, and reduce the pool of infectious tuberculosis that contributes to transmission (Corbett 2010; Marks 2019), is important to further reduce tuberculosis incidence and mortality, and reach the global tuberculosis-related goals of the End TB Strategy (WHO 2015). Therefore, more active approaches are needed to increase tuberculosis case detection, and systematic screening for tuberculosis disease is a possible means of achieving this (Burke 2021; Lönnroth 2013).

Screening

In the guidelines on tuberculosis screening developed by WHO and partners, systematic screening for tuberculosis disease is defined as "the systematic identification of people with presumed tuberculosis disease, in a predetermined target group, using tests, examinations, or other procedures which can be applied rapidly" (WHO 2013). Screening tests sort out apparently well people who probably have a disease from people who probably do not and are not intended to be diagnostic. People with positive or suspicious findings must be referred for diagnosis and necessary treatment (Wilson 1968). Screening is offered systematically to predetermined groups, and not only in response to a specific request or complaint by an individual seeking care (Lönnroth 2013; WHO 2013). While it often refers to screening outside of health facilities, it can be offered to those who seek health care (with or without signs and symptoms compatible with tuberculosis) and those who do not. The two main goals of systematic screening for tuberculosis disease are (1) better health outcomes for people with tuberculosis, through earlier detection and treatment; and (2) more effective reduction of tuberculosis transmission and incidence through shortening the average duration of tuberculosis infectiousness (Burke 2021; Lönnroth 2013; WHO 2013).

Index test(s)

This review focused on symptom and chest radiography (CXR) screening. For symptom screening, individuals are questioned about the presence of one or more symptoms considered suggestive of pulmonary tuberculosis, which are respiratory symptoms such as persistent cough, haemoptysis, or systemic symptoms including weight loss, fever, night sweats, and fatigue (Maher 2009). Symptom questions may be asked by health workers or by trained lay workers. CXR as a screening tool involves having participants undergo one posterior-anterior CXR recording. Different radiography technologies exist: conventional CXR (producing a 36 cm x 43 cm film), digital radiography, and mass miniature radiography (MMR) (Kerley 1942). In addition, CXR classification systems may distinguish between presence or absence of signs of lung parenchyma abnormalities, or signs suggestive of tuberculosis, or additionally for signs of possible extra-pulmonary tuberculosis such as pleural abnormalities or cardiomegaly. Signs suggestive of tuberculosis require interpretation by specialist CXR readers (usually radiologists or pulmonologists), while the presence of any (pulmonary) radiographic abnormality can be more easily interpreted by healthcare workers with a general medical background (for example, medical officers, clinical officers, and radiographers) (van't Hoog 2011; WHO 2011).

Systematic screening for tuberculosis disease may be offered with either symptom or CXR screening, or with symptom and CXR screening combined in parallel or sequentially (Figure 1; Hayen 2010; van't Hoog 2014a). Sequential (or serial) screening means that people are initially screened for symptoms, and are subsequently offered CXR if they have one or more symptoms.

Parallel screening implies that both symptom and CXR screening are offered at the same time, and people found to have symptoms, or CXR abnormalities, or both, are eligible for further microbacteriological examination. This is practised in tuberculosis prevalence surveys, for example, to achieve the highest sensitivity possible, while at the same time avoiding the need for laboratory investigations on all people being screened (WHO 2011).

Figure 1. Algorithms composed of one or more screening methods and one or more confirmatory tests. In panel A one screening tool is applied (e.g. symptoms) and screen positives are further evaluated by one confirmatory test with high sensitivity and high specificity (e.g. Xpert MTB/RIF). In panel B one screening tool is applied (e.g. symptoms) and screen positives are further evaluated by a confirmatory test with low sensitivity (e.g. sputum smear microscopy), and persons with a negative test receive a second test or procedure (e.g. clinical diagnosis, or sputum culture). In panel C two screening tools are applied (e.g. symptoms and chest radiography) and screen positives on either one or on both are further evaluated with a confirmatory test. In panel D two screening tools are applied sequentially. Screen positives on the first screen (e.g. symptoms) undergo a second screen (e.g. CXR) and if

also positive on the second a confirmatory test is applied. The single confirmatory test in panels C and D could also be replaced by two steps as in panel B (van't Hoog 2014a).

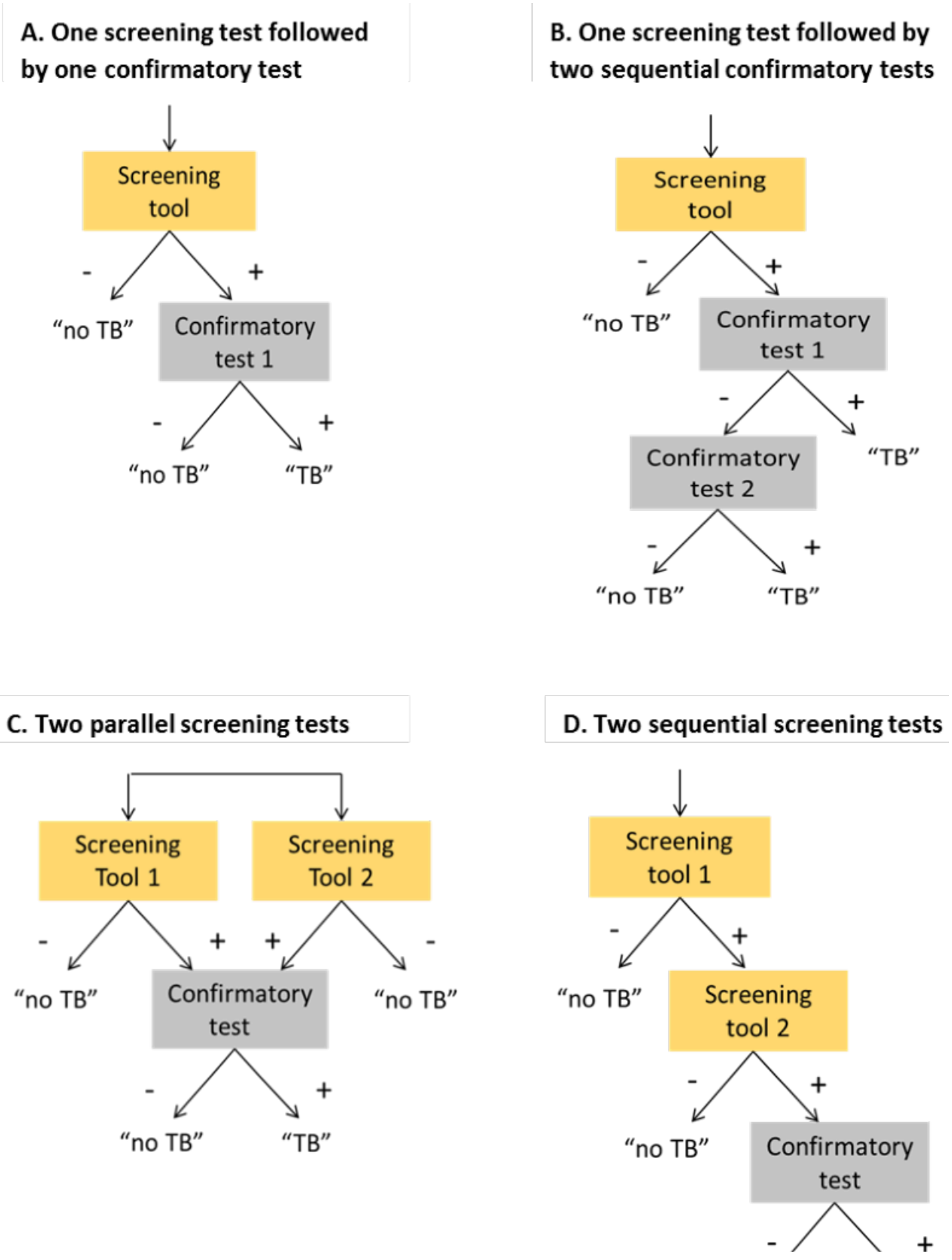
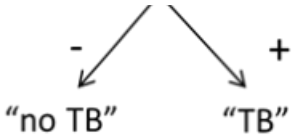


Figure 1. (Continued)



Clinical pathway

In a tuberculosis screening programme, the screening test(s) are offered as part of a diagnostic algorithm that also includes one or more confirmatory tests. Individuals with a positive screen are offered further confirmatory testing to establish a tuberculosis diagnosis. So called "true screen positives" are people rightfully referred for confirmatory testing as they have tuberculosis disease, and "false screen positives" are people who are referred for confirmatory testing but do not have tuberculosis disease. Individuals with a positive screen, but negative confirmatory test would not necessarily be declared disease-free, but may be advised on further examination or follow-up if warranted by the findings on screening (for example, severity of symptoms or the CXR finding; Okada 2012). People with a negative screen would not be further evaluated. This group includes both the "true screen negatives" who do not have tuberculosis and "false screen negatives", who will not be evaluated further, although they do have tuberculosis. The confirmatory test may be a nucleic acid amplification test (NAAT) such as the Xpert® MTB/RIF assay (Cepheid, Sunnyvale, CA), sputum smear microscopy, and in more resourced settings, mycobacterial culture, either liquid or solid. These are also the reference standards for the purpose of this review. People who have a negative confirmatory test result may be started on tuberculosis treatment after further clinical evaluation and a trial of broad spectrum antibiotics, or CXR, or both.

Reference standards

The reference standard in this review includes microbiological confirmation of pulmonary tuberculosis disease, for which one or several test methods may be used:

Mycobacterial culture

Confirmation of mycobacterial growth in cultured sputum followed by mycobacterial speciation to demonstrate *M tuberculosis* presence is considered to be the reference standard. Culture on liquid medium is believed to be the most sensitive, although prior to the availability of automated reading of mycobacterial growth inhibitor tubes (MGIT culture), culture on solid medium (Löwenstein-Jensen (LJ)) was the mainstay, and may still be the only available method in resource-constrained settings. MGIT culture increases the recovery of mycobacteria by 11% to 18% compared to LJ culture, but MGIT culture alone may have slightly lower specificity due to higher contamination rates (Chien 2000; Hanna 1999; Somoskövi 2000; Whitelaw 2009). The yield of mycobacterial culture also increases if two or three specimens per patient are tested (Monkongdee 2009).

Nucleic acid amplification tests (NAAT)

The Xpert® MTB/RIF assay is recommended by WHO as an initial diagnostic test for tuberculosis detection in sputum in adults with signs and symptoms of pulmonary tuberculosis (WHO 2020a). Xpert Ultra and Truenat MTB or MTB Plus may also be used (WHO 2020a). Xpert MTB/RIF is longest available and has, compared to culture, summary sensitivity of 85% (95% credible interval (CrI) 82% to 88%) and specificity 98% (95% CrI 97% to 98%) in a systematic review of 70 studies with high-certainty evidence (Horne 2019).

Sputum smear microscopy

Sputum smear microscopy is still a commonly available tuberculosis diagnostic test, although no longer recommended by WHO as the primary diagnostic method (WHO 2020a). Sputum smear microscopy detects the presence of acid fast bacilli (AFB), which is considered indicative of *M tuberculosis* in high tuberculosis-incidence settings. Compared to culture, the sensitivity of the Ziehl-Neelsen method (ZN) of sputum smear microscopy shows wide variation, and is between 50% and 70% in many studies (Steingart 2006a; Steingart 2006b). Direct ZN microscopy specificity is 98% (95% confidence interval (CI) 97% to 99%) (Cattamanchi 2010; Steingart 2006a; Steingart 2006b). Sputum smears may also be positive due to the presence of AFB that are not *M tuberculosis,* or to artefacts. Auramine-stained fluorescence microscopy (FM) sensitivity has on average a 10% higher sensitivity than ZN, but with slightly reduced specificity (Steingart 2006a). Processing sputum by centrifugation and various chemicals, including bleach and sodium hydroxide (NaOH), shows varying levels of increase in the sensitivity of sputum smear microscopy compared with the direct smear method, and similar or slightly lower specificity (Cattamanchi 2010; Steingart 2006b).

Other types of tuberculosis disease

Other types of tuberculosis disease not covered by this review include tuberculosis in sites outside the lungs, collectively called extra-pulmonary tuberculosis (EPTB), a condition that may affect almost every other organ and which constitutes 16% of new and relapse tuberculosis cases globally (WHO 2020), and culture-negative pulmonary tuberculosis, characterized by clinical disease, highly suggestive CXR abnormalities not explained by other causes, but with a negative sputum culture (Maher 2009). Clinical diagnosis and commencement of empirical tuberculosis treatment is often practised in settings where mycobacterial culture is not part of routine diagnosis for people with presumed pulmonary tuberculosis disease who have negative sputum smears or often in the case of presumed EPTB. Clinical algorithms that include a trial of antibiotics and a CXR if the trial was not successful generally have very low sensitivity, while tuberculosis diagnosis based on CXR alone has low specificity (Soto 2011; Swai 2011; van Cleeff 2003). In this review, we did not consider clinically diagnosed tuberculosis as an acceptable reference standard because of the lack of a uniform definition, poor and variable accuracy of clinical algorithms, and the varying ability to establish differential diagnoses across settings. EPTB and culture-negative pulmonary tuberculosis may be detected through screening, especially in high-income countries, but are not a primary focus of screening

programmes in other settings due to diagnostic challenges and low probability of transmission. Also, we did not consider serological tests in our review and they are not recommended for diagnosis of tuberculosis disease by the WHO (Steingart 2007).

Rationale

This review aimed to contribute to updated WHO guidelines on systematic screening for tuberculosis, which provide guidance about whether, when, whom, and how to screen for tuberculosis disease (WHO 2013; WHO 2021a). We compiled evidence about the accuracy of the most frequently available screening tools, and where possible generated summary estimates of the sensitivity and specificity of symptoms, CXR, and combinations of those if used as screening tools for tuberculosis disease. The accuracy of the screening tools and the confirmatory tests, as well as the tuberculosis prevalence in the screened population, will determine the potential yield of any screening programme and the burden on individuals and the health service. The latter includes the required amount of confirmatory tests and possibly diagnostic tests and procedures for other conditions. The WHO tuberculosis screening guidelines aim to provide evidence-based recommendations on populations that would benefit from screening, and on the choice of diagnostic algorithms (combinations of one or more screening test(s)) and confirmatory test(s)) in different populations and settings (Lönnroth 2013; WHO 2013). Therefore the yield of tuberculosis, the positive and negative predictive values, and the requirements in terms of diagnostic tests for different diagnostic algorithms have been calculated for different levels of tuberculosis prevalence as part of the guideline development process (van't Hoog 2014a; WHO 2021a). This information should help policy-makers choose the best diagnostic algorithm option for their specific setting, taking into account the background tuberculosis prevalence, available resources, and other logistical considerations (for example, the availability of radiology or NAAT equipment). The summary estimates of sensitivity and specificity of symptom and CXR screening from this review can inform such calculations and recommendations.

This review addresses tuberculosis screening among HIV-negative people and people with unknown HIV status (a proportion of whom may be HIV-infected). The risk of developing tuberculosis disease among the 38 million people living with HIV was 18 (range 15 to 21) times higher than in the rest of the global population (WHO 2020). In resource-limited settings, tuberculosis disease is a common and often undiagnosed cause of death (Gupta 2015). The sensitivity of sputum smear microscopy and Xpert MTB/RIF is lower in HIV-infected individuals with presumed tuberculosis (Getahun 2007; Horne 2019). Therefore people living with HIV should be regularly and systematically screened for tuberculosis disease according to a clinical algorithm, at each visit to a healthcare facility (WHO 2018). In addition, people living with HIV who are unlikely to have active tuberculosis and should be offered preventive treatment, regardless of their antiretroviral therapy (ART) status. Screening algorithms for people living with HIV have been defined based on other systematic reviews (Getahun 2011; Hamada 2018). The recommended algorithm to rule out tuberculosis disease used to be the Four Symptom Screen, meaning absence of any of the four symptoms of current cough, fever, weight loss, or night sweats and, if available, absence of abnormal radiographic findings (Getahun 2011; WHO 2018). The latest WHO guidelines recommend the Four Symptom Screen, C-reactive protein, CXR and molecular WHO-recommended rapid diagnostic tests, depending on the kind of sub-population and availability of resources (WHO 2021a). If tuberculosis screening among people with unknown HIV status is combined with HIV testing, then people who are HIV-positive should be referred for HIV care and treatment if they are not yet enrolled.

OBJECTIVES

To assess the sensitivity and specificity of questioning for the presence of one or more tuberculosis symptoms, or symptom combinations, or both; CXR; and combinations of these as screening tools for detecting bacteriologically confirmed pulmonary tuberculosis disease in people considered eligible for tuberculosis screening who are HIV-negative or whose HIV status is unknown.

Secondary objectives

To investigate heterogeneity, as far as data allow, in relation to:

- background epidemiology (prevalence of tuberculosis and of HIV among the study population);
- risk groups targeted (for example, migrants, occupational groups, prisoners, or the general population);
- reference standard (culture, Xpert MTB/RIF, sputum smear microscopy);
- screen test definition;
- representativeness of the study design and study population for intended screening practice;
- demographic characteristics of study participants (age, sex);
- geographic area and economic region.

We did not intend to make a formal comparison of the accuracy of screening tests as part of this review. As part of a WHO tuberculosis screening guideline development and update, diagnostic algorithms composed of screening and diagnostic methods were compared in decision models (van't Hoog 2014a), and an update of this modelling study will be part of the updated screening guidelines (WHO 2021a).

METHODS

Criteria for considering studies for this review

Types of studies

We have outlined any changes to the methods outlined in the published protocol (van't Hoog 2014b) in the 'Differences between protocol and review' section.

We included cross-sectional studies or observational cohort studies where a series of participants were tested with symptom screening, or CXR, or both, and the reference standard. We included studies in which participants were randomized to different screening tests and all participants were verified by the same reference standard. In randomized studies comparing screening strategies, we regarded each arm as a separate cohort. Case control studies were not included because of their potential to introduce bias in diagnostic accuracy estimates (Rutjes 2006). Studies in which participants with a negative screen undergo the reference standard are seldom conducted due to intense resource requirements. Thus, studies with the primary objective of evaluating the accuracy of a screening

test are rare. Therefore, we also included studies with a different primary objective that could potentially provide relevant data for our objectives. An example of such studies was community tuberculosis prevalence surveys for which the primary goal was measuring prevalence. Moreover, studies conducted as baseline measurements of a tuberculosis incidence cohort or randomized trial in which people with prevalent tuberculosis needed to be excluded at baseline were eligible.

We only included studies from which diagnostic two-by-two tables could be generated for a specific screen (symptom definition or CXR finding, or a combination). That is, studies that reported data from which we could extract true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). We included studies in which not all participants were subjected to the reference standard, a common design in tuberculosis prevalence surveys whereby it is assumed that people without tuberculosis suggestive symptoms and without CXR abnormalities do not have tuberculosis disease. Since this design feature may lead to biased accuracy estimates, we addressed this in the quality assessment as described below under Assessment of methodological quality.

We excluded studies in which screening was applied, but the number of tuberculosis cases identified by the reference standard was zero. For cohort studies, we only considered data from tuberculosis cases that were identified from investigations initiated at the time the screening was applied. Incident cases that arise after the screening were not considered, unless the study evaluated screening methods to identify the incident cases at the time of case identification.

We included studies published from 1992 onwards because from that time point the WHO-recommended directly observed treatment short-course (DOTS) strategy was implemented, which has led to improvements in passive case detection and standardized treatment (Dye 1998). Prior to DOTS, case detection was generally lower and any screening would result in predominant detection of people with more advanced tuberculosis disease. Since this epidemiological situation differs from the situation afterwards, the results from older studies are not as relevant. Moreover, older studies frequently screened using MMR, which was an exclusion criteria as explained below (Index tests).

Participants

Included participants were individuals eligible for systematic screening and not known to have tuberculosis disease at the time of screening. We included all types of populations, so study populations varied from the general population in an area with high tuberculosis rates (for example, in mass case finding or tuberculosis prevalence surveys) to specific target populations with much higher tuberculosis prevalence than the general population. Examples of specific populations are studies that target household members of a patient diagnosed with tuberculosis, studies in homeless populations, prison inmates, as well as studies about screening for immigration or occupational purposes (for example, among gold miners) where the goal may be to exclude people with active disease rather than early disease detection. We included studies regardless of whether the participants were screened for the first time or only once, or a population enrolled in longitudinal screening programmes with repeated screening rounds at predetermined intervals, which may be a potential source of heterogeneity.

The review focused on adults (15 years and older), but studies that combined adults and children were included if adults were a majority. We excluded studies focusing on young children (0 to 5 years old) or paediatric tuberculosis only because the clinical presentation of tuberculosis in young children differs from the presentation in adults and older children. Extrapulmonary disease is more common in children, for example. If the lungs are affected, young children more often have paucibacillary disease, and obtaining a sputum specimen can be difficult. Furthermore, clinical presentation including CXR findings are often part of the reference standard (Graham 2012; Luabeya 2012). A Cochrane Review on screening tests for active pulmonary tuberculosis in children has been published (Vonasek 2021). We excluded studies of HIV-infected people only.

We excluded studies that evaluated symptoms, or CXR, or both in a typical passive case detection setting. This applies to clinical settings where patients report to a health facility due to illness and have symptoms and signs that warrant tuberculosis investigations according to national or global guidelines for passive case detection (TB CARE I. 2014). We included studies in an out-patient context among people who would not be considered a presumed tuberculosis case by such guidelines (for example, attendants of diabetic clinics, antenatal clinics). Studies evaluating CXR screening in a pre-selected symptomatic population were outside the scope of this review and were excluded (e.g. Burgess 2001; Masur 2017); we report CXR index tests in populations defined as eligible for screening regardless of the presence of symptoms. Prevalence surveys, in which the reported bacteriologically positive tuberculosis cases included some people who had already started tuberculosis treatment, were included and described, since the proportion of identified cases to whom this applies is usually small (Hoa 2010; van't Hoog 2011b).

Index tests

For index tests based on symptom questions, we selected studies that evaluated one or more author-defined symptoms or symptom combinations, and categorized those into three commonly used and recognizable definitions:

- 'Cough for two or more weeks': Presence of cough lasting for two or more weeks. Prolonged cough is an important component of the definition of presumed tuberculosis in clinical guidelines (TB CARE I. 2014), recommending that all individuals with a cough lasting for two or more weeks require examination for tuberculosis. A small number of relatively older studies reported on cough lasting three or more weeks, which we considered obsolete as an index test based on these guidelines (TB CARE I. 2014). Contrary to an earlier version of this review (van't Hoog 2013a), we did not mix data on cough lasting for three or more weeks with ≥ 2 weeks. Index test definitions that were described as productive 'cough lasting two or more weeks', or 'cough lasting for two or more weeks or haemoptysis' are included in cough ≥ 2 weeks.
- 'Any cough': Presence of cough irrespective of its duration.
- 'Any tuberculosis symptom': Presence of at least one symptom positive out of a combination of at least three symptom questions, that should include cough and two or more systemic symptoms such as fever, night sweats, and weight loss (Getahun 2011).

Of studies reporting symptom combinations that did not meet the above three definitions, we narratively describe the accuracy of (i) combination of cough ≥ 2 weeks and in addition at least one non-cough symptom, and (ii) the combination of two out of several symptoms.

For CXR index tests, we included studies that used conventional radiography (large films, chemical development), digital radiography, or computed radiography (which is an 'upgrade' that allows the production of digital radiographs by conventional X-ray equipment). We excluded studies using MMR only, since this method is not expected to be used for future screening purposes, and has lower sensitivity compared to conventional CXR (Kerley 1942). With respect to the classification of abnormalities, we included all author-defined classification systems, and categorized these into the following:

- 'Any CXR abnormality', which in some studies implied any pulmonary abnormality, and in other studies also included cardiac abnormalities, or also abnormalities in other visible organs. In the analysis this distinction was dealt with as a source of heterogeneity.
- 'CXR abnormalities suggestive of tuberculosis'.

In addition, we reported on index tests that combine symptom questions with CXR in parallel.

Data on CXR reading by computer software (Melendez 2017; Muyoyeta 2017) were not included in this review, as they were addressed in a separate study for the tuberculosis screening guidelines (WHO 2021a). Data on index tests that combined symptoms and physical examination (e.g. a score used in prisons; Morasert 2018) were also outside the scope of this review.

Target conditions

The target condition of tuberculosis disease screening was bacteriologically confirmed pulmonary tuberculosis disease, characterized by the presence of *M tuberculosis* in sputum. We included studies that, in their definition of a bacteriologically positive tuberculosis case, allowed for the inclusion of people with one positive sputum culture or smear or Xpert MTB/RIF only but without symptoms or CXR abnormalities. Although this may reflect an early stage of disease or infectiousness along the tuberculosis spectrum (Pai 2016), it may also reflect laboratory cross contamination. The latter two are not primary targets of screening programmes and inclusion of these two states as a tuberculosis case in a screening tool evaluation would underestimate the sensitivity of the screening tool. We commented on such studies in the Assessment of methodological quality (Appendix 1). We excluded studies on tuberculosis infection only.

Reference standards

The reference standard was defined as any author-defined combination of mycobacterial culture (on solid or liquid medium), sputum smear microscopy, Xpert MTB/RIF, or other NAATs. When not all participants had received the reference standard as defined above, a study was included but it was considered to be at high risk of bias or to have reduced applicability, as described in the Assessment of methodological quality. In our earlier version of this review (van't Hoog 2013a), we included studies with sputum smear-positive cases only and discussed those separately. Such studies were excluded from this review (Mahomed 2013; Masur 2017; Sebhatu 2007), as NAATs are now recommended instead of sputum smear microscopy as the primary test to diagnose tuberculosis (WHO 2020), and the number of such studies was small.

Search methods for identification of studies

Electronic searches

We searched the databases MEDLINE (OVID), Embase (OVID), LILACS (Latin American and Caribbean Health Science Information database, BIREME), and HTA (Health Technology Assessment) from January 1992 to 10 December 2018 to identify titles and abstracts of peer-reviewed papers using the search terms listed in Appendix 2. We included combinations of three domains: (i) "tuberculosis" and related terms, (ii) terms related to "screening", "survey", "sensitivity", "specificity", and (iii) search terms related to the reference standard, "bacterial culture", "microscopy" (Appendix 2). We did not use a diagnostic search filter.

Searching other resources

We checked reference lists of relevant reviews and studies, searched websites of the WHO Global Tuberculosis Programme, and asked experts for relevant studies and unpublished reports, which included the WHO repository of national prevalence survey reports and summaries of those (WHO 2021b).

Data collection and analysis

Selection of studies

We screened for potentially eligible studies using broad criteria: (i) the publication was original research; and (ii) titles, abstracts, or key words suggested that symptom or CXR screening, or active case finding for tuberculosis took place in humans and (iii) data to determine the accuracy of a screening tool may be available. Two authors reviewed all titles and abstracts independently for eligibility. Studies were included for further assessment if they met the inclusion criteria. There was no language restriction. We developed a database of all articles, including full references and abstracts, initially in Reference Manager (v12) (Reference Manager 12), and later in Covidence (Covidence), for studies from an updated search from 2014 onwards. We obtained full-text articles of these studies and two authors assessed study eligibility using the predefined inclusion and exclusion criteria. The authors resolved any disagreements through discussion and, if necessary, with a third author.

Data extraction and management

We developed electronic data-extraction forms in Google forms and Microsoft Excel (Microsoft Excel), which were pilot tested by several authors. Two authors extracted all relevant data from the included studies independently, and discussed inconsistencies to reach consensus. For studies searched before 2014, one author extracted all relevant data and a second author checked the data extraction. The two authors discussed inconsistencies to obtain consensus.

The data extraction form included the following variables:

- Authors, publication year, journal.
- Study level characteristics: country in which the study was conducted (classified according to economic region: low, middle or high income); setting, risk groups included (occupational,

general population, immigrants), urban or rural setting; study design; method of participant selection; number of participants enrolled; number of participants for whom results were available.

- Study participant characteristics: age (mean or median), sex (% female), HIV prevalence among study participants (if data on HIV prevalence were not available in the study report(s) and the study closely resembled the general population, we used data from the Joint United Nations Programme on HIV and AIDS (UNAIDS) reports for the same year or country national reports on HIV), the proportion of participants with a reported history of previous tuberculosis, and the proportion reporting current smoking, if available. If data on current smoking were not available, these data were taken from the Tobacco Atlas year closest to the study, as indicated in Characteristics of included studies.
- First or one-off screening versus repeated screening at regular intervals. If the latter, we recorded whether the same participants were included more than once in the analysis.
- Prevalence of target condition (pulmonary tuberculosis, sputum smear-positive, bacteriologically positive) in the population. In addition, we also recorded the tuberculosis case notification rate (per 100,000 population) in the study population or, if unavailable in the region or country, either from the introduction or methods sections of the publication, or elsewhere.
- Stage of infection: proportion of the true tuberculosis cases included in the report that are bacteriologically positive, but have no signs of active disease (are asymptomatic and have no CXR abnormalities, and no bacteriological confirmation at a second point in time).
- Treatment status: number and proportion of true tuberculosis cases who were already on tuberculosis treatment at the time of screening.
- Reference standard: culture and type of medium (solid or liquid), microscopy and type (light or fluorescence), Xpert MTB/RIF or other NAAT; number of samples per individual tested, number of positive samples required for positive diagnosis, definition of positivity, other criteria included in the reference standard. We also included an exact narrative of the definition and recorded the definitions of the classifications for each included publication.
- Index tests: author definition of the index test, total number of symptoms asked for; radiography equipment type, CXR classification definition, type or reader (expert, radiologist, or pulmonologist; general medical officer; clinical officer; nurse; radiographer; other).
- Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) items (Appendix 1).
- Details of outcomes: the number of TP, TN, FP, and FN; number of participants missing or unavailable test results.

Assessment of methodological quality

Two review authors independently assessed the methodological quality of included studies using the tailored QUADAS-2 instrument (Whiting 2011). We assessed each of the four domains (Patient Selection, Index Test, Reference Standard, Flow and Timing) in terms of risk of bias, and the first three domains in terms of concerns regarding applicability to the review question. The tool with signalling questions tailored to this review can be found in Appendix 1. We resolved any disagreements through discussions to reach consensus. If needed, we reached consensus through discussion with a third author.

We identified a number of quality concerns in the Reference Standard domain that are specific to this review. We considered a number of microbacteriological reference standards as equal, and of high quality: 1) mycobacterial culture followed by mycobacterial speciation; 2) Xpert MTB/RIF or other NAAT, and 3) two positive smears but only in studies where participants were tested with sputum smear and culture (or sputum smear and Xpert MTB/RIF or other NAATs) and a small proportion (≥ 10%) of cases were defined based on two positive smears due to contaminated or negative or missing culture results. We added review-specific signalling questions to better capture potential incorporation and verification biases in the assessment of studies in which not all study participants were verified with the microbacteriological reference standard. A common study design that we included was that of national tuberculosis prevalence surveys (WHO 2011), in which participants with pre-defined symptoms or CXR abnormalities, or both, are further examined with microbiological tests and classified as either reference standard positive or negative based on microbiological test results. However, participants without pre-defined symptoms and without CXR abnormalities were assumed not to have tuberculosis and were defined as reference standard negative without further microbiological confirmation. We judged such designs as being at high risk of bias for the Reference Standard domain. This incorporation bias likely results in overestimation of sensitivity. The specificity of studies with incorporation bias is probably affected to a lesser degree, due to the high proportions of people who tested negative for tuberculosis in this type of study. We considered inclusion of a high proportion of tuberculosis cases who had one positive sputum culture, or Xpert MTB/RIF, or smear at one point in time without any symptom or CXR abnormality or a positive confirmation test at a second time point to be an applicability concern. We classified the concern as low if the proportion of less applicable cases was ≤ 10%.

In the Flow and Timing domain, we assessed whether all study participants who were supposed to have had a microbiological test according to the authors' study design, actually had the required test(s) and had results. We considered the risk of bias low when ≥ 95% had results, and high risk of bias when < 90% had results available. We found this to be reasonable because of the low prevalence of the target condition in almost all studies, implying that the absolute number of missing test results quickly outnumbers the absolute number of true cases identified. In cases where between < 95% and ≥ 90% of results were available, we made a study-specific judgement.

With respect to the Index Test domain, a potential concern was that in studies with a national prevalence survey design, the index tests (both symptom and CXR) served as criteria to determine if microbiological sputum examination would take place. This could bias the interpretation of the index tests if implicit judgements influence the interpretation. However, in these large national surveys both the interview and CXR interpretation procedures are highly standardized, and are not meant to clinically judge a tuberculosis diagnosis, and the study staff are well trained. Therefore, in these cases, we judged the risk of bias to be low.

Statistical analysis and data synthesis

We reported the number of included studies and participants for each index test definition, and the distribution of study populations. We generated diagnostic two-by-two tables, from which we calculated sensitivities and specificities for each index test with 95% CIs; we presented these figures in paired forest plots for each study. In addition, we used a receiver operating characteristic (ROC) plot of sensitivity versus 1-specificity to display the data for each index test. We performed meta-analyses of pairs of sensitivity and specificity using bivariate random-effects methods (Reitsma 2005) for the index test definitions described above. The bivariate model was preferred because all index tests produced binary results for which an implicit threshold was assumed. We developed the bivariate model in Stata(Stata). We evaluated subgroups and screen definitions for which we were unable to provide meaningful summary estimates of sensitivity and specificity using descriptive methods.

Investigations of heterogeneity

We examined the forest plots and ROC plots visually for heterogeneity. We analyzed potential sources of heterogeneity for each index test, one by one, as categorical covariates in a hierarchical mixed-model (using Stata XTMELOGIT), and examined whether the variable modified sensitivity, specificity, or both. We used the data described under Data extraction and management to create variables for investigations of heterogeneity. We first visually examined continuous variables in scatterplots and recoded these to categorical variables if the plots suggested a possible association with sensitivity or specificity. In presentation, we divided the variables by study-level characteristics, study participant characteristics, and index test characteristics. Some variables were combined or not used due to overlap with other variables or missing data (for example, almost all studies in a general population were mixed urban-rural). We dealt with study methodology assessments by QUADAS-2 in sensitivity analyses.

We treated study-level characteristics as follows:

- The study designs and populations largely overlapped and were combined in one binary variable reflecting studies in the general population, which were almost all tuberculosis prevalence surveys, versus studies in specific populations, of which the majority were routine screening or special active tuberculosis case finding activities. The specific populations comprised a wide range and were not further subdivided as this would have resulted in very small numbers per category.
- The prevalence of tuberculosis disease in the study population was calculated from the two-by-two tables and categorized as < 0.5%, ≥ 0.5% and < 2%, ≥ 2%.
- Geographic region, captured as WHO region, was categorized as Africa (approximately half of the studies), versus Other (of which the majority were studies from the Asia and Pacific regions, and < 5% were from the Eastern Mediterranean and the Americas regions combined).
- Economic region was categorized as low-income country (LIC), lower-middle income country (LMIC), and upper-middle income country (UMIC). One study from a high-income country and was added to UMIC.

Study participant characteristics were categorized as follows:

- HIV prevalence among the study population was recoded as low (≥ 1%), high (> 1%) based on the UNAIDS definition of a general epidemic (UNAIDS 2015), or very high (> 5%).
- Mean/median age of the study populations was categorized as < 35 years versus ≥ 35 years, based on the median.
- Smoking prevalence was divided into < 10% versus ≥ 10%, based on the median.
- The proportion of study participants reporting previous tuberculosis treatment was divided into three categories: < 5%, ≥ 5%, or missing.

For the index test 'any tuberculosis symptom' we examined whether the number of symptoms asked modified accuracy, and for the index test 'any CXR abnormality' whether the accuracy was modified by the definition: 'any pulmonary abnormalities' or also abnormalities in other visible organs. For CXR index tests we also examined if accuracy was modified by type of interpreter, categorized as specialist versus a category combining medical officer, clinical officer or other clinician. We examined possible associations between categorical covariates with the Fisher's exact test.

Sensitivity analyses

To explore whether the summary results were robust to methodological challenges, we performed sensitivity analyses using the QUADAS-2 judgements. We assessed whether the summary estimates and 95% CIs changed when excluding studies with quality concerns. We examined the robustness of the results for the choice of the microbiological methods used in the reference standard. We also examined whether exclusion of studies with a partially repeatedly screened population (a minority) altered the summary estimates in a sensitivity analysis (and not as a heterogeneity investigation) as the number of such studies was very low. Six studies in populations with high HIV prevalence reported accuracy separately for the HIV-positive and HIV-negative study participants. The HIV-negative populations of these studies were included for the primary analysis. In sensitivity analyses, the two-by-two data of the HIV-negative data were replaced with those of the mixed HIV-negative and HIV-positive population. Results of the sensitivity analyses are shown only if a noticeable change in the estimates was observed.

Assessment of reporting bias

We did not assess the risk of reporting bias in the included studies.

Certainty of the body of evidence

We assessed the certainty of evidence using the GRADE methodology (Schünemann 2020a; Schünemann 2020b).

Certainty of evidence and search date

After completion of the review based on the above search we explored if adding an updated search would potentially reach a higher certainty of evidence and therefore more firm conclusions. We conducted an updated search in MEDLINE and Embase on 2 July 2021 and restricted the search to records that included tuberculosis terms, DTA terms, and terms for our reference standard in title or abstract. This set would most likely include studies of high quality. We then assessed the restricted set using the review's eligibility criteria and of the studies that were eligible for inclusion, we

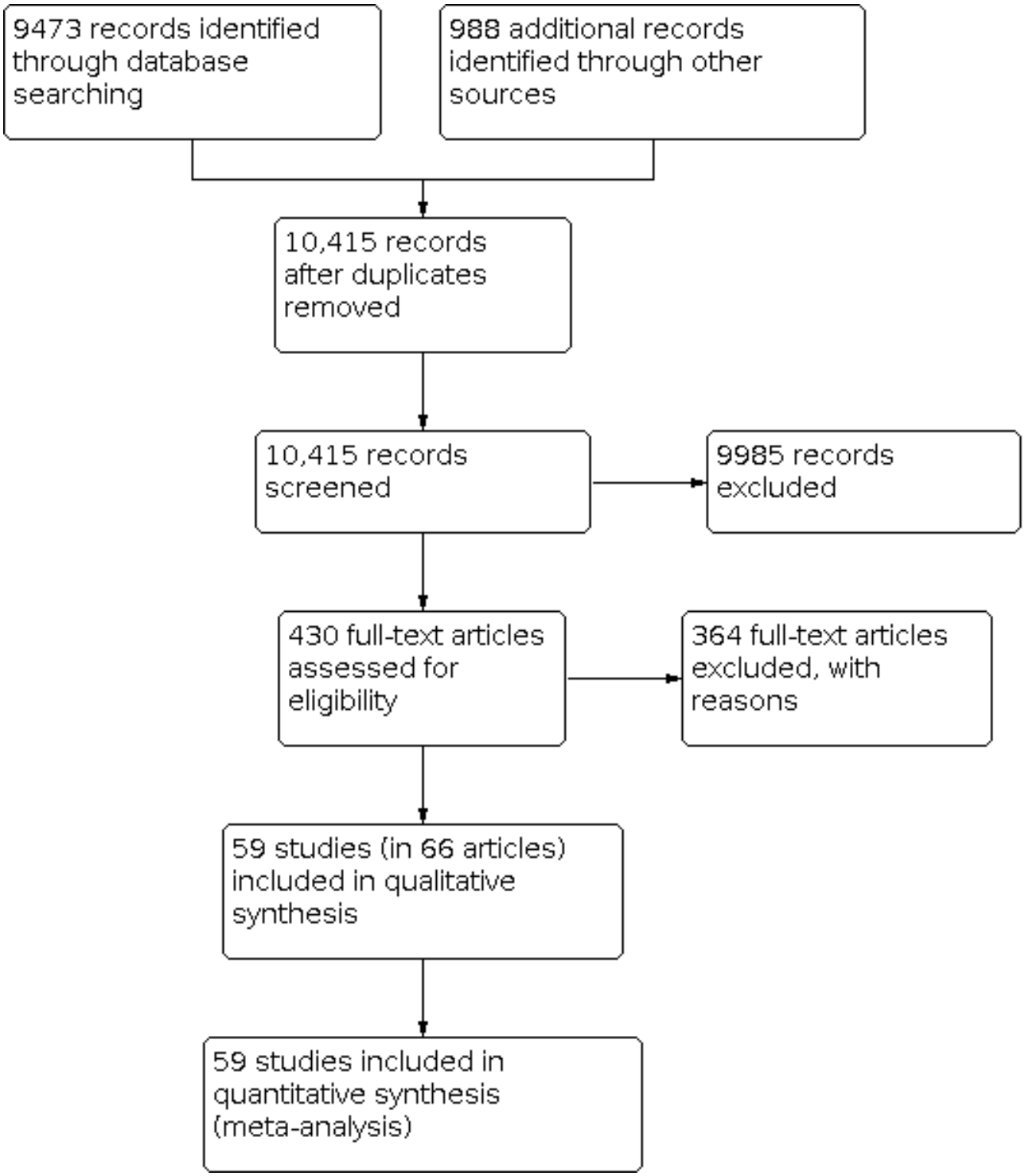
assessed risks of bias and applicability concerns in the QUADAS-2 domains.

RESULTS

Results of the search

The electronic database search identified 9473 titles and abstracts (Figure 2). Through expert consultation we identified 31 reports on national tuberculosis prevalence surveys as eligible for inclusion (of which eight were already captured in the search of the electronic databases), and we identified 957 potential relevant articles through reference checking. After removal of duplicates, we assessed 10,415 titles and abstracts, of which we identified 430 (4%) for full-text review. In the full-text assessment 364 articles were excluded. The most common reasons for exclusion were that the study design did not allow evaluation of the sensitivity and specificity of a screening test, usually because one screening method was used (often symptoms) and only screen positives were further evaluated by a reference standard. In 175 (48%) publications the data required to construct a diagnostic two-by-two table were not or insufficiently reported. In total, 66 publications and reports provided data on 59 studies. Two publications provided data on multiple studies (Claassens 2017a; Morishita 2017a), either in different geographic areas, or different populations.

Figure 2. Study flow diagram.



Most studies (n = 36; 61%) were conducted in the general population, followed by six (10%) in (household) contacts, five (8%) in prison inmates, and the remainder (n = 12; 21%) in a wide variety of special populations like migrants, homeless, or occupational settings, each represented by one or two studies. All but seven studies (12%) involved adults only, either from 15 or 18 years and older. For the seven studies that included adults and children accuracy data could not be disaggregated by age, but in all studies children < 15 years of age were a minority.

Most studies included in the meta-analyses provided data on multiple index tests. Forty-eight studies reported on one or more symptom index tests. Thirty-seven studies reported on CXR as a screening index test, either read for any abnormality (n = 23), which may or may not have included other visible organs in addition to

the lungs, or read for abnormalities suggestive of tuberculosis (n = 19), or both (n = 5). In three instances when data were available to construct a 2 x 2 table more than once for the same index test (CXR) in (almost) the same study population, we selected the most complete (Kapata 2016; Melendez 2017) or most recent report (MoH Cambodia 2005; MoH Cambodia 2012), or made a random choice (van't Hoog 2011).

The additional search results to 2 July 2021 were subject to the same screening and assessment procedure, and we have detailed the results in Appendix 3. Of the seven additional studies identified, all were small, at high risk of bias, and outside the main comparison in this analysis. It was clear these would make no material difference to the review findings or grading of the evidence, and we did not included these additional studies in this edition of the review. The full details of the assessment are in Appendix 3.

Methodological quality of included studies

Figure 3 and Figure 4 show the review authors' judgements on risk of bias and applicability concerns for CXR index tests as a group. We judged the risk of bias to be high for > 90% of studies in the Reference Standard domain, and for over half of the studies in the Flow and Timing domain. High risk of bias was less common in the domains Patient Selection and Index Test. We judged applicability concerns low in > 80% of studies in all three domains. For CXR index tests, no study was judged to be at low risk of bias and low applicability concerns in all domains. The most common reasons for high risk of bias were incorporation bias and verification bias. Similarly, for symptom index tests as a group, high risk of bias judgements were most frequent in the Reference Standard and Flow and Timing domains (Figure 5; Figure 6), and we judged only one study to be at low risk of bias and to have low applicability concerns in all domains.

Figure 3. Risk of bias and applicability concerns graph for CXR index tests: review authors' judgements about each domain presented as percentages across included studies.

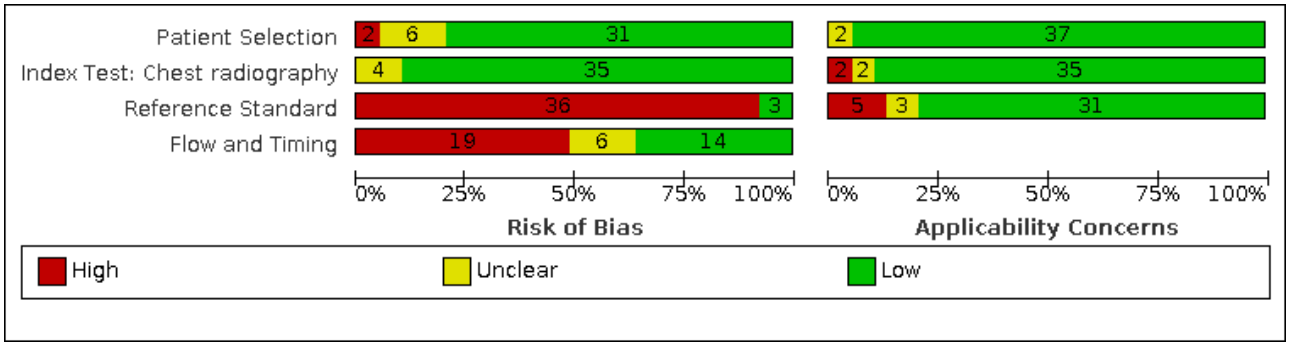


Figure 4. Risk of bias and applicability concerns summary for CXR index tests: review authors' judgements about each domain for each included study.

Figure 4. (Continued)

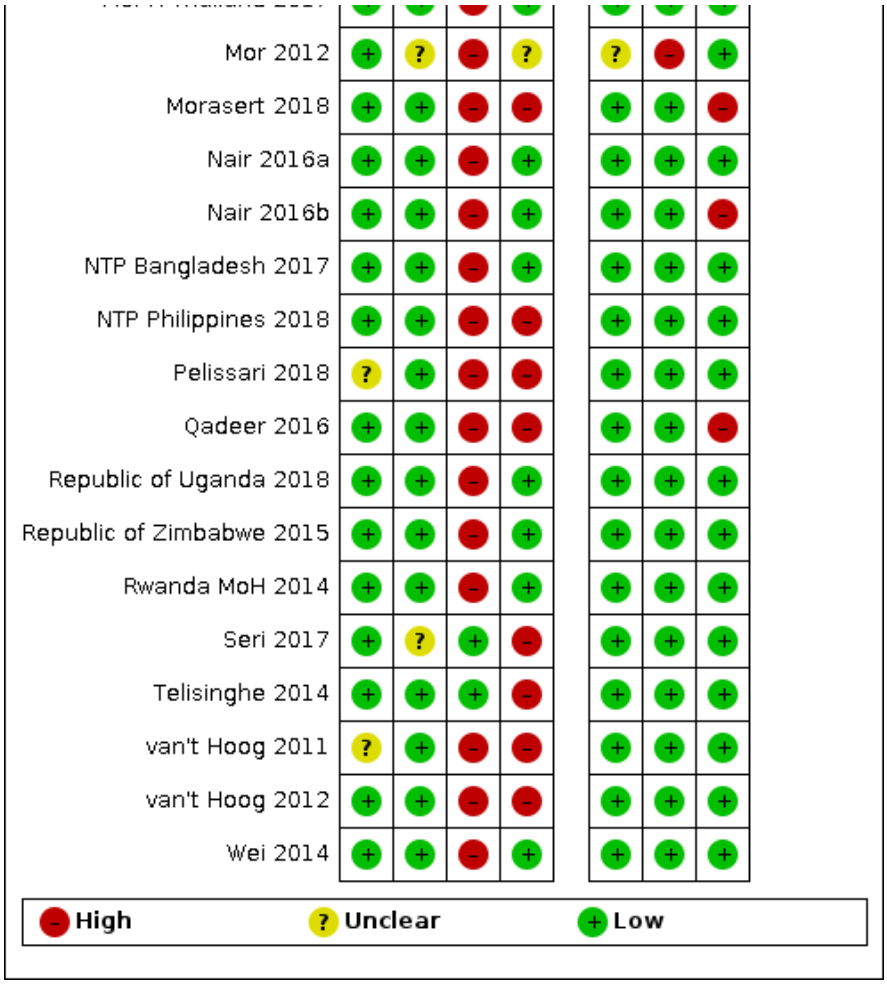


Figure 5. Risk of bias and applicability concerns graph for symptom question index tests: review authors' judgements about each domain presented as percentages across included studies.

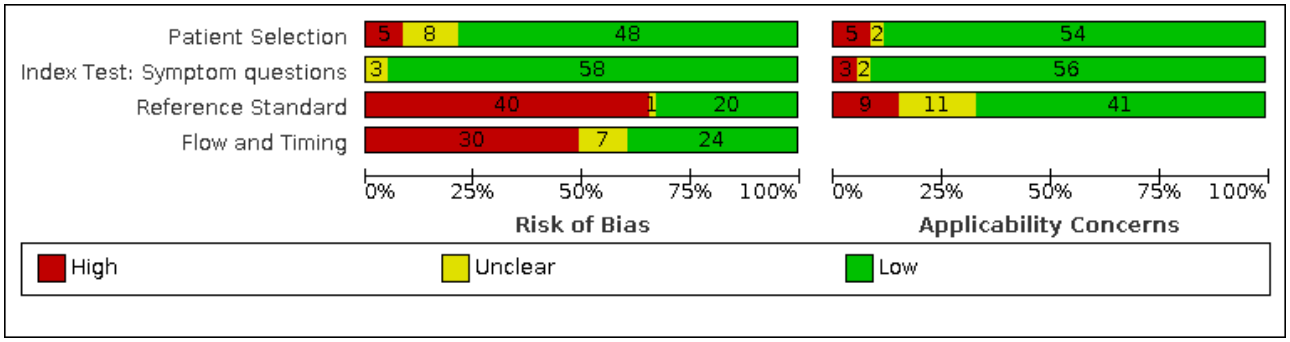


Figure 6. Risk of bias and applicability concerns summary for symptom question index tests: review authors' judgements about each domain for each included study.

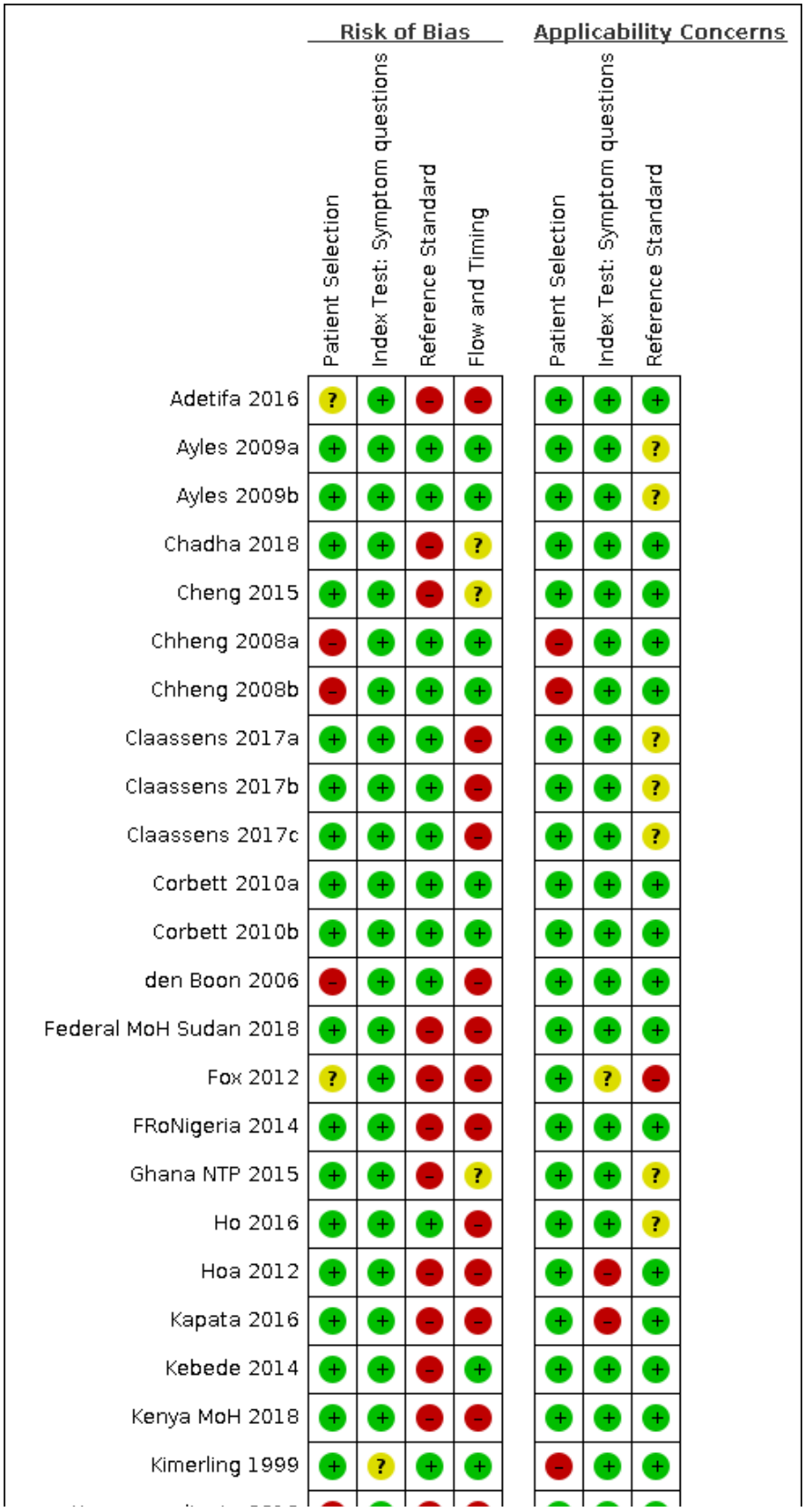


Figure 6. (Continued)

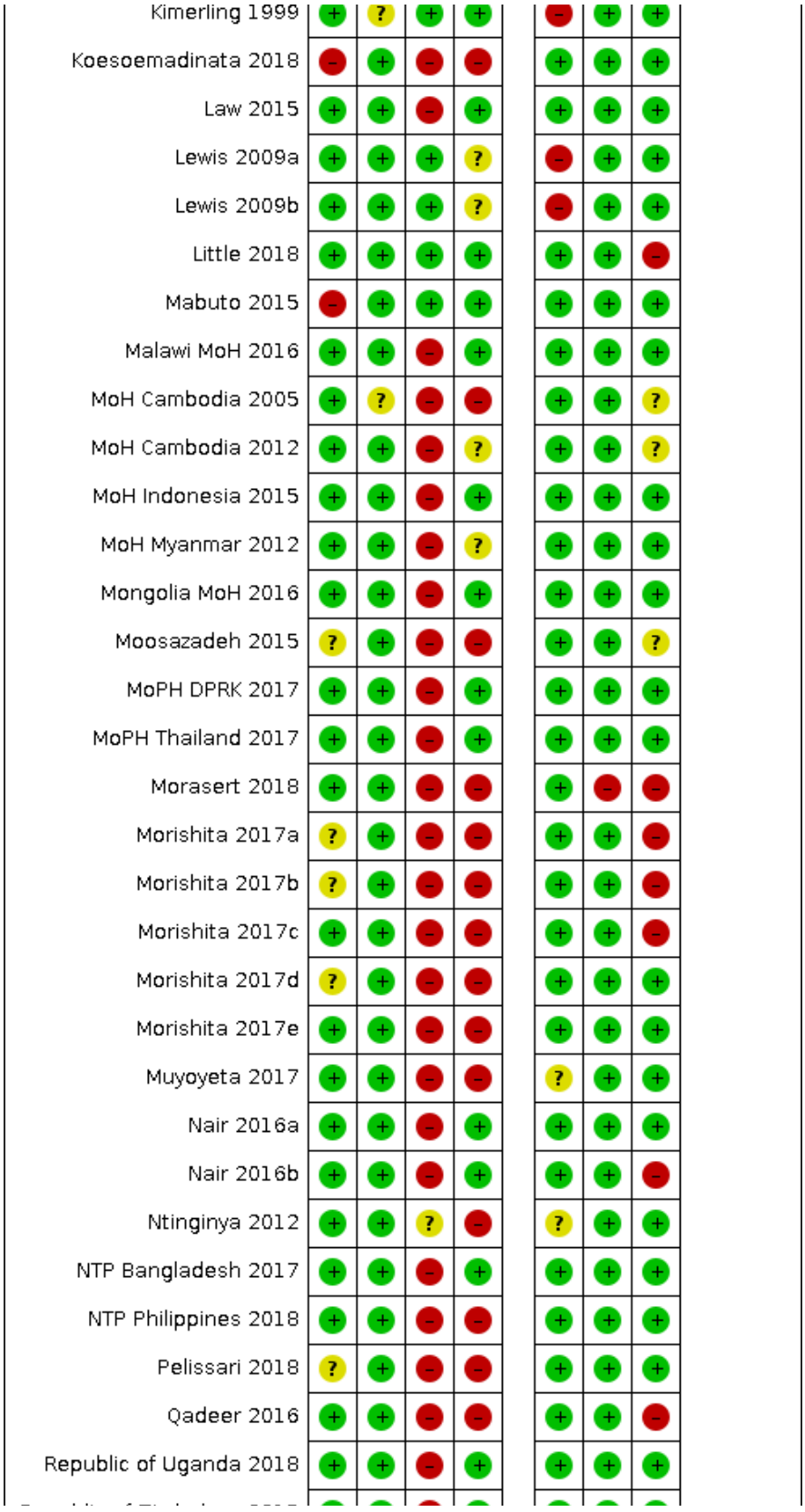
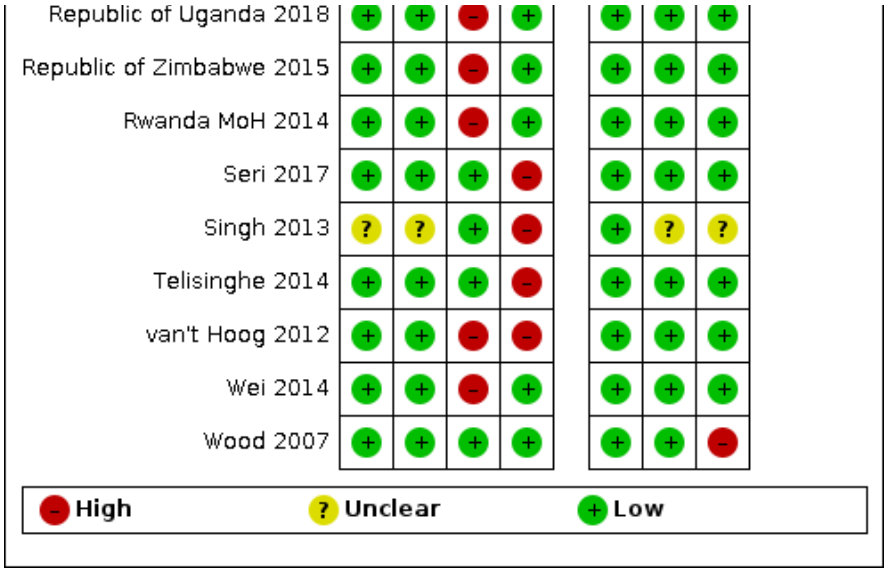


Figure 6. (Continued)



More granular graphs of risk of bias and applicability concerns (displaying judgements separately for each index test) are presented for symptom index tests in Figure 7, CXR index tests in Figure 8, and for CXR and symptoms in parallel in Figure 9.

Figure 7. QUADAS-2 summary of symptom index tests.

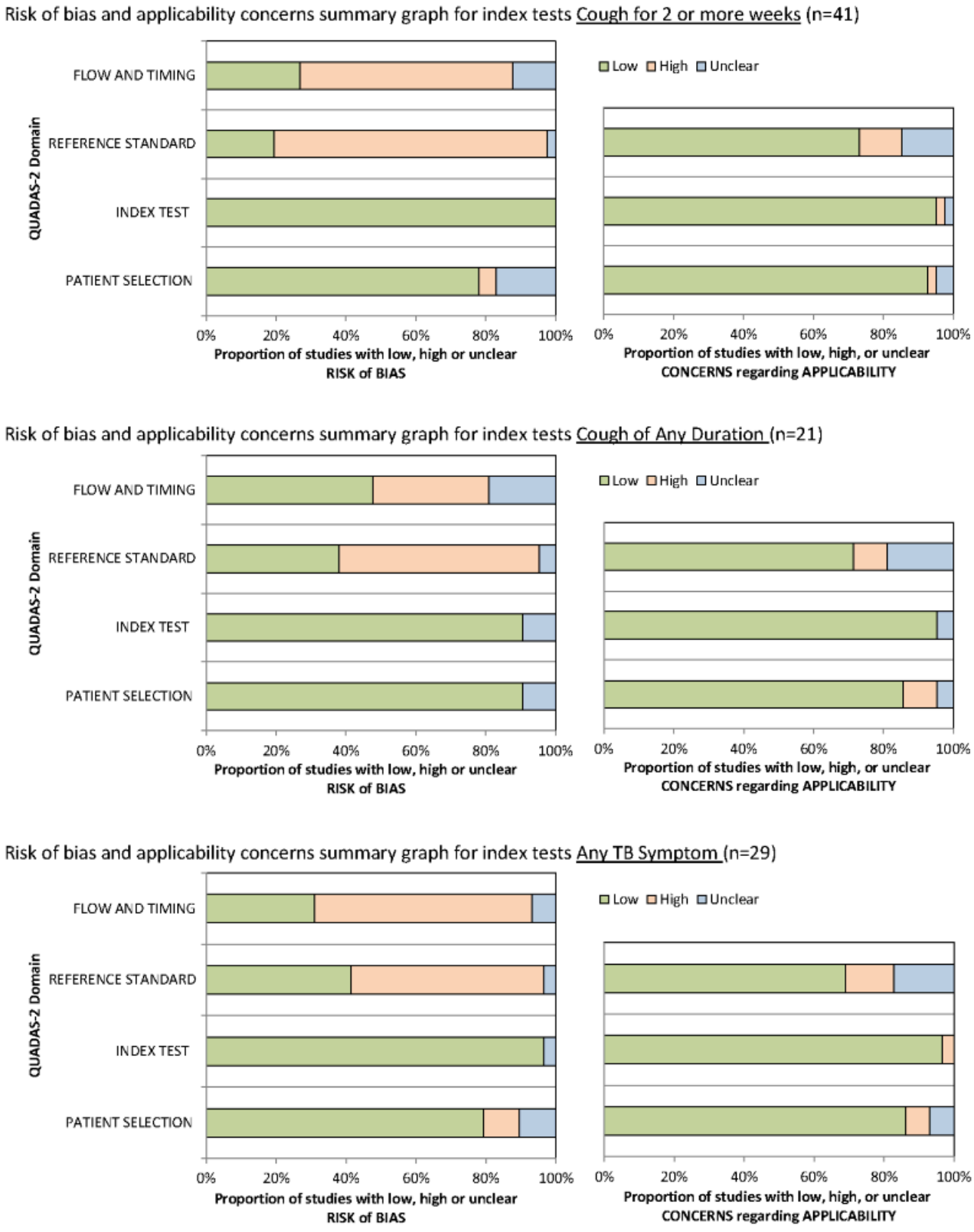


Figure 7. (Continued)

RISK of BIAS CONCERNS regarding APPLICABILITY

Figure 8. QUADAS-2 summary for each CXR index test.

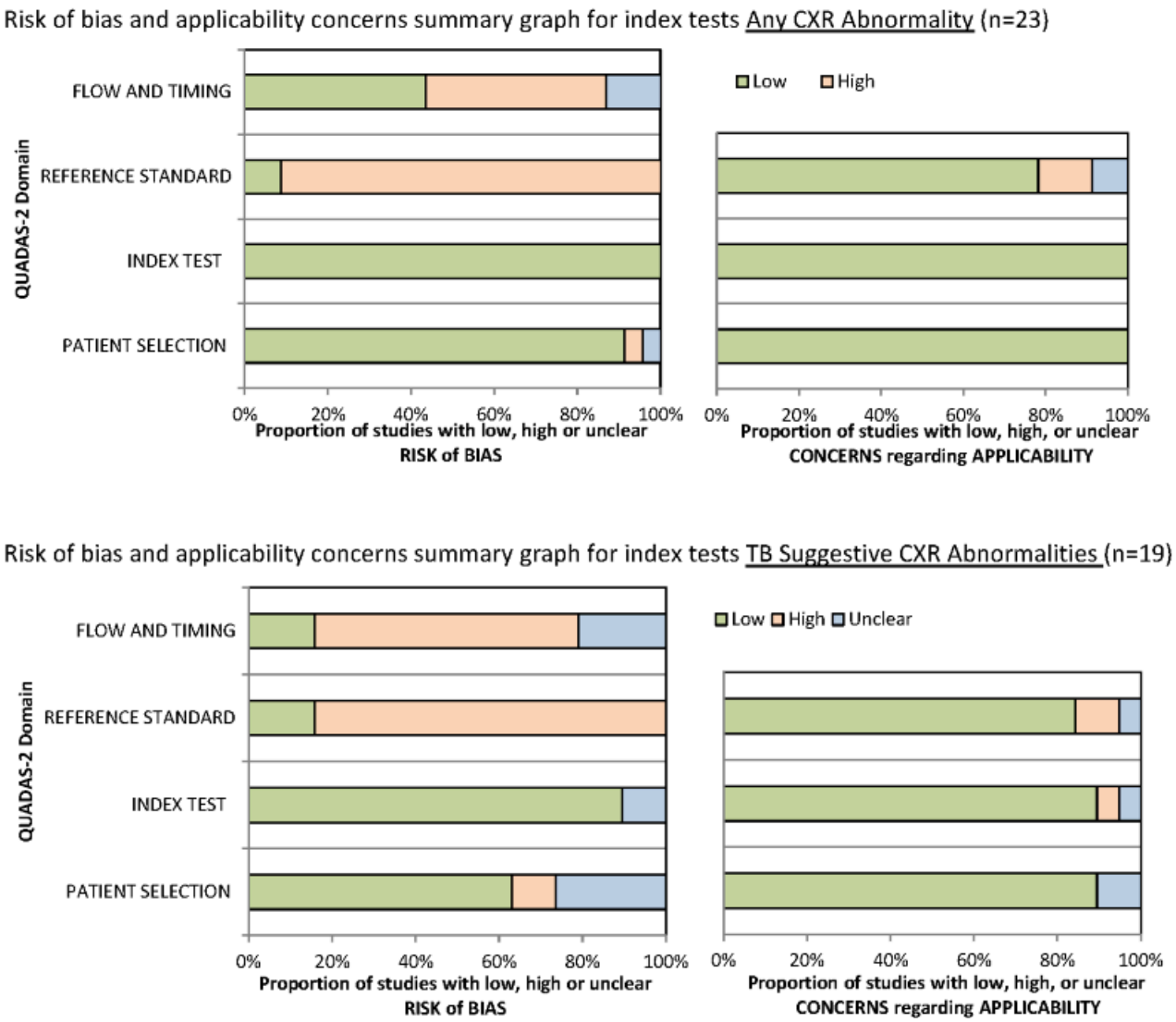
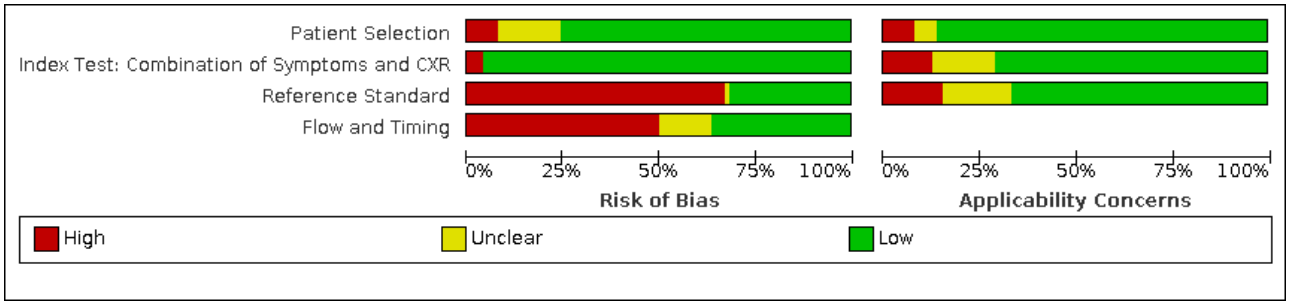


Figure 9. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies.



Findings

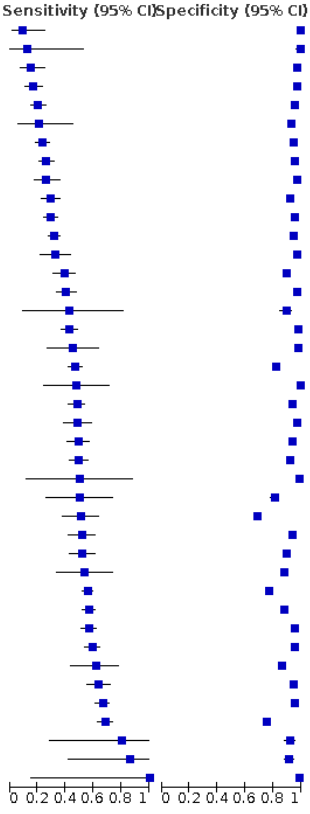
Index tests based on symptom questions

Cough for two or more weeks

Forty-one studies contributed to the index test cough ≥ 2 weeks, of which 26 (63%) were in the general population and 15 in other populations (Figure 10). In total, 7179 participants contributed to the estimation of sensitivity, which ranged between 10% and 100%. For this index test, the most extreme values were in studies with smaller sample sizes and wider 95% CIs, as well as in studies with larger numbers of tuberculosis cases. Specificity was more homogenous, ranging from 68% to 99%, to which 1,540,179 participants contributed.

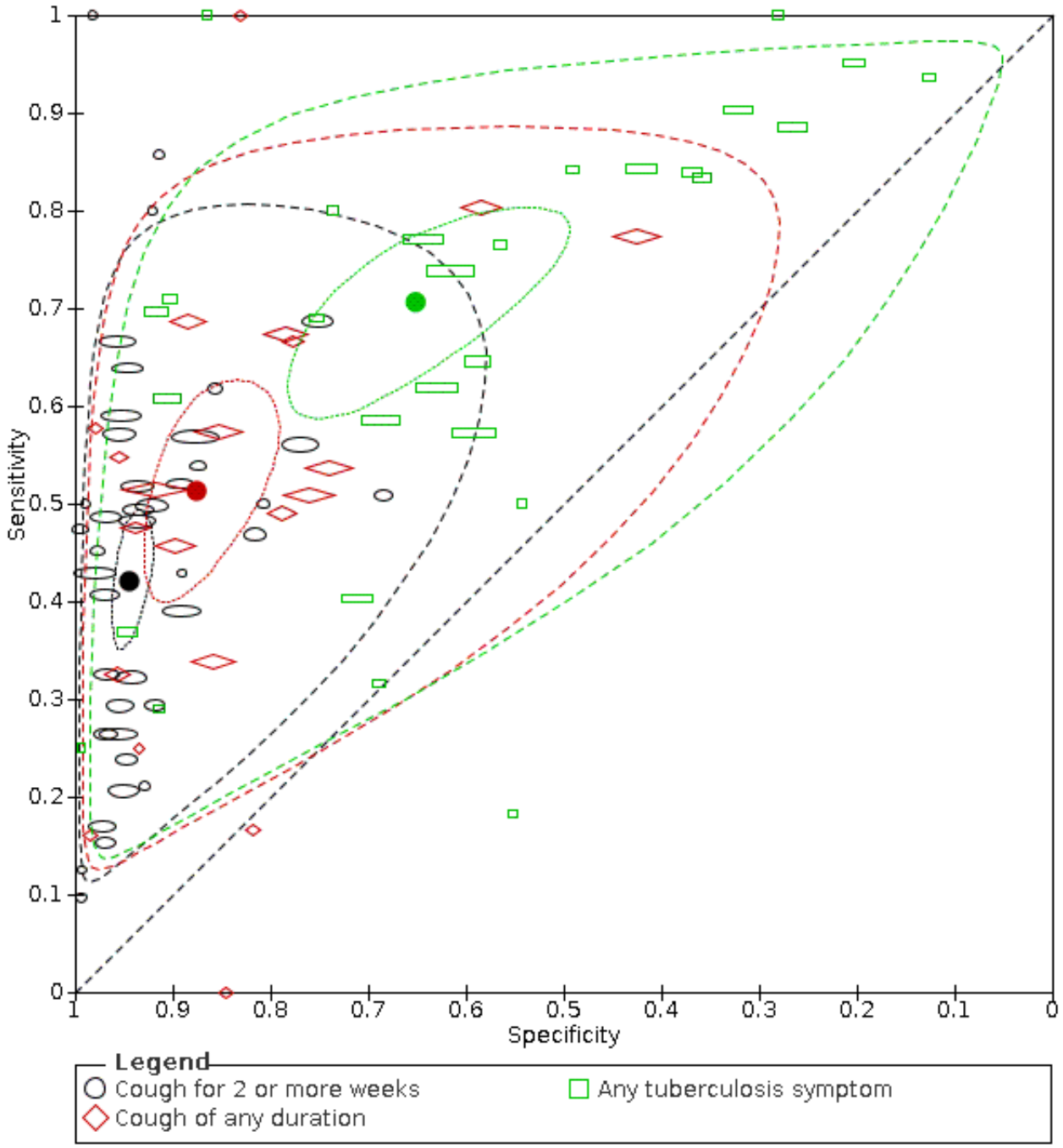
Figure 10. Forest plot of studies providing data on index test 'Cough for 2 or more weeks'. CI: confidence interval; FN: false negative; FP: false positive; TN: true negative; TP: true positive.

Study	TP	FP	FN	TN	Population	Sensitivity (95% CI)	Specificity (95% CI)
Lewis 2009b	3	9	28	1338	Occupational (gold miners)	0.10 [0.02, 0.26]	0.99 [0.99, 1.00]
Nair 2016a	1	1	7	142	Diabetes patients	0.13 [0.00, 0.53]	0.99 [0.96, 1.00]
Claassens 2017d	11	719	61	22473	HIV-negative subpopulation	0.15 [0.08, 0.26]	0.97 [0.97, 0.97]
MoPH Thailand 2017	24	1751	118	60643	General population	0.17 [0.11, 0.24]	0.97 [0.97, 0.97]
Mongolia MoH 2016	51	2495	197	47451	General population	0.21 [0.16, 0.26]	0.95 [0.95, 0.95]
Telisinghe 2014	4	44	15	570	Prison inmates	0.21 [0.06, 0.46]	0.93 [0.90, 0.95]
Claassens 2017b	72	802	230	14345	HIV-negative subpopulation	0.24 [0.19, 0.29]	0.95 [0.94, 0.95]
Hoa 2012	71	4129	198	89360	General population	0.26 [0.21, 0.32]	0.96 [0.95, 0.96]
Republic of Zimbabwe 2015	28	1040	78	32544	General population	0.26 [0.18, 0.36]	0.97 [0.97, 0.97]
Pelissari 2018	53	828	128	9317	Prison inmates	0.29 [0.23, 0.36]	0.92 [0.91, 0.92]
MoH Cambodia 2012	92	1712	222	35391	General population	0.29 [0.24, 0.35]	0.95 [0.95, 0.96]
NTP Philippines 2018	150	2665	316	43558	General population	0.32 [0.28, 0.37]	0.94 [0.94, 0.94]
Adetifa 2016	25	1366	52	41657	General population	0.32 [0.22, 0.44]	0.97 [0.97, 0.97]
Ho 2016	66	4710	103	38556	General population	0.39 [0.32, 0.47]	0.89 [0.89, 0.89]
Ghana NTP 2015	82	1887	120	59637	General population	0.41 [0.34, 0.48]	0.97 [0.97, 0.97]
Muyoyeta 2017	3	28	4	228	(Household) contacts	0.43 [0.10, 0.82]	0.89 [0.85, 0.93]
Cheng 2015	135	5157	180	247322	General population	0.43 [0.37, 0.49]	0.98 [0.98, 0.98]
Corbett 2010b	14	166	17	6924	HIV-negative subpopulation	0.45 [0.27, 0.64]	0.98 [0.97, 0.98]
Morishita 2017c	177	1059	201	4696	Prison inmates	0.47 [0.42, 0.52]	0.82 [0.81, 0.83]
Wei 2014	9	299	10	53950	General population	0.47 [0.24, 0.71]	0.99 [0.99, 1.00]
Kenya MoH 2018	147	3990	158	58755	General population	0.48 [0.42, 0.54]	0.94 [0.93, 0.94]
Federal MoH Sudan 2018	51	2612	54	80293	General population	0.49 [0.39, 0.59]	0.97 [0.97, 0.97]
Republic of Uganda 2018	79	2635	81	38359	General population	0.49 [0.41, 0.57]	0.94 [0.93, 0.94]
Law 2015	118	3079	119	35679	General population	0.50 [0.43, 0.56]	0.92 [0.92, 0.92]
Moosazadeh 2015	3	7	3	655	(Household) contacts	0.50 [0.12, 0.88]	0.99 [0.98, 1.00]
Seri 2017	9	153	9	640	Prison inmates	0.50 [0.26, 0.74]	0.81 [0.78, 0.83]
Morishita 2017d	32	659	31	1423	Indigenous population	0.51 [0.38, 0.64]	0.68 [0.66, 0.70]
Kebede 2014	57	2968	53	43619	General population	0.52 [0.42, 0.61]	0.94 [0.93, 0.94]
van't Hoog 2012	64	2200	59	18243	General population	0.52 [0.43, 0.61]	0.89 [0.89, 0.90]
den Boon 2006	14	314	12	2171	General population	0.54 [0.33, 0.73]	0.87 [0.86, 0.89]
Morishita 2017e	425	5083	334	16968	Adolescents	0.56 [0.52, 0.60]	0.77 [0.76, 0.78]
MoH Indonesia 2015	242	8310	184	59208	General population	0.57 [0.52, 0.62]	0.88 [0.87, 0.88]
MoPH DPRK 2017	194	2750	146	57593	General population	0.57 [0.52, 0.62]	0.95 [0.95, 0.96]
Qadeer 2016	197	4866	137	99431	General population	0.59 [0.53, 0.64]	0.95 [0.95, 0.95]
Morishita 2017b	21	227	13	1364	Urban poor communities	0.62 [0.44, 0.78]	0.86 [0.84, 0.87]
FRoNigeria 2014	92	2381	52	41661	General population	0.64 [0.55, 0.72]	0.95 [0.94, 0.95]
Chadha 2018	223	3851	112	83344	General population	0.67 [0.61, 0.72]	0.96 [0.95, 0.96]
Morishita 2017a	195	3138	89	9485	Rural poor communities	0.69 [0.63, 0.74]	0.75 [0.74, 0.76]
Ntinginya 2012	4	17	1	197	(Household) contacts	0.80 [0.28, 0.99]	0.92 [0.88, 0.95]
Koesoemadinata 2018	6	29	1	310	Diabetes patients	0.86 [0.42, 1.00]	0.91 [0.88, 0.94]
Fox 2012	2	10	0	533	(Household) contacts	1.00 [0.16, 1.00]	0.98 [0.97, 0.99]



The summary sensitivity of cough ≥ 2 weeks was 42.1% (95% CI 36.6% to 47.7%; very low-certainty evidence because of very serious risk of bias and serious inconsistency), and the summary specificity was 94.4% (95% CI 92.6% to 95.8%; high-certainty evidence) (Figure 11; Summary of findings 1).

Figure 11. Summary ROC plot of index tests: 'Any tuberculosis symptom', 'Cough of any duration', 'Cough for 2 or more weeks'. The plot shows summary estimates, 95% confidence (dotted lines) and 95% prediction intervals (dashed lines).

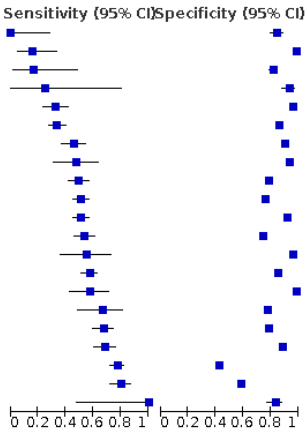


Cough of any duration (any cough)

Twenty-one studies contributed to any cough, of which 15 (71%) were in the general population (Figure 12). In total, 2734 participants contributed to the estimation of sensitivity, which ranged between 0% and 100%, again with the most extreme values from studies with smaller sample sizes and wider 95% CIs. Specificity was more homogenous, ranging from 43% to 99%, and 768,291 participants contributed.

Figure 12. Forest plot studies providing data on index test 'Cough of any duration'. CI: confidence interval; FN: false negative; FP: false positive; TN: true negative; TP: true positive.

Study	TP	FP	FN	TN	Population	Sensitivity (95% CI)	Specificity (95% CI)
Little 2018	0	41	11	227	(Household) contacts	0.00 [0.00, 0.28]	0.85 [0.80, 0.89]
Lewis 2009b	5	19	26	1328	Occupational (gold miners)	0.16 [0.05, 0.34]	0.99 [0.98, 0.99]
Wood 2007	2	134	10	605	General population	0.17 [0.02, 0.48]	0.82 [0.79, 0.85]
Kimerling 1999	1	8	3	115	Homeless and/or drug users	0.25 [0.01, 0.81]	0.93 [0.88, 0.97]
Republic of Zimbabwe 2015	35	1422	72	32207	General population	0.33 [0.24, 0.42]	0.96 [0.96, 0.96]
Mongolia MoH 2016	84	7007	164	42939	General population	0.34 [0.28, 0.40]	0.86 [0.86, 0.86]
MoPH Thailand 2017	65	6330	77	56064	General population	0.46 [0.37, 0.54]	0.90 [0.90, 0.90]
Rwanda MoH 2014	19	2618	21	40463	General population	0.47 [0.32, 0.64]	0.94 [0.94, 0.94]
Pelissari 2018	89	2137	92	8008	Prison inmates	0.49 [0.42, 0.57]	0.79 [0.78, 0.80]
MoH Myanmar 2012	155	12113	149	38950	General population	0.51 [0.45, 0.57]	0.76 [0.76, 0.77]
Cheng 2015	162	20481	153	231998	General population	0.51 [0.46, 0.57]	0.92 [0.92, 0.92]
Ho 2016	91	11209	78	32057	General population	0.54 [0.46, 0.62]	0.74 [0.74, 0.75]
Corbett 2010b	17	306	14	6784	HIV-negative subpopulation	0.55 [0.36, 0.73]	0.96 [0.95, 0.96]
Kenya MoH 2018	175	9130	130	53615	General population	0.57 [0.52, 0.63]	0.85 [0.85, 0.86]
Singh 2013	30	29	22	1527	(Household) contacts	0.58 [0.43, 0.71]	0.98 [0.97, 0.99]
Ayles 2009b	24	1250	12	4380	General population	0.67 [0.49, 0.81]	0.78 [0.77, 0.79]
Republic of Uganda 2018	108	8798	52	32196	General population	0.68 [0.60, 0.75]	0.79 [0.78, 0.79]
FRoNigeria 2014	99	5053	45	38989	General population	0.69 [0.61, 0.76]	0.89 [0.88, 0.89]
MoH Cambodia 2012	243	21312	71	15791	General population	0.77 [0.72, 0.82]	0.43 [0.42, 0.43]
van't Hoog 2012	99	8455	24	11982	General population	0.80 [0.72, 0.87]	0.59 [0.58, 0.59]
Ntinginya 2012	5	36	0	178	(Household) contacts	1.00 [0.48, 1.00]	0.83 [0.77, 0.88]



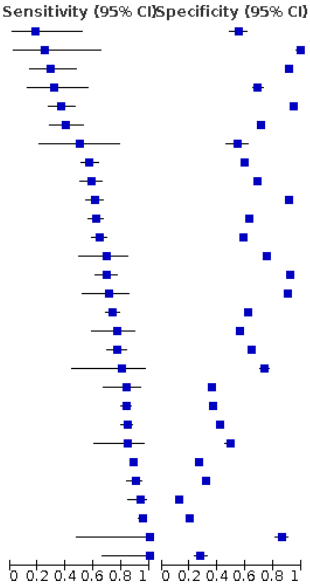
The summary sensitivity was 51.3% (95% CI 42.8% to 59.7%; very low-certainty evidence because of very serious risk of bias and serious inconsistency), and the summary specificity was 87.6% (95% CI 81.6% to 91.8%; low-certainty evidence because of serious inconsistency and serious imprecision) (Figure 11; Summary of findings 2).

Any symptom positive out of three or more tuberculosis-related symptom questions, including cough (any tuberculosis symptom)

Twenty-nine studies contributed data to any tuberculosis symptom, of which approximately half (n = 15) were in a general population. In total, 4180 participants contributed to the estimation of sensitivity, which ranged from 18% to 100% (Figure 13). Again, the most extreme values originated from studies with small numbers of participants, which had wide and overlapping 95% CIs. However, among larger studies, variation was also considerable. Estimations of specificity represent 506,712 participants. Specificity varied from 13% to 99%, with narrow 95% CIs for all studies.

Figure 13. Forest plot of studies providing data on index test 'Any tuberculosis symptom'. CI: confidence interval; FN: false negative; FP: false positive; TN: true negative; TP: true positive.

Study	TP	FP	FN	TN	Population	Sensitivity (95% CI)	Specificity (95% CI)
Little 2018	2	120	9	148	(Household) contacts	0.18 [0.02, 0.52]	0.55 [0.49, 0.61]
Nair 2016a	2	1	6	142	Diabetes patients	0.25 [0.03, 0.65]	0.99 [0.96, 1.00]
Lewis 2009b	9	117	22	1230	Occupational (gold miners)	0.29 [0.14, 0.48]	0.91 [0.90, 0.93]
Telisinghe 2014	6	191	13	423	Prison inmates	0.32 [0.13, 0.57]	0.69 [0.65, 0.73]
Republic of Zimbabwe 2015	39	1794	67	31790	General population	0.37 [0.28, 0.47]	0.95 [0.94, 0.95]
Claassens 2017d	29	6702	43	16490	HIV-negative subpopulation	0.40 [0.29, 0.53]	0.71 [0.71, 0.72]
Muyoyeta 2017	6	76	6	90	(Household) contacts	0.50 [0.21, 0.79]	0.54 [0.46, 0.62]
Mongolia MoH 2016	142	20373	106	29573	General population	0.57 [0.51, 0.63]	0.59 [0.59, 0.60]
Ho 2016	99	13545	70	29721	General population	0.59 [0.51, 0.66]	0.69 [0.68, 0.69]
Kapata 2016	161	4292	104	41542	General population	0.61 [0.55, 0.67]	0.91 [0.90, 0.91]
MoH Myanmar 2012	188	18922	116	32141	General population	0.62 [0.56, 0.67]	0.63 [0.63, 0.63]
Claassens 2017b	195	6240	107	8907	HIV-negative subpopulation	0.65 [0.59, 0.70]	0.59 [0.58, 0.60]
den Boon 2006	20	637	9	1941	General population	0.69 [0.49, 0.85]	0.75 [0.74, 0.77]
Malawi MoH 2016	92	2623	40	28824	General population	0.70 [0.61, 0.77]	0.92 [0.91, 0.92]
Corbett 2010b	22	686	9	6404	HIV-negative subpopulation	0.71 [0.52, 0.86]	0.90 [0.90, 0.91]
Kenya MoH 2018	225	24109	80	38636	General population	0.74 [0.68, 0.79]	0.62 [0.61, 0.62]
Morishita 2017b	26	693	8	898	Urban poor communities	0.76 [0.59, 0.89]	0.56 [0.54, 0.59]
FRoNigeria 2014	111	15701	33	28341	General population	0.77 [0.69, 0.84]	0.64 [0.64, 0.65]
Mabuto 2015	8	178	2	498	Occupational - Other	0.80 [0.44, 0.97]	0.74 [0.70, 0.77]
Ayles 2009b	30	3614	6	2016	General population	0.83 [0.67, 0.94]	0.36 [0.35, 0.37]
Morishita 2017c	317	3632	61	2123	Prison inmates	0.84 [0.80, 0.87]	0.37 [0.36, 0.38]
MoH Cambodia 2005	229	12673	43	9215	General population	0.84 [0.79, 0.88]	0.42 [0.41, 0.43]
Seri 2017	16	471	3	453	Prison inmates	0.84 [0.60, 0.97]	0.49 [0.46, 0.52]
Morishita 2017e	672	16200	87	5851	Adolescents	0.89 [0.86, 0.91]	0.27 [0.26, 0.27]
van't Hoog 2012	111	13878	12	6565	General population	0.90 [0.84, 0.95]	0.32 [0.31, 0.33]
Morishita 2017d	59	1820	4	262	Indigenous population	0.94 [0.85, 0.98]	0.13 [0.11, 0.14]
Morishita 2017a	270	10055	14	2568	Rural poor communities	0.95 [0.92, 0.97]	0.20 [0.20, 0.21]
Ntinginya 2012	5	29	0	185	(Household) contacts	1.00 [0.48, 1.00]	0.86 [0.81, 0.91]
Chheng 2008b	9	261	0	102	VCT attendants	1.00 [0.66, 1.00]	0.28 [0.24, 0.33]



The summary sensitivity was 70.6% (95% CI 61.7% to 78.2%; very low-certainty evidence because of very serious risk of bias and serious inconsistency), and specificity was 65.1% (95% CI 53.3% to 75.4%; low-certainty evidence because of serious inconsistency and serious imprecision) (Figure 11; Summary of findings 3).

Definition of any tuberculosis symptom (i.e. which symptom questions were asked) did not explain the variation. The number of symptoms asked ranged between 3 and 9 (median 5). In 16 studies with 3 to 5 symptom questions asked, sensitivity was 73.6% (95% CI 61.1% to 83.2%) and 67.0% (95% CI 54.6% to 77.3%) in 13 studies in which 6 to 8 questions were asked (P = 0.13). Specificity was not associated with number of symptom questions (65.3% for 3 to 5 and 65.2% for 6 to 8 questions).

Investigations of heterogeneity

Of the study-level and participant-level characteristics examined as possible explanations for heterogeneity of symptom index test(s), Table 1 shows the variables that were statistically significant modifiers of either sensitivity, or specificity, or both. At the study level, sensitivity of all three symptom index tests differed by country income level and was highest in studies from LICs compared to LMICs, and lowest in UMICs. Yet, the pattern in specificity was not consistent. The specificity of cough ≥ 2 weeks was higher in studies with lower tuberculosis prevalence (P = 0.03), but this trend was not statistically significant for any cough and any tuberculosis symptom. Of participant-level characteristics, the proportion reporting prior tuberculosis treatment modified sensitivity of both cough ≥ 2 weeks and any cough. Among studies reporting a higher proportion of prior tuberculosis treatment, sensitivity was lower, but for any tuberculosis symptom this was not statistically significant. Studies with high smoking prevalence showed higher sensitivity of any tuberculosis symptom, compared to lower smoking prevalence (76.2%, 95% CI 66.7% to 83.6%) versus 58.4% (95% CI 42.7% to 72.7%), P = 0.069), but this pattern was not found in cough ≥ 2 weeks and any cough. With respect to overlap between variables shown in Table 1, a higher proportion of previous tuberculosis was more common in the UMICs for all three symptom index tests (P < 0.05, Fisher's exact test). For any tuberculosis symptom, a higher proportion of previous tuberculosis overlapped with high smoking prevalence (P = 0.01). HIV prevalence did not modify sensitivity or specificity of any of the symptom index tests.

Sensitivity analyses symptom index tests

Exclusion of studies at unclear or high risk of bias, or with applicability concerns in the Patient Selection domain (Figure 7) did not change the summary estimates for cough ≥ 2 weeks and any cough. However, for any tuberculosis symptom, this resulted in a lower sensitivity and a higher specificity (Table 2). High risk of bias in the Reference Standard domain was present in the majority of studies. Restriction to studies with low risk of bias for this domain reduced the summary sensitivity estimate of all three symptom index tests and increased specificity. High risk of bias in the Flow and Timing domain was also common and restriction to studies with low risk of bias for this domain showed a similar pattern. Similar to CXR index tests, restriction to studies with a microbiological reference standard on at least 20% of participants lowered the sensitivity of cough ≥ 2 weeks and any cough, but less so for any tuberculosis symptom. Restriction to studies that included culture in the microbiological reference standard (with or without smear microscopy) resulted in a lower sensitivity and higher specificity of any tuberculosis symptom. Restriction to studies with Xpert MTB/RIF included in the reference standard but not culture (n = 7) resulted in a higher sensitivity and considerably lower specificity compared to the overall summary summary. Nevertheless, data are not shown as 5/7 of these studies were from one publication (Morishita 2017a), and shared other characteristics and biases, making it hard to attribute differences in accuracy to the reference standard.

Six studies from African countries with a high prevalence of HIV provided accuracy data on symptom index tests (any tuberculosis symptom n = 6; any cough n = 3, cough ≥ 2 weeks n = 4) separately for the HIV-negative and HIV-positive populations. We included the HIV-negative populations in our primary analyses given our review objectives. In additional sensitivity analyses, these studies were replaced with the data of their full study populations, thus also changing the HIV prevalence for these studies. The effect of this decision on summary estimates was minimal (data not shown). Furthermore, this decision did not change our previous finding regarding HIV prevalence not significantly modifying the sensitivity or specificity of symptom index tests.

Other symptom index test definitions

Six studies provided data on cough for three or more weeks (≥ 3 weeks). However, this definition was obsolete after clinical guidelines were changed to indicate that all persons with cough ≥ 2 weeks require further examination for tuberculosis. In two studies providing data both on cough ≥ 2 weeks and cough ≥ 3 weeks, enabling a within-study comparison of these two tests, the latter had slightly lower sensitivity and slightly higher specificity (Appendix 4).

A number of included studies provided data on other symptom combinations than those described above, mostly from post-hoc analyses comparing different combinations of the symptoms asked in the respective studies. The definitions of these other symptom combinations were quite heterogenous and therefore not further pooled in a meta-analysis. The most common combinations were cough ≥ 2 weeks and in addition at least one non-cough symptom, or a combination of two out of several symptoms. These are shown in Table 3, to give an impression how the sensitivity and specificity of such combinations lie in between those of cough ≥ 2 weeks and any tuberculosis symptom in the same studies (n = 10). Any comparisons between these tests or test combinations should be interpreted with caution, as a formal comparison was outside the scope of this review and these informal comparisons may reflect (very) low-certainty evidence. For additional combinations that were occasionally reported, we refer to the respective publications (Claassens 2017a; Corbett 2010b; Lewis 2009b).

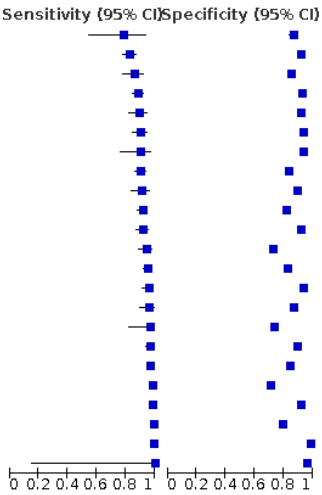
Chest radiography index tests

Any CXR abnormality

Twenty-three studies contributed to this index test, of which all but three were conducted in the general population (Figure 14). In total, 4532 participants contributed to the estimation of sensitivity, which ranged between 70% and 100%. The most extreme values were in studies with smaller sample sizes and wider 95% CIs. Specificity ranged from 71% to 99%, to which 1,034,525 participants contributed.

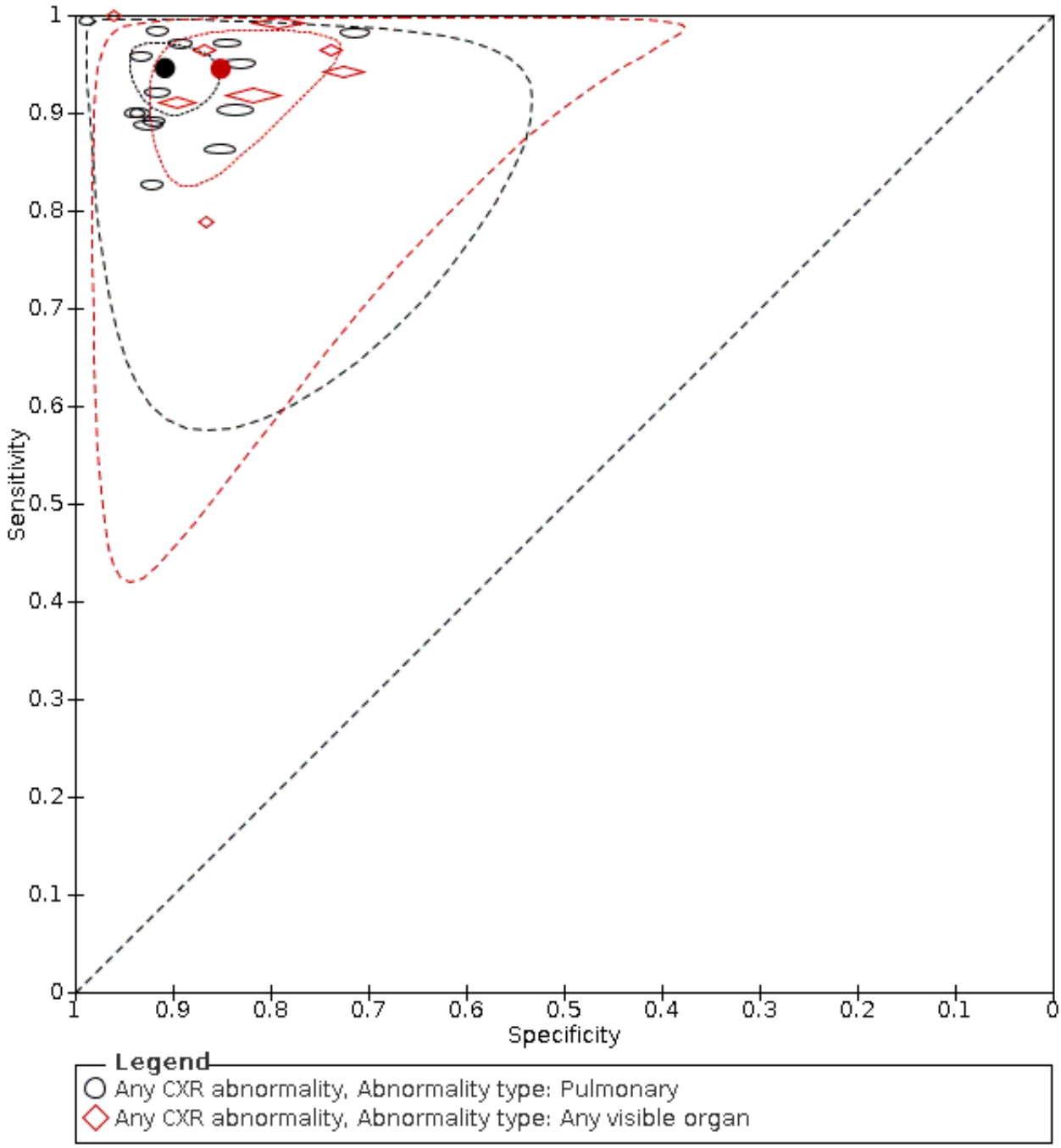
Figure 14. Forest plot of studies providing data on index test 'Any CXR abnormality'. CI: confidence interval; FN: false negative; FP: false positive; TN: true negative; TP: true positive.

Study	TP	FP	FN	TN	Population	Sensitivity (95% CI)	Specificity (95% CI)
Telisinghe 2014	15	82	4	532	Prison inmates	0.79 [0.54, 0.94]	0.87 [0.84, 0.89]
Kapata 2016	219	3539	46	41829	General population	0.83 [0.78, 0.87]	0.92 [0.92, 0.92]
Federal MoH Sudan 2018	88	11573	14	66255	General population	0.86 [0.78, 0.92]	0.85 [0.85, 0.85]
Qadeer 2016	299	7562	38	93353	General population	0.89 [0.85, 0.92]	0.93 [0.92, 0.93]
Kebede 2014	98	3720	12	42718	General population	0.89 [0.82, 0.94]	0.92 [0.92, 0.92]
Republic of Uganda 2018	144	2706	16	38137	General population	0.90 [0.84, 0.94]	0.93 [0.93, 0.94]
Rwanda MoH 2014	36	2616	4	40413	General population	0.90 [0.76, 0.97]	0.94 [0.94, 0.94]
NTP Bangladesh 2017	251	16126	27	82155	General population	0.90 [0.86, 0.94]	0.84 [0.83, 0.84]
Republic of Zimbabwe 2015	92	3339	9	29061	General population	0.91 [0.84, 0.96]	0.90 [0.89, 0.90]
Kenya MoH 2018	274	11275	24	50911	General population	0.92 [0.88, 0.95]	0.82 [0.82, 0.82]
Ghana NTP 2015	152	5006	13	54547	General population	0.92 [0.87, 0.96]	0.92 [0.91, 0.92]
van't Hoog 2012	113	5229	7	13867	General population	0.94 [0.88, 0.98]	0.73 [0.72, 0.73]
MoH Indonesia 2015	401	10801	21	53115	General population	0.95 [0.92, 0.97]	0.83 [0.83, 0.83]
MoPH Thailand 2017	136	4157	6	57644	General population	0.96 [0.91, 0.98]	0.93 [0.93, 0.93]
Morasert 2018	81	509	3	3414	Prison inmates	0.96 [0.90, 0.99]	0.87 [0.86, 0.88]
den Boon 2006	28	674	1	1905	General population	0.97 [0.82, 1.00]	0.74 [0.72, 0.76]
Law 2015	230	4167	7	34591	General population	0.97 [0.94, 0.99]	0.89 [0.89, 0.90]
Mongolia MoH 2016	238	7643	7	41633	General population	0.97 [0.94, 0.99]	0.84 [0.84, 0.85]
NTP Philippines 2018	430	11716	8	29289	General population	0.98 [0.96, 0.99]	0.71 [0.71, 0.72]
MoH Cambodia 2012	306	3103	5	33807	General population	0.98 [0.96, 0.99]	0.92 [0.91, 0.92]
MoH Myanmar 2012	290	10347	2	39602	General population	0.99 [0.98, 1.00]	0.79 [0.79, 0.80]
MoPH DPRK 2017	333	695	2	58619	General population	0.99 [0.98, 1.00]	0.99 [0.99, 0.99]
Fox 2012	2	20	0	523	(Household) contacts	1.00 [0.16, 1.00]	0.96 [0.94, 0.98]



The summary sensitivity was 94.7% (95% CI 92.2% to 96.4%; very low-certainty evidence because of very serious risk of bias and serious inconsistency), and the summary specificity was 89.1% (95% CI 85.6% to 91.8%; low-certainty evidence because of serious inconsistency and serious imprecision) (Figure 15; Summary of findings 4). Specificity of the index test definition that included abnormalities in other visible organs was 85.1% (95% CI 78.3% to 90.1%, n = 8), and specificity of pulmonary abnormalities only was 90.8% (95% CI 87.0% to 93.5%, n = 15). (Note: the two definitions were not compared in the same studies). This index test definition was not a statistically significant modifier of specificity (P = 0.096) (Figure 15). Sensitivity did not differ between index test definitions (94.5% and 95.6% respectively).

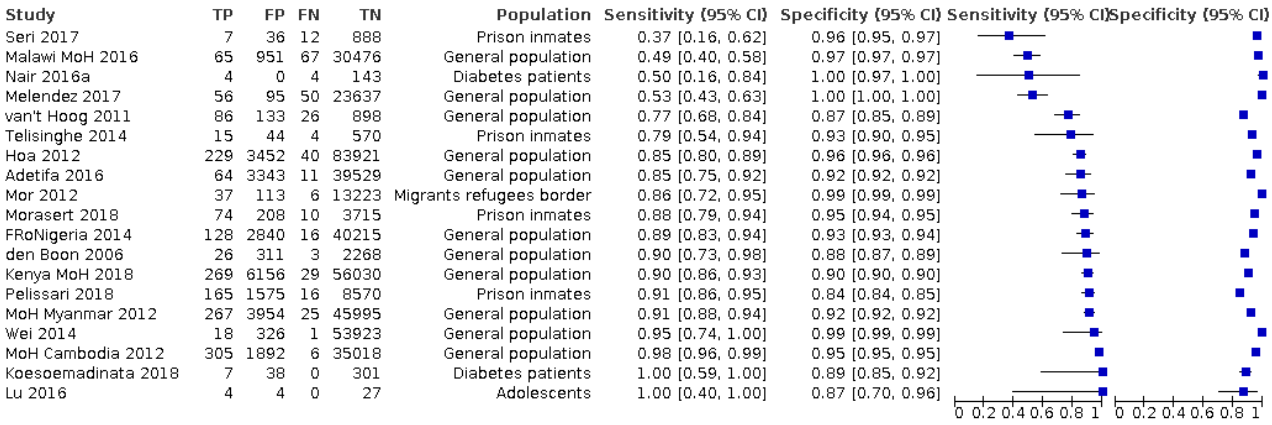
Figure 15. Summary ROC plot of index test 'Any CXR abnormality', stratified by studies restricting to pulmonary abnormalities only versus studies that also included extrapulmonary abnormalities like cardiomegaly. The plot shows summary estimates, 95% confidence (dotted lines) and 95% prediction intervals (dashed lines).



CXR abnormalities suggestive of tuberculosis

Nineteen studies contributed to this index test, of which 11 were in the general population and four in prison inmates (Figure 16). In total, 2152 participants contributed to the estimation of sensitivity, which ranged from 37% to 100%, with the most extreme values mostly from smaller studies. Specificity ranged from 84% to 100%, to which 464,818 participants contributed. Studies with low sensitivity had high specificity.

Figure 16. Forest plot of studies providing data on index test 'CXR abnormalities suggestive of tuberculosis'. CI: confidence interval; FN: false negative; FP: false positive; TN: true negative; TP: true positive.



The summary sensitivity of CXR abnormalities suggestive of tuberculosis was 84.8% (95% CI 76.7% to 90.4%; low-certainty evidence because of serious risk of bias and serious inconsistency), and the summary specificity was 95.6% (95% CI 92.6% to 97.4%; high-certainty evidence) (Summary of findings 5).

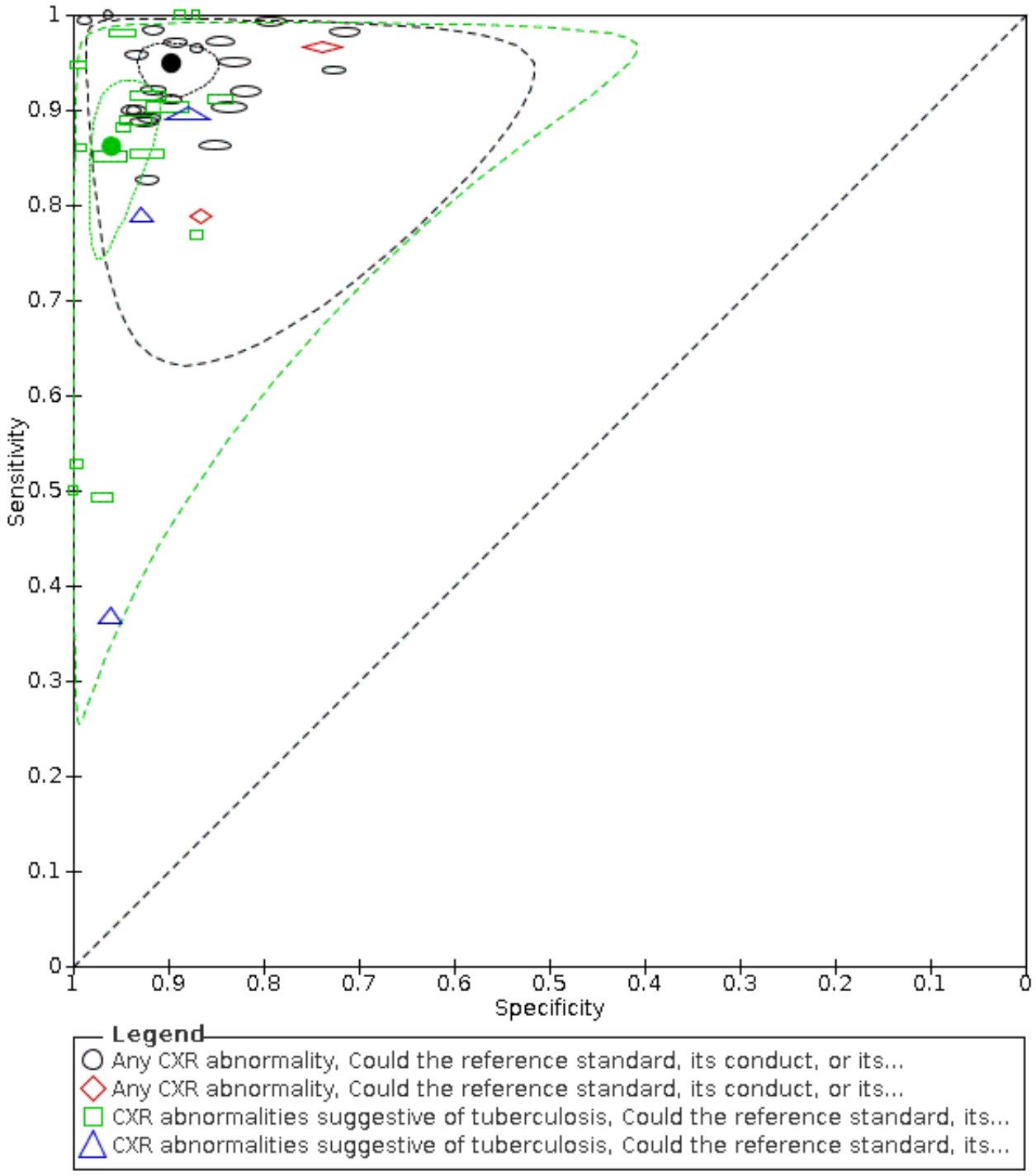
Investigations of heterogeneity

Of all study-level and participant-level characteristics examined as possible explanations for heterogeneity, Table 4 shows the variables that were statistically significant modifiers of either sensitivity, or specificity, or both. For both CXR index tests, sensitivity was lower in studies in the Africa region compared to other regions combined, but specificity was not associated with geographic region. In studies with lower tuberculosis prevalence, the sensitivity of both CXR index tests was lower, and specificity higher, although the former was only statistically significant for any CXR abnormality, and the latter only for CXR abnormalities suggestive of tuberculosis. The sensitivity of any CXR abnormality was lower in studies with a lower smoking prevalence among the population, and in studies with on average a younger population, but specificity did not differ. For CXR abnormalities suggestive of tuberculosis, sensitivity was also lower in studies with a lower prevalence of smoking, but this difference was not statistically significant. The categorical variables shown in Table 4 overlapped considerably, especially for the any CXR abnormality group: compared to the African region, studies from other regions had a higher median or mean age (P = 0.04) and smoking prevalence (P = 0.005). Also the median or mean age, tuberculosis prevalence and smoking were associated with each other (P = 0.03, P = 0.02, P = 0.05 respectively). Among the CXR abnormalities suggestive of tuberculosis group, the higher age and prevalence of smoking categories were more common in the other regions but were not statistically significant. HIV prevalence was higher among studies from the African region, for both CXR index tests (P = 0.04, P = 0.06). For CXR abnormalities suggestive of tuberculosis, sensitivity was lower with increasing HIV prevalence, but not statistically significant. CXR interpretation by specialists was less frequent in the African region. For Any CXR abnormality two of 10 studies with available data reported specialists, compared to 8/12 studies in the Asian/other regions; P = 0.04). For CXR abnormalities suggestive of tuberculosis the frequency of specialist interpreters did differ by region (5/8 in the African versus 6/9 in the Asian/other region s; P = 0.57). The sensitivity of specialists was slightly higher for both CXR index tests, but the difference in sensitivity and in specificity between type of CXR interpreter was not statistically significant.

Sensitivity analyses CXR index tests

Of the results of sensitivity analyses based on QUADAS-2 domains and other study design features, Table 5 shows the results that differed substantially from the primary meta-analytic estimates of any of the CXR index tests. Of the QUADAS-2 domains (Figure 8), risk of bias in the Reference Standard domain (absence of verification with a microbiological reference test) is an important issue, as we explained previously. We expect this feature to mainly bias sensitivity. The number of studies with low risk of bias in the Reference Standard domain was too small for model convergence for summary estimates. The ROC plot suggested lower accuracy in studies with a low risk of bias in the Reference Standard domain (Figure 17).

Figure 17. Summary ROC plot of index tests 'Any CXR abnormality' and 'CXR abnormalities suggestive of tuberculosis' each stratified by high versus low risk of bias in the QUADAS-2 reference standard domain (see Footnote). The plot shows summary estimates, 95% confidence (dotted lines) and 95% prediction intervals (dashed lines). The number of studies with low risk was too low for meta-analyses, hence a summary estimate is not shown.
Footnote: **The full QUADAS-2 question referred to in the legend is: Could the reference standard, its conduct, or its interpretation have introduced bias?**



As a different approach to examining the effect of excluding studies at high risk of bias, we restricted the analysis to seven studies in which 20% or more of participants had a microbiological reference standard. This decision did not change summary sensitivity,

but resulted in a lower specificity, and more so for any CXR abnormality. Restriction to studies with low applicability concerns in the Reference Standard domain slightly reduced summary sensitivity for abnormalities suggestive of tuberculosis, but not for any CXR abnormality. When restricting to studies with low risk in the Flow and Timing domain, summary specificity was slightly higher for both CXR index tests compared to the primary meta-analysis. Summary sensitivity of CXR abnormalities suggestive of tuberculosis was lower, for which only three studies were available. Although some sensitivity analyses resulted in a (slightly) different summary sensitivity or specificity, the CIs of all these results largely overlapped with those of the primary meta-analyses due to the loss of precision.

Chest radiography and symptom questions in parallel

Twenty-five studies contributed to the index test that combines CXR and symptom questions in parallel. This index test is diverse in its definition. Almost half of the studies reported on CXR abnormalities suggestive of tuberculosis, the other half on pulmonary abnormalities, and two on any CXR abnormality (Figure 18). These were combined with questions on cough, mostly cough ≥ 2 weeks, while a few studies used cough ≥ 3 weeks or any cough. Some studies used a combination of symptoms or a score based on different symptom questions. Table 6 shows each study with available data for accuracy of 'parallel CXR and symptom index test' alongside the accuracy of CXR in the same study, as an informal comparison, reflecting very low-certainty evidence. Generally, adding symptoms in parallel to CXR screening might increase sensitivity, and reduce specificity. For the addition of cough ≥ 2 weeks to any (pulmonary) CXR abnormality (n = 9) the median increase in sensitivity and decrease in specificity was 1.6% (range -9.4 to +11%) and -3.2% (range -0.3 to -6.2%) respectively. However this varied between studies in this subgroup and at the same time overlap with studies applying other symptom and CXR definitions (Table 6; Figure 19). We meta-analysed these parallel index tests separately for subgroups of more homogenous definitions of CXR classification and symptom questions applied, but then combined them into one overall estimate, as the summary sensitivity and sensitivity of the subgroups were very similar to each other (Table 7). The summary sensitivity of all studies (n = 25) was 99.6% (95% CI 98.3% to 99.9%) and summary specificity was 84.2% (95% CI 81.1% to 87.0%). In all but one study, we judged the risk of bias in the Reference Standard domain to be high, mostly due to incorporation bias as participants without CXR abnormalities and without symptoms were classified as not having tuberculosis and ineligible for further confirmation by microbiological sputum tests. Restricting the analysis to studies in which at least 20% of all participants received microbacteriological testing (Table 7) widened the confidence interval for sensitivity (99.0%, 95% CI 93.5% to 99.8%) and substantially reduced specificity to 74.5% (95% CI 70.3% to 78.3%). These comparisons are informal and the certainty of evidence was not assessed. We did not further investigate for heterogeneity, since the studies on this index test are the same studies included in the analyses of CXR index tests described above.

Figure 18. Forest plot of studies providing data on index test 'Parallel CXR and symptom screening'. CI: confidence interval; FN: false negative; FP: false positive; TN: true negative; TP: true positive.

Study	TP	FP	FN	TN	Abnormality type	Parallel-CXR-symptom combi	Sensitivity (95% CI)	Specificity (95% CI)
FRoNigeria 2014	144	4544	0	39498	Sugg. TB	(prolonged) cough	1.00 [0.97, 1.00]	0.90 [0.89, 0.90]
MoH Cambodia 2005	272	3301	0	18859	Sugg. TB	(prolonged) cough	1.00 [0.99, 1.00]	0.85 [0.85, 0.86]
Koesoemadinata 2018	7	63	0	276	Sugg. TB	(prolonged) cough	1.00 [0.59, 1.00]	0.81 [0.77, 0.85]
Pelissari 2018	181	3318	0	6827	Sugg. TB	(prolonged) cough	1.00 [0.98, 1.00]	0.67 [0.66, 0.68]
Kenya MoH 2018	301	9020	4	53725	Sugg. TB	(prolonged) cough	0.99 [0.97, 1.00]	0.86 [0.85, 0.86]
MoH Myanmar 2012	299	11936	12	39120	Sugg. TB	(prolonged) cough	0.96 [0.93, 0.98]	0.77 [0.76, 0.77]
Hoa 2012	253	7245	16	86244	Sugg. TB	(prolonged) cough	0.94 [0.91, 0.97]	0.92 [0.92, 0.92]
Malawi MoH 2016	132	3300	0	28147	Sugg. TB	several symptoms	1.00 [0.97, 1.00]	0.90 [0.89, 0.90]
Morasert 2018	83	817	1	3106	Sugg. TB	several symptoms	0.99 [0.94, 1.00]	0.79 [0.78, 0.80]
den Boon 2006	26	0	3	0	Sugg. TB	several symptoms	0.90 [0.73, 0.98]	Not estimable
Telisinghe 2014	16	219	3	395	Sugg. TB	several symptoms	0.84 [0.60, 0.97]	0.64 [0.60, 0.68]
MoH Indonesia 2015	426	15020	0	52498	Pulmonary	(prolonged) cough	1.00 [0.99, 1.00]	0.78 [0.77, 0.78]
MoPH DPRK 2017	340	4462	0	55881	Pulmonary	(prolonged) cough	1.00 [0.99, 1.00]	0.93 [0.92, 0.93]
Rwanda MoH 2014	40	4707	0	38381	Pulmonary	(prolonged) cough	1.00 [0.91, 1.00]	0.89 [0.89, 0.89]
Republic of Uganda 2018	160	4852	0	36142	Pulmonary	(prolonged) cough	1.00 [0.98, 1.00]	0.88 [0.88, 0.88]
Qadeer 2016	339	10120	1	95441	Pulmonary	(prolonged) cough	1.00 [0.98, 1.00]	0.90 [0.90, 0.91]
Mongolia MoH 2016	245	9365	3	40696	Pulmonary	(prolonged) cough	0.99 [0.97, 1.00]	0.81 [0.81, 0.82]
NTP Philippines 2018	448	13517	18	32706	Pulmonary	(prolonged) cough	0.96 [0.94, 0.98]	0.71 [0.70, 0.71]
Federal MoH Sudan 2018	96	12553	15	70346	Pulmonary	(prolonged) cough	0.86 [0.79, 0.92]	0.85 [0.85, 0.85]
Ghana NTP 2015	167	6189	35	55335	Pulmonary	(prolonged) cough	0.83 [0.77, 0.88]	0.90 [0.90, 0.90]
Kapata 2016	265	5883	0	39951	Pulmonary	several symptoms	1.00 [0.99, 1.00]	0.87 [0.87, 0.87]
MoPH Thailand 2017	142	5908	0	56486	Pulmonary	several symptoms	1.00 [0.97, 1.00]	0.91 [0.90, 0.91]
NTP Bangladesh 2017	278	20316	0	78116	Pulmonary	several symptoms	1.00 [0.99, 1.00]	0.79 [0.79, 0.80]
van't Hoog 2012	119	6319	4	14114	Any visible organ	(prolonged) cough	0.97 [0.92, 0.99]	0.69 [0.68, 0.70]
Republic of Zimbabwe 2015	103	4533	4	29096	Any visible organ	several symptoms	0.96 [0.91, 0.99]	0.87 [0.86, 0.87]

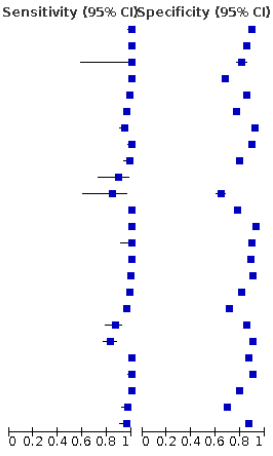
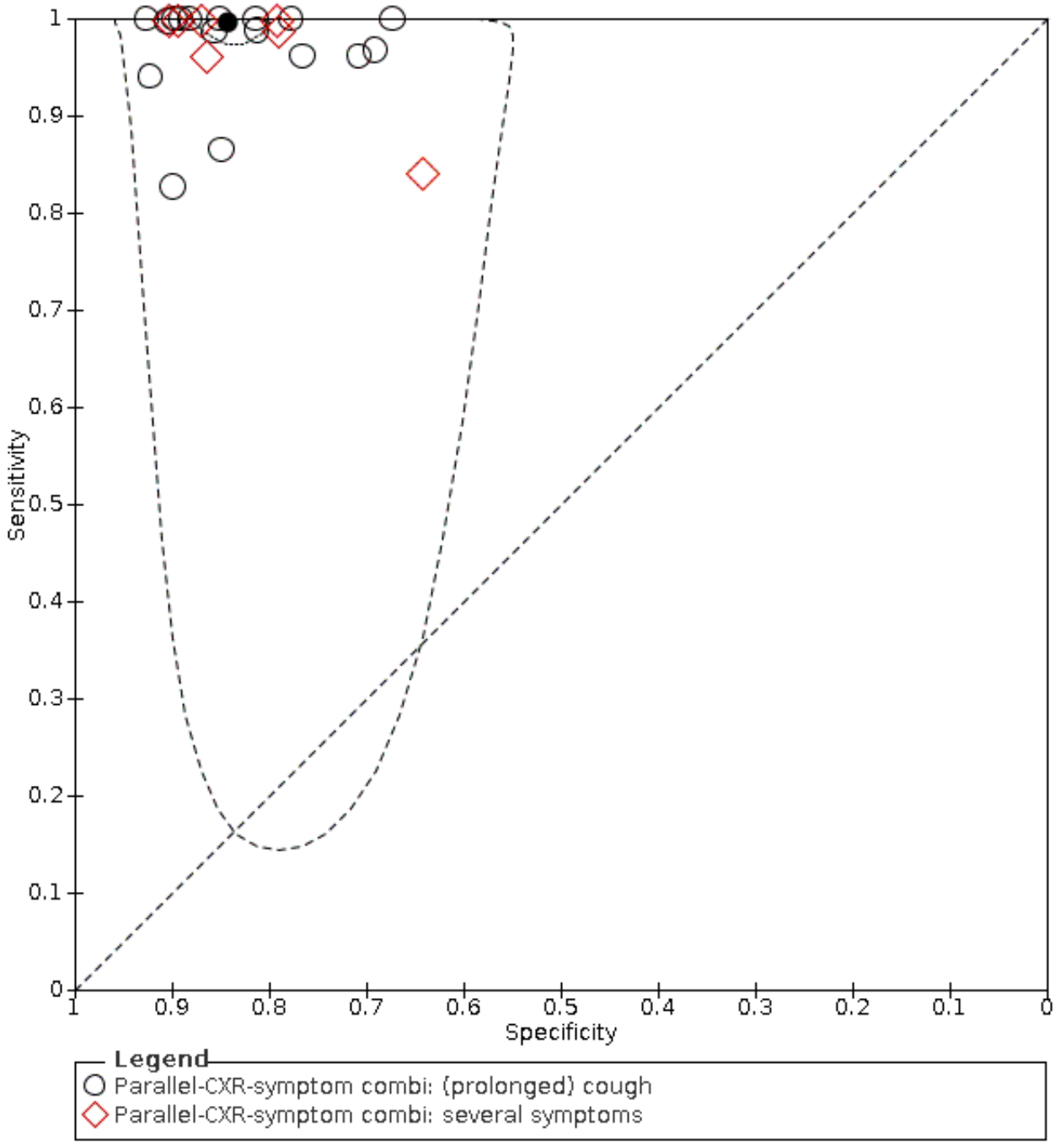


Figure 19. Summary ROC plot of studies providing data on Index Test 'Parallel CXR and symptom screening'. The plot shows the summary estimate, 95% confidence (dotted lines) and 95% prediction intervals (dashed lines). Circles denote studies that applied (prolonged) cough as the symptom screen, and squares denote studies that applied another symptom combination as the symptom screen.



Certainty of evidence and search date

The explorative search to assess if an updated search would potentially reach a higher certainty of evidence, resulted in approximately 3000 references of which 262 references remained after restricting the inclusion to tuberculosis terms, DTA terms, and terms for our reference standard in title or abstract. Assessment of title and abstract by two review authors independently resulted in 40 articles for full-text selection of which 12 studies fulfilled our inclusion criteria. See Characteristics of studies awaiting classification for the characteristics of each study and Appendix 3 for an assessment of their potential to add to the certainty of evidence. In short, five of these are studies that would replace

reports that are already in the review and not add data. Of the remaining studies two had a very low (10 and 2) number of tuberculosis cases to contribute to sensitivity estimates. The remaining studies had concerns about high risk of bias in at least two QUADAS-2 domains, suggesting that an update would not increase the certainty of evidence. In the current results the large number of prevalence surveys already provide us with a considerable amount of data to estimate specificity, and more data will not further increase the certainty of evidence. The uncertainty in the estimates for sensitivity will not be resolved, as the additional studies are small in sample size and tuberculosis cases.

DISCUSSION

Summary of main results

The objective of this review was to assess the sensitivity and specificity of index tests comprising tuberculosis symptoms, CXR abnormalities, or both, as screening tools to detect bacteriologically confirmed pulmonary tuberculosis disease in people considered eligible for tuberculosis screening, who are HIV-negative or whose HIV status is unknown. Moreover, the review aimed to investigate heterogeneity in relation to regional, epidemiological, participant, and test characteristics. Of 59 studies providing data on one or more of the index test definitions, almost all had methodological quality concerns. In the majority of studies, we considered the risk of bias in the Reference Standard domain to be high. Studies considered to be at high risk of bias in the Flow and Timing domain were also common.

The three most common symptom index tests, cough ≥ 2 weeks (41 studies), any cough (21 studies), and any tuberculosis symptom (29 studies), showed a summary sensitivity of 42.1% (95% CI 36.6% to 47.7%), 51.3% (95% CI 42.8% to 59.7%), and 70.8% (95% CI 61.5% to 78.6%; all very low-certainty evidence), and a specificity of 94.4% (95% CI 92.6% to 95.8%; high-certainty evidence), 87.6% (95% CI 81.6% to 91.8%; low-certainty evidence), and 63.7% (95% CI 51.7% to 74.3%; low-certainty evidence), respectively. The data on symptom index tests were highly heterogenous. The studies on any tuberculosis symptom were the most heterogeneous, but had the lowest number of variables explaining this variation. Symptom index tests also showed regional variation; sensitivity was modified by country income level, but not WHO region, with higher sensitivity in studies from LICs compared to LMICs, and lower sensitivity in studies from UMICs. The proportion of people with a history of previous tuberculosis among the screened participants modified the sensitivity of both cough index tests, and the specificity of the cough ≥ 2 weeks index test. A difference in the prevalence of HIV infection between the studies did not provide a statistically significant explanation of heterogeneity.

CXR index tests were the most sensitive screening tools. The summary sensitivity of any CXR abnormality (23 studies) was 94.7% (95% CI 92.2% to 96.4%; very low-certainty evidence) and 84.8% (95% CI 76.7% to 90.4%; low-certainty evidence) for CXR abnormalities suggestive of tuberculosis (19 studies), and specificity was 89.1% (95% CI 85.6% to 91.8%; low-certainty evidence) and 95.6% (95% CI 92.6% to 97.4%; high-certainty evidence), respectively. Presence of heterogeneity reduced the certainty of evidence more so for sensitivity than for specificity, and could be partly explained by regional variation. In studies in the African WHO region, sensitivity was lower compared to other regions combined, but specificity was not. This is likely due to differences in regional distributions of other modifying factors, such as the prevalence of smoking, the age distribution, and tuberculosis prevalence among the study populations, and possibly the availability of medical specialist to interpret CXRs.

Parallel combinations of CXR and symptom questions (25 studies) had heterogenous definitions but similar sensitivities and specificities (informal comparison; certainty of evidence not assessed). The addition of cough ≥ 2 weeks, whether to any (pulmonary) CXR abnormality or to tuberculosis suggestive CXR abnormalities resulted in a similar summary sensitivity and specificity, 99.2% (95% CI 96.8% to 99.8) and 84.9% (95% CI 81.2% to 88.1%) (n = 15).

Our summary estimates of sensitivity of all index tests should be interpreted with great caution. As shown in our sensitivity analyses, it may have been overestimated, especially due to incorporation bias. For screening programmes, the observed heterogeneity implies that the expected sensitivity and - to a lesser extent - specificity of an index test in a specific setting cannot be predicted with great precision.

Our analyses of heterogeneity highlighted regional differences as the most consistent explanation for variation. CXR index tests showed lower sensitivity in the WHO Africa region compared to others (mostly SEARO/WPRO regions), and symptom index tests varied by country income level, with higher sensitivity in LICs. A hypothesized explanation for this may be poverty-related differences in access to care and care seeking. In the previous report (van't Hoog 2013a), we also found regional variation with higher sensitivity of symptom index tests in African compared to Asian countries, which we then attributed to coinciding differences in HIV prevalence in the population. Our current results suggest that regional differences in other factors such as smoking, age, and tuberculosis prevalence among the screened population may explain the differences more than differences in overall HIV prevalence.

Strengths and weaknesses of the review

One of the strengths of this review is the comprehensive search strategy, which included not only four electronic databases and reference checking, but also an extensive search of grey literature. Other strengths are the wide representation of the study population and the detailed investigation of heterogeneity and sensitivity analyses. In our analyses we combined studies from all populations, based on the rationale that differences in sensitivity and specificity between different population categories will be reflected in differences in epidemiological and demographic factors, which we explored in our analyses of heterogeneity. The large number of studies allowed analyses exploring a large number of covariates as sources of heterogeneity, which highlighted regional differences as the most consistent explanation for variation.

Our analyses of heterogeneity are, however, somewhat limited by the fact that the variables reflect study-level variation, and the analysis approach implies stratification by one variable at the time. Our results do not predict individual participant-level variation. CXR abnormalities are more common among individuals who smoke, are older, and after a previous tuberculosis episode (MoH Myanmar 2012; Pinsky 2006), thus reducing the specificity of CXR as a screening tool for tuberculosis disease. However, to

predict the specificity in a particular target population (for instance, high average age, high smoking prevalence, and frequent previous tuberculosis) an individual patient data (IPD) meta-analysis would be needed (Getahun 2011). The conduct of the index test, like the exact definition of CXR abnormalities and background of the interpreter may matter more than the limited power of our analyses could demonstrate. Our review did not address differences in quality assurance of index tests, nor in background and training of interviewers administering the symptom questionnaires. A next update may incorporate more index test details in the analyses. Another statistical analysis limitation may be that many studies applied cluster sampling, requiring adjustment in the analysis, with larger standard errors and slightly wider CIs as a result, compared to the binomial exact CIs used in the analysis in Review Manager. While small study-level differences can be noted in some studies (e.g. den Boon 2006; van't Hoog 2012), an overall impact on our summary estimates is expected to be small.

Another limitation of this review is that studies and reports published from 2019 onwards were not included. The results as presented have informed WHO's updated tuberculosis screening guidelines (WHO 2021a). While an updated search will yield more studies, we doubt that our findings or the certainty of evidence would change, for several reasons. The perfect study for our objective is hardly ever conducted due to resource implications, and as demonstrated in Figure 4 and Figure 6, we judged almost all studies to have one or more concerns with respect to bias or applicability. The majority of studies in this review were national tuberculosis prevalence surveys, which share a design as explained above, that considers the trade-off between the objective to measure the prevalence of tuberculosis disease, which requires large sample sizes, and cost and logistical feasibility. For the purpose of our review, this design implies incorporation bias, potential bias in the Flow and Timing domain, and potentially overestimated sensitivity. While the magnitude of bias may vary depending on the type of CXR abnormalities and symptoms used in these prevalence surveys (Law 2020; WHO 2011), more studies of this type will add to the problem of incorporation bias. The estimates of specificity from national tuberculosis prevalence surveys with the presence of incorporation bias by design are, however, still reliable, as the numbers misclassified as negative for tuberculosis will be very small compared to the large numbers of persons who test negative for tuberculosis. For the summary estimates of specificity, the certainty of evidence has increased compared to the earlier report (van't Hoog 2013a). For summary sensitivity, however, the certainty of evidence remained very low, despite the larger sample (Table 8). Lastly, the definitions of the index tests are subject to human interpretation, which may vary between studies and between individuals. Inter- and intra-reader variation in CXR reading is well known and, generally speaking, agreement between readers is modest (den Boon 2005; van't Hoog 2011). We do not expect technological advancements to overcome variation in human interpretation of the index tests as defined in this review. Computer-assisted reading of digital CXRs (Melendez 2017; Muyoyeta 2017), as an alternative for human CXR interpretation, is being addressed in other reviews (WHO 2021a). As explained in the results and Appendix 3, we explored what an updated search would add to the certainty of evidence, and concluded that besides a considerable amount of time, an update would not increase the certainty of evidence as it would add a few more studies with the same level of heterogeneity and risk of bias. The prevalence surveys already provide us with a considerable amount of data to estimate specificity, and more data will not further increase the certainty of evidence. The uncertainty in the estimates for sensitivity will not be resolved, as the additional studies are small in sample size and tuberculosis cases.

Another important reason for our serious doubt about performing an update is that the tuberculosis field is urgently waiting for the results of this review to be published. Waiting until the update is ready would not be of benefit to the field. The data from the review, as they are now, are included in the WHO 2021a update of the tuberculosis screening guidelines. We think that a more logical moment for an update would be the next update of the guideline (in five to six years). In that way the data from the Cochrane Review and the WHO guideline are the same.

Certainty of the evidence

Despite the strengths of the review, the certainty of evidence remained very low for the summary sensitivity estimates of all index tests, except for CXR abnormalities suggestive of tuberculosis, for which the certainty of evidence was low. Serious inconsistency, as discussed above, contributed to a very low certainty of evidence, as the range in point estimates was wide to very wide. Furthermore, the (very) serious risk of bias in the Reference Standard domain reduced the certainty of evidence, as more than half of the studies did not require all participants to undergo microbacteriological testing, but classified participants as tuberculosis-negative based on results of CXR and symptoms (incorporation bias). Also in the Flow and Timing domain we considered the risk of bias to be high for over half of the studies, mostly because of the extent of missing test results for participants who required microbacteriological testing based on the protocol. Our sensitivity analyses confirmed the direction of this bias, since restricting analyses to studies with low risk of bias in these QUADAS-2 domains resulted in considerably lower summary sensitivity estimates. The high sensitivity of the parallel combination of CXR and symptom questions should be interpreted with even greater caution as in most of these studies sensitivity was (close to) 100%. The sensitivity was 100% by definition due to the fact that participants without CXR abnormalities and without symptoms were defined by this parallel combination as 'not having tuberculosis' and were ineligible for further confirmation by microbiological sputum tests (Law 2020; WHO 2011). However, we do not present results from sensitivity analyses as primary results, because any choice of subset would be somewhat arbitrary and remain biased due to other concomitant factors. For specificity, we considered the certainty of evidence to be high for CXR abnormalities suggestive of tuberculosis and for cough ≥ 2 weeks, and low for the other index test definitions. The latter because the confidence interval around the number of false positives was such that the proportion of the population requiring follow-up testing can vary by almost a factor two, which has serious implications for the estimation of resources (Summary of findings 1; Summary of findings 2; Summary of findings 3; Summary of findings 4; Summary of findings 5).

Comparison with other systematic reviews

Earlier systematic reviews assessing symptoms and CXR as tuberculosis screening tools in adults present subsets of studies that are also included in the current review (Assefa 2019b; van't Hoog 2013a). Other systematic reviews on screening tools

for pulmonary tuberculosis in adults address either different tests: Xpert MTB/RIF and Xpert Ultra (Shapiro 2021) and C-reactive protein (Yoon 2017), or different populations: HIV-infected adults (Hamada 2018) and children (Vonasek 2021). In the recent Cochrane Review assessing the accuracy of symptom screening, CXR, or Xpert MTB/RIF for screening of childhood tuberculosis (Vonasek 2021), the summary sensitivity of one of the symptom screens, ≥ one of cough, fever, or poor weight gain in tuberculosis contacts, seems high (89%) compared to the most sensitive symptom screen in our review (71%). Also the summary specificity of CXR with any abnormality in children who were tuberculosis contacts was high (99%). However, the 95% confidence intervals are wide as the paediatric review included fewer studies. Moreover, comparisons between adult and paediatric pulmonary tuberculosis are challenging because the diseases are so different. Restriction of the reference standard to bacteriological confirmation in our review, implying underdiagnosis of tuberculosis, and additional verification bias, resulted in overestimated sensitivity, as shown in our analyses. The paediatric review applied a composite reference standard including clinical pulmonary tuberculosis of which symptoms and CXR are inherent components. If agreement between the index test and the reference standard increases, accuracy will be overestimated due to incorporation bias (Vonasek 2021).

Applicability of findings to the review question

We had fewer concerns about applicability of the included studies to the review question compared to concerns about bias. The applicability of the definition of the target condition was a possible concern in 15% (CXR) to 30% (symptoms) of studies. The target condition of tuberculosis screening in this review was bacteriologically confirmed pulmonary tuberculosis, for which we allowed as a reference standard any author-defined combination of mycobacterial culture (on solid or liquid medium), sputum smear microscopy, Xpert MTB/RIF, or other NAATs, but also required some evidence of active tuberculosis disease, like symptoms, CXR abnormalities, or bacteriological confirmation on repeated sputum samples. Inclusion of cases defined on clinical grounds only implied an applicability concern, as well as inclusion of cases on the basis of a single positive sputum culture at one point in time without any symptom or CXR abnormality. As an illustration of the latter: in one study from which data for different case definitions could be extracted, of 14 cases as defined by the study only seven cases met the case definition that was applicable to the review (Nair 2016a; Nair 2016b). A few studies applied more stringent clinical and microbacteriological follow-up procedures for individuals with one positive culture result (Corbett 2010b; Lewis 2009b), which may better reflect the tuberculosis disease spectrum (Pai 2016) and the clinical reality in which diagnosing tuberculosis disease with certainty can be challenging. Our reference standard definition is quite strict, and would not capture, for instance, extrapulmonary tuberculosis without concurrent presence of pulmonary tuberculosis, in which case the sputum-based microbacteriological tests will be negative. The increasing availability of diagnostic methods to test extrapulmonary samples may call for an expansion of the reference standard definition in future updates.

While mycobacterial culture is the most sensitive method for mycobacteriological confirmation, in practice the use of NAATs such as Xpert MTB/RIF, Xpert Ultra, and TrueNat to confirm tuberculosis disease are recommended (WHO 2020a). Screening programmes may ultimately be more interested in the accuracy of the whole diagnostic algorithm, namely the screening method and a NAAT to make a final diagnosis (Figure 1), which was outside the scope of this review.

In the majority of included studies, populations were screened for tuberculosis for the first time, or as a one-off activity, which limits the applicability of the results to repeated screening situations. Ultimately, a major goal of tuberculosis screening is to reduce transmission of *M tuberculosis* by reducing the prevalence of tuberculosis disease and tuberculosis infection, which requires repeated screening at regular intervals (Corbett 2010; Marks 2019). In populations that are screened at regular intervals, the sensitivity of screening tests may be lower (Lewis 2009a), due to the fact that people are detected at an earlier stage of disease. We did not classify this population characteristic as an applicability concern and the number of included studies providing data from longitudinal screening programmes was too small to investigate this as a source of heterogeneity. With respect to participant selection, we considered national tuberculosis prevalence surveys, which are highly standardized in design and conduct and aim for a representative sample, equally as applicable as screening programmes that may be more pragmatic and may encourage participation of more symptomatic individuals, possibly lowering specificity (e.g. Morishita 2017a). In practice the execution of screening programmes will also vary. Lastly, our findings on CXR index tests are not applicable to the use of CXR screening as a second, sequential screening test to persons pre-selected on the presence of symptoms. In such populations, CXR specificity will likely be lower (Burgess 2001; Chadha 2018; Masur 2017). The sensitivity of a sequential symptom screen followed by CXR for symptom screen positives will be restricted by the sensitivity of the symptom index test. Such an algorithm would, however, be less resource intensive (van't Hoog 2014a), and requires further study.

While concerns about the applicability of findings to the review questions are generally low, the applicability of the review results in helping screening programmes choose a screening method requires additional considerations. Empirical studies suggest that various screening tests can reduce prevalence in settings with a high tuberculosis burden when applied at regular intervals (Corbett 2010a; Marks 2019). For further guidance on the choice of screening methods we refer to the WHO guidelines (WHO 2021a).

AUTHORS' CONCLUSIONS

Implications for practice

The summary estimates of the sensitivity and specificity of symptom screening and chest radiography (CXR) screening (and a combination of the two) can inform the choice of screening and diagnostic algorithms in settings where systematic screening for tuberculosis disease is being implemented. The high sensitivity of CXR index tests (based on low- to very low-certainty evidence), also with symptom questions in parallel, suggests a high yield of persons with tuberculosis disease. However, in addition to the evidence presented in this review, the implementation of screening approaches needs tailoring to the local epidemiological situation and health systems resources available. Additional considerations will determine the design of screening and diagnostic algorithms, such as the availability and accessibility of CXR facilities or the resources to fund them, and the need for more or fewer

diagnostic tests to confirm the diagnosis, depending on screening test specificity. Lower specificity has resource implications and implications for individuals who screen positive, but do not have tuberculosis as they are subjected to unnecessary tests and expenses. Furthermore, CXR may not be appropriate for all people who are screened, including pregnant women and young children. Symptom screening with lower sensitivity, on the other hand, may imply delayed diagnosis for individuals who screen negative but actually have tuberculosis, and possibly ongoing transmission. For further guidance we refer to the WHO guidelines to which this review's results contributed (WHO 2021a).

The review findings should be interpreted with caution due to methodological study limitations and (regional) variation in sensitivity and specificity. Due to unexplained heterogeneity, the sensitivity and specificity of an index test in a specific setting cannot be predicted with great precision. This should be borne in mind when planning for and implementing tuberculosis screening programmes.

Implications for research

Given the limitations of tuberculosis screening test accuracy estimates in predicting the value of screening tests in practice, empirical studies are needed on the accuracy, feasibility, and cost-effectiveness of screening and diagnostic algorithms in different settings, and ultimately on their impact when applied at repeated intervals on reducing transmission and on individual patient outcomes. In the short term, to further contribute to screening programme decisions about the use of CXR and its projected resource needs, based on sensitivity and specificity, a meta-analysis to determine the accuracy of CXR as a second, sequential screening method in populations that are first pre-selected, preceded by symptom screening, would be valuable. Factors to address in a next update of this review include conduct and quality aspects of the index tests, and further analyses around reference standard definitions. The latter to explore differences in accuracy due to inclusion of clinically diagnosed and extrapulmonary tuberculosis versus microbiologically confirmed pulmonary tuberculosis only, and the use of newer microbiological tests, on which data availability will hopefully increase. Furthermore we recommend an individual patient data (IPD) meta-analysis and meta-regression of factors associated with heterogeneity of especially the specificity of CXR index tests.

ACKNOWLEDGEMENTS

The Cochrane Infectious Diseases Group (CIDG) and Diagnostic Test Accuracy (DTA) Editor is Dr Karen Steingart.

The editorial base of the Cochrane Infectious Diseases Group is funded by UK aid from the UK government for the benefit of low- and middle-income countries (project number 300342-104). The views expressed do not necessarily reflect the UK government's official policies.

We acknowledge the participants of the WHO tuberculosis screening guideline development groups in 2012, headed by Knut Lönnroth, and 2020, headed by Cecily Miller, for their comments on the protocol, and on the results. We thank René Spijker for his assistance with search strategy development and execution. We also thank Dr Irwin Law who provided valuable information and insights on national tuberculosis prevalence surveys. We thank Dave Sinclair, Knut Lönnroth, Frank Cobelens, and Ellen Mitchell for their input in the earlier version of this review.

REFERENCES

References to studies included in this review

Adetifa 2016 *{published data only}***10.2471/blt.14.151670**

* Adetifa IM, Kendall L, Bashorun A, Linda C, Omoleke S, Jeffries D, et al.A tuberculosis nationwide prevalence survey in Gambia, 2012. *Bulletin of the World Health Organization* 2016;**94**(6):433-41. [DOI: 10.2471/blt.14.151670] [Ade2089]

Ministry of Health and Social Welfare The Gambia.The Gambian Survey of Tuberculosis Prevalence (GAMSTEP). April 2014. Ministry of Health and Social Welfare The Gambia. The Gambian Survey of Tuberculosis Prevalence (GAMSTEP). April 2014.

Ayles 2009a *{published data only}*

Ayles H, Schaap A, Nota A, Sismanidis C, Tembwe R, De Haas P, et al.Prevalence of tuberculosis, HIV and respiratory symptoms in two Zambian communities: implications for tuberculosis control in the era of HIV. *PLOS One* 2009;**4**(5):e5602.

Ayles 2009b *{published data only}*

Ayles H, Schaap A, Nota A, Sismanidis C, Tembwe R, De Haas P, et al.Prevalence of tuberculosis, HIV and respiratory symptoms in two Zambian communities: implications for tuberculosis control in the era of HIV. *PLOS One* 2009;**4**(5):e5602.

Chadha 2018 *{published data only}***10.1016/j.ijtb.2018.05.015**

Chadha VK, Anjinappa SM, Rade K, Baskaran D, Narang P, Kolappan C, et al.Sensitivity and specificity of screening tools and smear microscopy in active tuberculosis case finding. *Indian Journal of Tuberculosis* 2018;**66**(1):99-104.

Cheng 2015 *{published data only}***http://dx.doi.org/10.1371/journal.pone.0127725**

* Cheng J, Wang L, Zhang H, Xia Y.Diagnostic value of symptom screening for pulmonary tuberculosis in China. *PLOS One* 2015;**10**(5):e0127725. [DOI: 10.1371/journal.pone.0127725]

Wang L, Zhang H, Ruan Y, Chin DP, Xia Y, Cheng S, et al.Tuberculosis prevalence in China, 1990-2010; a longitudinal analysis of national survey data. *Lancet* 2014;**383**(9934):2057-64. [DOI: 10.1016/S0140-6736(13)62639-2]

Chheng 2008a *{published data only}*

Chheng P, Tamhane A, Natpratan C, Tan V, Lay V, Sar B, et al.Pulmonary tuberculosis among patients visiting a voluntary confidential counseling and testing center, Cambodia. *International Journal of Tuberculosis & Lung Disease* 2008;**12**(3 Suppl 1):54-62.

Chheng 2008b *{published data only}*

Chheng P, Tamhane A, Natpratan C, Tan V, Lay V, Sar B, et al.Pulmonary tuberculosis among patients visiting a voluntary confidential counseling and testing center, Cambodia. *International Journal of Tuberculosis & Lung Disease* 2008;**12**(3 Suppl 1):54-62.

Claassens 2017a *{published data only}***10.1371/journal.pone.0172881**

Ayles H, Muyoyeta M, Du Toit E, Schaap A, Floyd S, Simwinga M, et al.Effect of household and community interventions on the burden of tuberculosis in southern Africa: the ZAMSTAR community-randomised trial. *Lancet* 2013;**382**(9899):1183-94. [DOI: 10.1016/S0140-6736(13)61131-9]

* Claassens MM, van Schalkwyk C, Floyd S, Ayles H, Beyers N.Symptom screening rules to identify active pulmonary tuberculosis: findings from the Zambian South African Tuberculosis and HIV/AIDS Reduction (ZAMSTAR) trial prevalence surveys. *PLOS One* 2017;**12**(3):e0172881. [DOI: 10.1371/journal.pone.0172881]

Claassens 2017b *{published data only}***10.1371/journal.pone.0172881**

Ayles H, Muyoyeta M, Du Toit E, Schaap A, Floyd S, Simwinga M, et al.Effect of household and community interventions on the burden of tuberculosis in southern Africa: the ZAMSTAR community-randomised trial. *Lancet* 2013;**382**(9899):1183-94. [DOI: 10.1016/S0140-6736(13)61131-9]

* Claassens MM, van Schalkwyk C, Floyd S, Ayles H, Beyers N.Symptom screening rules to identify active pulmonary tuberculosis: findings from the Zambian South African Tuberculosis and HIV/AIDS Reduction (ZAMSTAR) trial prevalence surveys. *PLOS One* 2017;**12**(3):e0172881. [DOI: 10.1371/journal.pone.0172881]

Claassens 2017c *{published data only}***10.1371/journal.pone.0172881**

Ayles H, Muyoyeta M, Du Toit E, Schaap A, Floyd S, Simwinga M, et al.Effect of household and community interventions on the burden of tuberculosis in southern Africa: the ZAMSTAR community-randomised trial. *Lancet* 2013;**382**(9899):1183-94. [DOI: 10.1016/S0140-6736(13)61131-9]

* Claassens MM, van Schalkwyk C, Floyd S, Ayles H, Beyers N.Symptom screening rules to identify active pulmonary tuberculosis: Findings from the Zambian South African Tuberculosis and HIV/AIDS Reduction (ZAMSTAR) trial prevalence surveys. *PLOS One* 2017;**12**(3):e0172881. [DOI: 10.1371/journal.pone.0172881]

Claassens 2017d *{published data only}***10.1371/journal.pone.0172881**

Ayles H, Muyoyeta M, Du Toit E, Schaap A, Floyd S, Simwinga M, et al.Effect of household and community interventions on the burden of tuberculosis in southern Africa: the ZAMSTAR community-randomised trial. *Lancet* 2013;**382**(9899):1183-94. [DOI: 10.1016/S0140-6736(13)61131-9]

* Claassens MM, van Schalkwyk C, Floyd S, Ayles H, Beyers N.Symptom screening rules to identify active pulmonary tuberculosis: findings from the Zambian South African Tuberculosis and HIV/AIDS Reduction (ZAMSTAR) trial prevalence surveys. *PLOS One* 2017;**12**(3):e0172881. [DOI: 10.1371/journal.pone.0172881]

Corbett 2010a *{published data only}*

Corbett EL, Bandason T, Cheung YB, Makamure B, Dauya E, Munyati SS, et al.Prevalent infectious tuberculosis in Harare, Zimbabwe: burden, risk factors and implications for control.

International Journal of Tuberculosis and Lung Disease 2009;**13**(10):1231-7.

* Corbett EL, Zezai A, Cheung YB, Bandason T, Dauya E, Munyati SS, et al.Provider-initiated symptom screening for tuberculosis in Zimbabwe: diagnostic value and the effect of HIV status. *Bulletin of the World Health Organization* 2010;**88**(1):13-21.

Corbett 2010b *{published data only}*

Corbett EL, Bandason T, Cheung YB, Makamure B, Dauya E, Munyati SS, et al.Prevalent infectious tuberculosis in Harare, Zimbabwe: burden, risk factors and implications for control. *International Journal of Tuberculosis and Lung Disease* 2009;**13**(10):1231-7.

* Corbett EL, Zezai A, Cheung YB, Bandason T, Dauya E, Munyati SS, et al.Provider-initiated symptom screening for tuberculosis in Zimbabwe: diagnostic value and the effect of HIV status. *Bulletin of the World Health Organization* 2010;**88**(1):13-21.

den Boon 2006 *{published data only}*

* den Boon S, White NW, van Lill SW, Borgdorff MW, Verver S, Lombard CJ, et al.An evaluation of symptom and chest radiographic screening in tuberculosis prevalence surveys. *International Journal of Tuberculosis and Lung Disease* 2006;**10**(8):876-82.

den Boon S, van Lill SW, Borgdorff MW, Enarson DA, Verver S, Bateman ED, et al.High prevalence of tuberculosis in previously treated patients, Cape Town, South Africa. *Emerging Infectious Diseases* 2007;**13**(8):1189-94.

Federal MoH Sudan 2018 *{unpublished data only}*

Federal Ministry of Health Sudan.Sudan TB Prevalence Survey Report 2013 – 2014. Report 2018. [SUD5025] [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 217]

Fox 2012 *{published data only}*

Fox GJ, Nhung NV, Sy DN, Lien LT, Cuong NK, Britton WJ, et al.Contact investigation in households of patients with tuberculosis in Hanoi, Vietnam: a prospective cohort study. *PLOS One* 2012;**7**(11):e49880.

FRoNigeria 2014 *{published data only}*

Federal Republic of Nigeria.Report FIRST National TB Prevalence Survey 2012, Nigeria. Report 2014. [#5008] [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 177]

Ghana NTP 2015 *{unpublished data only}*

Ghana National TB Control Programme.Ghanian National Population Based Tuberculosis Prevalence Survey in 2013. Report 2015. [GHA5015] [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 121]

Ho 2016 *{published data only}***10.1016/s1473-3099(16)30067-6**

Ho J, Nguyen PTB, Nguyen TA, Tran KH, Van Nguyen S, Nguyen NV, et al.Reassessment of the positive predictive value and specificity of Xpert MTB/RIF: a diagnostic accuracy study in the context of community-wide screening for tuberculosis. *Lancet Infectious Diseases* 2016;**16**(9):1045-51. [DOI: 10.1016/s1473-3099(16)30067-6]

Hoa 2012 *{published data only}*

* Hoa NB, Cobelens FGJ, Sy DN, Nhung NV, Borgdorff MW, Tiemersma EW.Yield of interview screening and chest X-ray abnormalities in a tuberculosis prevalence survey. *International Journal of Tuberculosis and Lung Disease* 2012;**16**(6):762-7.

Hoa NB, Sy DN, Nhung NV, Tiemersma EW, Borgdorff MW, Cobelens FGJ.National survey of tuberculosis prevalence in Viet Nam [Enquete nationale sur la prevalence de la tuberculose au Viet Nam]. *Bulletin of the World Health Organization* 2010;**88**(4):273-80.

Kapata 2016 *{published data only}*

Government of the Republic of Zambia.National Tuberculosis Prevalence survey 2013 2014 Technical Report. Report. [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 257]

* Kapata N, Chanda-Kapata P, Ngosa W, Metitiri M, Klinkenberg E, Kalisvaart N, et al.The prevalence of tuberculosis in Zambia: results from the First National TB Prevalence Survey, 2013–2014. *PLOS One* 2016;**11**(1):e0146392. [DOI: 10.1371/journal.pone.0146392]

Kebede 2014 *{published data only}*

Federal Democratic Republic of Ethiopia Ministry of Health, Ethiopian Health and Nutrition Research Institute.First Ethiopian National Population Based Tuberculosis Prevalence Survey. Report 2011. [#5004] [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 105]

* Kebede AH, Alebachew Z, Tsegaye F, Lemma E, Abebe A, Agonafir M, et al.The first population-based national tuberculosis prevalence survey in Ethiopia, 2010-2011. *International Journal of Tuberculosis and Lung Disease* 2014;**18**(6):635-9. [DOI: 10.5588/ijtld.13.0417] [5007]

Kenya MoH 2018 *{unpublished data only}*

Republic of Kenya Ministry of Health.Kenya National Tuberculosis Survey. Report 2018. [KEN5017] [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 137]

Kimerling 1999 *{published data only}*

Kimerling ME, Shakes CF, Carlisle R, Lok KH, Benjamin WH, Dunlap NE.Spot sputum screening: evaluation of an intervention in two homeless shelters. *International Journal of Tuberculosis and Lung Disease* 1999;**3**(7):613-9.

Koesoemadinata 2018 *{published data only}*

Koesoemadinata RC, Kranzer K, Livia R, Susilawati N, Annisa J, Soetedjo NNM, et al.Computer-assisted chest radiography reading for tuberculosis screening in people living with diabetes mellitus. *International Journal of Tuberculosis and Lung Disease* 2018;**22**(9):1088-94. [Covidence#2411]

Law 2015 *{published data only}*10.1111/tmi.12536

Law I, Sylavanh P, Bounmala S, Nzabintwali F, Paboriboune P, Iem V, et al.The first national tuberculosis prevalence survey of Lao PDR (2010-2011). *Tropical Medicine & International Health* 2015;**20**(9):1146-54. [DOI: 10.1111/tmi.12536]

Lewis 2009a *{published data only}*

Lewis JJ, Charalambous S, Day JH, Fielding KL, Grant AD, Hayes RJ, et al.HIV infection does not affect active case finding of tuberculosis in South African gold miners. *American Journal of Respiratory and Critical Care Medicine* 2009;**180**(12):1271-8.

Lewis 2009b *{published data only}*

Lewis JJ, Charalambous S, Day JH, Fielding KL, Grant AD, Hayes RJ, et al.HIV infection does not affect active case finding of tuberculosis in South African gold miners. *American Journal of Respiratory and Critical Care Medicine* 2009;**180**(12):1271-8.

Little 2018 *{published data only}*10.1186/s12879-018-3193-7

Little KM, Msandiwa R, Martinson N, Golub J, Chaisson R, Dowdy D.Yield of household contact tracing for tuberculosis in rural South Africa. *BMC Infectious Diseases* 2018;**18**(1):299.

Lu 2016 *{published data only}*10.1016/j.jrid.2016.01.004

Lu X, Wang X, Li D, Wang X, Wu J.Comparative study of DR and CT in the application of close contacts screening for tuberculosis outbreaks. *Radiology of Infectious Diseases* 2016;**3**(1):34-9. [DOI: 10.1016/j.jrid.2016.01.004]

Mabuto 2015 *{published data only}*10.1186/s12889-015-1592-x

Mabuto T, Zwane E, Chihota V, Gresak G, Charalambous S, Churchyard GJ, et al.Tuberculosis active case finding: uptake and diagnostic yield among minibus drivers in urban South Africa. *BMC Public Health* 2015;**15**:242. [DOI: 10.1186/s12889-015-1592-x]

Malawi MoH 2016 *{unpublished data only}*

Malawi Ministry of Health.Malawi Tuberculosis Prevalence Survey. Report 2016. [MAL5019] [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 153]

Melendez 2017 *{published data only}*10.5588/ijtld.16.0851

Kapata N, Chanda-Kapata P, Ngosa W, Metitiri M, Klinkenberg E, Kalisvaart N, et al.The prevalence of tuberculosis in Zambia: results from the First National TB Prevalence Survey, 2013–2014. *PLOS One* 2016;**11**(1):e0146392. [DOI: 10.1371/journal.pone.0146392]

Melendez J, Philipsen RHHM, Chanda-Kapata P, Sunkutu V, Kapata N, van Ginneken B.Automatic versus human reading of chest X-rays in the Zambia National Tuberculosis Prevalence Survey. *International Journal of Tuberculosis and Lung Disease* 2017;**21**(8):880-6. [DOI: 10.5588/ijtld.16.0851]

MoH Cambodia 2005 *{published and unpublished data}*

Ministry of Health Cambodia.Report National TB Prevalence Survey, 2002 Cambodia. Report 2005.

MoH Cambodia 2012 *{published data only}*

Mao TE, Okada K, Yamada N, Peou S, Ota M, Saint S, et al.Cross-sectional studies of tuberculosis prevalence in Cambodia between 2002 and 2011. *Bulletin of the World Health Organization* 2014;**92**(8):573–81. [DOI: 10.2471/BLT.13.131581] [#3263]

Ministry of Health Kingdom of Cambodia.Second National TB Prevalence Survey Cambodia, 2011. openjicareport.jica.go.jp/pdf/12120325.pdf 2012. [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 81]

MoH Indonesia 2015 *{unpublished data only}*

Ministry of Health, Republic of Indonesia.Indonesia National Tuberculosis Prevalence Survey 2013-2014. Report 2015. [IND5016] [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 129]

MoH Myanmar 2012 *{published data only}*

Ministry of Health Department of Health Government of Myanmar.Report on National TB Prevalence Survey 2009-2010, Myanmar. who.int/tb/advisory_bodies/impact_measurement_taskforce/prevalencesurveymyanmar_2009-10report.pdf (accessed 5 Sep 2012).

Mongolia MoH 2016 *{unpublished data only}*

Mongolia Ministry of Health.Report of the First National Tuberculosis Prevalence Survey in Mongolia 2014-2015. Report 2016. [MON5020] [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 129]

Moosazadeh 2015 *{published data only}*

Moosazadeh M, Khanjani N, Parsaee M.The prevalence of latent tuberculosis infection and smear positive pulmonary tuberculosis in people with household close contact with tuberculosis in north of Iran. *Iran Journal of Medical Sciences* 2015;**40**(2):161-5.

MoPH DPRK 2017 *{unpublished data only}*

Ministry of Public Health DPRK .Report of DPRK National TB Prevalence Survey (2015-2016). Report 2017. [DPR5012] [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 97]

MoPH Thailand 2017 *{published data only}*

Tuberculosis Office, Department of Disease Control, Ministry of Public Health.2012-2013 National Tuberculosis prevalence survey 2012-2013. Report 2017. [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 225] [THA5026]

Mor 2012 *{published data only}*

Mor Z, Leventhal A, Weiler-Ravell D, Peled N, Lerman Y.Chest radiography validity in screening pulmonary tuberculosis in immigrants from a high-burden country. *Respiratory Care* 2012;**57**(7):1137-44.

Morasert 2018 *{published data only}*10.5588/ijtld.17.0654

Morasert T, Worapas W, Kaewmahit R, Uphala W.Prevalence and risk factors associated with tuberculosis disease in Suratthani Central Prison, Thailand. *International Journal of Tuberculosis and Lung Disease* 2018;**22**(10):1203-9.

Morishita 2017a *{published data only}*10.1371/journal.pone.0171310

Morishita F, Garfin AM, Lew W, Oh KH, Yadav RP, Reston JC, et al.Bringing state-of-the-art diagnostics to vulnerable populations: the use of a mobile screening unit in active case finding for tuberculosis in Palawan, the Philippines. *PLOS One* 2017;**12**(2):e0171310. [DOI: 10.1371/journal.pone.0171310]

Morishita 2017b *{published data only}*10.1371/journal.pone.0171310

Morishita F, Garfin AM, Lew W, Oh KH, Yadav RP, Reston JC, et al.Bringing state-of-the-art diagnostics to vulnerable populations: the use of a mobile screening unit in active case finding for tuberculosis in Palawan, the Philippines. *PLOS One* 2017;**12**(2):e0171310. [DOI: 10.1371/journal.pone.0171310]

Morishita 2017c *{published data only}*10.1371/journal.pone.0171310

Morishita F, Garfin AM, Lew W, Oh KH, Yadav RP, Reston JC, et al.Bringing state-of-the-art diagnostics to vulnerable populations: the use of a mobile screening unit in active case finding for tuberculosis in Palawan, the Philippines. *PLOS One* 2017;**12**(2):e0171310. [DOI: 10.1371/journal.pone.0171310]

Morishita 2017d *{published data only}*10.1371/journal.pone.0171310

Morishita F, Garfin AM, Lew W, Oh KH, Yadav RP, Reston JC, et al.Bringing state-of-the-art diagnostics to vulnerable populations: the use of a mobile screening unit in active case finding for tuberculosis in Palawan, the Philippines. *PLOS One* 2017;**12**(2):e0171310. [DOI: 10.1371/journal.pone.0171310]

Morishita 2017e *{published data only}*10.1371/journal.pone.0171310

Morishita F, Garfin AM, Lew W, Oh KH, Yadav RP, Reston JC, et al.Bringing state-of-the-art diagnostics to vulnerable populations: the use of a mobile screening unit in active case finding for tuberculosis in Palawan, the Philippines. *PLOS One* 2017;**12**(2):e0171310. [DOI: 10.1371/journal.pone.0171310]

Muyoyeta 2017 *{published data only}*10.1186/s12879-017-2388-7

Muyoyeta M, Kasese NC, Milimo D, Mushanga I, Ndhlovu M, Kapata N, et al.Digital CXR with computer aided diagnosis versus symptom screen to define presumptive tuberculosis among household contacts and impact on tuberculosis diagnosis. *BMC Infectious Diseases* 2017;**17**(1):301.

Nair 2016a *{published data only}*10.1186/s40248-016-0058-z

Nair A, Guleria R, Kandasamy D, Sharma R, Tandon N, Singh UB, et al.Prevalence of pulmonary tuberculosis in young adult patients with Type 1 diabetes mellitus in India. *Multidisciplinary Respiratory Medicine* 2016;**11**:22. [DOI: 10.1186/s40248-016-0058-z]

Nair 2016b *{published data only}*10.1186/s40248-016-0058-z

Nair A, Guleria R, Kandasamy D, Sharma R, Tandon N, Singh UB, et al.Prevalence of pulmonary tuberculosis in young adult patients with Type 1 diabetes mellitus in India. *Multidisciplinary Respiratory Medicine* 2016;**11**:22. [DOI: 10.1186/s40248-016-0058-z]

Ntinginya 2012 *{published data only}*

Ntinginya EN, Squire SB, Millington KA, Mtafya B, Saathoff E, Heinrich N, et al.Performance of the Xpert MTB/RIF assay in an active case-finding strategy: a pilot study from Tanzania. *International Journal of Tuberculosis and Lung Disease* 2012;**16**(11):1468-70.

NTP Bangladesh 2017 *{published data only}*

National TB Control Programme (Dr Mahmudur Rahman).National Tuberculosis Prevalence Survey, Bangladesh, 2015-2016. Report 2017. [BAN5010] [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 73]

NTP Philippines 2018 *{unpublished data only}*

National TB Control Programme, Department of Health/ Philippine Council for Health Research and Development (PCHRD)/Foundation for the Advancement of Clinical Epidemiology (FACE, Inc).National Tuberculosis Prevalence Survey 2016. Report 2018. [PHI5023] [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 201]

Pelissari 2018 *{published data only}*10.5588/ijtld.17.0526

Pelissari DM, Kuhleis DC, Bartholomay P, Barreira D, Oliveira CLP, de Jesus RS, et al.Prevalence and screening of active tuberculosis in a prison in the South of Brazil. *International Journal of Tuberculosis and Lung Disease* 2018;**22**(10):1166-71.

Qadeer 2016 *{published data only}*10.1371/journal.pone.0148293

Qadeer E, Fatima R, Yaqoob A, Tahseen S, Ul Haq M, Ghafoor A, et al.Population based National Tuberculosis Prevalence Survey among Adults (>15 Years) in Pakistan, 2010-2011. *PLOS One* 2016;**11**(2):e0148293. [DOI: 10.1371/journal.pone.0148293]

Republic of Uganda 2018 *{unpublished data only}*

The Republic of Uganda.The Uganda National Tuberculosis Prevalence Survey, 2014-2015. Survey Report. Available at health.go.ug/cause/the-uganda-national-tuberculosis-prevalence-survey-2014-2015-survey-report/. [UGA5027]

Republic of Zimbabwe 2015 *{unpublished data only}*

Republic of Zimbabwe – Ministry of Health and Child Care.Report of the First National Population-based Tuberculosis Prevalence Survey. Report 2015. [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 265] [ZIM5031]

Rwanda MoH 2014 *{unpublished data only}*

Republic of Rwanda Ministry of Health.Report of the First National Pulmonary Tuberculosis Prevalence Survey in Rwanda. Report 2014. [RWA5024] [Summary available in WHO 2021b (who.int/publications/i/item/9789240022430) page 209]

Seri 2017 *{published data only}*10.1371/journal.pone.0181995

Séri B, Koffi A, Danel C, Ouassa T, Blehoué MA, Ouattara E, et al.Prevalence of pulmonary tuberculosis among prison inmates: a cross-sectional survey at the Correctional and Detention Facility of Abidjan, Côte d'Ivoire. *PLOS One* 2017;**12**(7):e0181995. [DOI: 10.1371/journal.pone.0181995]

Singh 2013 *{published data only}*

Singh J, Sankar MM, Kumar S, Gopinath K, Singh N, Mani K, et al.Incidence and prevalence of tuberculosis among household contacts of pulmonary tuberculosis patients in a peri-urban population of South Delhi, India. *PLOS One* 2013;**8**(7):e69730.

Telisinghe 2014 *{published data only}*10.1371/journal.pone.0087262

Hanifa Y, Telisinghe L, Fielding KL, Malden JL, Churchyard GJ, Grant AD, et al.The diagnostic accuracy of urine lipoarabinomannan test for tuberculosis screening in a South African correctional facility. *PLOS One* 2015;**10**(5):e0127956. [DOI: 10.1371/journal.pone.0127956] [#1406]

* Telisinghe L, Fielding KL, Malden JL, Hanifa Y, Churchyard GJ, Grant AD, et al.High tuberculosis prevalence in a South African prison: the need for routine tuberculosis screening. *PLOS One* 2014;**9**(1):e87262. [DOI: 10.1371/journal.pone.0087262]

van't Hoog 2011 *{published data only}*

van't Hoog AH, Meme HK, van Deutekom H, Mithika AM, Olunga C, Onyino F, et al.High sensitivity of chest radiograph reading by clinical officers in a tuberculosis prevalence survey. *International Journal of Tuberculosis and Lung Disease* 2011;**15**(10):1308-14.

van't Hoog 2012 *{published data only}*

van't Hoog AH, Laserson KF, Githui WA, Meme HK, Agaya JA, Odeny LO, et al.High prevalence of pulmonary tuberculosis and inadequate case finding in rural western Kenya. *American Journal of Respiratory & Critical Care Medicine* 2011;**183**(9):1245-53.

* van't Hoog AH, Meme HK, Laserson KF, Agaya JA, Muchiri BG, Githui WA, et al.Screening strategies for tuberculosis prevalence surveys: the value of chest radiography and symptoms. *PLOS One* 2012;**7**(7):e38691.

Wei 2014 *{published data only}*

Wei X, Zhang X, Yin J, Walley J, Beanland R, Zou G, et al.Changes in pulmonary tuberculosis prevalence: evidence from the 2010 population survey in a populous province of China. *BMC Infectious Diseases* 2014;**14**:21.

Wood 2007 *{published data only}*

Wood R, Middelkoop K, Myer L, Grant AD, Whitelaw A, Lawn SD, et al.Undiagnosed tuberculosis in a community with high HIV prevalence: implications for tuberculosis control. *American Journal of Respiratory & Critical Care Medicine* 2007;**175**(1):87-93.

References to studies excluded from this review

Abboud 2013 *{published data only}*

Abboud S, Weiss F, Siegel E, Jeudy J.TB or Not TB: interreader and intrareader variability in screening diagnosis on an iPad versus a traditional display. *Journal of the American College of Radiology* 2013;**10**(1):42-4.

Abebe 2011 *{published data only}*

Abebe DS, Bjune G, Ameni G, Biffa D, Abebe F.Prevalence of pulmonary tuberculosis and associated risk factors in Eastern Ethiopian prisons. *International Journal of Tuberculosis and Lung Disease* 2011;**15**(5):668-73.

Abubakar 2010 *{published data only}*

Abubakar I, Story A, Lipman M, Bothamley G, van Hest R, Andrews N, et al.Diagnostic accuracy of digital chest radiography for pulmonary tuberculosis in a UK urban population. *European Respiratory Journal* 2010;**35**(3):689-92.

Aerts 2000 *{published data only}*

Aerts A, Habouzit M, Mschiladze L, Malakmadze N, Sadradze N, Menteshashvili O, et al.Pulmonary tuberculosis in prisons of the ex-USSR state Georgia: results of a nation-wide prevalence survey among sentenced inmates. *International Journal of Tuberculosis and Lung Disease* 2000;**4**(12):1104-10.

Al 2013 *{published data only}*

Al Hosani FI, Yahia GA.Prevalence of pulmonary tuberculosis among expatriates subjected to medical visa screening in Abu Dhabi, United Arab Emirates. *Journal of Epidemiology and Global Health* 2013;**3**(1):23-30.

Alamo 2012 *{published data only}*

Alamo ST, Kunutsor S, Walley J, Thoulass J, Evans M, Muchuro S, et al.Performance of the new WHO diagnostic algorithm for smear-negative pulmonary tuberculosis in HIV prevalent settings: a multisite study in Uganda. *Tropical Medicine & International Health* 2012;**17**(7):884-95.

Alcântara 2012 *{published data only}*

Alcântara CCS, Kritski AL, Ferreira VG, Façanha MC, Pontes RS, Mota RS, et al.Factors associated with pulmonary tuberculosis among patients seeking medical attention at referral clinics for tuberculosis [Fatores associados à tuberculose pulmonar em pacientes que procuraram serviços de saúde de referência para tuberculose]. *Jornal Brasileiro de Pneumologia* 2012;**38**(5):622-9.

Aldridge 2010 *{published data only}*

Aldridge RW, Story A, Stagg H, Lipman M, Knight J, Taubman D, et al.Sensitivity and specificity of mobile digital chest radiography for the diagnosis of active pulmonary tuberculosis. a cohort study in high risk groups in London. *Thorax* 2010;**65**:A6.

Ali 2012 *{published data only}*

Ali H, Zeynudin A, Mekonnen A, Abera S, Ali S.Smear Posetive Pulmonary Tuberculosis (PTB) prevalence amongst patients at Agaro Teaching Health Center, South West Ethiopia. *Ethiopian Journal of Health Sciences* 2012;**22**(1):71-6.

Alsedà 2003 *{published data only}*

Alsedà M, Godoy P.Study investigating infection in contacts of tuberculosis patients in a semi-urban area [Estudio de contactos de enfermos tuberculosos en un área semiurbana]. *Enfermedades Infecciosas y Microbiologia Clinica* 2003;**21**(6):281-6.

Arenas 2008 *{published data only}*

Arenas NE, Torres E, Durango CJ, Cuervo LI, Coronado SM, Gómez A.Detecting active tuberculosis in Calarc -Quind¡o, Colombia, during 2005 [Búsqueda activa de individuos con tuberculosis pulmonar y extrapulmonar en Calarcá-Quindío, Colombia-2005]. *Revista de Salud Pública* 2008;**10**(2):279-89.

Arpaz 2003 *{published data only}*

Arpaz S, Keskin S, Kiter G, Sezgin N, Budin D, Algan A, et al.Evaluation of Nazilli tuberculosis dispensary activities executed between June 1996-May 2000: pre-DOT situation. *Tuberkuloz ve toraks* 2003;**51**(3):289-97.

Arzhaeva 2009 *{published data only}*

Arzhaeva Y, Hogeweg L, de Jong PA, Viergever MA, van Ginneken B.Global and local multi-valued dissimilarity-based classification: application to computer-aided detection of tuberculosis. *Medical Image Computing & Computer-Assisted Intervention* 2009;**12**(Pt 2):724-31.

Assefzadeh 2009 *{published data only}*

Assefzadeh M, Barghi RGh, Shahidi ShS.Tuberculosis case--finding and treatment in the central prison of Qazvin province, Islamic Republic of Iran. *Eastern Mediterranean Health Journal* 2009;**15**(2):258-63.

Bai 2001 *{published data only}*

Bai GH, Kim SJ, Lee EK, Lew WJ.Incidence of pulmonary tuberculosis in Korean civil servants: second study, 1992-1994. *International Journal of Tuberculosis and Lung Disease* 2001;**5**(4):346-53.

Balasangameshwara 1993 *{published data only}*

Balasangameshwara VH, Chakraborty AK.Validity of case-finding tools in a national tuberculosis programme. *Tubercle and Lung Disease* 1993;**74**(1):52-8.

Balasubramanian 1995 *{published data only}*

Balasubramanian R, Sadacharam K, Selvaraj R, Xavier T, Gopalan BN, Shanmugam M, et al.Feasibility of involving literate tribal youths in tuberculosis case-finding in a tribal area in Tamil Nadu. *Tubercle and Lung Disease* 1995;**76**(4):355-9.

Balasubramanian 2004 *{published data only}*

Balasubramanian R, Garg R, Santha T, Gopi PG, Subramani R, Chandrasekaran V, et al.Gender disparities in tuberculosis: report from a rural DOTS programme in south India. *International Journal of Tuberculosis and Lung Disease* 2004;**8**(3):323-32.

Banda 1998 *{published data only}*

Banda HT, Harries AD, Welby S, Boeree MJ, Wirima JJ, Subramanyam VR, et al.Prevalence of tuberculosis in TB suspects with short duration of cough. *Transactions of the Royal Society of Tropical Medicine & Hygiene* 1998;**92**(2):161-3.

Banda 2009 *{published data only}*

Banda HT, Gausi F, Harries AD, Salaniponi FM.Prevalence of smear-positive pulmonary tuberculosis among prisoners in Malawi: a national survey. *International Journal of Tuberculosis and Lung Disease* 2009;**13**(12):1557-9.

Banu 2010 *{published data only}*

Banu S, Hossain A, Uddin MK, Uddin MR, Ahmed T, Khatun R, et al.Pulmonary tuberculosis and drug resistance in Dhaka central jail, the largest prison in Bangladesh. *PLOS One* 2010;**5**(5):e10759.

Basta 2005 *{published data only}*

Basta PC.Tuberculosis Between the Native People Suruí of Rondonia, Amazonia, Brazil [A tuberculose entre o povo indígena Suruí de Rondônia, Amazônia, Brasil] [PhD thesis]. Rio de Janeiro (Brazil): Escola Nacional de Saúde Pública Sergio Arouca, Fundação Oswaldo Cruz, 2005.

Basta 2006 *{published data only}*

Basta PC, Coimbra CE Jr, Escobar AL, Santos RV, Alves LC, Fonseca Lde S.Survey for tuberculosis in an indigenous population of Amazonia: the Surui of Rondonia, Brazil. *Transactions of the Royal Society of Tropical Medicine & Hygiene* 2006;**100**(6):579-85.

Bates 2013 *{published data only}*

Bates M, Ahmed Y, Chilukutu L, Tembo J, Cheelo B, Sinyangwe S, et al.Use of the Xpert(R) MTB/RIF assay for diagnosing pulmonary tuberculosis comorbidity and multidrug-resistant TB in obstetrics and gynaecology inpatient wards at the University Teaching Hospital, Lusaka, Zambia. *Tropical Medicine & International Health* 2013;**18**(9):1134-40.

Becerra 2005 *{published data only}*

Becerra MC, Pachao-Torreblanca IF, Bayona J, Celi R, Shin SS, Kim JY, et al.Expanding tuberculosis case detection by screening household contacts. *Public Health Reports* 2005;**120**(3):271-7.

Beser 1993 *{published data only}*

Beser E.The prevalence of pulmonary tuberculosis using different methods in group screenings. *East African Medical Journal* 1993;**70**(12):768-71.

Beyanga 2018 *{published data only}*

Beyanga M, Kidenya BR, Gerwing-Adima L, Ochodo E, Mshana SE, Kasang C.Investigation of household contacts of pulmonary tuberculosis patients increases case detection in Mwanza City, Tanzania. *BMC Infectious Diseases* 2018;**18**(110):1102018.

Bhat 2009 *{published data only}*

Bhat J, Rao VG, Gopi PG, Yadav R, Selvakumar N, Tiwari B, et al.Prevalence of pulmonary tuberculosis amongst the tribal population of Madhya Pradesh, central India. *International Journal of Epidemiology* 2009;**38**(4):1026-32.

Bhatia 2002 *{published data only}*

Bhatia S, Dranyi T, Rowley D.Tuberculosis among Tibetan refugees in India. *Social Science & Medicine* 2002;**54**(3):423-32.

Binkin 1996 *{published data only}*

Binkin NJ, Zuber PL, Wells CD, Tipple MA, Castro KG.Overseas screening for tuberculosis in immigrants and refugees to the United States: current status. *Clinical Infectious Diseases* 1996;**23**(6):1226-32.

Bjerregaard-Andersen 2010 *{published data only}*

Bjerregaard-Andersen M, da Silva ZJ, Ravn P, Ruhwald M, Andersen PL, Sodemann M, et al.Tuberculosis burden in an urban population: a cross sectional tuberculosis survey from Guinea Bissau. *BMC Infectious Diseases* 2010;**10**:96.

Bloss 2012 *{published data only}*

Bloss E, Makombe R, Kip E, Smit M, Chirenda J, Gammino VM, et al.Lessons learned during tuberculosis screening in public medical clinics in Francistown, Botswana. *International Journal of Tuberculosis & Lung Disease* 2012;**16**(8):1030-2.

Bonvin 1992 *{published data only}*

Bonvin L, Zellweger JP.Mass miniature X-ray screening for tuberculosis among immigrants entering Switzerland. *Tubercle & Lung Disease* 1992;**73**(6):322-5.

Borgdorff 2000 *{published data only}*

Borgdorff MW, Nagelkerke NJ, Dye C, Nunn P.Gender and tuberculosis: a comparison of prevalence surveys with notification data to explore sex differences in case detection. *International Journal of Tuberculosis & Lung Disease* 2000;**4**(2):123-32.

Borra 2009 *{published data only}*

Borra H, Si TM, Morris MJ, Waibel KH, Arroyo RA, Merrill GA, et al. In:Reliability of tuberculosis screening tests in patients receiving tumor necrosis factor antagonist therapy in a United States rheumatology clinic. American College of Rheumatology/ Association of Rheumatology Health Professional Annual Scientific Meeting (ACR/ARHP) 09; 2010 Nov 9; Atlanta (GA). Vol. 60. 2009:137.

Braden 1995 *{published data only}*

Braden CR.Infectiousness of a university student with laryngeal and cavitary tuberculosis. Investigative team. *Clinical Infectious Diseases* 1995;**21**(3):565-70.

Bucher 1994 *{published data only}*

Bucher H, Weinbacher M, Gyr K.Compliance of asylum seekers with an expanded border health screening program in the Basel-Stadt canton 1992-1993. *Schweizerische Medizinische Wochenschrift* 1994;**124**(38):1660-5.

Burgess 2001 *{published data only}*

Burgess AL, Fitzgerald DW, Severe P, Joseph P, Noel E, Rastogi N, et al.Integration of tuberculosis screening at an HIV voluntary counselling and testing centre in Haiti. *AIDS* 2001;**15**(14):1875-9.

Cain 2010 *{published data only}*

Cain KP, McCarthy KD, Heilig CM, Monkongdee P, Tasaneeyapan T, Kanara N, et al.An algorithm for tuberculosis screening and diagnosis in people with HIV. *New England Journal of Medicine* 2010;**362**(8):707-16.

Callister 2002 *{published data only}*

Callister MEJ, Barringer J, Thanabalasingam ST, Gair R, Davidson RN.Pulmonary tuberculosis among political asylum seekers screened at Heathrow Airport, London, 1995-9. *Thorax* 2002;**57**(2):152-6.

Carbonara 2005 *{published data only}*

Carbonara S, Babudieri S, Longo B, Starnini G, Monarca R, Brunetti B, et al.Correlates of Mycobacterium tuberculosis infection in a prison population. *European Respiratory Journal* 2005;**25**(6):1070-6.

Castro 2011 *{published data only}*

Castro CB, Costa PA, Ruffino-Netto A, Maciel EL, Kritski AL.Assessment of a clinical score for screening suspected pulmonary tuberculosis cases. *Revista de Saude Publica* 2011;**45**(6):1110-6.

Catanzaro 2013 *{published data only}*

Catanzaro A, Daley C.A summary of the third global interferon-release assay symposium. *Infection Control and Hospital Epidemiology* 2013;**34**(6):619-24.

Chadha 2012 *{published data only}*

Chadha VK, Kumar P, Anjinappa SM, Singh S, Narasimhaiah S, Joshi MV, et al.Prevalence of pulmonary tuberculosis among adults in a rural sub-district of South India. *PLOS One* 2012;**7**(8):e42625.

Chakraborty 1995 *{published data only}*

Chakraborty AK, Suryanarayana HV, Murthy VV, Murthy MS, Shashidhara AN.Prevalence of tuberculosis in a rural area by an alternative survey method without prior radiographic screening of the population. *Tubercle and Lung Disease* 1995;**76**(1):20-4.

Chang 2010 *{published data only}*

Chang CM, Lee NY, Lee HC, Wu CJ, Chen PL, Lee CC, et al.Positive tuberculin skin tests in nursing home residents in Southern Taiwan. *Archives of Gerontology & Geriatrics* 2010;**51**(3):e129-32.

Chan-Yeung 2007 *{published data only}*

Chan-Yeung M, Dai DL, Cheung AH, Chan FH, Kam KM, Tam CM, et al.Tuberculin skin test reaction and body mass index in old age home residents in Hong Kong. *Journal of the American Geriatrics Society* 2007;**55**(10):1592-7.

Chen 2011 *{published data only}*

Chen CY, Sheng WH, Cheng A, Tsay W, Huang SY, Tang JL, et al.Clinical characteristics and outcomes of Mycobacterium tuberculosis disease in adult patients with hematological malignancies. *BMC Infectious Diseases* 2011;**11**:324.

Chen 2012 *{published data only}*

Chen W, Xia Y, Li X, Zhou L, Li C, Wan K, et al.A tuberculosis outbreak among senior high school students in China in 2011. *Journal of International Medical Research* 2012;**40**(5):1830-9.

Chen 2013 *{published data only}*

Chen SC, Chen KL, Chen KH, Chien ST, Chen KT.Updated diagnosis and treatment of childhood tuberculosis. *World Journal of Pediatrics* 2013;**9**(1):9-16.

China 2004 *{published data only}*

China Tuberculosis Control Collaboration.The effect of tuberculosis control in China. *Lancet* 2004;**364**(9432):417-22.

Churchyard 1999 *{published data only}*

Churchyard GJ, Kleinschmidt I, Corbett EL, Mulder D, De Cock KM.Mycobacterial disease in South African gold miners in the era of HIV infection. *International Journal of Tuberculosis & Lung Disease* 1999;**3**(9):791-8.

Churchyard 2010 *{published data only}*

Churchyard GJ, Fielding KL, Lewis JJ, Chihota VN, Hanifa Y, Grant AD.Symptom and chest radiographic screening for infectious tuberculosis prior to starting isoniazid preventive therapy: yield and proportion missed at screening. *AIDS* 2010;**24**(Suppl 5):S19-27.

Claassens 2013 *{published data only}*

Claassens M, van Schalkwyk C, den Haan L, Floyd S, Dunbar R, van Helden P, et al.High prevalence of tuberculosis and insufficient case detection in two communities in the Western Cape, South Africa. *PLOS One* 2013;**8**(4):e58689.

Claassens 2013a *{published data only}*

Claassens MM, Jacobs E, Cyster E, Jennings K, James A, Dunbar R, et al.Tuberculosis cases missed in primary health care facilities: should we redefine case finding? *International Journal of Tuberculosis and Lung Disease* 2013;**17**(5):608-14.

Clarke 2003 *{published data only}*

Clarke M, Dick J, Zwarenstein M, Diwan V.DOTS for temporary workers in the agricultural sector. An exploratory study in tuberculosis case detection. *Curationis* 2003;**26**(4):66-71.

Coelho 2011 *{published data only}*

Coelho Filho JC, Caribé MA, Caldas SCC, Martins Netto E.Is tuberculosis difficult to diagnose in childhood and adolescence? [A tuberculose na infância e na adolescência é difícil de diagnosticar?]. *Jornal Brasileiro De Pneumologia* 2011;**37**(3):1-5.

Corbett 2004 *{published data only}*

Corbett EL, Charalambous S, Moloi VM, Fielding K, Grant AD, Dye C, et al.Human immunodeficiency virus and the prevalence of undiagnosed tuberculosis in African gold miners. *American Journal of Respiratory & Critical Care Medicine* 2004;**170**(6):673-9.

Costa 2010 *{published data only}*

Costa JT, Silva R, Sá R, Cardoso MJ, Ribeiro C, Nienhaus A.Comparison of interferon-gamma release assay and tuberculin test for screening in healthcare workers. *Revista Portuguesa de Pneumologia* 2010;**16**(2):211-21.

Costa 2011 *{published data only}*

Costa JC, Silva R, Ferreira J, Nienhaus A.Active tuberculosis among health care workers in Portugal. *Jornal Brasileiro de Pneumologia* 2011;**37**(5):636-45.

Costantino 2010 *{published data only}*

Costantino F, Faure G, Carvalho MD, Rat AC, Dintinger H, Loeuille D, et al. In:High level of disease activity in chronic inflammatory rheumatisms increases the rate of indeterminate interferon-gamma-release assay results for latent tuberculosis infection detection. American College of Rheumatology/ Association of Rheumatology Health Professional Annual Scientific Meeting (ACR/ARHP) 10; 16 Oct 2009 Philadelphia (PA). Vol. 62. 2010:768.

Crampin 2011 *{published data only}*

Crampin A, Kasimba S, Mwaungulu NJ, Dacombe R, Floyd S, Glynn JR, et al.Married to M. tuberculosis: risk of infection and disease in spouses of smear-positive tuberculosis patients. *Tropical Medicine & International Health* 2011;**16**(7):811-8.

Cruz-Hervert 2012 *{published data only}*

Cruz-Hervert LP, García-García L, Ferreyra-Reyes L, Bobadilla-del-Valle M, Cano-Arellano B, Canizales-Quintero S, et al.Tuberculosis in ageing: high rates, complex diagnosis and poor clinical outcomes. *Age and Ageing* 2012;**41**(4):488-95.

Cuevas 2011a *{published data only}*

Cuevas LE, Al-Sonboli N, Lawson L, Yassin MA, Arbide I, Al-Aghbari N, et al.LED fluorescence microscopy for the diagnosis of pulmonary tuberculosis: a multi-country cross-sectional evaluation. *PLOS Medicine* 2011;**8**(7):e1001057.

Cuevas 2011b *{published data only}*

Cuevas LE, Yassin MA, Al-Sonboli N, Lawson L, Arbide I, Al-Aghbari N, et al.A multi-country non-inferiority cluster randomized trial of frontloaded smear microscopy for the diagnosis of pulmonary tuberculosis. *PLOS Medicine* 2011;**8**(7):e1000443.

Dara 2013 *{published data only}*

Dara M, Gushulak BD, Posey DL, Zellweger JP, Migliori GB.The history and evolution of immigration medical screening for tuberculosis. *Expert Review of Anti-Infective Therapy* 2013;**11**(2):137-46.

Datta 2001 *{published data only}*

Datta M, Radhamani MP, Sadacharam K, Selvaraj R, Rao DL, Rao RS, et al.Survey for tuberculosis in a tribal population in North Arcot District. *International Journal of Tuberculosis and Lung Disease* 2001;**5**(3):240-9.

de Vries 2007 *{published data only}*

de Vries G, van Hest RA, Richardus JH.Impact of mobile radiographic screening on tuberculosis among drug users and homeless persons. *American Journal of Respiratory and Critical Care Medicine* 2007;**176**(2):201-7.

Eckhoff 2000 *{published data only}*

Eckhoff CT.Evaluation of a clinical index among adult contacts of children with tuberculosis in rural Haiti. *International Journal of Tuberculosis and Lung Disease* 2000;**4**(12):1143-8.

Forman 2003 *{published data only}*

Forman PD, Kinney C.Evaluation of a tuberculosis screening questionnaire for use in an Alaskan homeless population. *Alaska Medicine* 2003;**45**(4):94-101.

Fox 2011 *{published data only}*

Fox GJ, Sy D, Nhung NV, Lien LT, Cuong NK, Britton WJ, et al.Prevalence of tuberculosis among household contacts of smear positive tuberculosis patients in four districts in Ha Noi,

Viet Nam. *American Journal of Respiratory and Critical Care Medicine* 2011;**183**:A1843.

Fox 2012a *{published data only}*

Fox GJ, Sy DNS, Nhung NV, Lien LT, Cuong NK, Britton WJ, et al.Active case-finding of tuberculosis among household contacts in Hanoi, Vietnam - a 12-month prospective study. *American Journal of Respiratory and Critical Care Medicine* 2012;**185**:A3248.

Gopi 2003 *{published data only}*

Balasubramanian R, Garg R, Santha T, Gopi PG, Subramani R, Chandrasekaran V, et al.Gender disparities in tuberculosis: report from a rural DOTS programme in south India. *International Journal of Tuberculosis and Lung Disease* 2004;**8**(3):323-32.

* Gopi PG, Subramani R, Radhakrishna S, Kolappan C, Sadacharam K, Devi TS, et al.A baseline survey of the prevalence of tuberculosis in a community in south India at the commencement of a DOTS programme. *International Journal of Tuberculosis and Lung Disease* 2003;**7**(12):1154-62.

Santha T, Renu G, Frieden TR, Subramani R, Gopi PG, Chandrasekaran V, et al.Are community surveys to detect tuberculosis in high prevalence areas useful? Results of a comparative study from Tiruvallur District, South India. *International Journal of Tuberculosis and Lung Disease* 2003;**7**(3):258-65.

Gopi 2008a *{published data only}*

Gopi PG, Subramani R, Chandrasekaran V, Santha T, Narayanan PR.Impact of improved treatment success on the prevalence of TB in a rural community based on active surveillance. *Indian Journal of Tuberculosis* 2008;**55**(1):22-7.

Gopi 2008b *{published data only}*

Gopi PG, Subramani R, Narayanan PR.Evaluation of different types of chest symptoms for diagnosing pulmonary tuberculosis cases in community surveys. *Indian Journal of Tuberculosis* 2008;**55**(3):116-21.

Hamid 2012 *{published data only}*

Hamid S, Hussain SA, Ali I.Comparative analysis of case screening with varying cough duration and sputum samples for diagnosis of tuberculosis in patients attending the OPD at a tertiary care hospital at Srinagar, India. *Nigerian Journal of Clinical Practice* 2012;**15**(4):430-5.

Hanrahan 2013 *{published data only}*

Hanrahan CF, Selibas K, Deery CB, Dansey H, Clouse K, Bassett J, et al.Time to treatment and patient outcomes among TB suspects screened by a single point-of-care xpert MTB/RIF at a primary care clinic in Johannesburg, South Africa. *PLOS One* 2013;**8**(6):e65421.

Henostroza 2016 *{published data only}***http://dx.doi.org/10.1186/s12879-016-1460-z**

Henostroza G, Harris JB, Kancheya N, Nhandu V, Besa S, Musopole R, et al.Chest radiograph reading and recording system: evaluation in frontline clinicians in Zambia. *BMC Infectious Diseases* 2016;**16**:136. [DOI: http://dx.doi.org/10.1186/s12879-016-1460-z]

Hoffmann 2013 *{published data only}*

Hoffmann CJ, Variava E, Rakgokong M, Masonoke K, van der Watt M, Chaisson RE, et al.High prevalence of pulmonary tuberculosis but low sensitivity of symptom screening among HIV-infected pregnant women in South Africa. *PLOS One* 2013;**8**(4):e62211.

Hong 1993 *{published data only}*

Hong YP, Kim SJ, Kwon DW, Chang SC, Lew WJ, Han YC.The sixth Nationwide Tuberculosis Prevalence Survey in Korea, 1990. *Tubercle and Lung Disease* 1993;**74**(5):323-31.

Hong 1998 *{published data only}*

Hong YP, Kim SJ, Lew WJ, Lee EK, Han YC.The seventh nationwide tuberculosis prevalence survey in Korea, 1995. *International Journal of Tuberculosis and Lung Disease* 1998;**2**(1):27-36.

Horie 2007 *{published data only}*

Horie T, Lien LT, Tuan LA, Tuan PL, Sakurada S, Yanai H, et al.A survey of tuberculosis prevalence in Hanoi, Vietnam. *International Journal of Tuberculosis & Lung Disease* 2007;**11**(5):562-6.

Kim 2012 *{published data only}*

Kim L, Heilig CM, McCarthy KD, Phanuphak N, Chheng P, Kanara N, et al.Symptom screen for identification of highly infectious tuberculosis in people living with HIV in Southeast Asia. *Journal of Acquired Immune Deficiency Syndromes* 2012;**60**(5):519-24.

Leung 2005 *{published data only}*

Leung CC, Chan CK, Tam CM, Yew WW, Kam KM, Au KF, et al.Chest radiograph screening for tuberculosis in a Hong Kong prison. *International Journal of Tuberculosis & Lung Disease* 2005;**9**(6):627-32.

Mahomed 2013 *{published data only}*

Mahomed H, Ehrlich R, Hawkridge T, Hatherill M, Geiter L, Kafaar F, et al.Screening for TB in high school adolescents in a high burden setting in South Africa. *Tuberculosis* 2013;**93**(3):357-62.

Masur 2017 *{published data only}***http://dx.doi.org/10.4269/ajtmh.16-0674**

Masur J, Koenig SP, Julma P, Ocheretina O, Durán-Mendicuti MA, Fitzgerald DW, et al.Active tuberculosis case finding in Haiti. *American Journal of Tropical Medicine and Hygiene* 2017;**97**(2):433-5. [DOI: 10.4269/ajtmh.16-0674]

Mathez 2007 *{published data only}*

Mathez C, Bangala Y, Bady P, Zellweger JP.Active screening for pulmonary tuberculosis by chest x-ray among immigrants at the Swiss border. *Swiss Medical Weekly* 2007;**137**(45-6):649-54. [#3550] [PMID: https://pubmed.ncbi.nlm.nih.gov/18027111/]

Mor 2013 *{published data only}*

Mor Z, Leventhal A, Diacon AH, Finger R, Schoch OD.Tuberculosis screening in immigrants from high-prevalence countries: interview first or chest radiograph first? A pro/con debate. *Respirology* 2013;**18**(3):432-8.

Santha 2005 *{published data only}*

Santha T, Garg R, Subramani R, Chandrasekaran V, Selvakumar N, Sisodia RS, et al.Comparison of cough of 2 and 3 weeks to improve detection of smear-positive tuberculosis cases among out-patients in India. *International Journal of Tuberculosis & Lung Disease* 2005;**9**:61-8.

Sebhatu 2007 *{published data only}*

Sebhatu M, Kiflom B, Seyoum M, Kassim N, Negash T, Tesfazion A, et al.Determining the burden of tuberculosis in Eritrea: a new approach. *Bulletin of the World Health Organization* 2007;**85**(8):593-9.

Senkoro 2016 *{published data only}***10.5588/ijtld.15.0340**

Senkoro M, Mfinanga S, Egwaga S, Mtandu R, Kamara DV, Basra D, et al.Prevalence of pulmonary tuberculosis in adult population of Tanzania: a national survey, 2012. *International Journal of Tuberculosis and Lung Disease* 2016;**20**(8):1014-21. [DOI: 10.5588/ijtld.15.0340]

Tupasi 2000 *{published data only}*

Tupasi TE, Radhakrishna S, Quelapio MI, Villa ML, Pascual ML, Rivera AB, et al.Tuberculosis in the urban poor settlements in the Philippines. *International Journal of Tuberculosis & Lung Disease* 2000;**4**:4-11.

References to studies awaiting assessment

Assefa 2019 *{published data only}*

Assefa D, Belachew F, Wondimagegn G, Klinkenberg E.Missed pulmonary tuberculosis: a cross sectional study in the general medical inpatient wards of a large referral hospital in Ethiopia. *BMC Infectious Diseases* 2019;**19**(1):60.

Bekken 2020 *{published data only}*

Bekken GK, Ritz C, Selvam S, Jesuraj N, Hesseling AC, Doherty TM, et al.Identification of subclinical tuberculosis in household contacts using exposure scores and contact investigations. *BMC Infectious Diseases* 2020;**20**(1):96.

Bonsu 2020 *{published data only}*

Bonsu F, Addo KK, Alebachew Z, Gyapong J, Badu-Peprah A, Gockah R, et al.National population-based tuberculosis prevalence survey in Ghana, 2013. *International Journal of Tuberculosis and Lung Disease* 2020;**24**(3):321-8.

Hamda 2020 *{published data only}*

Hamda SG, Tshikuka JG, Joel D, Setlhare V, Monamodi G, Mbeha B, et al.Contribution of Xpert MTB/RIF to tuberculosis case finding among pregnant women in Botswana. *Public Health in Action* 2020;**10**(2):76-81.

Lansang 2021 *{published data only}*

Lansang MAD, Alejandria MM, Law I, Juban NR, Amarillo MLE, Sison OT, et al.High TB burden and low notification rates in the Philippines: the 2016 national TB prevalence survey. *PLOS One* 2021;**16**(6):e0252240.

Li 2019 *{published data only}*

Li B, Zhang X, Guo J, Wang J, Pianduo B, Wei X, et al.Prevalence of pulmonary tuberculosis in Tibet Autonomous Region, China, 2014. *International Journal of Tuberculosis and Lung Disease* 2019;**23**(7):735-40.

Migambi 2020 *{published data only}*

Migambi P, Gasana M, Uwizeye CB, Kamanzi E, Ndahindwa V, Kalisvaart N, et al.Prevalence of tuberculosis in Rwanda: results of the first nationwide survey in 2012 yielded important lessons for TB control. *PLOS One* 2020;**15**(4):e0231372.

Nalunjogi 2021 *{published data only}*

Nalunjogi J, Mugabe F, Najjingo I, Lusiba P, Olweny F, Mubiru J, et al.Accuracy and incremental yield of the chest X-ray in screening for tuberculosis in Uganda: a cross-sectional study. *Tuberculosis Research & Treatment Print* 2021;**2021**(0):6622809.

Nguyen HV 2020 *{published data only}*

Nguyen HV, Tiemersma EW, Nguyen HB, Cobelens FGJ, Finlay A, Glaziou P, et al.The second national tuberculosis prevalence survey in Vietnam. *PLOS One* 2020;**15**(4):e0232142. [DOI: https://doi.org/10.1371/journal.pone.0236532]

Nguyen TBP 2020 *{published data only}*

Nguyen TBP, Nguyen TA, Luu BK, Le TTO, Nguyen VS, Nguyen KC, et al.A comparison of digital chest radiography and Xpert MTB/RIF in active case finding for tuberculosis. *International Journal of Tuberculosis & Lung Disease* 2020;**24**(9):934-40.

Pasipamire 2020 *{published data only}*

Pasipamire M, Broughton E, Mkhontfo M, Maphalala G, Simelane-Vilane B, Haumba S.Detecting tuberculosis in pregnant and postpartum women in Eswatini. *African Journal of Laboratory Medicine* 2020;**9**(1):a837.

Reichler 2020 *{published data only}*

Reichler MR, Hirsch C, Yuan Y, Khan A, Dorman SE, Schluger N, et al.Predictive value of TNF-alpha, IFN-gamma, and IL-10 for tuberculosis among recently exposed contacts in the United States and Canada. *BMC Infectious Diseases* 2020;**20**(1):553.

Additional references

Achkar 2011

Achkar JM, Jenny-Avital ER.Incipient and subclinical tuberculosis: defining early disease states in the context of host immune response. *Journal of Infectious Diseases* 2011;**204**(Suppl 4):S1179-86.

Assefa 2019b

Assefa Y, Woldeyohannes S, Gelaw YA, Hamada Y, Getahun H.Screening tools to exclude active pulmonary TB

in high TB burden countries: systematic review and meta-analysis. *International Journal of Tuberculosis and Lung Disease* 2019;**23**(6):728-34. [DOI: 10.5588/ijtld.18.0547]

Burke 2021

Burke RM, Nliwasa M, Feasey HRA, Chaisson LH, Golub JE, Naufal F, et al.Community-based active case-finding interventions for tuberculosis: a systematic review. *Lancet Public Health* 2021;**6**(5):e283-99. [DOI: 10.1016/S2468-2667(21)00033-5]

Cattamanchi 2010

Cattamanchi A, Davis JL, Pai M, Huang L, Hopewell PC, Steingart KR.Does bleach processing increase the accuracy of sputum smear microscopy for diagnosing pulmonary tuberculosis? *Journal of Clinical Microbiology* 2010;**48**(7):2433-9.

Chien 2000

Chien HP, Yu MC, Wu MH, Lin TP, Luh KT.Comparison of the BACTEC MGIT 960 with Lowenstein-Jensen medium for recovery of mycobacteria from clinical specimens. *International Journal of Tuberculosis and Lung Disease* 2000;**4**(9):866-70.

Corbett 2010

Corbett EL, Bandason T, Duong T, Dauya E, Makamure B, Churchyard GJ, et al.Comparison of two active case-finding strategies for community-based diagnosis of symptomatic smear-positive tuberculosis and control of infectious tuberculosis in Harare, Zimbabwe (DETECTB): a cluster-randomised trial. *Lancet* 2010;**376**(9748):1244-53.

Covidence [Computer program]

Veritas Health Innovation Covidence.Melbourne, Australia: Veritas Health Innovation, accessed 1 January 2019. Available at covidence.org.

den Boon 2005

den Boon S, Bateman ED, Enarson DA, Borgdorff MW, Verver S, Lombard CJ, et al.Development and evaluation of a new chest radiograph reading and recording system for epidemiological surveys of tuberculosis and lung disease. *International Journal of Tuberculosis and Lung Disease* 2005;**9**(10):1088-96.

den Boon 2008

den Boon S, Verver S, Lombard CJ, Bateman ED, Irusen EM, Enarson DA, et al.Comparison of symptoms and treatment outcomes between actively and passively detected tuberculosis cases: the additional value of active case finding. *Epidemiology and Infection* 2008;**136**(10):1342-9.

Dye 1998

Dye C, Garnett GP, Sleeman K, Williams BG.Prospects for worldwide tuberculosis control under the WHO DOTS strategy. Directly observed short-course therapy. *Lancet* 1998;**352**(9144):1886-91.

Getahun 2007

Getahun H, Harrington M, O'Brien R, Nunn P.Diagnosis of smear-negative pulmonary tuberculosis in people with HIV infection or AIDS in resource-constrained settings: informing urgent policy changes. *Lancet* 2007;**369**(9578):2042-9.

Getahun 2011

Getahun H, Kittikraisak W, Heilig CM, Corbett EL, Ayles H, Cain KP, 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 Medicine* 2011;**8**(1):e1000391. [DOI: 10.1371/journal.pmed.1000391] [PMID: 21267059]

Golub 2005

Golub JE, Mohan CI, Comstock GW, Chaisson RE.Active case finding of tuberculosis: historical perspective and future prospects. *International Journal of Tuberculosis and Lung Disease* 2005;**9**(11):1183-203.

Graham 2012

Graham SM, Ahmed T, Amanullah F, Browning R, Cardenas V, Casenghi M, et al.Evaluation of tuberculosis diagnostics in children: 1. Proposed clinical case definitions for classification of intrathoracic tuberculosis disease. Consensus from an expert panel. *Journal of Infectious Diseases* 2012;**205**(Suppl 2):S199-208.

Gupta 2015

Gupta RK, Lucas SB, Fielding KL, Lawn SD.Prevalence of tuberculosis in post-mortem studies of HIV-infected adults and children in resource-limited settings: a systematic review and meta-analysis. *AIDS* 2015;**29**(15):1987-2002. [DOI: 10.1097/QAD.0000000000000802] [PMID: 26266773]

Hamada 2018

Hamada Y, Lujan J, Schenkel K, Ford N, Getahun H.Sensitivity and specificity of WHO's recommended four-symptom screening rule for tuberculosis in people living with HIV: a systematic review and meta-analysis. *Lancet HIV* 2018;**5**(9):e515-23. [DOI: 10.1016/S2352-3018(18)30137-1] [PMID: PMID: 30139576]

Hanna 1999

Hanna BA, Ebrahimzadeh A, Elliott LB, Morgan MA, Novak SM, Rusch-Gerdes S, et al.Multicenter evaluation of the BACTEC MGIT 960 system for recovery of mycobacteria. *Journal of Clinical Microbiology* 1999;**37**(3):748-52.

Hayen 2010

Hayen A, Macaskill P, Irwig L, Bossuyt P.Appropriate statistical methods are required to assess diagnostic tests for replacement, add-on, and triage. *Journal of Clinical Epidemiology* 2010;**63**(8):883-91.

Hoa 2010

Hoa NB, Sy DN, Nhung NV, Tiemersma EW, Borgdorff MW, Cobelens FGJ.National survey of tuberculosis prevalence in Viet Nam. *Bulletin of the World Health Organization* 2010;**88**(4):273-80.

Horne 2019

Horne DJ, Kohli M, Zifodya JS, Schiller I, Dendukuri N, Tollefson D, et al.Xpert MTB/RIF and Xpert MTB/RIF Ultra for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database of Systematic Reviews* 2019, Issue 6. Art. No: CD009593. [DOI: 10.1002/14651858.CD009593.pub4]

Houben 2016

Houben RM, Dodd PJ.The global burden of latent tuberculosis infection: a re-estimation using mathematical modelling. *PLOS Medicine* 2016;**13**(10):e1002152. [DOI: 10.1371/journal.pmed.1002152] [PMID: 27780211]

Kerley 1942

Kerley P, Major T.Technique in mass miniature radiography. *British Journal of Radiology* 1942;**15**(180):346-7.

Law 2020

Law I, Floyd K, African TB Prevalence Survey Group.National tuberculosis prevalence surveys in Africa, 2008–2016: an overview of results and lessons learned. *Tropical Medicine & International Health* 2020;**25**(11):1308-27. [DOI: 10.1111/tmi.13485]

Lawn 2011

Lawn SD, Zumla AI.Tuberculosis. *Lancet* 2011;**378**(9785):57-72.

Luabeya 2012

Luabeya KK, Mulenga H, Moyo S, Tameris M, Sikhondze W, Geldenhuys H, et al.Diagnostic features associated with culture of Mycobacterium tuberculosis among young children in a vaccine trial setting. *Pediatric Infectious Disease Journal* 2012;**31**(1):42-6.

Lönnroth 2013

Lönnroth K, Corbett E, Golub J, Godfrey-Faussett P, Uplekar M, Weil D, et al.Systematic screening for active tuberculosis: rationale, definitions and key considerations. *International Journal of Tuberculosis and Lung Disease* 2013;**17**(3):289-98.

Maher 2009

Maher D.Clinical features and index of suspicion in adults (HIV-negative and HIV-positive). In: Tuberculosis. A Comprehensive Clinical Reference. Philadelphia: Saunders, 2009:164-8.

Marks 2019

Marks GB, Nguyen NV, Nguyen PTB, Nguyen TA, Nguyen HB, Tran KH, et al.Community-wide screening for tuberculosis in a high-prevalence setting. *New England Journal of Medicine* 2019;**381**(14):1347-57. [DOI: 10.1056/NEJMoa1902129]

Microsoft Excel [Computer program]

Microsoft Excel.Redmond, Washington: Microsoft Corporation, 2010.

Monkongdee 2009

Monkongdee P, McCarthy KD, Cain KP, Tasaneeyapan T, Nguyen HD, Nguyen TN, et al.Yield of acid-fast smear and mycobacterial culture for tuberculosis diagnosis in people with human immunodeficiency virus. *American Journal of Respiratory Critical Care Medicine* 2009;**180**(9):903-8.

Okada 2012

Okada K, Onozaki I, Yamada N, Yoshiyama T, Miura T, Saint S, et al.Epidemiological impact of mass tuberculosis screening: a 2-year follow-up after a national prevalence survey. *International Journal of Tuberculosis and Lung Disease* 2012;**16**(12):1619-24.

Onozaki 2015

Onozaki I, Law I, Sismanidis C, Zignol M, Glaziou P, Floyd K.National tuberculosis prevalence surveys in Asia, 1990-2012: an overview of results and lessons learned. *Tropical Medicine & International Health* 2015;**20**(9):1128-45. [DOI: 10.1111/tmi.12534]

Pai 2016

Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, et al.Tuberculosis. *Nature Reviews Disease Primers* 2016;**2**:16076. [DOI: 10.1038/nrdp.2016.76]

Pinsky 2006

Pinsky PF, Freedman M, Kvale P, Oken M, Caporaso N, Gohagan J.Abnormalities on chest radiograph reported in subjects in a cancer screening trial. *Chest* 2006;**130**(3):688-93. [DOI: 10.1378/chest.130.3.688]

Reference Manager 12 [Computer program]

Reference Manager.Version 12. Philadelphia, PA: Thomson Reuters, 2010. www.refman.com/.

Reitsma 2005

Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH.Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005;**58**(10):982-90.

Rieder 1999

Rieder HL.Tuberculosis. Etiologic epidemiology: risk factors for disease given that infection has occurred. In: Epidemiologic Basis of Tuberculosis Control. 1st edition. Paris: International Union Against Tuberculosis and Lung Disease, 1999:63-119.

Rutjes 2006

Rutjes AW, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PM.Evidence of bias and variation in diagnostic accuracy studies. *CMAJ* 2006;**174**(4):469-76.

Schünemann 2020a

Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al.GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. *Journal of Clinical Epidemiology* 2020;**122**:129-41. [DOI: 10.1016/j.jclinepi.2019.12.020]

Schünemann 2020b

Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al.GRADE guidelines: 21 part 2. Test accuracy: inconsistency, imprecision, publication bias, and other domains for rating the certainty of evidence and presenting it in evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology* 2020;**122**:142-52. [DOI: 10.1016/j.jclinepi.2019.12.021]

Shapiro 2021

Shapiro AE, Ross JM, Yao M, Schiller I, Kohli M, Dendukuri N, et al.Xpert MTB/RIF and Xpert Ultra assays for screening for pulmonary tuberculosis and rifampicin resistance in adults, irrespective of signs or symptoms. *Cochrane Database of*

Systematic Reviews 2021, Issue 3. Art. No: CD013694. [DOI: 10.1002/14651858.CD013694.pub2]

Somoskövi 2000

Somoskövi A, Ködmön C, Lantos A, Bártfai Z, Tamási L, Füzy J, et al.Comparison of recoveries of mycobacterium tuberculosis using the automated BACTEC MGIT 960 system, the BACTEC 460 TB system, and Lowenstein-Jensen medium. *Journal of Clinical Microbiology* 2000;**38**(6):2395-7.

Soto 2011

Soto A, Solari L, Gotuzzo E, Acinelli R, Vargas D, Van der Stuyft P.Performance of an algorithm based on WHO recommendations for the diagnosis of smear-negative pulmonary tuberculosis in patients without HIV infection. *Tropical Medicine & International Health* 2011;**16**(4):424-30.

Sreeramareddy 2009

Sreeramareddy CT, Panduru KV, Menten J, Van den Ende J.Time delays in diagnosis of pulmonary tuberculosis: a systematic review of literature. *BMC Infectious Diseases* 2009;**9**:91.

Stata [Computer program]

Stata: Statistical Software.Version 11. College Station, TX, USA: StataCorp, 2010. Available at www.stata.com.

Steingart 2006a

Steingart KR, Henry M, Ng V, Hopewell PC, Ramsay A, Cunningham J, et al.Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. *Lancet Infectious Diseases* 2006;**6**(9):570-81.

Steingart 2006b

Steingart KR, Ng V, Henry M, Hopewell PC, Ramsay A, Cunningham J, et al.Sputum processing methods to improve the sensitivity of smear microscopy for tuberculosis: a systematic review. *Lancet Infectious Diseases* 2006;**6**(10):664-74.

Steingart 2007

Steingart KR, Henry M, Laal S, Hopewell PC, Ramsay A, Menzies D, et al.Commercial serological antibody detection tests for the diagnosis of pulmonary tuberculosis: a systematic review. *PLOS Medicine* 2007;**4**(6):e202.

Swai 2011

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

TB CARE I. 2014

TB CARE I.International Standards for Tuberculosis Care. 3rd edition. The Hague: TB CARE I, 2014.

Tiemersma 2011

Tiemersma EW, van der Werf MJ, Borgdorff MW, Williams BG, Nagelkerke NJ.Natural history of tuberculosis: duration and fatality of untreated pulmonary tuberculosis in HIV negative patients: a systematic review. *PLOS One* 2011;**6**(4):e17601.

UNAIDS 2015

UNAIDS.UNAIDS Terminology Guidelines 2015. www.unaids.org/en/resources/documents/2015/2015_terminology_guidelines (accessed 11 November 2020).

van Cleeff 2003

van Cleeff MR, Kivihya-Ndugga L, Githui W, Nganga L, Odhiambo J, Klatser PR.A comprehensive study of the efficiency of the routine pulmonary tuberculosis diagnostic process in Nairobi. *International Journal of Tuberculosis and Lung Disease* 2003;**7**(2):186-9.

van't Hoog 2011b

van't Hoog AH, Laserson KF, Githui WA, Meme HK, Agaya JA, Odeny LO, et al.High prevalence of pulmonary tuberculosis and inadequate case finding in rural western Kenya. *American Journal of Respiratory and Critical Care Medicine* 2011;**183**(9):1245-53.

van't Hoog 2013

van't Hoog AH, Marston BJ, Ayisi JG, Agaya JA, Muhenje O, Odeny LO, et al.Risk factors for inadequate TB case finding in Rural Western Kenya: a comparison of actively and passively identified TB patients. *PLOS One* 2013;**8**(4):e61162.

van't Hoog 2014a

van't Hoog AH, Onozaki I, Lonnroth K.Choosing algorithms for TB screening: a modelling study to compare yield, predictive value and diagnostic burden. *BMC Infectious Diseases* 2014;**14**:532.

Vonasek 2021

Vonasek B, Ness T, Takwoingi Y, Kay AW, van Wyk SS, Ouellette L, et al.Screening tests for active pulmonary tuberculosis in children. *Cochrane Database of Systematic Reviews* 2021, Issue 6. Art. No: CD013693. [DOI: 10.1002/14651858.CD013693.pub2]

Whitelaw 2009

Whitelaw A, Sturm WA.Microbiological testing for *Mycobactrium tuberculosis*. In: SchaafHS, ZumlaAI, editors(s). Tuberculosis. A Comprehensive Clinical Reference. Philadelphia: Saunders, 2009:164-8.

Whiting 2011

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al.QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529-36.

WHO 2011

World Health Organization.Tuberculosis Prevalence Surveys: A Handbook. 2nd edition. Geneva: World Health Organization, 2011. [ISBN 978 92 4 15481]

WHO 2013

World Health Organization.Systematic Screening for Active Tuberculosis. Principles and Recommendations. WHO/HTM/TB/2013.04 edition. Geneva: World Health Organization, 2013.

WHO 2015

World Health Organization.Implementing The End of TB Strategy: The Essentials. WHO/HTM/TB/2015.31 edition. Geneva: World Health Organization, 2015. [ISBN 978 92 4 150993 0]

WHO 2018

World Health Organization.Latent tuberculosis infection: updated and consolidated guidelines for programmatic management. 2018. https://apps.who.int/iris/bitstream/handle/10665/260233/9789241550239-eng.pdf (accessed 12 Oct 2020).

WHO 2020

World Health Organization.Global tuberculosis control: WHO report 2020. www.who.int/tb/data (accessed 22 March 2021).

WHO 2020a

World Health Organization.WHO consolidated guidelines on tuberculosis. Module 3: diagnosis – rapid diagnostics for tuberculosis detection. 2020. who.int/publications/i/item/who-consolidated-guidelines-on-tuberculosis-module-3-diagnosis---rapid-diagnostics-for-tuberculosis-detection (accessed 22 March 2021).

WHO 2021a

World Health Organization.WHO consolidated guidelines on tuberculosis. Module 2: Screening – Systematic screening for tuberculosis disease. who.int/publications/i/item/9789240022676 (accessed 22 March 2021).

WHO 2021b

World Health Organization.National tuberculosis prevalence surveys 2007-2016. who.int/publications/i/item/9789240022430 (accessed 19 Jan 2022).

Wilson 1968

Wilson JMG, Jungner G.Principles and practice of screening for disease. In: Originally appearing in: Commission on Chronic Illness (1957) Chronic illness in the United States: Volume I. Prevention of Chronic Illness, Cambridge, Massachusetts, Harvard University Press, p.45) (accessed at http://whqlibdoc.who.int/php/WHO_PHP_34.pdf). Geneva, Switzerland: World Health Organization, 1968.

Yoon 2017

Yoon C, Chaisson LH, Patel SM, Allen IE, Drain PK, Wilson D, et al.Diagnostic accuracy of C-reactive protein for active pulmonary tuberculosis: a meta-analysis. *International Journal of Tuberculosis and Lung Disease* 2017;**21**(9):1013-9. [DOI: 10.5588/ijtld.17.0078]

References to other published versions of this review

van't Hoog 2013a

van't Hoog AH, Langendam MW, Mitchell E, Cobelens FG, Sinclair D, Leeflang MMG, et al.A systematic review of the sensitivity and specificity of symptom- and chest-radiography screening for active pulmonary tuberculosis in HIV-negative persons and persons with unknown HIV status. March 2013. who.int/tb/Review2Accuracyofscreeningtests.pdf?ua=1 (accessed 1 November 2021).

van't Hoog 2014b

van't Hoog AH, Langendam M, Mitchell E, Cobelens FG, Sinclair D, Leeflang M, et al.Symptom- and chest-radiography screening for active pulmonary tuberculosis in HIV-negative adults and adults with unknown HIV status. *Cochrane Database of Systematic Reviews* 2014, Issue 1. Art. No: CD010890. [DOI: 10.1002/14651858.CD010890]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Adetifa 2016

Study characteristics	
Patient Sampling	Cross-sectional design; prevalence survey. Multi-stage cluster sampling. 80 clusters were randomly selected from the entire country, procedure equivalent to proportional to size. Within clusters permanent residents were eligible, and visitors who had arrived in the household 4 weeks or more before. Of 55,832 eligible individuals, 43,100 (77.2%) participated.
Patient characteristics and setting	General population in The Gambia, both rural and urban, of 15 years and older Median age: 28 years 59% female

Adetifa 2016 *(Continued)*

	HIV prevalence: not reported; the HIV/AIDS prevalence among adults aged (15 to 49) in The Gambia is 1.8% (1.5% to 2.2%) (2015 estimate unaids.org) History of previous TB treatment: not reported Smoking prevalence: not reported. Tobacco use among adults was approximately 10% in 2015 (19.3% in men, 0.8% in women; https://tobaccoatlas.org/)
Index tests	Symptom questions were asked: cough, fever, chest pain, night sweats, shortness of breath, anorexia, weight loss, haemoptysis Data extracted on: 1) Cough for 2 or more weeks 2) Study-specific combination of symptoms out of several: cough less than 2 weeks + any 2 symptoms positive out of: weight loss, fever, night sweats, chest pain, shortness of breath, anorexia, haemoptysis), or cough for 2 weeks or more, or 3 or more other symptoms 3) Chest radiography suggestive of tuberculosis
Target condition and reference standard(s)	Bacteriological tests: smear microscopy (FM) and liquid culture, on 2 sputum samples Case definition: at least 1 smear positive or 1 culture positive Participants with bacteriological test: 12.8%; those without, by design, were assumed not to have TB Bacteriological confirmation on 2 samples, or 1 positive sample and CXR abnormalities
Flow and timing	Of 5948 individuals supposed to receive a bacteriological reference standard (by design, i.e. suspected of TB) 429 did not get a test or had no result. This number (429) is very high compared to the prevalence of TB in this study group (77 cases), therefore we judged high risk of bias.
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	

Adetifa 2016 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Unclear		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Adetifa 2016 *(Continued)*

Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Ayles 2009a

Study characteristics	
Patient Sampling	Cross-sectional survey of TB and HIV prevalence. Random sampling of households in enumeration areas.
Patient characteristics and setting	General population of 15 years and above of 2 political divisions (1 urban, 1 rural) in Lusaka Province, Zambia; recruited in 2005 64% between 15 and 34 years old 54% female HIV prevalence 28.6% TB prevalence 870/100,000 History of previous TB treatment: 4% Smoking not reported. World Tobacco Atlas 3rd ed. (2008): daily tobacco use approximately 10% in adults (18.0% for men, 2.1% for women)
Index tests	Symptom screens: - Cough of any duration - Cough for 3 or more weeks or haemoptysis (programme definition of 'TB suspect' at the time of survey) Any TB symptom (out of 6): Questions asked: currently coughing - if yes, between 1 and 21 days or > 21 days; shortness of breath; fever; night sweats; weight loss; chest pain
Target condition and reference standard(s)	Bacteriological tests: 1 sputum sample cultured (2 MGIT tubes); ZN microscopy on positive cultures Incorporation: 100% had bacteriological test Case definition: 1 positive culture irrespective of symptoms/signs of active disease. ± 7% asymptomatic, CXR findings incomplete. Moderate applicability concern. Since the reference standard is based on one sputum sample only, some misclassification is possible.
Flow and timing	Samples were collected at the same time as symptom interview

Ayles 2009a *(Continued)*

	223 (2.7%) of 8325 consented participants without culture result for various reasons and 25 (0.3%) who did not submit a sputum sample were excluded from the analysis
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		

Ayles 2009a *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Yes		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Low risk	

Ayles 2009b

Study characteristics			
Patient Sampling	Cross-sectional survey of TB and HIV prevalence; random sampling of households in enumeration areas. Same as Ayles 2009a, but HIV-negative subset only.		
Patient characteristics and setting	See Ayles 2009a; general population, but HIV-negative subset only; HIV prevalence: 0%		
Index tests	Symptom screening; same as Ayles 2009a		
Target condition and reference standard(s)	See Ayles 2009a		
Flow and timing	See Ayles 2009a		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		

Ayles 2009b *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Yes		

Ayles 2009b *(Continued)*

Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	Low risk

Chadha 2018

Study characteristics	
Patient Sampling	Cross-sectional - prevalence survey. Cluster sampling then adults aged greater than or equal to 15 years in 5 areas. For the review data from Gujarat is used since MMR screening was used in the other 4 areas. Of 97,483 eligible residents, 87,530 (89.8%) participated
Patient characteristics and setting	General population 15 years and older, in Gujarat, India, a LMIC, enrolled in 2011; mixed urban and rural TB prevalence 275/100,000 population, incidence: 204/100,000 Distrubution of age, % female, HIV prevalence and smoking not provided. In 2016, India's HIV prevalence in adults aged 15 to 49 was approximately 0.2% (https://www.unaids.org/en/regionscountries/countries/india; accessed 20 April 2021) 2.2% of participants reported previous TB treatment Smoking - Tobacco Atlas 3rd edition 2009: current cigarette use approximately 15% (27.6% for males, 1% for females)
Index tests	Symptom screen: cough greater than or equal to 2 weeks CXR any pulmonary abnormality in persons pre-screened with symptoms (outside scope of this review)
Target condition and reference standard(s)	Bacteriological tests: 2 samples for microscopy and culture on solid medium Incorporation: 9.8% (8546/87,530) received a reference standard; by design, participants without symptoms or CXR abnormality were assumed not to have TB Case definition: culture positivity (one or more colonies of *M tuberculosis* confirmed for speciation by biochemical tests) on any of the 2 specimens OR any of the 2 specimens exhibited AFB on microscopy, or both No confirmed TB cases without signs of active disease at the time of screening
Flow and timing	Of 9553 sputum eligible, 8546 (89.5%) had sputum examined; not clear if all with results had results on 2 samples
Comparative	

Chadha 2018 *(Continued)*

Notes	CXR (digital, mobile): any pulmonary abnormality; 2 x 2 data for CXR can only be obtained of a subset of the population of persons with symptoms

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Unclear		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		

Chadha 2018 *(Continued)*

Did the study design require all patients to receive a bacteriological reference standard?	No	
Incorporation bias avoided?	No	
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Unclear	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Unclear	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Unclear risk

Cheng 2015

Study characteristics	
Patient Sampling	Cross-sectional design. National prevalence survey in 2010. Multi-stage stratified sampling to randomly select 176 investigation points from within the 31 mainland provinces. Within investigation points all residents were eligible.
Patient characteristics and setting	General population in China, urban and rural, of 15 years and above Mean/median age and sex distribution of study population not reported TB prevalence 442 per 100,000 HIV/AIDS prevalence in general population 0.037% in 2014 (https://www.unaids.org/sites/default/files/country/documents/CHN_narrative_report_2015.pdf) Proportion with previous TB treatment: not reported Daily tobacco use in population 15 and above in 2015, approximately 23% (https://tobaccoatlas.org; accessed December 2019); 44.8% of males, 2% of females

Cheng 2015 *(Continued)*

Index tests	Symptom screens: - Cough greater than or equal to 2 weeks (with or without haemoptysis) - Cough or haemoptysis of any duration - Cough greater than or equal to 3 weeks - Suspected symptom defined by China NTP = cough, expectoration last for longer than 2 weeks or haemoptysis
Target condition and reference standard(s)	Bacteriological tests: microscopy and culture (LJ) - not sure if it was ZN or FM microscopy; 3 samples were collected Case definition: a bacteriologically positive TB patient was defined as an individual with at least one positive smear or culture. Positive smear had at least one acid-fast bacillus identified within 100 fields under microscopy, and positive culture had at least one colony of *Mycobacterium tuberculosis* complex isolated by using Löwenstein-Jensen medium. Incorporation bias?: 18% of participants received bacteriological tests
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Cheng 2015 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?		Low concern
DOMAIN 2: Index Test (Chest radiography)		
DOMAIN 2: Index Test (Combination of Symptoms and CXR)		
DOMAIN 3: Reference Standard		
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes	
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes	
Did the study design require all patients to receive a bacteriological reference standard?	No	
Incorporation bias avoided?	No	
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Unclear	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Could the patient flow have introduced bias?		Unclear risk

Chheng 2008a

Study characteristics	
Patient Sampling	All adults who received routine pretest counselling and underwent HIV testing at a VCCT were eligible. Of 1220 subjects tested and referred for TB screening, 583 (48%) approached the study team located in the same building. Of these, 79 (14%) refused to participate in screening.

Chheng 2008a *(Continued)*

Patient characteristics and setting	Individuals aged 19 years and older attending the community voluntary confidential counselling and testing (VCCT) centre of the Battambang, Cambodia, who decided themselves to go for TB screening, March to September 2005 Median age 31 years (IQR 25 to 39) 38% female HIV prevalence 25% among participants; 6.2% had a prior history of TB treatment TB prevalence in region 508/100,000 population Smoking not reported; Tobacco Atlas 1st ed. (2002): 37% of adults
Index tests	Symptom screens - Cough of 3 weeks or more - Any TB symptom out of 6: cough 3 weeks, haemoptysis, fever, loss of appetite, night sweats, rapid weight loss (self-reported) Additional questions were asked on HIV/AIDS symptoms - Other combination: fever or haemoptysis or weight loss
Target condition and reference standard(s)	Bacteriological tests: 3 sputum samples for FM smear microscopy and solid (LJ) culture All participants had bacteriological testing Case definition: any single sputum culture-positive for *M tuberculosis* or at least 2 sputum smears positive for AFB (excluding MOTT positives), irrespective of clinical signs/symptoms All cases have at least a TB symptom (Table 1)
Flow and timing	Of the 496 participants, 469 (95%) submitted at least 2 sputum samples, and 430 (87%) submitted 3 samples for examination
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	

Chheng 2008a *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Unclear		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Chheng 2008b

Study characteristics			
Patient Sampling	Same as Chheng 2008a; HIV-negative subset only		
Patient characteristics and setting	Same as Chheng 2008a; HIV-negative subset only		
Index tests	Same as Chheng 2008a		
Target condition and reference standard(s)	Same as Chheng 2008a		
Flow and timing	Same as Chheng 2008a		
Comparative			
Notes			

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			

Chheng 2008b *(Continued)*

Is the bacteriological reference standard likely to correctly classify the target condition?	Yes	
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes	
Did the study design require all patients to receive a bacteriological reference standard?	Yes	
Incorporation bias avoided?	Yes	
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Unclear	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Claassens 2017a

Study characteristics	
Patient Sampling	Baseline of an intervention trial (ZAMSTAR trial). TB/HIV prevalence survey in South Africa. 8 communities were selected. Within these clusters all individuals aged 18 years who stayed in households in the previous 24 hours were asked to participate. The manuscript reports data for a 'train' and a 'test' dataset, which are combined.
Patient characteristics and setting	General urban population of 15 years and older in Cape Town region, South Africa in 2010 Age: approximately half below age 30 62.4% female Among screened population:

Claassens 2017a *(Continued)*

	- TB prevalence 2338 per 100,000 - HIV prevalence 18% - Smoking prevalence 8% - Previous TB treatment reported 9%
Index tests	Symptoms asked for: current cough, how many weeks of cough, currently producing phlegm/sputum or blood, current shortness of breath, sweating at night or fever, and weight loss within the past month Symptom index tests: - Cough 2 or more weeks - Any 1 TB symptom positive (out of 7) The manuscript reports on other, post-hoc combinations
Target condition and reference standard(s)	Bacteriological tests: all participants provided one sputum sample for culture. The sample was split and cultures in 2 MGIT tubes. Smear microscopy was only done if a culture was positive. Case definition: a positive culture for *Mycobacterium tuberculosis* All participants had a bacteriological test Of the cases 40% did not report symptoms; CXR abnormalities were not checked/reported in all cases
Flow and timing	Based on Figure 3 in Ayles et al. paper: 29% of participants excluded because they did not have an evaluable sample
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			

Claassens 2017a *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Claassens 2017b

Study characteristics	
Patient Sampling	Baseline of an intervention trial (ZAMSTAR trial). TB/HIV prevalence survey in South Africa. 8 communities were selected. Within these clusters all individuals aged 18 years who stayed in households in the previous 24 hours were asked to participate. The manuscript reports data for HIV-negative participants for a 'train' and a 'test' dataset, which are combined.
Patient characteristics and setting	Same as Claassens 2017a except for - HIV prevalence 0%
Index tests	Symptoms asked for: current cough, how many weeks of cough, currently producing phlegm/sputum or blood, current shortness of breath, sweating at night or fever, and weight loss within the past month Symptom index tests: Cough 2 or more weeks, or any one TB symptom positive The manuscript reports on other, post-hoc combinations
Target condition and reference standard(s)	Bacteriological tests: all participants provided one sputum sample for culture. The sample was split and cultures in 2 MGIT tubes. Smear microscopy was only done if a culture was positive. Case definition: a positive culture for *Mycobacterium tuberculosis* All participants had a bacteriological test Of the cases 40% did not report symptoms CXR abnormalities were not checked/reported in all cases
Flow and timing	Based on Figure 3 in Ayles et al. paper: 29% of participants excluded because they did not have an evaluable sample
Comparative	
Notes	Note on Index test domain: We consider Cough > 2 weeks, and Any one TB symptom positive as low risk. The other combinations in this manuscript are post-hoc analyses on screening rules, meaning these symptom combinations imply a higher risk of bias.

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		

Claassens 2017b *(Continued)*

Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		

Claassens 2017b *(Continued)*

Could the patient flow have introduced bias?	High risk

Claassens 2017c

Study characteristics	
Patient Sampling	Baseline of an intervention trial (ZAMSTAR trial). TB/HIV prevalence survey in Zambia. 16 communities were selected. Within these clusters all individuals aged 18 years who stayed in households in the previous 24 hours were asked to participate.
Patient characteristics and setting	General rural population of 15 years and older in 2 districts of Lusaka Province, Zambia in 2010 Age: approximately half below age 30 62.2% female Among screened population: - TB prevalence 555 per 100,000 - HIV prevalence 17% - Smoking prevalence 8% - Previous TB treatment reported 9%
Index tests	Symptoms asked for: current cough, how many weeks of cough, currently producing phlegm/sputum or blood, current shortness of breath, sweating at night or fever and weight loss within the past month Symptom index tests: - Cough 2 or more weeks - Any one TB symptom positive (out of 7) The manuscript reports on other, post-hoc combinations
Target condition and reference standard(s)	Bacteriological tests: all participants provided one sputum sample for culture. The sample was split and cultures in 2 MGIT tubes. Smear microscopy was only done if a culture was positive. Case definition: a positive culture for *Mycobacterium tuberculosis* All participants had a bacteriological test Of the cases 40% did not report symptoms CXR abnormalities were not checked/reported in all cases
Flow and timing	Based on Figure 3 in Ayles et al paper: 29% of participants excluded because they did not have an evaluable sample
Comparative	
Notes	

Claassens 2017c *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear

Claassens 2017c *(Continued)*

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Could the patient flow have introduced bias?		High risk

Claassens 2017d

Study characteristics	
Patient Sampling	Baseline of an intervention trial (ZAMSTAR trial). TB/HIV prevalence survey. 8 communities were selected. Within these clusters all individuals aged 18 years who stayed in households in the previous 24 hours were asked to participate. The manuscript reports data for HIV-negative participants for a 'train' and a 'test' dataset, which are combined.
Patient characteristics and setting	Same as Claassens 2017c except for: - HIV prevalence 0%
Index tests	Symptoms asked for: current cough, how many weeks of cough, currently producing phlegm/sputum or blood, current shortness of breath, sweating at night or fever, and weight loss within the past month Symptom index tests: Cough 2 or more weeks, or any one TB symptom positive The manuscript reports on other, post-hoc combinations
Target condition and reference standard(s)	Bacteriological tests: all participants provided one sputum sample for culture. The sample was split and cultures in 2 MGIT tubes. Smear microscopy was only done if a culture was positive. Case definition: a positive culture for *Mycobacterium tuberculosis* All participants had a bacteriological test Of the cases 40% did not report symptoms CXR abnormalities were not checked/reported in all cases
Flow and timing	Based on Figure 3 in Ayles et al paper: 29% of participants excluded because they did not have an evaluable sample
Comparative	

Claassens 2017d *(Continued)*

Notes	Note on Index test domain: We consider Cough > 2 weeks, and Any one TB symptom positive as low risk. The other combinations in this manuscript are post-hoc analyses on screening rules, meaning these symptom combinations imply a higher risk of bias.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Claassens 2017d *(Continued)*

Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes
Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	High risk

Corbett 2010a

Study characteristics	
Patient Sampling	Population-based cross-sectional survey of HIV and TB prevalence on a random sample of adults in high-density suburbs, enrolled at home. Participants who declined HIV testing (10%) were excluded since the analysis was stratified by HIV status. Risk of bias from this exclusion judged to be low.
Patient characteristics and setting	General population of 16 years and older in urban suburbs in Harare, Zimbabwe, 2005 to 2006 TB prevalence 700 per 100,000; HIV prevalence 20.7% 60.7% female history of previous TB treatment 3.3% Current smoker: 8.6% Median age 25 years (IQR 20 to 35) in HIV-negative population and 32 (26 to 38) in HIV-positive population
Index tests	Symptom screens: current cough of any duration or severity; chronic cough defined as 2 weeks duration or more Any one symptom out of 5: cough, haemoptysis during the previous year, self-reported fever or "hot body", night sweats, and a subjective report of weight loss
Target condition and reference standard(s)	Bacteriological tests: all participants were screened by sputum culture (LJ) on 1 sample. Participants who tested positive on sputum culture during screening or reported any TB symptom, or both, were followed up to confirm or exclude TB, based on 2 additional sputum specimens for ZN microscopy and solid (LJ) culture, and clinical assessment including CXR. Of asymptomatic individuals these 2 samples were pooled into 1 culture. Case definition: definite case defined as 2 positive cultures, or 1 positive culture with clinical evidence of TB (individuals with one positive culture, but without clinical evidence were not considered a case)
Flow and timing	1.4% of participants were unable to give sputum and excluded from the analysis, a very small percentage Reference standard procedures differed in 2 ways: i) asymptomatic participants with negative screening culture had 1 culture only; ii) dur-

Corbett 2010a *(Continued)*

	ing follow-up to confirm cases 2 samples were either pooled in 1 culture or not according to symptoms
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	

Corbett 2010a *(Continued)*

Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Yes		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Low risk	

Corbett 2010b

Study characteristics			
Patient Sampling	See Corbett 2010a; HIV-negative subset		
Patient characteristics and setting	HIV-negative subset of the population in Corbett 2010a Median age 25 years (IQR 20 to 35) 59% female History of previous TB treatment 9.1% 8.6% current smoker		
Index tests	See Corbett 2010a		
Target condition and reference standard(s)	See Corbett 2010a		
Flow and timing	See Corbett 2010a		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		

Corbett 2010b *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Yes		

Corbett 2010b *(Continued)*

Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	No	
Could the patient flow have introduced bias?		Low risk

den Boon 2006

Study characteristics	
Patient Sampling	Cross-sectional community lung health survey - randomized 15% sample of all addresses in study area (839 addresses). All persons at selected addresses were eligible for investigation. If no consent, the neighbouring household was eligible. Persons 15 years and older completed questionnaires and had CXR and sputum samples collected. 75% of persons with interview had a CXR and of those 45% gave a sputum sample. All participants who had a CXR are included in the analysis for this review. The authors report accuracy on the subset who gave a sputum sample, only 1/3 of those who consented. This was adjusted for this review, to limit bias.
Patient characteristics and setting	General urban population in 2 urban communities in Cape Town, South Africa, in 2002 Participants were 15 years or above, median age between 35 and 44 years (43% is between 15 and 44) 51% females Previous TB treatment reported by 9.7% HIV prevalence was not measured but estimated to be 12.4% or less Smoking not reported; Tobacco Atlas 1st ed. (2002): 26.5% adult smoking
Index tests	Symptom screening, CXR screening, and a combination. The symptom questionnaire contained questions on the presence and duration of symptoms of cough, haemoptysis, weight loss, night sweats, and fever. Prolonged cough was defined as cough for 2 or more weeks. Conventional CXR was used at a health facility to which participants had to travel. CXRs were interpreted by an experienced pulmonologist using a standardized reporting form to assess all CXRs for abnormalities consistent with TB (including parenchymal, pleural, and central structure abnormalities) and for any other abnormalities, including other visible organs. A stratified sample of 31% of the CXRs was re-read by a second reader. Of symptom and CXR screening combined, data is reported on the sensitivity only, of any TB symptom or CXR abnormalities consistent with TB or both.
Target condition and reference standard(s)	Bacteriological tests: 1 sputum sample was requested of all participants, was tested with ZN smear microscopy and solid (LJ) culture. A 2nd sputum sample was requested when the first sample as positive on smear microscopy or culture, from which 26 cases were defined in the main study. Case definition: bacteriologically confirmed TB case = a person with either 2 positive smears or 2 positive cultures, or 1 positive smear and 1 positive culture or 1 positive smear or culture, and CXR with TB-realted abnormalities, or a positive sputum smear or culture result from specimen collected at the health

den Boon 2006 *(Continued)*

	centre within 2 months after sputum collection in the previous survey. Of 29 cases in the analysis, 3 did not meet the definition of a bacteriologically confirmed TB cases but were defined based on one sample result only. Of the cases 3.4% had no signs of active disease (asymptomatic, normal CXR), a low percentage.
Flow and timing	Although the authors report accuracy in a subset or participants who gave a sputum sample, which was 1/3 of the total number who consented to participate, we included participants who had a CXR in the analysis for this review, which was 75% of consenting participants. Of those with CXR 45% gave a sputum sample. We assume that participants who did not give sputum did not have TB (i.e. reference standard negative).
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		

den Boon 2006 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		

den Boon 2006 *(Continued)*

Were all patients included in the analysis?	No
Could the patient flow have introduced bias?	High risk

Federal MoH Sudan 2018

Study characteristics	
Patient Sampling	Cross-sectional national prevalence survey. Multi-stage stratified cluster sampling using PPS of 114 clusters of 800 participants on average, stratified by urban vs rural. 4 clusters cancelled for security reasons, and 1 removed due to a protocol violation. Of 96,979 eligible residents, 83,202 (86%) participated Resident definition: household members who resided in the selected household for the past 6 months, and visitors who spent at least 3 weeks in the household prior to the census
Patient characteristics and setting	General population 15 years and older, in Sudan, a LMIC, enrolled in 2013 to 2014 TB prevalence (bacteriological confirmation, year/survey) 183/100,000 population, incidence (2018): 71/100,000 Median age: approximately 53% were 34 or younger 56.7% female HIV prevalence in the general population aged 15 to 49 years was 0.2% in 2013 (UNAIDS http://aidsinfo.unaids.org; accessed April 2017) Information on smoking and reported previous TB treatment not provided Tobacco Atlas: adult smoking (15+ year-olds) using tobacco daily: 2015 < 1% (1.3% for men, 0.4% female; https://tobaccoatlas.org/country/sudan/; accessed April 2021)
Index tests	1. Symptom screen: cough for 2 weeks or more 2. CXR (mobile digital): an abnormal chest X-ray meant any lung (including pleura) abnormality detected on interpretation by the medical officer (e.g. opacities, cavitation, fibrosis, pleural effusion, calcification, or any unexplained or suspicious shadow). Congenital abnormalities, normal variants and bony abnormalities including fractures are excluded by definition as were findings such as increased heart size and other heart-related abnormalities. Abnormal CXR or positive symptoms (cough for 2 weeks or more) or both
Target condition and reference standard(s)	Bacteriological tests: 2 samples for direct FM microscopy and Ogawa culture Incorporation: 16.9% received a reference standard. By design, participants with cough of 2 or more weeks or a CXR abnormality or both were eligible for sputum examination Participants not eligible for sputum or without a sputum result were assumed not to have TB Case definition: bacteriologically positive TB case = definite case = MTB confirmed by culture

Federal MoH Sudan 2018 *(Continued)*

	13.5% (15/111) of true (confirmed) TB cases were symptom-screen negative and had CXR findings that were unavailable/unknown
Flow and timing	Of those eligible for sputum examination, 80% had at least 1 culture result available; of those 65% only had the 2 required samples, therefore high risk of bias
Comparative	
Notes	5 more symptoms were asked about (cough ≥ 2 weeks, haemoptysis, sputum production, chest pain, fever), but no 2 x 2 data provided. Participants positive on these screens were sent for bacteriological examination. Assuming the study staff are well trained, risk of bias for the index test interpretation is considered low. CXR: Same argument as for symptoms; CXR threshold: implicit threshold. Intentional over-reading of CXR was encouraged so that no suspected cases were left out.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Federal MoH Sudan 2018 *(Continued)*

If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		

Federal MoH Sudan 2018 *(Continued)*

Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Fox 2012

Study characteristics	
Patient Sampling	Cohort of household contacts of patients with smear-positive pulmonary TB ('index patients') were enrolled between September 2009 and January 2011. Index patients were defined as the first person in the household to be diagnosed with TB. Household contacts of any age were eligible for inclusion if they lived in the same dwelling as the index patient during the 2 months prior to their diagnosis. Index patients were then encouraged to bring their household members to the clinics as soon as possible after their initial diagnosis. Contacts were asked to come to the clinic and recruited of 212 index cases (out of 545; 39%). This may have introduced bias.
Patient characteristics and setting	Household contacts of smear-positive TB patients in Hanoi, Vietnam; recruited in 2009 (n = 545) All ages were eligible; mean age 36.3 years; 16.9% were under the age of 15 years 59.6% female 14.2% reported current smoking 2% reported history of previous TB treatment HIV prevalence not reported; HIV prevalence (aged 15 to 49) estimated at 0.43% in 2009 (Vietnam HIV/AIDS Estimates and Projections 2007-2012. MOH, 2009)
Index tests	Symptom questions: having cough or sputum for 2 weeks or more, haemoptysis within the last month or both. Not known if more questions were asked. This combination of symptoms is different from the screen definitions prolonged cough or any symptom, but resembles mostly prolonged cough. CXR taken at a hospital, classified as any abnormality (includes other visible organs) vs normal
Target condition and reference standard(s)	Only 'suspects' were eligible for bacteriology: reported having cough or sputum for 2 weeks or more, haemoptysis within the last month, or any abnormalities on the chest radiograph. Others were assumed not to have TB. 24 suspects were identified 24/545 = 4.4%, of whom 13 gave sputum (13/545 = 2.4%) Of suspects 3 or more sputum samples were examined using Ziehl-Neelsen smear microscopy and solid culture on Lowenstein-Jensen medium. Case definition: smear-positive pulmonary TB was defined by the presence of at least 1 positive smear in combination with an abnormal chest radiograph, or 1 positive smear plus a positive culture. Smear-negative TB was diagnosed if contacts had radiographic changes consistent with TB, no re-

Fox 2012 *(Continued)*

	sponse to broad-spectrum antibiotics and a response to anti-tuberculous drug treatment (i.e. bacteriologically negative TB)
Flow and timing	Sputum followed directly on symptom interview and CXR. Of 24 suspects expected to give sputum, 13 gave sputum (13/24 = 54%); 46% were unable
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			

Fox 2012 *(Continued)*

DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Unclear		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

FRoNigeria 2014

Study characteristics	
Patient Sampling	Cross-sectional - national TB prevalence survey. Multi-stage sampling method was used (first geographical zones; then 70 clusters PPS). Clusters were enumeration areas with on average 700 eligible individuals. Due to an unstable security situation, field operations could not take place in the two States of Borno and Yobe. 3 enumeration areas were replaced.
Patient characteristics and setting	General population of 15 years and older, rural and urban, recruited in 2012 (n = 44,186) 52.8% were 15 to 34 years old

FRoNigeria 2014 *(Continued)*

	59% female HIV prevalence was not measured in the survey (4.1% in the general population) Smoking prevalence not reported 1.2% reported history of previous TB treatment Smoking in adults: Tobacco Atlas 3rd edition 2009: 4.5% to 5% cigarette smoking (9.1% in males, 0.2% in females)
Index tests	6 symptom questions were asked: cough + duration, sputum, haemoptysis, chest pain, body weight loss, fever. Persons with cough of 14 or more days were eligible to provide sputum samples. Was this known to the interviewers? Same for CXR. CXR: portable mobile X-ray units (MinXray), a computed radiography system (CR) equipped for digital images. Classification: any pulmonary abnormality vs no pulmonary abnormality, defined as any abnormal shadow in the lung field and mediastinum, or pleural effusion except pleural thickness or small single calcification. CXR with any pulmonary abnormality implied eligibility for sputum collection. Index tests: - Cough of any duration - Cough for 14 days or longer - Any one TB symptom out of 6 - CXR abnormalities suggestive of TB, defined as any abnormal shadow in the lung field and mediastinum, or pleural effusion except pleural thickness or small single calcification - CXR abnormality suggestive of TB and/or cough for 14 days or longer in parallel
Target condition and reference standard(s)	Persons with TB suggestive abnormalities had to give sputum; 10.6% were eligible for sputum submission, others are assumed to be TB-negative Bacteriological tests: 2 sputum samples (one spot and one morning) were obtained and were stained for AFB microscopy using ZN stain and both specimens were processed and cultured on LJ medium for *Mycobacterium tuberculosis* Case definition: smear-positive culture-positive (n = 75; definite), or smear-positive culture-negative (n = 32; probable), or S- C+ (n = 35; definite) or S- C+ (n =2; probable?). Total 144.
Flow and timing	10.6% were eligible for sputum submission. Of those 2.8% did not submit, and 12% submitted only 1 sample (rather than 2). Overall 79.2% had lab results available out of those eligible for lab testing. All participants are included in the analysis. Those without lab tests done or without results are considered to be free of active TB.
Comparative	
Notes	

FRoNigeria 2014 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

FRoNigeria 2014 *(Continued)*

DOMAIN 3: Reference Standard		
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes	
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes	
Did the study design require all patients to receive a bacteriological reference standard?	No	
Incorporation bias avoided?	No	
Could the reference standard, its conduct, or its interpretation have introduced bias?	High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?	High risk	

Ghana NTP 2015

Study characteristics	
Patient Sampling	Cross-sectional national prevalence survey. Multi-stage stratified cluster sampling using PPS of 98 clusters of 650 participants on average, stratified by urban vs rural. Of 67,757 eligible individuals 61,726 (91%) participated. Resident definition: permanent residents who lived in the household at least 1 day in the past 14 days, or visitors who lived in the household at least 7 days in the past 14 days.
Patient characteristics and setting	General population 15 years and older, in Ghana, a LMIC, enrolled in 2013

Ghana NTP 2015 *(Continued)*

	TB prevalence (bacteriologically confirmed, survey) 356/100,000 population, incidence (2018): 148/100,000 Median age between 25 and 34 years 55.4% female HIV prevalence in the general population aged 15 to 49 years was 1.5% in 2013 (http://aidsinfo.unaids.org/; accessed April 2017) Smoking: 5% of survey participants 7.8% of participants reported previous TB treatment
Index tests	Symptom screen: cough for ≥ 2 weeks CXR screen (direct digital): any pulmonary abnormality (abnormalities in the lung, pleura, and/or mediastinum) Combined cough for ≥ 2 weeks or CXR any pulmonary abnormality or both
Target condition and reference standard(s)	Bacteriological tests: 2 samples for concentrated ZN microscopy and culture on LJ and MGIT media; Xpert MTB/RIF for smear-positive specimens and if cultures were contaminated. Only MGIT was used for case definition. Incorporation: 12.2% received a reference standard. By design, participants with cough of 2 or more weeks or a CXR abnormality or both were eligible for sputum examination. Participants not eligible for sputum or without a sputum result were assumed not to have TB. Case definition: definite: smear-positive and culture positive in at least one sample OR smear-positive, Xpert-positive and CXR consistent with TB. Then smear-negative and culture or Xpert positive in 2 samples OR smear-negative and culture/Xpert-positive in one sample and CXR consistent with TB. 17.3% of true (confirmed) TB cases were asymptomatic and no information on CXR provided.
Flow and timing	Of those eligible for sputum examination, 91% had at least 1 culture result available. Over 700 had not, which is quite a lot compared to 202 cases, but hard to judge the magnitude of bias.
Comparative	
Notes	Index tests: risk of bias - although CXR and symptom determined if participants were eligible to give sputum in the survey, risk of bias on index test was considered low, assuming staff were well trained. Other symptoms asked about were cough, chest pain, weight loss, fever, night sweats, but no 2 x 2 data.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Ghana NTP 2015 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		

Ghana NTP 2015 *(Continued)*

Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Ho 2016

Study characteristics	
Patient Sampling	Cross-sectional study; field staff visited every household in 60 villages selected at random from throughout the province
Patient characteristics and setting	General rural population of 15 years and older in Ca Mau province in Vietnam; year of data collection not reported Median age 32 (IQR 15 to 47) 54% female HIV prevalence not reported; HIV prevalence among adults (15 to 49) in 2016: 0.4% (https://www.unaids.org) TB prevalence 367 per 100,000 Smoking not reported. Using tobacco daily in 2015: approximately 20%. (38.7% for male, 0.9% for female; https://tobaccoatlas.org; accessed April 2020) % with history of previous TB treatment: not reported
Index tests	1) Cough on day of screening (classified as cough of any duration)

Ho 2016 *(Continued)*

	2) Cough lasting 2 weeks or more 3) At least 1 symptom on day of screening out of 5 questions 4) Any symptom for the past 2 weeks or more Although the threshold for 3) and 4) was not pre-specified, the authors probably did not compose thresholds in a data-driven way
Target condition and reference standard(s)	Bacteriological tests: 1 sputum sample was tested with Xpert MTB/RIF (note that culture and CXR were also used but data not extractable) Case definition: Xpert-positive on 1 sputum sample
Flow and timing	A large proportion (47%) did not cough up sputum, but were included in the analysis under the (biased) assumption: 'all who did not cough up sputum did not have TB'
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			

Ho 2016 *(Continued)*

DOMAIN 3: Reference Standard		
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes	
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes	
Did the study design require all patients to receive a bacteriological reference standard?	Yes	
Incorporation bias avoided?	Yes	
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Unclear
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

Hoa 2012

Study characteristics	
Patient Sampling	Cross-sectional TB prevalence survey with multi-stage cluster sampling, stratified by area (urban, rural, remote). In total, 105,000 adults divided over 70 clusters were selected with probability proportional to size; 103,924 were eligible and 94,179 participated
Patient characteristics and setting	General population of 15 years and older, urban and rural, recruited in 2006 in Vietnam Median age between 35 and 44 years (42.0% were between 15 and 34 years) 54.7% female

Hoa 2012 *(Continued)*

	History of TB treatment: 0.5% HIV prevalence not reported; HIV prevalence in adults 15 to 49 years 0.5% (UNAIDS 2006: http://data.unaids.org/pub/report/2006/2006_gr_en.pdf) Smoking information not provided; Tobacco Atlas 1st edition (2002): 27% of adult population current tobacco use
Index tests	1) Persistent productive cough, meaning current cough for more than 2 weeks with sputum production. Individuals were also asked about TB history, and individuals who were currently receiving anti-tuberculosis treatment or had received treatment in the 2 years preceding the survey were not interviewed about cough. 2) All participants underwent CXR screening by 70 × 70 mm mass miniature radiography (photofluorography) or digital X-ray units (mobile). Not stated what proportion had what type of CXR. X-ray images were scored as showing no abnormalities (normal), abnormalities suggestive of TB, or other abnormalities (including other visible organs). All abnormal CXR images were reread in central X-ray reading units. 3) Persistent productive cough or CXR suggestive of TB or both
Target condition and reference standard(s)	Only persons with persistent productive cough or abnormal CXR or both on screening were requested to submit 3 sputum specimens for bacteriological examination; 8.1% received bacteriological testing Bacteriological tests: 2 samples tested for ZN microscopy, 1 for LJ culture Case definition: 1 positive culture or 2 positive smears or 1 positive smear and CXR abnormal
Flow and timing	6% to 10% were unable to produce sputum and assumed as TB-negative
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	

Hoa 2012 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	

Hoa 2012 *(Continued)*

Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

Kapata 2016

Study characteristics	
Patient Sampling	Cross-sectional national prevalence survey. Multi-stage stratified cluster sampling using PPS of 66 selected clusters (CSAs) in 49 districts in all the 10 provinces of Zambia. On average 825 persons per cluster. Resident definition: slept in the household 24 hours prior to the census. Of 54,830 eligible residents, 46,099 (84%) participated.
Patient characteristics and setting	General population 15 years and older, in Zambia, a LMIC, enrolled in 2013 to 2014 TB prevalence (bacteriologically confirmed, year/survey), 638/100,000 population, incidence (2013): 427/100,000 Median age 32 years (IQR 22 to 47) 56.4% female HIV prevalence: 30,584 (33.7%) of 46,099 participants were HIV-tested of whom 6.7% were HIV-positive Smoking and previous TB treatment information n/a (only in presumptive TB cases) In Zambia of adults 15+ using tobacco daily in 2015: 16% in men, 2.5% in women (https://tobaccoatlas.org/country/zambia/; accessed May 2020)
Index tests	Symptom screen: any one of the following 3 symptoms for at least 2 weeks: cough, fever and/or chest pain. The definition is somewhat different: it cannot be classified as cough > 2 weeks, but is also quite different from most 'any TB symptom' studies as there are only 3 symptoms that are required to be present for at least 2 weeks.

Kapata 2016 *(Continued)*

	CXR: direct digital, an abnormal chest X-ray was defined as one showing any lung abnormality (e.g. opacities, cavitation, fibrosis, pleural effusion, calcification, any unexplained or suspicious shadow) and heart abnormalities detected on interpretation by a qualified medical officer.
Target condition and reference standard(s)	Bacteriological tests: 2 samples (spot, morning): concentrated preparation, FM (LED, auramine stain) and culture on LJ media and MGIT media. Xpert MTB/RIF testing was performed for all smear-positive or if the only sample collected yielded inconclusive results after a case definition meeting. Incorporation: 13.3% received a reference standard. By design, participants with TB symptoms or a CXR abnormality or both were eligible for sputum examination. Participants not eligible for sputum or without a sputum result were assumed not to have TB. Case definition: a bacteriologically confirmed pulmonary TB case was a participant with smear-positive and/or culture-positive pulmonary TB (or Xpert MTB/RIF positive).
Flow and timing	91.3% of eligible had a result. 60.5% had results for FM and MGIT on 2 samples. Not clear what happened to the 'undefined' category at the bottom of Figure 1. Classified as TB negative?
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Kapata 2016 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Kapata 2016 *(Continued)*

Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Kebede 2014

Study characteristics	
Patient Sampling	Cross-sectional national tuberculosis prevalence survey. Stratified, multi-stage sampling of 85 clusters/kebeles, selected from 3 strata, proportional to the population size living in each area (Urban, Rural, Pastoralist).
Patient characteristics and setting	General population in Ethiopia of 15 years and older, 2010 to 2011 (n = 46,697) Median age 33.6 years 53% female Smoking, information not provided; Tobacco Atlas 3rd ed. 2009: current cigarette smoking: 6.9% in men, 0.5% in women Previous history of TB treatment: in past 5 years (excluding those currently on treatment): 1.6% HIV prevalence in study population not reported, in the general population HIV prevalence is 2.3%
Index tests	Symptoms: - Cough for 14 days or more (considered eligible for sputum examinations in survey) - Any 1 symptom out of: cough and duration, fever for 2 or more weeks, weight loss of more than 3 kg in the last 4 weeks, night sweats for 2 or more weeks, cervical lymph node swelling CXRs with conventional portable CXR equipment. Classified as any pulmonary abnormality vs no pulmonary abnormality (field reading). Abnormal was defined as: any abnormality in lung field or mediastinum, including cavities, infiltrates, pleural effusion, hilar or mediastinal lymphadenopathy, pulmonary nodules, interstitial abnormalities, and healed TB.
Target condition and reference standard(s)	Bacteriological test: 2 samples examined with fluorescence microscopy, 1 of them cultured on solid (LJ) medium Incorporation: 12.6% had a reference standard. Persons without symptoms and without CXR abnormality were considered not to have TB

Kebede 2014 *(Continued)*

	Case definition: bacteriologically confirmed TB: includes definite and probable TB. Definite TB case = study participant with 1 culture-positive specimen with *Mycobacterium tuberculosis* and at least 1 of following conditions: AFB smear-positive, culture-positive in another specimen, CXR consistent with TB by audited central reading. Probable TB: AFB smear-positive no culture (TB) positive AND no isolation of MOTT. Of 110 cases 6 (5%) have 1 positive smear only.
Flow and timing	From the total eligible individuals for sputum examination: 5868 (97%) individuals had at least 1 smear result, 5606 (92%) individuals had both smear results, and 5807 (99%) individuals had 1 culture, of which 5503 (95%) had a culture result
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Kebede 2014 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?			
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Kenya MoH 2018

Study characteristics	
Patient Sampling	Cross-sectional national prevalence survey. Multi-stage stratified cluster sampling using PPS of 100 clusters of 720 participants on average, stratified by urban vs rural. 1 cluster was excluded for security reasons. Of 76,291 eligible residents, 63,050 (83%) participated Resident definition: lived in the selected cluster for at least 30 consecutive days prior to the census
Patient characteristics and setting	General population 15 years and older, in Kenya, a LMIC, enrolled in 2015 to 2016 TB prevalence (bacteriologically confirmed, year/survey) 558/100,000 population, incidence (2018): 292/100,000 Median age between 25 and 34 years 59% female HIV prevalence in the general population aged 15 to 49 years was 5.6% (UNAIDS http://aidsinfo.unaids.org/; accessed January 2018) Smoking status and history of previous TB treatment among survey participants not collected Tobacco Atlas using tobacco daily in 2015, adults: approximately 8%; 14.9% of males, 1% female (https://tobaccoatlas.org/country/kenya/; accessed April 2020)
Index tests	Symptom screens: - Cough for ≥ 2 weeks - Cough of any duration - Any TB symptom out of 7: cough, chest pain, shortness of breath, weight loss, fever, fatigue, drenching night sweats CXR: direct digital, mobile - TB abnormalities (field reading: infiltrate or consolidation, nodules, cavitary lesion, pleural effusion, hilar or mediastinal lymphadenopathy, linear or interstitial disease (in children) - Any CXR abnormality: as above and also musculoskeletal, cardiac, pleural abnormality, diaphragmatic cp angle blunting, solitary calcified nodules or node CXR TB abnormality or cough for ≥ 2 weeks or both Persons with CXR TB abnormality or cough for ≥ 2 weeks or both were asked to submit sputum
Target condition and reference standard(s)	Bacteriological tests: 2 sputum samples for direct FM microscopy and LJ culture; sample for Xpert MTB/RIF Incorporation: 14.1% received a reference standard. By design, participants with cough of 2 or more weeks or a CXR abnormality or both were eligible for sputum examination. Participants not eligible

Kenya MoH 2018 *(Continued)*

	for sputum or without a sputum result were assumed not to have TB. Case definition: bacteriologically positive TB case = definite = MTB confirmed by culture and/or Xpert 1.3% of true (confirmed) TB cases were symptom-screen negative and chest X-ray exempted
Flow and timing	A total of 9715 were eligible to provide sputum and of these it appears that 9121 (94%) did
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Kenya MoH 2018 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Kimerling 1999

Study characteristics	
Patient Sampling	Screening campaign in 2 homeless shelters. 2 screening rounds were held, 10 months apart, in each location starting in 1996: screenings were done during the evening hours before the assignment of beds to registered clients, which was the only time that their movements could be controlled.
Patient characteristics and setting	Homeless and/or drug users attending 2 shelters in USA Alabama; adults; screening was in 1996 and 1997 Adult males only (n = 127) Mean age 40.8 years (SD 9.4) 9 were screened twice, but did not have TB TB prevalence in study population 310/100,000 population HIV prevalence, history of previous TB, and smoking in population not reported
Index tests	Productive cough (any duration)
Target condition and reference standard(s)	Bacteriological tests: 1 sputum sample, tested with smear microscopy for AFB using Kinyoun's modification, and cultured on solid media (Middlebrook 7H-11 and Löwenstein-Jensen) Case definition: 1 positive culture (with or without smear) Sputum was requested from all participants All (4) cases had CXR abnormalities
Flow and timing	120/127 (94.5%) participants provided a sputum specimen for bacteriological testing; 8 gave a second specimen during subsequent screenings
Comparative	
Notes	

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			High

Kimerling 1999 *(Continued)*

DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Koesoemadinata 2018

Study characteristics	
Patient Sampling	People with diabetes were recruited consecutively from the endocrine outpatients clinic at Hasan Sadikin General Hospital and 25 community health centres. Cross-sectional design. 2 x 2 table available for subset with DICOM files available, which was 346 (43.6%) individuals. DICOM files that were missing had been deleted before being accessed to free up limited electronic storage space on the routine system.
Patient characteristics and setting	People with diabetes aged greater than or equal to 18 years of age; in Bandung, Indonesia, a LMIC; recruited in 2013 Mean age 59.3 years (SD 10.2) 56.4% female TB prevalence in study population 2601/100,000; incidence in Indonesia 319/100,000 HIV prevalence among adults is generally low (at 0.4%) and people with diabetes are not considered to be a high-risk population for HIV in Indonesia 11% had history of previous TB Smoking information not provided
Index tests	Symptom screening: cough for 2 or more weeks CXR, digital, read by radiologist for abnormalities suggestive of TB (probable or possible active TB)
Target condition and reference standard(s)	Bacteriological tests: 2 samples for ZN microscopy and MGIT culture and Xpert MTB/RIF Incorporation: 20% received a reference standard; by design, participants without were assumed not to have TB Case definition: active pulmonary TB - smear, culture, or Xpert-positive or if a pulmonologist decided to treat them for TB No TB cases without signs of active disease at the time of screening
Flow and timing	Not clear who had Xpert testing. Of 70 persons required to give sputum by design, 15 did not (21%) and for an additional 2 (3%) culture was not done.
Comparative	
Notes	Data on CAD were not used

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			

Koesoemadinata 2018 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		

Koesoemadinata 2018 *(Continued)*

Did the study design require all patients to receive a bacteriological reference standard?	No	
Incorporation bias avoided?	No	
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Unclear	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Unclear	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

Law 2015

Study characteristics	
Patient Sampling	Cross-sectional design. National prevalence survey Lao PDR. Participants were selected from 50 clusters of 800 people each. All eligible persons were invited and only those > 15 years who lived there for at least 14 days were eligible.
Patient characteristics and setting	General population of 15 years and older, mixed urban and rural; data collected in 2010 Median age 36 (IQR 24 to 50) 54.8% female HIV prevalence low (0.3% in adults aged 15 to 49 in 2018 https://www.unaids.org) TB prevalence 595 per 100,000 Smoking prevalence not reported; 46.9% of adults using tobacco daily in 2015 (https://tobaccoatlas.org) % with history of previous TB treatment: not provided

Law 2015 *(Continued)*

Index tests	Symptom screen: cough for 2 weeks or more in the past month OR haemoptysis in the past month CXR (conventional mobile); any pulmonary abnormality
Target condition and reference standard(s)	2 sputum samples were tested with smear microscopy (ZN) and with solid culture Incorporation: 16.1% had a reference standard, by design Case definitions: smear-positive OR culture-positive, divided into 'definite' smear-positive pulmonary TB: at least one positive smear plus a positive culture, and 'probable' smear-positive pulmonary TB: at least one positive smear plus a CXR consistent with TB
Flow and timing	Even though there were screen positives who did not receive a bacteriological reference standard, the proportion of these people was small. Also the proportion of people who did not have the same reference standard was small.
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			

Law 2015 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Low risk	

Lewis 2009a

Study characteristics			

Lewis 2009a *(Continued)*

Patient Sampling	Annual medical examination of all employees of a gold mining industry to determine fitness to work. Alternate miners attending their annual medical examination were sampled; permanent employees were included and contractors were excluded.
Patient characteristics and setting	An occupational health centre of a single gold mining company in the Free State Province, South Africa from July 2000 to January 2001. Since all employees require annual screening, repeat screening of the same individuals is common. Only males aged ≥ 20 years, median age 41 years (IQR 20 to 61); n = 1955 TB prevalence in study population 2609 per 100,000; HIV prevalence 29% 9.8% previously treated for TB; 27% had some signs of silicosis Smoking prevalence not provided; Tobacco Atlas 1st edition (2002): 42% in male population in South Africa
Index tests	- Symptom trio: at least 1 of new or worsening cough (any duration), night sweats, or weight loss - New or worsening cough of any duration - New or worsening cough of cough > 2 weeks - New or worsening cough of cough > 3 weeks - At least 1 symptom out of 6 symptom questions asked: new or worsening cough and duration, new or worsening sputum production, haemoptysis, night sweats, fever and duration of fever, and weight loss of more than 5 kg in previous 6 months (mini-CXR was done but not evaluated as a screen in this review). The authors report on 10 combinations.
Target condition and reference standard(s)	Bacteriological tests of all participants: 2 sputum specimens for FM microscopy and LJ culture were collected. Individuals with predefined symptoms, a new or changing radiological lesion on mini-chest radiograph compared with the previous chest radiograph, a positive sputum smear or positive culture were further investigated, with 3 sputum specimens for FM microscopy, LJ culture, and organism identification (and a standard size chest radiograph). Case definition: definite: 1 positive culture and compatible clinical or radiological features, or 2nd positive sputum culture; presumed: compatible clinical features, no response to antibiotics, response to anti-TB treatment, and were smear-positive or new radiological evidence (no presumed cases were identified).
Flow and timing	The first 2 sputum samples, collected during screening, were considered as a screening. If positive on smear, culture, or symptoms or MMR was positive, then the suspect was further evaluated with more sputum cultures, full CXR. Therefore not all cases were culture-positive at the time of screening, but had e.g. an abnormal MMR and a positive culture at the time of follow-up. There is some time in between the first screening and the review.
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns

Lewis 2009a *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			

Lewis 2009a *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Yes		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Lewis 2009b

Study characteristics			
Patient Sampling	See Lewis 2009a; HIV-negative participants only		
Patient characteristics and setting	See Lewis 2009a; HIV-negative participants only		
Index tests	See Lewis 2009a		
Target condition and reference standard(s)	See Lewis 2009a		
Flow and timing	See Lewis 2009a		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		Low risk	

Lewis 2009b *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Yes		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Little 2018

Study characteristics	
Patient Sampling	Routine screening of household contacts, defined as any person living on the same residential plot who shared either the same residential structure or frequent meals with the index case, of consecutively enrolled index cases (recently diagnosed with TB in public clinics) were home visited and eligible to participate
Patient characteristics and setting	Household contacts of all ages, in Vhembe District, a rural municipality in Limpopo Province, South Africa, a UMIC; enrolled in 2013 TB prevalence (bacteriologically confirmed, year) 3940/100,000 population, incidence: 350/100,000 Median age 26 years, IQR 17 to 50; 24% were children under 15 72% female HIV prevalence 16% Smoking: 4% of participants 10% of participants reported previous TB treatment
Index tests	Symptom screens: - Cough of any duration - At least 1 TB symptom out of 6: cough, fever, fatigue, loss of appetite, weight loss, night sweats
Target condition and reference standard(s)	Bacteriological tests: 1 sample for FM microscopy and liquid culture Incorporation: 95.7% received a reference standard; all were eligible Case definition: newly diagnosed TB, confirmed by smear or culture The majority (82%) of TB cases were asymptomatic with no information on CXR (not done). The case definition is based on 1 positive culture and only 2/11 cases reported 1 or more symptoms.
Flow and timing	3 persons with prevalent TB were excluded from analysis, which makes sense. Only 12/279 = 4.3% could not provide sputum (considered TB-negative), which is a small percentage.
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Little 2018 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Little 2018 *(Continued)*

Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Lu 2016

Study characteristics	
Patient Sampling	Contact tracing of close contacts in a school after an outbreak of smear-positive tuberculosis. Probably the whole class of the index case was selected, but this was not clearly described.
Patient characteristics and setting	Adolescents 15 to 17 years old in a middle school in Dalian, China; data collected in 2014 Mean age 15.6 years (SD 0.5) 48.6% female TB prevalence in the population not reported; in the general population (see Cheng 2015) 442/100,000 HIV prevalence not reported, but generally very low in China (HIV/AIDS prevalence in general population 0.037% in 2014; https://www.unaids.org/sites/default/files/country/documents/CHN_narrative_report_2015.pdf) Information on smoking and history of previous TB treatment not reported. In China in 2015 < 1% of 10 to 14-year olds smoked (1.5% in boys, 0.3% in girls) https://tobaccoatlas.org/country/china/; accessed May 2020)
Index tests	CXR. Direct radiography. Posterior-anterior chest DR with digital radiography. Classification of pulmonary tuberculosis signs by Xiwei Lu's standards. Abnormal shadow detection on DR. Based on DR screening and clinical examination 8 TB cases were diagnosed. All 8 had bacteriology of whom 4 had positive bacteriology. The CXR may have been used primarily to diagnose tuberculosis, but this is not clear.
Target condition and reference standard(s)	Bacteriological tests: sputum smear, culture, and Xpert MTB/RIF were carried out. Unclear how many samples were collected/tested: of all 35 students, or only 8 + 6 = 14 with CXR (DR) or CT abnormalities. Unclear which microscopy and culture methods were applied. Case definition: authors report that 1 case had a positive culture, 3 had positive Xpert. We take that as reference standard.

Lu 2016 *(Continued)*

	Incorporation bias: 40% received the reference standard (probably 14/35 had bacteriology; not clear how many samples were tested per person).		
Flow and timing			
Comparative			
Notes			

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Unclear		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	No		
Did the study design require all patients to receive a bacteriological reference standard?	No		

Lu 2016 *(Continued)*

Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Unclear		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Unclear		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Mabuto 2015

Study characteristics	
Patient Sampling	Screening of minibus drivers in portable gazebos in minibus parking bays; voluntary participation
Patient characteristics and setting	Occupational setting: minibus drivers in central Johannesburg, urban South Africa; data collected in 2011 Median age 37 years (IQR 30 to 45); 3% females Among screened population: - TB prevalence not reported - HIV prevalence approximately 14.5% - Smoking 38% - History of previous TB treatment 7%
Index tests	1 or more of 4 TB symptoms: cough, fever, night sweats, weight loss (WHO 4-symptom tool for HIV-positive)
Target condition and reference standard(s)	Bacteriological tests: a single spot sputum sample was collected for fluorescence microscopy and culture on MGIT. Speciation using Hain GenoType MTBDRplus and the Hain GenoType Mycobacterium CM.

Mabuto 2015 *(Continued)*

	Case definition: definite TB = a sputum culture positive for *M tuberculosis* Incorporation: all participants eligible for reference test 2 of 10 cases were asymptomatic; no information on CXR provided
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		

Mabuto 2015 *(Continued)*

Incorporation bias avoided?	Yes	
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Yes	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Malawi MoH 2016

Study characteristics	
Patient Sampling	Cross-sectional national prevalence survey. Multi-stage stratified cluster sampling using PPS of 74 clusters of 500 participants on average, stratified by urban, semi-urban, rural. Of 39,026 eligible residents, 31,579 (81%) participated. Resident definition: slept in the household for at least 14 days prior to the census.
Patient characteristics and setting	General population 15 years and older, in Malawi, a LIC; enrolled in 2013 to 2014 TB prevalence (bacteriologically confirmed, year/survey) 452/100,000 population, incidence (2018): 181/100,000 Median age between 25 and 34 years 58.5% female HIV prevalence: all participants were asked if they had ever been tested for HIV and 63% disclosed their status, 9.3% HIV-positive Smoking (current): 9.6% of survey participants 0.7% of participants reported previous TB treatment
Index tests	Symptom screen: any of the following symptoms for at least 1 week: cough, sputum production, haemoptysis, chest pain, weight loss, night sweats, fatigue, fever or shortness of breath

Malawi MoH 2016 *(Continued)*

	CXR: TB-related abnormalities (conventional mobile equipment) Combination of the presence of the above symptom screen, CXR screen, or both
Target condition and reference standard(s)	Bacteriological tests: 2 samples for FM microscopy and LJ culture Incorporation: 10.6% received a reference standard. By design, participants with any of the 9 symptoms for ≥ 1 week or a TB abnormality on CXR or both were eligible for sputum examination. Participants not eligible for sputum or without a sputum result were assumed not to have TB. Case definition: bacteriologically positive TB case = definite = MTB confirmed by culture and/or Xpert
Flow and timing	Low risk of bias since 97.5% of the 3432 who were eligible for sputum submitted at least 1, and mostly 2, samples and had a result
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Malawi MoH 2016 *(Continued)*

If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		

Malawi MoH 2016 *(Continued)*

Could the patient flow have introduced bias?	Low risk

Melendez 2017

Study characteristics	
Patient Sampling	Cross-sectional design. A subset of the national prevalence survey in Zambia conducted in 2013/2014. Random sampling of 66 clusters in all provinces. Within a cluster all residents were eligible. The subset comprised 52.7% of survey participants, and were those of whom a CXR could be collected and results linked at the time of study. Unclear risk of bias because of the unavailability of many CXRs.
Patient characteristics and setting	General population, urban and rural, of 15 years and older Median age 36 (SD 17) 56.2% was female In study population: - HIV prevalence 6.8% - TB prevalence 455 per 100,000 - Information on smoking and history of TB treatment not provided In Zambia of adults 15+ using tobacco daily in 2015: 16% in men, 2.5% in women (https://tobaccoatlas.org/country/zambia/; accessed May 2020)
Index tests	Digital CXR. index tests: Human reader (central reading by radiologist), abnormalities consistent with TB (Melendez 2017b) Of other index tests reported in this publication the data are taken from Kapata 2016 , which reports on the same original study.
Target condition and reference standard(s)	Bacteriological test: 2 samples tested with liquid culture; positive cultures were speciated with ZN stain and Capillia test Case definition: culture-positive for *M tuberculosis* Incorporation: 16% had a reference standard
Flow and timing	427 index test positives did not receive culture at the end; these persons were excluded from the analysis
Comparative	
Notes	
Methodological quality	

Melendez 2017 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (Symptom questions)			
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Melendez 2017 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Could the patient flow have introduced bias?		High risk

MoH Cambodia 2005

Study characteristics	
Patient Sampling	Cross-sectional - national prevalence survey cluster sampling, 42 clusters selected, persons 10 years or older were eligible. Those who could not appear at the survey site due to sickness, handicaps or old age were home-visited by the census team. For disease prevalence survey (n = 23,084), total 22,160 consenting participants. They interviewed survey participants or their parents in case of children and collected information on name, sex, age, TB history, TB symptoms and health-seeking behaviours. All participants aged 10 and above with suspected TB symptoms such as coughs lasting for 3 weeks or more and/or blood contained sputum were asked to submit sputum specimens regardless of the results of the X-ray screening. Those who could not appear at the survey site due to sickness, handicaps or old age were home-visited by the census team. Transportation to the examination site was arranged for those who could not afford themselves. All those deemed to have an abnormal chest radiograph in the lung field or mediastinum more than a single small calcification nodule or pleural adhesion at the cost phrenic angle were asked to proceed to the Bacteriological Examination Unit immediately. Even when abnormal findings consisted of shadows not compatible with TB such as bronchiolectasis or bronco-pneumonia in the lower lobe, sputum examinations were requested. When they did not agree to visit the site, at least sputum examinations were carried out if they were identified as TB suspects by symptoms. All were interviewed for symptoms and n = 22,012 (99.3%) had a CXR. 2 sputum specimens were collected.
Patient characteristics and setting	General population of 10 years and older; mixed urban and rural; in 2002 (n = 22,160) 47.7% is between 10 and 24 years old 54.2% female History of previous TB treatment reported by 2.1% HIV prevalence 1% (from another survey) Smoking prevalence not reported; Tobacco Atlas 1st ed. 2002: 37% adult smoking
Index tests	Cough any duration Cough for 3 weeks or more

MoH Cambodia 2005 *(Continued)*

	Any symptom out of 8 TB-related shadows or other lung disease Cough for ≥ 3 weeks and/or TB-related shadows or other lung disease Symptom questions asked (8): cough, sputum, sputum with blood, chest pain, lost weight, fatigue, fever, sweat at night, other; and number of days. Since cough for 3 or more weeks determined who was examined by sputum or not, this may have affected the questioning. Conventional mobile CXR. Abnormalities were defined as: TB-related shadow, or other lung disease, i.e. an abnormal chest radiograph in the lung field or mediastinum more than a single small calcification nodule or pleural adhesion at the cost phrenic angle. For this review classified as abnormalities consistent with TB vs not consistent, other lung disease was included in the definition (not exactly 'any abnormality' either since small calcifications and pleural adhesions were considered as abnormal for the classification).
Target condition and reference standard(s)	Only persons with cough of 3 or more weeks, haemoptysis or CXR abnormalities were selected for sputum examination; 14.9% had a reference standard Bacteriological test: 2 specimens tested with ZN microscopy and solid culture (Ogawa and Kudoh mediums) Case definition: 1 or more cultures positive for *M tuberculosis* (with or without positive smear), or 2 positive smear results, or 1 positive smear result with an X-ray result consistent with active tuberculosis Applicability concern because for some smear-positive culture negative cases CXR was part of the case definition
Flow and timing	Some participants with symptoms but no CXR abnormalities were also excluded from sputum collection (n = 419), > 10%
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			

MoH Cambodia 2005 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	No		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		

MoH Cambodia 2005 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Unclear
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

MoH Cambodia 2012

Study characteristics	
Patient Sampling	Cross-sectional - prevalence survey. Target area was whole area of Cambodia. 62 clusters were selected by the population proportionate multi-stage cluster sampling. The study target population included all persons aged 15 years or older who had resided at the selected survey sites for 2 weeks or longer at the time of survey, except for those meeting the exclusion criteria (persons living at military and diplomatic compounds, hospitals and hotels).
Patient characteristics and setting	General population 15 years and older (n = 37,417); the same areas were included in the survey of 2002 (> 10 years earlier) 52.4% were between 15 and 34 years old 54% female 4% reported history of previous TB Smoking percentage not reported; Tobacco Atlas: adult smoking(15+ years old) % using tobacco daily in 2015: 32.1% males, 2.4% females https://tobaccoatlas.org/country/cambodia/; accessed May 2012) HIV prevalence not reported; 0.7% in general population in 2012 (https://aidsdatahub.org/sites/default/files/publication/UNDP_Report_about_HIV_Social_Protection_Schemes.pdf; accessed May 2012)

MoH Cambodia 2012 *(Continued)*

Index tests	8 symptoms questions: cough + duration, sputum, haemoptysis, chest pain, loss of weight, fatigue, fever, night sweats Conventional mobile CXR. The field CXR reader and/or the second reader (the team leader) screened the subjects for eligibility for sputum collection immediately (if any abnormal shadow in the lung field and mediastinum, or pleural effusion except pleural thickness or small single calcification) CXR reading classifications: abnormalities consistent with TB vs not consistent (final reading (includes active and healed TB)); any pulmonary abnormality vs no pulmonary abnormality (field reading)
Target condition and reference standard(s)	Bacteriological tests: 2 samples for FM microscopy and culture; culture method not clear Incorporation: 12.3% received a reference standard. By design, eligibility for sputum examination was if cough for 2 weeks or longer or haemoptysis was present and/or on CXR any abnormal shadow in the lung field or mediastinum other than a single small calcification nodule with a size less than 10 mm or pleural adhesion at costophrenic angle(s). Others were assumed not to have TB. Case definition: bacteriologically positive TB case includes definite (*M tuberculosis* confirmed by culture) and probable (smear-positive on 2 smears, or 1 smear and CXR suggestive of active TB; and MTB not confirmed by culture) The percentage of true (confirmed) TB cases had no signs of active disease at the time of screening not reported
Flow and timing	3.5% of those expected to give sputum, did not give sputum; % with missing results unclear
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			

MoH Cambodia 2012 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

MoH Cambodia 2012 *(Continued)*

Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Unclear risk

MoH Indonesia 2015

Study characteristics	
Patient Sampling	Cross-sectional national prevalence survey. Multi-stage stratified cluster sampling using PPS of 156 clusters of 500 participants on average, stratified by island and urban vs rural. Of 76,576 eligible residents, 67,944 (89%) participated. Resident definition: lived in the household for at least 1 month prior to the census
Patient characteristics and setting	General population 15 years and older, in Indonesia, a LMIC, enrolled in 2013-2014 TB prevalence (bacteriologically confirmed, year/survey) 759/100,000 population, incidence (2018): 316/100,000 Median age between 35 and 44 years 53.4% female HIV prevalence in the general population aged 15 to 49 years was 0.4% (UNAIDS http://aidsinfo.unaids.org/; accessed April 2017) Smoking: 40.4% of participants in a KAP survey among persons who screened positive 3.2% of participants reported previous TB treatment
Index tests	Symptom screen: cough for ≥ 2 weeks and/or haemoptysis (survey screen) CXR (digital), abnormality in lung or pleura Combination: cough for ≥ 2 weeks and/or haemoptysis OR abnormal CXR (any pulmonary abnormality) OR both
Target condition and reference standard(s)	Bacteriological tests: 2 sputum samples for ZN microscopy and LJ culture (2 samples cultured in 52 clusters, 1 sample in 104 clusters). Xpert MTB/RIF for smear-positive samples or inconclusive culture. Incorporation: 21.7% received a reference standard. By design, participants not eligible for sputum or without a sputum result were assumed not to have TB.

MoH Indonesia 2015 *(Continued)*

	Case definition: definite: MTB confirmed by culture and/or Xpert; probable: for 6 out of 7, cultures were identified by niacin but not MPT64, with chest X-ray suggestive of TB. One case was a pregnant participant who was Xpert-positive, but culture contaminated.
Flow and timing	Of the 67,944 participants, 15,446 were eligible to give sputum and, of these, 15,141 submitted at least 1 sputum (98%) and 14,604 submitted 2 specimens (95%). At least 1 culture result available 14,773 (96%).
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

MoH Indonesia 2015 *(Continued)*

DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Unclear		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

MoH Myanmar 2012

Study characteristics			

MoH Myanmar 2012 *(Continued)*

Patient Sampling	Cross-sectional national TB prevalence survey 2009-2010. Multi-staged cluster sampling of 70 clusters, providing 57,607 persons 15 years and older selected. 51,367 (89%) agreed to participate. All were interviewed and eligible for CXR.
Patient characteristics and setting	General population (n = 51,367) 15 years or older; median between 35 and 44 56% female TB prevalence in study population 613/100,000 History of previous TB treatment reported by 2.9% Smoking reported by 28.7% HIV prevalence in study population not measured (9 participants self-reported HIV infection). In 1999 a peak prevalence in Myanmar of ± 2.7% in 1999 according to HIV prevalence data in antenatal care clinics (http://www.unaids.org/en/dataanalysis/monitoringcountryprogress/2010progressreportssubmittedbycountries/myanmar_2010_country_progress_report_en.pdf).
Index tests	Symptoms screens: - Cough of any duration - Cough 3 or more weeks - Any 1 symptom out of 8: cough, expectoration, blood in sputum, weight loss, low-grade fever, chest pain, others, within the last month, and duration - TB suspect means a person who has cough more than 3 weeks and/or have haemoptysis CXR screening (conventional, mobile). Field readers (general medical officer) and specialized readers at central level. Classifications: - Abnormalities consistent with TB vs not consistent, at central level - Any abnormality vs normal (includes other visible organs) at field reading level Symptom and CXR combination: cough 3 weeks or/and CXR suggestive of TB
Target condition and reference standard(s)	Bacteriological test: 2 samples for solid culture (Ogawa) and FM smear microscopy Incorporation: persons with cough 3 or more weeks and/or haemoptysis or any lesion in the lung fields or mediastinum on CXR, and those without CXR were eligible for sputum collection. Those were 12,235 (23%). Participants ineligible for sputum are assumed not to have TB. Case definition: 2 positive smears or 1 positive smear with TB-CXR, or 1 positive culture

MoH Myanmar 2012 *(Continued)*

	Of all cases 12 (4%) were among those without CXR, who did not report symptoms
Flow and timing	Of 12,235 participants eligible for sputum collection, 11,587 gave sputum (99.3%); 433 (4%) had 1 sputum result only, instead of 2
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			

MoH Myanmar 2012 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	Unclear		
Could the patient flow have introduced bias?		Unclear risk	

Mongolia MoH 2016

Study characteristics			

Mongolia MoH 2016 *(Continued)*

Patient Sampling	Cross-sectional national prevalence survey. Multi-stage stratified cluster sampling using PPS of 96 clusters of 600 (city)/500 (other) participants on average, stratified by city, provincial centre, rural. Of 60,031 eligible residents, 50,309 (84%) participated. Resident definition: slept in the household for 14 days prior to the census.
Patient characteristics and setting	General population 15 years and older, in Mongolia, a LMIC, enrolled in 2014 to 2015 TB prevalence (bacteriologically confirmed, year/survey) 560/100,000 population, incidence (2018): 428/100,000 Median age between 35 and 44 years 60% female In 2014, the prevalence of HIV in the general population aged 15 to 49 years was < 0.1% (UNAIDS http://aidsinfo.unaids.org/; accessed April 2017) Smoking: 24.4% of survey participants 4.2% of participants reported previous TB treatment
Index tests	A symptom questionnaire with 9 questions was applied: cough, sputum production, haemoptysis, weight loss, fever, chest pain, short breath, loss of appetite, night sweats Symptom screens: - Cough of any duration - Cough for ≥ 2 weeks (eligible for sputum in survey) - Any TB symptom out of 9 CXR (direct digital): any lung abnormality CXR or cough for ≥ 2 weeks or both
Target condition and reference standard(s)	Bacteriological tests: 2 sputum samples for FM microscopy and Ogawa culture; Xpert MTB/RIF done for smear-positive specimens only Incorporation: 18.9% received a reference standard. By design, participants with cough of 2 or more weeks or a CXR abnormality or both were eligible for sputum examination. Participants not eligible for sputum or without a sputum result were assumed not to have TB. Case definition: bacteriologically positive TB case includes definite n = 245 (MTB confirmed by culture and/or Xpert), and probable n = 3 (MTB not bacteriologically confirmed but CXR suggestive of TB, or scanty culture-positive and clinical confirmation 3/248 (1.2%) cases were symptom-negative and CXR exempt
Flow and timing	Of 10,359 who were eligible for sputum 9546 (92%) submitted at least 1 specimen and 9473 (91%) submitted 2; 9527 (92%) had at least 1 culture result available

Mongolia MoH 2016 *(Continued)*

Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Mongolia MoH 2016 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Moosazadeh 2015

Study characteristics	
Patient Sampling	Cross-sectional design, contact tracing of family members of smear-positive PTB cases, sampling was based on census. A household close contact was someone who has lived for more than 30 days in the same house as an index case.
Patient characteristics and setting	Household contacts of 6 years and older, in Mazandaran province, Iran, a UMIC, enrolled in 2010 TB prevalence (study) 900/100,000 population, incidence: 14/100,000

Moosazadeh 2015 *(Continued)*

	Mean age 36.9 years (SD 19.6) 55.5% female Information on HIV, smoking and history of previous TB not provided UNAIDS: HIV prevalence - the percentage of people living with HIV - among adults (15 to 49 years) was 0.1% in 2018 (https://www.unaids.org/en/regionscountries/countries/islamicrepublicofiran; accessed April 2021) Smoking Atlas, 3rd edition: cigarette use in men 24.0%, in women 1.9%
Index tests	Cough more than 2 weeks
Target condition and reference standard(s)	Bacteriological tests: 2 samples for ZN smear microscopy on all patients, culture only if one smear was positive Incorporation: 100% received a bacteriological reference standard as above Case definition: bacteriologically positive TB case (1 positive sputum smear AND 1 positive culture) OR at least 2 positive sputum smears The percentage of true (confirmed) TB cases without signs of active disease at the time of screening is not clear. 3 SSM+ had cough. 3 SSM+ had weight loss. Not clear if those are the same persons.
Flow and timing	Differential verification: smear-positives received culture; of smear-negatives only 10%
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			

Moosazadeh 2015 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	No		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Yes		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

MoPH DPRK 2017

Study characteristics	
Patient Sampling	Cross-sectional national prevalence survey. Multi-stage stratified cluster sampling using PPS of 100 clusters of 700 participants on average, stratified by urban vs rural. 5 clusters were replaced for reasons of inaccessibility. Of 71,877 eligible residents, 60,683 (84.4%) participated. Resident definition: registered in the living administrative unit for at least 2 weeks before the census.
Patient characteristics and setting	General population 15 years and older, in Democratic People's Republic of Korea, a LIC, enrolled in 2015 to 2016 TB prevalence (bacteriologically confirmed, year/survey) 641/100,000 population, incidence (2018): 513/100,000 Median age between 40 and 44 years 55.5% female Information on HIV not provided, as there have been no HIV cases detected in the last decade Smoking: 12.7% of survey participants 2.4% of participants reported previous TB treatment
Index tests	Symptom screen: cough for ≥ 15 days and/or haemoptysis Chest X-ray (conventional) any pulmonary abnormality; abnormal chest radiograph in the lung field or mediastinum other than a single small calcification nodule with a size less than 10 mm or pleural adhesion at cost-phrenic angle(s) Combination of cough for ≥ 15 days and/or haemoptysis or CXR abnormality as above or both
Target condition and reference standard(s)	Bacteriological tests: 2 samples for FM microscopy and LJ culture Incorporation: 7.6% received a reference standard. By design, participants with cough for ≥ 15 days and/or haemoptysis or a CXR abnormality or both were eligible for sputum examination. Participants not eligible for sputum or without a sputum result were assumed not to have TB. Case definition: bacteriologically positive TB case includes *M tuberculosis* confirmed by culture (positive or scanty); smear-positive culture negative as judged by the survey expert committee.
Flow and timing	A total of 4802 people were eligible for sputum examination, of whom 4462 (92.5%) submitted at least 1 sputum specimen and 4412 (93%) submitted 2 sputum specimens. At least 1 culture result available 4586 (95.5%). This is quite high, so low risk of bias.
Comparative	
Notes	On other symptom questions (cough, sputum, haemoptysis, breathless, fever, no appetite, ineffectualness, night sweating, weight loss) 2 x 2 data were not available

MoPH DPRK 2017 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

MoPH DPRK 2017 *(Continued)*

DOMAIN 3: Reference Standard		
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes	
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes	
Did the study design require all patients to receive a bacteriological reference standard?	No	
Incorporation bias avoided?	No	
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

MoPH Thailand 2017

Study characteristics	
Patient Sampling	Cross-sectional nationwide prevalence survey, except metropolitan Bangkok. Multi-stage stratified cluster sampling using PPS of 83 clusters of 900 participants on average, stratified by urban vs rural. Of 78,839 eligible residents, 62,536 (79%) participated. Resident definition: permanent residents according to household registration, or temporary residents or nonresidents who had slept in the household for at least 2 weeks prior to the census. The exclusion of Bangkok affects the national prevalence estimate, but does not necessarily imply bias or inappropriate exclusion for the purpose of accuracy of the screening tools.

MoPH Thailand 2017 *(Continued)*

Patient characteristics and setting	General population 15 years and older, in Thailand, a UMIC, enrolled in 2012 to 2013 TB prevalence (bacteriologically confirmed, year/survey) 242/100,000 population, incidence (2015): 172/100,000 Median age approximately 45 years 56.5% female HIV prevalence in the general population aged 15 to 49 years was 1.2% (UNAIDS http://aidsinfo.unaids.org/; accessed April 2017) Smoking and previous TB treatment not reported Smoking in 2012: 42.3% in males and 2.4% in females 15+ years (https://apps.who.int/gho/data/node.main.65; accessed 14 May 2020)
Index tests	Symptom screens: - Cough of any duration - Cough greater than or equal to 2 weeks In the survey a scoring system was used: total score ≥ out of cough ≥ 2 weeks (3 points), haemoptysis over the past month (3 points), cough < 2 weeks (2 points), weight loss in the past month (1 point), fever ≥ 1 week in the past 2 weeks (1 point), night sweats in the past month (1 point) CXR: any lung abnormality Combination as used in the survey: clinical score greater than or equal to 3 OR clinical score = 1 & CXR exempt OR abnormal CXR (any pulmonary abnormality) This combination was specific for the survey and may be different from a public health screening programme
Target condition and reference standard(s)	Bacteriological tests: 2 sputum samples for ZN microscopy and Ogawa culture Incorporation: 9.3% received a reference standard. By design, participants not eligible for sputum or without a sputum result were assumed not to have TB. Case definition: definite: MTB confirmed by culture; probable: MTB not confirmed by culture, but at least 1 smear-positive with chest X-ray suggestive of TB, or 2 smear-positive, or 1 smear-positive and confirmed as TB cases by referral health facilities. There are 12 (8.4%) probable cases.
Flow and timing	Of 6050 eligible to give sputum, 5988 did and 5821 (96.2%) had at least 1 sputum result available
Comparative	
Notes	
Methodological quality	

MoPH Thailand 2017 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 3: Reference Standard			

MoPH Thailand 2017 *(Continued)*

Is the bacteriological reference standard likely to correctly classify the target condition?	Yes	
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes	
Did the study design require all patients to receive a bacteriological reference standard?	No	
Incorporation bias avoided?	No	
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Mor 2012

Study characteristics	
Patient Sampling	All Jewish Ethiopian immigrants to Israel were screened for PTB in Addis Ababa since June 2001, 2 to 3 weeks before air-travel. All non-pregnant immigrants older than 1 year have a CXR. Each immigrant also completes a symptoms questionnaire (screened immigrants were asked to respond whether they suffered from prolonged cough, lasting 3 weeks, haemoptysis, chest pain, fever (38 °C), night sweats, and weight loss), and underwent physical examination and a one-step tuberculin skin test (TST)
Patient characteristics and setting	Jewish Ethiopian immigrants to Israel, screened in Ethiopia (n = 13,379) Age from 1 year and older, mean age and gender distribution not provider for population (only for TB cases: mean 34 (SD 25) years 53% female

Mor 2012 *(Continued)*

	TB prevalence in study population 321 per 100,000 HIV prevalence: not exactly reported, but approximately 2% Smoking not reported for study population; in Ethiopia, 2000: prevalence of smoking among persons 15+: 9.0% in males, 0.6% in females
Index tests	CXR (postero-anterior); all films are read by the radiography department in Carmel Hospital in Haifa, Israel, for CXR showing changes suggestive of PTB
Target condition and reference standard(s)	Bacteriological tests and incorporation: 3 samples for LJ and ZN requested from individuals who had previously been treated for tuberculosis, have a positive response in the symptoms questionnaire, and those whose CXR shows changes suggestive of PTB, resulting in 1.1% of participants being examined by bacteriological tests. Others are assumed not to have TB. Case definition: active PTB: symptomatic patient with pulmonary disease and confirmed MTB complex culture.
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (Symptom questions)			
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	

Mor 2012 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Unclear		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Morasert 2018

Study characteristics	
Patient Sampling	Cross-sectional study; all prison inmates during the study were included and screened using both a TB screening questionnaire and CXR
Patient characteristics and setting	Prison inmates. Age criteria not provided, in Suratthani Province, South Thailand, a UMIC, enrolled in 2015 TB prevalence (bacteriologically confirmed) 2096/100,000 population, incidence: 172/100,000

Morasert 2018 *(Continued)*

	Median age 32.9 years (SD 9.1) 15.3% female HIV prevalence in study population 7.5% Smoking: 70% of participants % with previous TB treatment not reported
Index tests	Symptom screens: score ≥ 3 points out of the prison screening questionnaire composed of 8 questions. Scores were rated as follows: 3 points for any history of previous anti-tuberculosis treatment, cough 72 weeks, haemoptysis in the past month or cervical lymphadenopathy 72 cm; 2 points for cough, 2 weeks; 1 point for intermittent or persistent fever in the past month, weight loss, 5% of body weight in the past month or night sweats in the past month; and 0 points for none of the above. CXR: i) any abnormality (study category 1, 2, 3) and ii) abnormality suggestive of TB (study category 2, 3) Symptom and CXR combined: i) score ≥ 3 points or CXR any abnormality or both, ii) score ≥ 3 points or abnormality suggestive of TB or both
Target condition and reference standard(s)	Bacteriological tests: 2 samples for ZN microscopy and Xpert MTB/RIF Incorporation: 17% received a reference standard. By design, participants with CXR- and QS- had no bacteriological test and were assumed not to have TB. Case definition included definite cases and probable. Definite case defined as sputum-positive on microscopy or Xpert. Probable case: CXR consistent with active TB (Category 2 or 3) but in whom the criteria for a definite case were not met, together with clinical and CXR improvement after treatment. 23/84 = 27% of cases were probable. Definite cases include high proportion with smear only.
Flow and timing	High proportion of those eligible for sputum collection did not have results. Of 25 transferred, no sputum collected. Of the 900 screening positives (QS+ and/or CXR+) 680 received AFB results and 426 Xpert.
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		

Morasert 2018 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	No		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	No		
Did the study design require all patients to receive a bacteriological reference standard?	No		

Morasert 2018 *(Continued)*

Incorporation bias avoided?	No	
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		High
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

Morishita 2017a

Study characteristics	
Patient Sampling	Mass screening campaign. Morishita 2017a includes the rural poor communities (n = 12,907), who voluntarily visited the mobile unit. The entire ACF targeted 5 vulnerable populations including: 1) residents in rural poor communities, 2) residents in urban poor communities, 3) prison inmates, 4) indigenous population, and 5) high school students. Selection differs somewhat per population. In the targeted prisons and high schools, all eligible persons were screened by the project. In rural and urban poor communities and indigenous population, only those who voluntarily visited the mobile unit were screened. Of 23 municipalities and one city in Palawan Province, 7 municipalities were selected based on health access and the level of TB case detection, and classified as rural poor communities. Additionally, of 66 barangays (the smallest administrative division) in Puerto Princesa City, three barangays were selected based on the income levels in the official statistics, and classified as urban poor communities. The project also targeted 2 major indigenous populations, inmates from 6 prison facilities (2 jails and the national prison that has 4 sub-colonies located in Puerto Princesa City), and students from 8 high schools.
Patient characteristics and setting	A rural poor population 15 years and older, Palawan Province and Puerto Princesa City, Philippines, a LMIC, enrolled in 2012 TB prevalence (study) 2100/100,000 population, incidence: 522/100,000 Overall: median age 46.0 years (IQR 33 to 59) 62.4% female

Morishita 2017a *(Continued)*

	In 2015, the prevalence of HIV in the general population aged 15 to 49 years was < 0.1% (UNAIDS http://aidsinfo.unaids.org/; accessed May 2017) Smoking: 29.0% of survey participants 9.4% of participants reported previous TB treatment
Index tests	Symptom screens: The report is a retrospective review of the screening data, so in theory the symptom combinations could have been decided post-hoc. However, the extracted index tests that (cough for 2 or more weeks, any TB symptoms) are so obvious, that bias is unlikely. Symptom questions: cough (any duration), cough (2 or more weeks), fever, night sweats, weight loss
Target condition and reference standard(s)	Bacteriological tests: 2 sputum samples for LED-FM microscopy and 1 for Xpert MTB/RIF Persons were clinically assessed by the physician, mostly based on abnormal CXR findings, and additionally on the presence of TB symptoms. TB suspects provided 2 spot sputum specimens, for LED-FM microscopy and Xpert MTB/RIF testing. Case definition: bacteriologically positive TB case a positive result by smear microscopy and/or Xpert (MTB detected), or strong clinical evidence if tests were negative. Incorporation: 16.1% received a reference standard by design, the screening physicians selected persons who would get a reference test based on presence of CXR abnormalities, and in addition clinical symptoms. Incorporation bias is likely considerable, as persons with a normal CXR seem not likely to have been investigated with bacteriological tests. 10/284 (3.5%) of TB cases were bacteriologically negative, but diagnosed clinically.
Flow and timing	Overall, for all populations: 1) the definition of who would be investigated with bacteriological tests (= suspected TB) is unclear, so we cannot assess if people were left out in any systematic way; 2) CXR missing in 7?; 3) of n = 5225 with suspected TB, 4204 (81.4%) had smear examination; 5165 (98.9%) had Xpert MTB/RIF
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	

Morishita 2017a *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it prespecified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	No		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	No		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			

Morishita 2017a *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Morishita 2017b

Study characteristics	
Patient Sampling	Mass screening campaign. Morishita 2017b includes the urban poor communities (n = 1625), who voluntarily visited the mobile unit. The entire ACF targeted 5 vulnerable populations including: 1) residents in rural poor communities, 2) residents in urban poor communities, 3) prison inmates, 4) indigenous population, and 5) high school students. Selection differs somewhat per population. In the targeted prisons and high schools, all eligible persons were screened by the project. In rural and urban poor communities and indigenous population, only those who voluntarily visited the mobile unit were screened. Of 23 municipalities and 1 city in Palawan Province, 7 municipalities were selected based on health access and the level of TB case detection, and classified as rural poor communities. Additionally, of 66 barangays (the smallest administrative division) in Puerto Princesa City, 3 barangays were selected based on the income levels in the official statistics, and classified as urban poor communities. The project also targeted 2 major indigenous populations, inmates from 6 prison facilities (2 jails and the national prison that has 4 subcolonies located in Puerto Princesa City), and students from 8 high schools.
Patient characteristics and setting	A urban poor population 15 years and older, Palawan Province and Puerto Princesa City, Philippines, a LMIC, enrolled in 2012 TB prevalence (study) 1800/100,000 population, incidence: 522/100,000 Overall: median age 40.0 years (IQR 26 to 52) 55.7% female In 2015, the prevalence of HIV in the general population aged 15 to 49 years was < 0.1% (UNAIDS http://aidsinfo.unaids.org/; accessed May 2017) Smoking: 35.9% of survey participants 7.4% of participants reported previous TB treatment
Index tests	Symptom screens:

Morishita 2017b *(Continued)*

	The report is a retrospective review of the screening data, so in theory the symptom combinations could have been decided post-hoc. However, the extracted index tests (cough for 2 or more weeks, any TB symptoms) are so obvious, that bias is unlikely. Symptom questions: cough (any duration), cough (2 or more weeks), fever, night sweats, weight loss
Target condition and reference standard(s)	Bacteriological tests: 2 sputum samples for LED-FM microscopy and 1 for Xpert MTB/RIF Persons were clinically assessed by the physician, mostly based on abnormal CXR findings, and additionally on the presence of TB symptoms. TB suspects provided 2 spot sputum specimens, for LED-FM microscopy and Xpert MTB/RIF testing. Case definition: bacteriologically positive TB case a positive result by smear microscopy and/or Xpert (MTB detected), or strong clinical evidence if tests were negative. Incorporation: 22.5% received a reference standard by design, the screening physicians selected persons who would get a reference test based on presence of CXR abnormalities, and in addition clinical symptoms. Incorporation bias is likely considerable, as persons with a normal CXR seem not likely to have been investigated with bacteriological tests. 4/34 (12%) of TB cases were bacteriologically negative, but diagnosed clinically.
Flow and timing	Overall, for all populations: 1) the definition of who would be investigated with bacteriological tests (= suspected TB) is unclear, so we cannot assess if people were left out in any systematic way; 2) CXR missing in 7?; 3) of n = 5225 with suspected TB, 4204 (81.4%) had smear examination; 5165 (98.9%) had Xpert MTB/RIF
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Morishita 2017b *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	No		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	No		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		

Morishita 2017b *(Continued)*

Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Morishita 2017c

Study characteristics	
Patient Sampling	Mass screening campaign. Morishita 2017c includes the prison inmates (n = 6133); all eligible persons were screened The entire ACF targeted 5 vulnerable populations including: 1) residents in rural poor communities, 2) residents in urban poor communities, 3) prison inmates, 4) indigenous population, and 5) high school students. Selection differs somewhat per population. In the targeted prisons and high schools, all eligible persons were screened by the project. In rural and urban poor communities and indigenous population, only those who voluntarily visited the mobile unit were screened. Of 23 municipalities and one city in Palawan Province, seven municipalities were selected based on health access and the level of TB case detection, and classified as rural poor communities. Additionally, of 66 barangays (the smallest administrative division) in Puerto Princesa City, 3 barangays were selected based on the income levels in the official statistics, and classified as urban poor communities. The project also targeted 2 major indigenous populations, inmates from 6 prison facilities (2 jails and the national prison that has 4 sub-colonies located in Puerto Princesa City), and students from 8 high schools.
Patient characteristics and setting	Prison inmates 15 years and older, Palawan Province and Puerto Princesa City, Philippines, n LMIC, enrolled in 2012 TB prevalence (study) 6000/100,000 population, incidence: 522/100,000 Overall: median age 41.0 years (IQR 33 to 49) 1.1% female In 2015, the prevalence of HIV in the general population aged 15 to 49 years was < 0.1% (UNAIDS http://aidsinfo.unaids.org/; accessed May 2017) Smoking: 82.4% of survey participants 15.2% of participants reported previous TB treatment
Index tests	Symptom screens: The report is a retrospective review of the screening data, so in theory the symptom combinations could have been decided post-hoc. However, the extracted index tests (cough for 2 or more weeks, any TB symptoms) are so obvious, that bias is unlikely. Symptom questions: cough (any duration), cough (2 or more weeks), fever, night sweats, weight loss
Target condition and reference standard(s)	Bacteriological tests: 2 sputum samples for LED-FM microscopy and 1 for Xpert MTB/RIF Persons were clinically assessed by the physician, mostly based on abnormal CXR findings, and additionally on the presence of TB symptoms. TB suspects provided 2 spot sputum specimens, for LED-FM microscopy and Xpert MTB/RIF testing.

Morishita 2017c *(Continued)*

	Case definition: bacteriologically positive TB case a positive result by smear microscopy and/or Xpert (MTB detected), or strong clinical evidence if tests were negative. Incorporation: 39% received a reference standard by design, the screening physicians selected persons who would get a reference test based on presence of CXR abnormalities, and in addition clinical symptoms. Incorporation bias is likely considerable, as persons with a normal CXR seem not likely to have been investigated with bacteriological tests. 12/378 (3.2%) of TB cases were bacteriologically negative, but diagnosed clinically.
Flow and timing	Overall, for all populations: 1) the definition of who would be investigated with bacteriological tests (= suspected TB) is unclear, so we cannot assess if people were left out in any systematic way; 2) CXR missing in 7?; 3) of n = 5225 with suspected TB, 4204 (81.4%) had smear examination; 5165 (98.9%) had Xpert MTB/RIF
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it prespecified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			

Morishita 2017c *(Continued)*

DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	No		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	No		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Morishita 2017d

Study characteristics			

Morishita 2017d *(Continued)*

Patient Sampling	Mass screening campaign. Morishita 2017d includes the indigenous population (n = 2145) who voluntarily visited the mobile unit. The entire ACF targeted 5 vulnerable populations including: 1) residents in rural poor communities, 2) residents in urban poor communities, 3) prison inmates, 4) indigenous population, and 5) high school students. Selection differs somewhat per population. In the targeted prisons and high schools, all eligible persons were screened by the project. In rural and urban poor communities and indigenous population, only those who voluntarily visited the mobile unit were screened. Of 23 municipalities and 1 city in Palawan Province, 7 municipalities were selected based on health access and the level of TB case detection, and classified as rural poor communities. Additionally, of 66 barangays (the smallest administrative division) in Puerto Princesa City, 3 barangays were selected based on the income levels in the official statistics, and classified as urban poor communities. The project also targeted 2 major indigenous populations, inmates from 6 prison facilities (2 jails and the national prison that has 4 subcolonies located in Puerto Princesa City), and students from 8 high schools.
Patient characteristics and setting	Indigenous 15 years and older, Palawan Province, Philippines, n LMIC, enrolled in 2012 TB prevalence (study) 2900/100,000 population, incidence: 522/100,000 Overall: median age 44.0 years (IQR 32 to 58) 64.1% female In 2015, the prevalence of HIV in the general population aged 15 to 49 years was < 0.1% (UNAIDS http://aidsinfo.unaids.org/; accessed May 2017) Smoking: 32.6% of survey participants 8.9% of participants reported previous TB treatment
Index tests	Symptom screens: The report is a retrospective review of the screening data, so in theory the symptom combinations could have been decided post-hoc. However, the extracted index tests (cough for 2 or more weeks, any TB symptoms) are so obvious, that bias is unlikely. Symptom questions: cough (any duration), cough (2 or more weeks), fever, night sweats, weight loss
Target condition and reference standard(s)	Bacteriological tests: 2 sputum samples for LED-FM microscopy and 1 for Xpert MTB/RIF Persons were clinically assessed by the physician, mostly based on abnormal CXR findings, and additionally on the presence of TB symptoms. TB suspects provided 2 spot sputum specimens, for LED-FM microscopy and Xpert MTB/RIF testing. Case definition: bacteriologically positive TB case a positive result by smear microscopy and/or Xpert (MTB detected), or strong clinical evidence if tests were negative. Incorporation: 15.1% received a reference standard by design, the screening physicians selected persons who would get a reference test based on presence of CXR abnormalities, and in addition clinical symptoms. Incorporation bias is likely considerable, as persons with a normal CXR seem not likely to have been investigated with bacteriological tests. None of the TB cases were bacteriologically negative, but diagnosed clinically.
Flow and timing	Overall, for all populations: 1) the definition of who would be investigated with bacteriological tests (= suspected TB) is unclear, so we cannot assess if people were left out in any systematic way; 2) CXR missing in 7?; 3) of n = 5225 with suspected TB, 4204 (81.4%) had smear examination; 5165 (98.9%) had Xpert MTB/RIF

Morishita 2017d *(Continued)*

Comparative
Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it prespecified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	No		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	No		

Morishita 2017d *(Continued)*

Did the study design require all patients to receive a bacteriological reference standard?	No	
Incorporation bias avoided?	No	
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

Morishita 2017e

Study characteristics	
Patient Sampling	Mass screening campaign. Morishita 2017e includes the high school students (n = 2293); all eligible persons were screened The entire ACF targeted 5 vulnerable populations including: 1) residents in rural poor communities, 2) residents in urban poor communities, 3) prison inmates, 4) indigenous population, and 5) high school students. Selection differs somewhat per population. In the targeted prisons and high schools, all eligible persons were screened by the project. In rural and urban poor communities and indigenous population, only those who voluntarily visited the mobile unit were screened. Of 23 municipalities and 1 city in Palawan Province, 7 municipalities were selected based on health access and the level of TB case detection, and classified as rural poor communities. Additionally, of 66 barangays (the smallest administrative division) in Puerto Princesa City, 3 barangays were selected based on the income levels in the official statistics, and

Morishita 2017e *(Continued)*

	classified as urban poor communities. The project also targeted 2 major indigenous populations, inmates from 6 prison facilities (2 jails and the national prison that has 4 sub-colonies located in Puerto Princesa City), and students from 8 high schools.
Patient characteristics and setting	High school students, Palawan Province and Puerto Princesa City, Philippines, a LMIC, enrolled in 2012 TB prevalence (study) 200/100,000 population, incidence: 522/100,000 Overall: median age 16.0 years (IQR 15 to 16) 58.3% female. In 2015, the prevalence of HIV in the general population aged 15 to 49 years was < 0.1% (UNAIDS http://aidsinfo.unaids.org/; accessed May 2017) Smoking: 13.2% of survey participants 1.1% of participants reported previous TB treatment
Index tests	Symptom screens: The report is a retrospective review of the screening data, so in theory the symptom combinations could have been decided post-hoc. However, the extracted index tests (cough for 2 or more weeks, any TB symptoms) are so obvious, that bias is unlikely. Symptom questions: cough (any duration), cough (2 or more weeks), fever, night sweats, weight loss
Target condition and reference standard(s)	Bacteriological tests: 2 sputum samples for LED-FM microscopy and 1 for Xpert MTB/RIF Persons were clinically assessed by the physician, mostly based on abnormal CXR findings, and additionally on the presence of TB symptoms. TB suspects provided 2 spot sputum specimens, for LED-FM microscopy and Xpert MTB/RIF testing. Case definition: bacteriologically positive TB case a positive result by smear microscopy and/or Xpert (MTB detected), or strong clinical evidence if tests were negative. Incorporation: 3.1% received a reference standard by design, the screening physicians selected persons who would get a reference test based on presence of CXR abnormalities, and in addition clinical symptoms. Incorporation bias is likely considerable, as persons with a normal CXR seem not likely to have been investigated with bacteriological tests. None of the TB cases were bacteriologically negative, but diagnosed clinically.
Flow and timing	Overall, for all populations: 1) the definition of who would be investigated with bacteriological tests (= suspected TB) is unclear, so we cannot assess if people were left out in any systematic way; 2) CXR missing in 7?; 3) of n = 5225 with suspected TB, 4204 (81.4%) had smear examination; 5165 (98.9%) had Xpert MTB/RIF
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Morishita 2017e *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	No		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	No		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			

Morishita 2017e *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

Muyoyeta 2017

Study characteristics	
Patient Sampling	Cross-sectional design; household contacts of consecutive index bacteriologically confirmed TB cases Inclusion of household contacts of 'certain patients diagnosed with TB' who presented at the government health facility. The index patients are consecutively enrolled, but for the household contacts this is unclear. Investigators seem to have included every eligible participant consecutively; only 919 of 4297 eligible participants presented to the health facilities and 54 people who did not get CXR (for unknown reasons) were further discounted. Unclear whether the 865 people includable in the analyses are representative of the 4297 eligible ones.
Patient characteristics and setting	Household contacts of all ages, in Lusaka, Zambia, a LMIC, enrolled in 2013 TB prevalence not reported, incidence (2010): 361/100,000 Median age 15 years (IQR 6 to 31) 59.7% female HIV prevalence 23.1% in study population subjected to CAD4TB; data on symptom screens available for a subset with different HIV prevalences Previous TB treatment information not provided Prevalence of smoking among those 15+ in 2015: 26.5% in males, 4.6% in females (unknown in 13 to 15 years) (https://apps.who.int/gho/data/node.main.65; accessed 14 May 2020)

Muyoyeta 2017 *(Continued)*

Index tests	Symptom screening: 1) Any TB symptom of 4: cough any duration OR night sweats OR weight loss OR fever 2) Cough equal to or more than 2 weeks Computer-assisted reading of digital CXR (CAD4TB version 1.08). Set threshold defining normal CXR (CAD less than 61) and abnormal CXR (CAD score greater than or equal to 61).
Target condition and reference standard(s)	Bacteriological tests: 1 sputum samples for Xpert if CXR positives, or 2 samples for FM microscopy (if CXR negatives) Incorporation: 100% received a reference standard Case definition: Xpert positive or LED fluorescent microscopy (FM) positive
Flow and timing	CXR positives got Xpert, while CXR negatives got FM smear (differential verification). Persons who did not submit a sputum were excluded from the analysis (partial verification). For the index tests 1) any symptoms and 2) cough equal to or more than 2 weeks; only a subpopulation was used based on HIV status.
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Muyoyeta 2017 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	No		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	No		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Nair 2016a

Study characteristics	
Patient Sampling	Cross-sectional study selecting all persons attending the diabetes clinic of 1 hospital in New Delhi

Nair 2016a *(Continued)*

Patient characteristics and setting	Diabetes patients attending a tertiary care hospital, in New Delhi, India, n LMIC, enrolled in 2014 TB prevalence 256/100,000 population, incidence: 211/100,000 Mean age 28.2 years (SD 11.2) 49.7% female HIV prevalence not measured in study. In India HIV prevalence among adults (15 to 49 years) was 0.2% (UNAIDS https://www.unaids.org/en/regionscountries/countries/india; accessed 14 April 2021) Smoking: 0.7% of participants 15.2% of participants reported previous TB treatment
Index tests	Symptom screening: - Cough for more than 2 weeks - Any symptom positive out of 4: cough > 2 weeks, fever, haemoptysis, noticeable weight loss during past 6 months CXR (digital) - TB abnormality: pre-defined by authors as: consolidation, fibrocavitary disease and hilar lymphadenopathy
Target condition and reference standard(s)	Bacteriological tests: 2 samples for LJ culture Incorporation: 16.6% received a bacteriological reference standard; by design, participants without were assumed not to have TB Case definition - primary definition: at least 1x culture-positive + evidence of active disease OR at least 2x culture-positive in samples collected at different time points
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	

Nair 2016a *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		

Nair 2016a *(Continued)*

Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Yes		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Nair 2016b

Study characteristics			
Patient Sampling	Same as Nair 2016a		
Patient characteristics and setting	Same as Nair 2016a		
Index tests	Same as Nair 2016a		
Target condition and reference standard(s)	Bacteriological tests: 2 samples for LJ culture Incorporation: 16.6% received a bacteriological reference standard. By design, participants without were assumed not to have TB. Case definition: alternative definition: at least 1x culture-positive (collected during screening). Individual patient data were available for 25 patients who received a bacteriological reference standard. We re-analyzed this dataset so that only those with a single culture-positive result would classify as active disease. Then 40% of TB cases are without signs of active disease at the time of screening.		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	

Nair 2016b *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		

Nair 2016b *(Continued)*

Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Yes	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Ntinginya 2012

Study characteristics	
Patient Sampling	Cross-sectional study. Household contacts aged ≥ 5 years were enrolled of 80 consecutive, smear-positive index cases diagnosed between 1 December 2010 and 31 May 2011. Contacts were defined as those who lived in the same house or plot as an index case and who also shared meals with the index case.
Patient characteristics and setting	Household contacts aged 5 years or older, in urban and rural districts of Mbeya, Tanzania, recruited between 27 April and 7 June 2011 (n = 219) Mean age 29.1 years (SD 17.6) 59.4% female HIV prevalence not reported Smoking information not provided. Prevalence of smoking among 15+ in 2012 in Tanzania: 29.2% in males, 4.0% in females (unknown in 13 to 15 years) (https://apps.who.int/gho/data/node.main.65; accessed 14 May 2020) History of previous TB treatment reported by 2.7% Prevalence of HIV among adults aged 15 to 49 years (%), 2010: 5.0 (4.3 to 5.7) (https://apps.who.int/gho/data/node.main.622?lang=en; accessed 14 May 2020)
Index tests	7 symptom questions were asked: cough, cough > 14 days, weight loss, night sweats, chest pain, haemoptysis, fatigue, loss of appetite The study reports on accuracy of different symptoms and combinations, of which we use: any cough, cough > 14 days, 1 or more TB symptom out of 7, and 2 or more TB symptoms
Target condition and reference standard(s)	All 219 household contacts were invited to produce up to 2 sputum samples for testing with ZN microscopy, LJ and MGIT culture, and Xpert MTB/RIF. A participant was diagnosed with TB if at least 1 sputum sample tested positive for *M tuberculosis* on solid (Löwenstein-Jensen) or liquid (BACTEC™ MGIT 960) culture. All 5 TB cases had at least 1 TB symptom.

Ntinginya 2012 *(Continued)*

	Incorporation: although all were invited for sputum examination, only 15% was examined; unclear why so low.
Flow and timing	Only 15% produced sputum. Not clear if non-producers were repeatedly encouraged. 5 contacts unable to produce a sputum sample were referred to the health service due to high suspicion of TB. 56 samples were collected of 33 persons, so some gave 1 sample only. Some samples could not be split, so Xpert was done on the pellet rather than on direct sputum.
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		

Ntinginya 2012 *(Continued)*

Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes	
Did the study design require all patients to receive a bacteriological reference standard?	Yes	
Incorporation bias avoided?	Unclear	
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

NTP Bangladesh 2017

Study characteristics	
Patient Sampling	Cross-sectional national prevalence survey. Multi-stage stratified cluster sampling using PPS. 125 clusters of 800 participants on average, stratified by urban vs rural. One cluster was replaced for security reasons. Resident definition: lived in the cluster for at least 2 weeks before the census.
Patient characteristics and setting	National prevalence survey of the general population of 15 years and older in Bangladesh, a LMIC, conducted in 2015 to 2016 TB prevalence (bacteriologically confirmed) 287/100,000 population, incidence (2018): 221/100,000 The prevalence of HIV in the general population aged 15 to 49 years was < 0.1% (UNAIDS http://aidsinfo.unaids.org/; accessed May 2017) Median age between 25 and 34 years

NTP Bangladesh 2017 *(Continued)*

	49% female Smoking: 19.3% of survey participants 1.9% had a prior history of TB treatment
Index tests	Symptom screen is scoring system: participants scored positive if total score ≥ 3: cough greater than or equal to 2 weeks, haemoptysis in last month scored 3 points; weight loss in last month, fever greater than or equal to 1 week in last month, night sweats in past month scored 1 point. CXR (portable digital direct radiography): any pulmonary abnormality (field reading) Symptom score or CXR abnormality or both The symptom screen and CXR determined if someone had to submit sputum so this may introduce bias, although with trained staff this should be minimal The symptom scoring and CXR algorithm was designed to screen within this TB prevalence survey, and may be less applicable to a public health screening programme
Target condition and reference standard(s)	Bacteriological tests: 2 sputum samples for FM microscopy and LJ culture; Xpert MTB/RIF on 1 sample Incorporation: 20.7% received a reference standard; by design, participants with symptom score below threshold or a CXR abnormality or both were eligible for sputum examination Participants not eligible for sputum or without a sputum result were assumed not to have TB Case definition of bacteriologically positive TB case: definite MTB = confirmed by culture and/or Xpert
Flow and timing	Of the 20,594 who were eligible to submit sputum 20,463 (99%) submitted at least 1 specimen, 20,010 (97%) submitted 2 samples, 20,378 (99%) had at least 1 culture result available, and 20,425 (99%) at least 1 Xpert result available, which is very high, so low risk
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		

NTP Bangladesh 2017 *(Continued)*

Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		

NTP Bangladesh 2017 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

NTP Philippines 2018

Study characteristics	
Patient Sampling	Cross-sectional national prevalence survey. Multi-stage stratified cluster sampling using PPS of 106 survey clusters in 4 strata – National Capital Region, regions 3 and 4-A; rest of Luzon; Visayas; and Mindanao. The target cluster size was 500 individuals. A total of 89,663 individuals from 19,707 households were enumerated in the survey census, of whom 61,466 (69%) were eligible and invited to participate. Of these, 46,689 (76%) participated. A small number of clusters was replaced because of security reasons, constituting low risk of bias. Resident definition: lived in the household for at least 2 weeks prior to the census.
Patient characteristics and setting	General population 15 years and older, in the Philippines, a LMIC, enrolled in 2016 TB prevalence (bacteriologically confirmed, survey) 1159/100,000 population, incidence (2014): 288/100,000 Median age between 35 and 44 years 55.3% female HIV prevalence not provided. In 2015, the prevalence of HIV in the general population aged 15 to 49 years was < 0.1% (UNAIDS http://aidsinfo.unaids.org/; accessed May 2017) Smoking: 39.7% of survey participants

NTP Philippines 2018 *(Continued)*

	5.5% of participants reported previous TB treatment
Index tests	Symptom screen: cough for ≥ 2 weeks and/or haemoptysis CXR: direct digital mobile, field reading for any pulmonary abnormality (any lung or mediastinum abnormality) Symptom and/or CXR: cough for ≥ 2 weeks and/or haemoptysis, or any pulmonary abnormality on field reading, or both
Target condition and reference standard(s)	Bacteriological tests: 2 samples for direct FM microscopy (2 samples), cultured on Ogawa medium (1) and Xpert MTB/RIF (1) Incorporation: 38.8% received a reference standard. By design, participants with cough of 2 or more weeks or a CXR abnormality or both were eligible for sputum examination. Participants not eligible for sputum or without a sputum result were assumed not to have TB. Case definition: bacteriologically positive TB case constitutes definite MTB confirmed by culture and/or Xpert. The percentage of true (confirmed) TB cases who were asymptomatic, no information on CXR available/exempted 3.9%. A small percentage, of whom some may in fact have had CXR abnormalities.
Flow and timing	A total of 18,597 participants (40%) were eligible for sputum examination; of these, 16,242 (87%) submitted at least 1 sputum specimen and 15,547 (84%) submitted 2 sputum specimens. Sputum specimens from 16,200 (87%) participants were tested with Xpert MTB/RIF. High-risk since there may be bias in index test false negatives if persons with mild TB were more unable to give sputum.
Comparative	
Notes	Questions on cough ≥ 2 weeks, haemoptysis, fever, weight loss, night sweats were asked but data for 2 x 2 table not reported. The presence of cough for 2 or more weeks determined whether someone would give sputum. This may influence interviewers to err on the positive side. We assumed that the study staff were well trained, so the symptom interview should be objective enough to minimize bias. Similarly for CXR, and symptom and/or CXR as a combined screen: an abnormal CXR determined eligibility for sputum examination, but we assumed that due to survey procedures and training the risk of bias is low; CXR threshold - not applicable: implicit threshold.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		

NTP Philippines 2018 *(Continued)*

Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		

NTP Philippines 2018 *(Continued)*

Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Pelissari 2018

Study characteristics	
Patient Sampling	Cross-sectional design. 2 screening rounds are described. Data are used from the first round. Participant selection: all new inmates were eligible (95% of participants, most likely consecutively enrolled); existing inmates from 3 blocks (5%) in the blocks (non-consecutive enrollment).
Patient characteristics and setting	Prison inmates, 18 years and older, in Porto Alegre Central Prison, Porto Alegre, capital of the state of Rio Grande do Sul, Brazil, a UMIC, enrolled in 2014 TB prevalence 1898/100,000 population, incidence 33.5/100,000 Median age approximately 28 (79% were between 18 and 34 year) 0% female HIV prevalence among participants 5.6% Smoking among participants not reported. In Brazil, 2015, prevalence of smoking any tobacco product among persons

Pelissari 2018 *(Continued)*

	aged ≥ 15 years, males: 19.3% (https://apps.who.int/gho/data/node.main.65; accessed 15 May 2020) 5.2% of participants reported previous TB treatment
Index tests	Symptom screening: - Any cough vs no cough - Cough more than 2 weeks vs 2 weeks or less (not clear how that was asked) CXR - conventional; abnormalities consistent with TB (suggestive TB alterations) Combination - cough (any duration) or CXR with suggestive TB alterations or both
Target condition and reference standard(s)	Bacteriological tests: 1 sputum sample for FM microscopy and Xpert MTB/RIF. Not clear when culture was done. If sample quality was poor, a new sample was collected from those still in the prison. Samples with insufficient sputum were prioritized for Xpert testing. Samples that were salivary or contained food particles were subjected to microscopy. Incorporation: 19.3% received a reference standard; participants without were assumed not to have TB. Case definition: confirmed patients were defined as those with at least 1 positive result on any 1 of the laboratory tests (Xpert, smear microscopy, or culture), except for those patients with a culture result of non-tuberculous mycobacteria.
Flow and timing	A large proportion of patients (close to 50%) who were eligible for the reference standard, did not get the reference standard for unclear reasons
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			

Pelissari 2018 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	No		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			

Pelissari 2018 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

Qadeer 2016

Study characteristics	
Patient Sampling	Cross-sectional national prevalence survey. Multi-stage stratified cluster sampling following WHO guidance in 95 clusters of 1400 participants on average. A few areas with high security concerns were excluded (6.4% of population). Of 131,1331 eligible residents, 15,915 (80.7%) participated Resident definition: slept in the household
Patient characteristics and setting	General population 15 years and older, in Pakistan, a LMIC, enrolled in 2010 TB prevalence 364/100,000 population, incidence: 231/100,000 Median age 30 years (IQR 21 to 45) 57.6% female HIV prevalence not measured. In Pakistan the percentage of people living with HIV among adults (15 to 49 years) was 0.1% in 2018 (https://www.unaids.org/en/regionscountries/countries/pakistan; accessed 14 April 2021) Smoking and history of previous TB treatment not reported In Pakistan, 2010, prevalence of smoking any tobacco product among persons aged ≥ 15 years, males 40.2%; females 4.2% (https://apps.who.int/gho/data/node.main.65; accessed 15 May 2020)
Index tests	Symptom screen: cough for 2 or more weeks CXR any pulmonary abnormality (field reading)

Qadeer 2016 *(Continued)*

	Parallel: cough for 2 or more weeks or CXR pulmonary abnormality or both
Target condition and reference standard(s)	Bacteriological tests: 3 sputum samples for ZN microscopy (SS) and 1 for modified Kudoh (solid) culture (C); if culture results were not available, but at least 1 smear was positive, a NAAT was performed on scraped smear material. Incorporation: 8.4% received a reference standard. By design, participants with cough of 2 or more weeks or a CXR abnormality or both were eligible for sputum examination. Participants not eligible for sputum or without a sputum result were assumed not to have TB. Case definition: SS+: C+ 5+ colonies; C+< 5 col & (SS+ or abnormal CXR); NAAT+ and SS+ / SS-: C+ 5+ colonies; C+< 5 col & (SS+ or abnormal CXR) 2 x 2 data is only available in publication for 207 definite smear-positive TB cases. Additional information for 341 bacteriologically positive cases received from authors.
Flow and timing	81% of eligible persons submitted at least one sputum smear specimen, and 73% one culture examined
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Qadeer 2016 *(Continued)*

DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?			
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		

Qadeer 2016 *(Continued)*

Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Republic of Uganda 2018

Study characteristics	
Patient Sampling	Cross-sectional national prevalence survey. Multi-stage stratified cluster sampling using PPS of 70 clusters of 580 participants on average, stratified by urban vs rural. Of 45,293 eligible residents, 41,154 (91%) participated. Resident definition: permanent residents who stayed at least 1 night in the household in the past 2 weeks; temporary visitors who arrived at least 2 weeks before census day.
Patient characteristics and setting	General population 15 years and older, in Uganda, a LIC, enrolled in 2014 to 2015 TB prevalence (bacteriologically confirmed, year/survey) 401/100,000 population, incidence (2008): 311/100,000 36% were between 15 and 24 years, 62% were between 15 and 34 years; median around 30 years 57.5% female HIV prevalence in the general population aged 15 to 49 years in 2014 was 7.1% (UNAIDS http://aidsinfo.unaids.org/; accessed April 2017) Smoking: 7.3% of survey participants 2.0% of participants reported previous TB treatment
Index tests	Symptom screens: - Cough of any duration - Cough of 2 weeks or longer CXR: conventional portable; field reading: abnormal lung fields (non-pulmonary abnormalities are excluded) Cough of 2 weeks or longer or CXR abnormal lung fields or both (this survey symptom screen defined qualification for sputum collection)
Target condition and reference standard(s)	Bacteriological tests: 2 sputum samples for direct ZN microscopy and concentrated LJ culture; Xpert MTB/RIF done on smear-positive specimens and/or if both samples were culture contaminated Incorporation: 11.6% received a reference standard. By design, participants with cough of 2 or more weeks or a CXR abnormality or both were eligible for sputum examination. Participants not eli-

Republic of Uganda 2018 *(Continued)*

	gible for sputum or without a sputum result were assumed not to have TB. Case definition of bacteriologically positive TB case: definite: MTB confirmed by culture and/or Xpert and smear-positive or MTB confirmed by culture and/or Xpert with chest X-ray consistent with TB.
Flow and timing	5142 participants (13%) were eligible for sputum examination, of whom 4844 (94%) submitted at least 1 sputum specimen and 4532 (88%) submitted 2 sputum specimens. 4758 had at least 1 culture result = 93% of those eligible for sputum, which is a high percentage so low risk of bias.
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Republic of Uganda 2018 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Republic of Zimbabwe 2015

Study characteristics	
Patient Sampling	Cross-sectional nationwide prevalence survey. Multi-stage stratified cluster sampling using PPS of 75 clusters of 600 participants on average, stratified by urban vs rural; 2 clusters were replaced for logistical reasons. Of 43,478 eligible residents, 33,736(78%) participated. Resident definition: permanent residents who had slept at least 1 night out of the last 14 days at the time of census; non-residents who had slept in the household for 14 days or more before the time of the census.
Patient characteristics and setting	General population 15 years and older, in Zimbabwe, a LMIC, enrolled in 2014 TB prevalence (bacteriologically confirmed, year/survey) 344/100,000 population, incidence (2013): 552/100,000 Median age: 55% were between 15 and 34 years 57.9% female HIV prevalence in the population aged 15 to 49 years: 15% (UNAIDS http://aidsinfo.unaids.org/; accessed April 2017) Smoking: among persons with presumptive TB, 32.9% of males were smokers and 14.6% were ex-smokers while only 1.3% of females were smokers. Prevalence of current tobacco smoking among persons aged 15 years and older, 2016, both sexes: 16.2% (https://apps.who.int/gho/data/view.main.GSWCAH20v; accessed 14 May 2020). 4.1% of participants reported previous TB treatment
Index tests	Symptom screen: - Any 1 of the following 3 symptoms: cough of any duration, haemoptysis in the past 12 months, drenching night sweats (used in survey) - Cough for 2 or more weeks - Cough of any duration CXR (direct digital): any lung abnormality (field reading) Combination: any 1 of the 3 TB symptoms or CXR pulmonary abnormality or both The symptom and CXR algorithm used to screen within this TB prevalence survey was different from most similar surveys, and may be different from a public health screening programme
Target condition and reference standard(s)	Bacteriological tests: 2 sputum samples for concentrated FM microscopy and culture on LJ and MGIT media; Xpert MTB/RIF on smear-positive specimens Incorporation: 16.9% received a reference standard; by design, participants not eligible for sputum or without a sputum result were assumed not to have TB

Republic of Zimbabwe 2015 *(Continued)*

	Case definition: definite: MTB = confirmed by culture and/or Xpert 4/107 cases were symptom negative and CXR exempt
Flow and timing	98% of participants eligible for sputum examination provided at least 1 specimen, 94% both specimens. 98% of participants had at least 1 culture result available.
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			

Republic of Zimbabwe 2015 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Rwanda MoH 2014

Study characteristics	
Patient Sampling	Cross-sectional national prevalence survey. Multi-stage stratified cluster sampling using PPS of 73 clusters of 610 participants on

Rwanda MoH 2014 *(Continued)*

	average, not stratified. Total participants: of 45,058 eligible, 4128 (96%) participated. Resident definition: lived in the household at least 1 month prior to the interview.
Patient characteristics and setting	General population 15 years and older, in Rwanda, a LIC, enrolled in 2012 TB prevalence (bacteriologically confirmed, year/survey) 119/100,000 population, incidence (2012): 86/100,000 Median age between 25 and 34 years; 56.8% is below 34 58% female The prevalence of HIV in the general population aged 15 to 49 years was estimated at 3.1% in 2012 (UNAIDS http://aidsinfo.unaids.org/; accessed April 2017). Among those eligible for sputum examination 4.9% were HIV-positive. Smoking: 12.9% of sub-study participants 1.3% of participants reported previous TB treatment
Index tests	Cough of any duration; CXR for any lung abnormality (direct digital, field reading); combination of cough of any duration or CXR for any lung abnormality or both. Of the other symptoms that were asked about no 2 x 2 tables can be obtained.
Target condition and reference standard(s)	Bacteriological tests: 2 samples for concentrated FM microscopy and LJ culture Incorporation: 10.6% received a reference standard. By design, participants with cough of any duration, or any lung abnormality on CXR or both were eligible for sputum examination. Participants not eligible for sputum or without a sputum result were assumed not to have TB. Case definition: bacteriologically confirmed TB: definite (n = 35): MTB confirmed by 2 cultures, or 1 culture and smear-positive, or 1 culture with chest X-ray suggestive of TB. Probable (n = 5): MTB not confirmed by culture but 2 smear-positive specimens or 1 smear-positive with chest X-ray suggestive of TB.
Flow and timing	A total of 4747 people were eligible for sputum examination, of whom 4700 (99%) submitted at least 1 sputum specimen and 4412 (93%) submitted 2 sputum specimens. At least 1 culture result available 4589 (97%). This is quite high, so low risk of bias.
Comparative	
Notes	The symptom screening question and/or CXR reading determined if someone would get a bacteriological test or not. Although this could introduce bias, the risk is considered low assuming the study staff were well trained.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Rwanda MoH 2014 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		

Rwanda MoH 2014 *(Continued)*

Did the study design require all patients to receive a bacteriological reference standard?	No	
Incorporation bias avoided?	No	
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Seri 2017

Study characteristics	
Patient Sampling	Cross-sectional study. All prisoners from 3 buildings (the men's long-term detention building, the women's detention building, and the infirmary inpatients ward) who signed the informed consent were included.
Patient characteristics and setting	Prison inmates, adults (minors were excluded) in Cote d'Ivoire, Adidjan, a LMIC, enrolled in 2015 The study to estimate prevalence was done 16 years into the screening programme, so persons may have been screened before TB prevalence 6257/100,000 population, incidence 148/100,000 Median age 31 years (IQR 26 to 37) 6.6% female HIV prevalence in study population 3.1% Smoking: 55% of participants 5% of participants reported previous TB treatment

Seri 2017 *(Continued)*

Index tests	Symptom screens - Cough lasting for more than 2 weeks (on subset of n = 811) - Any TB symptom positive out of 4 (cough lasting for more than 2 weeks, fever lasting for more than 2 weeks, recent loss of appetite, chest pain) CXR - abnormalities consistent with TB vs not consistent
Target condition and reference standard(s)	Bacteriological tests: 2 sputum samples for FM microscopy and 1 culture (LJ and MGIT); 2nd sample cultured if 1) current cough lasting more than 2 weeks, 2) abnormal chest X-ray images, 3) HIV seropositivity, and/or 4) positive microscopy of the first sputum sample Incorporation: by design, all participants were supposed to have a reference standard. 61% received a reference standard (see flow and timing). Participants without results were assumed not to have TB. Case definition: confirmed = positive culture for MTB complex; probable (positive microscopy with at least one AFB per field and or CXR suggestive of TB) and possible cases are not considered as cases in the 2 x 2 table Of the TB cases 3/19 = 16% had no symptoms and unknown whether they had CXR abnormalities
Flow and timing	40% did not have culture results (see Figure 1). Patients with certain symptom or CXR criteria had 2 samples cultured, increasing the chance of being TB-confirmed; others had only 1 sample cultured.
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			

Seri 2017 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No		

Seri 2017 *(Continued)*

Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Singh 2013

Study characteristics	
Patient Sampling	Contact tracing: a total of 508 open index cases and their 1792 household contacts were enrolled in a cohort study, from various Directly Observed Treatment Short-course (DOTS) centres of South Delhi region. Of these, only 432 index cases and 1608 (89.7%) contacts were included in the study, while the remaining were excluded either if the address of the index case was wrong or the address of the household contact was wrong.
Patient characteristics and setting	Household contacts of TB cases, recruited between May 2007 to May 2009 in South Delhi, India (n = 1608) No age restriction, 81% were older than 12 years; mean age 26.5 years (SD 15.9) 46.1% females HIV prevalence not reported Smoking not reported; Tobacco Atlas 1st ed 2002: 16.0% among adults History of previous TB treatment reported by 0.8% National HIV prevalence in India is estimated at 0.3% in 2008 (http://naco.gov.in/sites/default/files/Technical%20Report%20India%20HIV%20Estimates%202010.pdf; accessed 16 May 2020)
Index tests	Symptom questions asked: fever, cough, anorexia, breathlessness, weight loss, night sweats, fatigue, chest pain Definition of screens not provided, inadequate method of reporting Cough of any duration of relevance to this review
Target condition and reference standard(s)	Bacteriological test and incorporation: all participants were expected to provide 2 sputum samples for examination by ZN microscopy, 1 LJ and 1 MGIT 960 culture Case definition: culture-positive sample on MGIT 960, detected at the time of recruitment
Flow and timing	25% did not give sputum; unclear if all persons who gave sputum had 1 or 2 samples tested
Comparative	

Singh 2013 *(Continued)*

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			

Singh 2013 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?		Unclear
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

Telisinghe 2014

Study characteristics	
Patient Sampling	Cross-sectional study in a prison; random enrollment of a sample of offenders who had been incarcerated for at least 6 months ("currently incarcerated") and a consecutive sample of "newly sentenced" offenders. Individuals with an expected stay of less than 3 months in the study facility were excluded to ensure follow-up of medical records.
Patient characteristics and setting	Male prison inmates in Johannesburg, South Africa, enrolled between Sept 2009 and Oct 2010 Adults; median age 32 (IQR 27 to 37) TB incidence in South Africa in 2010: 981/100,000 In full study 25% HIV-positive, but data extracted from HIV-negative inmates only, in whom TB prevalence was 2800/100,000 Current smoking 57.9% History of previous TB 12.8% Some may have been screened before, with inconsistent methods
Index tests	Symptom questions asked: current cough and duration, fever, night sweats, or unintentional weight loss Symptom index tests: - Cough 2 or more weeks - Any 1 of cough 2 or more weeks, night sweats, or unintentional weight loss - WHO screening tool for PLHIV: any of current cough, fever, weight loss, or night sweats

Telisinghe 2014 *(Continued)*

	CXR type not stated, assessed by 2 readers using a standardized tool. Index tests: - CXR suggestive of TB - Any CXR abnormality - Combination of WHO 4 symptom tool for PLHIV or CXR suggestive of TB Authors also report on: Cough > 2 weeks or CXR suggestive of TB Any cough or CXR suggestive of TB Symptom combination (any of cough 2 or more weeks, night sweats or unintentional weight loss) or CXR suggestive of TB
Target condition and reference standard(s)	All participants provided 2 spot sputum specimens for FM smear microscopy and liquid mycobacterial culture (MGIT), speciation with the GenoType Mycobacterium CM kit. Case definition includes definite or probable. Definite cases were sputum culture-positive for *M tuberculosis* with compatible clinical or radiological features (as assessed by one or both readers), or additional microbiological confirmation (any grade of smear or further positive culture). Probable cases were those with one culture positive for *M tuberculosis* without compatible clinical or radiological features or smear-positive grade 1+ or more, with or without compatible clinical or radiological features (as assessed by one or both readers). 100% had culture 2/19 had 1 positive culture without symptoms or CXR abnormality
Flow and timing	Of all consenting participants (HIV+ and HIV-) 3 of 981 consenting participants were excluded because of missing sputum culture at enrolment, 39 "possible tb cases were excluded, 62 with missing CXR at enrolment and 21 with missing urine HIV test results were excluded; 125/981 = 12.7%, which could have introduced bias.
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	

Telisinghe 2014 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	

Telisinghe 2014 *(Continued)*

Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Yes	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Unclear	
Were all patients included in the analysis?	No	
Could the patient flow have introduced bias?		High risk

van't Hoog 2011

Study characteristics	
Patient Sampling	Cross-sectional regional prevalence survey. In 40 clusters, randomly sampled out of 105, all persons of 15 years and older were eligible; 91% participated. Of 19,216 CXRs with field reading, 1143 were selected for assessment by expert readers through stratified sampling: among CXRs of participants without TB 1031 were randomly selected; all CXRs of 123 persons identified with bacteriologically confirmed prevalent TB were eligible, and 112 (91%) were retrieved.
Patient characteristics and setting	Same as van't Hoog 2012: General rural population (n = 20,566) 15 years and older; median age 35 years 63% females TB prevalence 600 per 100,000 HIV prevalence 16.8% History of previous TB treatment: 2.1% Smoking: 10.4%
Index tests	Abnormality consistent with TB and any abnormality by radiologist or pulmonologist (data are used of expert reader #1)
Target condition and reference standard(s)	Same as van't Hoog 2012: Bacteriological test: 2 sputum samples requested of all participants for FM smear microscopy

van't Hoog 2011 *(Continued)*

	Incorporation: persons with symptoms (presence of cough > 7 days, and/or haemoptysis of any duration and/or 2 out of 3 of: fever >7 days, night sweats for >7 days, or weight loss) and/or any CXR abnormality were eligible for culture of 1 sputum sample on solid (LJ) and liquid (MGIT) medium, 32.3% of total. Missing and contaminated results are considered TB-negative. Case definition: 1 positive culture or 2 positive smears (unless MOTT) or 1 positive smear and CXR TB-abnormality)
Flow and timing	See van't Hoog 2012
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		

van't Hoog 2011 *(Continued)*

Did the study design require all patients to receive a bacteriological reference standard?	No	
Incorporation bias avoided?	No	
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

van't Hoog 2012

Study characteristics	
Patient Sampling	Cross-sectional regional prevalence survey. In 40 clusters, randomly sampled out of 105, all persons of 15 years and older were eligible; 91% participated.
Patient characteristics and setting	General rural population (n = 20,566) 15 years and older; median age 35 years 63% females TB prevalence 600 per 100,000 HIV prevalence 16.8% History of previous TB treatment: 2.1% Smoking: 10.4%
Index tests	Symptom questions (cough, haemoptysis, weight loss, fever, night sweats, and duration) and conventional CXR requested of all participants. CXRs field read by clinical officers for any abnormality.

van't Hoog 2012 *(Continued)*

	Questionnaire was administered, using handheld computers, with questions on the presence and duration of symptoms that are possibly suggestive of TB (cough, haemoptysis, weight loss, fever, night sweats). Duration of cough less than 2 weeks was recorded in days, and of 2 weeks and longer in weeks. Symptoms qualifying as a positive symptom screen were the presence of cough for more than 7 days, and/or haemoptysis of any duration and/or 2 out of 3 of the following symptoms: fever for 7 days, night sweats for 7 days, or weight loss resulting in a changed fit of clothes. Cough of any duration Cough for 2 or more weeks Any symptom of any duration or severity out of 5 CXR any abnormality CXR abnormalities suggestive of TB Cough for 2 or more weeks or any CXR abnormality
Target condition and reference standard(s)	Bacteriological test: 2 sputum samples requested of all participants for FM smear microscopy. Incorporation: persons with symptoms (presence of cough > 7 days, and/or haemoptysis of any duration and/or 2 out of 3 of: fever >7 days, night sweats for >7 days, or weight loss) and/or any CXR abnormality were eligible for culture of 1 sputum sample on solid (LJ) and liquid (MGIT) medium, 32.3% of total. Missing and contaminated results are considered TB-negative. Case definition: 1 positive culture or 2 positive smears (unless MOTT) or 1 positive smear and CXR TB-abnormality).
Flow and timing	Of all participants, 99% submitted 1 smear sample, 96% 2 smear samples. n = 7342 were eligible for solid culture (LJ) on 1 sample, 6646 (91%) had a culture result
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern

van't Hoog 2012 *(Continued)*

DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	No		
Incorporation bias avoided?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			

van't Hoog 2012 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	No	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

Wei 2014

Study characteristics	
Patient Sampling	Cross-sectional - population-based TB prevalence survey. Stratified multi-stage random sampling was used to select the 35 clusters within 17 prefectures, targeting residents of 15 years old or above who had lived in the selected clusters for more than 6 months. Military barracks and prisons located in the cluster were excluded.
Patient characteristics and setting	General population in Shandong province, China, of 15 years and older, urban and rural; recruited in 2010 (n = 54,268) Mean age 46 years 53% female HIV prevalence not reported for study population, but from other sources < 0.01% Smoking among TB suspects 31% History of previous TB: not reported
Index tests	Persistent cough (lasting 2 weeks or longer) NB patients were asked about symptoms suggestive of TB, such as persistent cough (lasting 2 weeks or longer), haemoptysis, weight loss and fever. Accuracy data are reported on persistent cough only. CXR type not reported. Classified as abnormalities consistent with TB vs not consistent (defined as any abnormal shadow in the lung field and mediastinum, or pleural effusion except pleural thickness or small single calcification).
Target condition and reference standard(s)	1.1% received the reference standard by design, classified as TB suspects; non-suspects are assumed not to have TB

Wei 2014 *(Continued)*

	Bacteriological test: 3 sputum samples for ZN microscopy and 2 for LJ culture were requested Case definition: those with positive smear or culture sputum specimens were classified as sputum bacteriologically confirmed cases
Flow and timing	9684 (98.6%) out of 9825 eligible for sputum examination and at least 1 culture available (see WHO country profile flow diagram), which is very high. Information about how many had 1, 2 or 3 samples tested is unclear.
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Wei 2014 *(Continued)*

DOMAIN 2: Index Test (Combination of Symptoms and CXR)		
Were the index test results interpreted without knowledge of the results of the reference standard?		
If a threshold was used, was it pre-specified?		
Could the conduct or interpretation of the index test have introduced bias?		
Are there concerns that the index test, its conduct, or interpretation differ from the review question?		
DOMAIN 3: Reference Standard		
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes	
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes	
Did the study design require all patients to receive a bacteriological reference standard?	No	
Incorporation bias avoided?	No	
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Yes	
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Unclear	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Wood 2007

Study characteristics		

Wood 2007 *(Continued)*

Patient Sampling	Cross-sectional regional prevalence survey. Simple random sampling from census sampling frame. Of 971 adults permanently residing on the identified plots, 762 (78%) consented to participate.
Patient characteristics and setting	General population of 5 years and older of an urban high-density residential area in South Africa (n = 1150) in 2005 Median age 27 years Gender distribution not provided HIV prevalence in study population 23% Smoking 27% History of previous TB treatment: 8%
Index tests	Cough of any duration NB TB symptoms asked for: cough, loss of appetite, weight loss, and night sweats. Study also reports on 2 or more symptoms, any TB symptom, of which only the sensitivity is available.
Target condition and reference standard(s)	Bacteriological test: 2 sputum samples requested and examined with FM microscopy and liquid culture (MGIT) Incorporation: 99.5% provided sample(s) Case definition: 2 positive smears or 2 positive cultures 67% of cases were asymptomatic
Flow and timing	99.5% provided sample(s)
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (Symptom questions)			

Wood 2007 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Chest radiography)			
DOMAIN 2: Index Test (Combination of Symptoms and CXR)			
DOMAIN 3: Reference Standard			
Is the bacteriological reference standard likely to correctly classify the target condition?	Yes		
Were the bacteriological reference standard results interpreted without knowledge of the results of the index test?	Yes		
Did the study design require all patients to receive a bacteriological reference standard?	Yes		
Incorporation bias avoided?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Were there among the true (confirmed) TB cases, persons who only had one positive smear or culture and no signs of active disease?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients who were supposed to receive a bacteriological reference standard (by design) actually have one, and have a result?	Yes		
Did patients subjected to a bacteriological reference standard all receive the same reference standard?	Unclear		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

ACF: active case finding
AFB: acid fast bacilli
CXR: chest radiography
CAD: computer assisted reading of digital radiographs

DR: digital radiography
FM: fluorescence microscopy
HIV: human immunodeficiency virus
IQR: interquartile range
LIC: low-income country
LJ: Löwenstein-Jensen
LMIC: lower middle-income country
MGIT: mycobacterial growth inhibitor tubes
MMR: mass miniature radiography
MOTT: Mycobacterium other than tuberculosis
MTB: *Mycobacterium tuberculosis*
NTP: national tuberculosis (control) programme
NAAT: nucleic acid amplification test
PLHIV: people living with HIV
PPS: probability proportional to size
PTB: pulmonary tuberculosis
SD: standard deviation
SSM: sputum smear microscopy
TB: tuberculosis
TST: tuberculin skin test
UMIC: upper middle-income country
VCCT: voluntary confidential counselling and testing (centre)
ZN: Ziehl-Neelsen

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Abboud 2013	Only agreement between readers is provided, no bacteriological results
Abebe 2011	Only prisoners with chronic cough were eligible, no further screen
Abubakar 2010	CXR reader agreement study
Aerts 2000	Only one screen applied and screen negatives are not further evaluated
Al 2013	Only one screen applied and screen negatives are not further evaluated
Alamo 2012	Population ineligible (children, HIV-infected or self-reported tuberculosis suspects only)
Alcântara 2012	Population ineligible (children, HIV-infected or self-reported tuberculosis suspects only)
Aldridge 2010	Only one screen applied and screen negatives are not further evaluated; cases in screen negatives were identified through surveillance register linkage
Ali 2012	No info on screening tests
Alsedà 2003	No full text
Arenas 2008	Only one screen applied and screen negatives are not further evaluated
Arpaz 2003	Population ineligible (children, HIV-infected or self-reported tuberculosis suspects only); description of routine tuberculosis cases finding and treatment over time
Arzhaeva 2009	Development of a method (CAD)
Assefzadeh 2009	Only one screen applied and screen negatives are not further evaluated

Study	Reason for exclusion
Bai 2001	Only one screen applied and screen negatives are not further evaluated; Cohort study looking for incident tuberculosis cases only
Balasangameshwara 1993	Population ineligible (children, HIV-infected or self-reported tuberculosis suspects only)
Balasubramanian 1995	Only one screen applied and screen negatives are not further evaluated
Balasubramanian 2004	Active tuberculosis cases not specified by symptom or CXR result
Banda 1998	Population ineligible. They are "tuberculosis suspects with short duration of cough" attending an outpatient department of a large hospital. The tuberculosis prevalence among participants was 35% and 55% had HIV-related illness, suggesting a passive case detection population.
Banda 2009	Only one screen applied and screen negatives are not further evaluated
Banu 2010	Only one screen applied and screen negatives are not further evaluated
Basta 2005	Only one screen applied and screen negatives are not further evaluated
Basta 2006	Only one screen applied and screen negatives are not further evaluated
Bates 2013	Patients with a cough were selected for inclusion and 28% had tuberculosis, but no information on symptoms is provided by bacteriology status
Becerra 2005	Only one screen applied and screen negatives are not further evaluated
Beser 1993	Only one screen applied and screen negatives are not further evaluated
Beyanga 2018	Not enough data for 2 x 2 table
Bhat 2009	Only one screen applied and screen negatives are not further evaluated
Bhatia 2002	Cohort study looking for incident tuberculosis cases only
Binkin 1996	Evaluation of a screening programme (not a screening tool)
Bjerregaard-Andersen 2010	Only one screen applied and screen negatives are not further evaluated
Bloss 2012	Only one screen applied and screen negatives are not further evaluated
Bonvin 1992	Only one screen applied and screen negatives are not further evaluated
Borgdorff 2000	No original data; analysis of reported survey data
Borra 2009	No cases of bacteriologically confirmed active tuberculosis identified; Population ineligible (children, HIV-infected or self-reported tuberculosis suspects only)
Braden 1995	Only one screen applied and screen negatives are not further evaluated
Bucher 1994	Full text not available
Burgess 2001	Evaluation of CXR screening in a pre-selected symptomatic population (outside the scope of this review)
Cain 2010	Population ineligible (children, HIV-infected or self-reported tuberculosis suspects only)

Study	Reason for exclusion
Callister 2002	Only one screen applied and screen negatives are not further evaluated
Carbonara 2005	No screening for symptoms and no sputum specimen
Castro 2011	Only one screen applied and screen negatives are not further evaluated; Population ineligible
Catanzaro 2013	Summary of conference on IGRA's; insufficient original data
Chadha 2012	Only one screen applied and screen negatives are not further evaluated; a subsample had symptom and MMR screening, but SSM+ in persons with negative screen are not mentioned
Chakraborty 1995	Population examined by bacteriology, but no data on symptoms
Chan-Yeung 2007	Only one screen applied and screen negatives are not further evaluated
Chang 2010	No cases of bacteriologically confirmed active tuberculosis identified, TST investigation only
Chen 2011	No information on symptoms/CXR TB-negatives among the population
Chen 2012	Not enough information for a 2 x 2 table
Chen 2013	Population ineligible, review, no original data; paediatric tuberculosis
China 2004	Not enough data in report to fill a 2 x 2 table
Churchyard 1999	Only one screen applied and screen negatives are not further evaluated
Churchyard 2010	Data allow only calculation of sensitivity, not specificity
Claassens 2013	Only one screen applied and screen negatives are not further evaluated
Claassens 2013a	Not enough information presented to fill a 2 x 2 table
Clarke 2003	Full text not available
Coelho 2011	Population ineligible
Corbett 2004	No data presented from which sensitivity and specificity of screening can be calculated
Costa 2010	CXR only in TST or IGRA positives
Costa 2011	Cohort study looking for incident tuberculosis cases only; unclear which patients received bacteriological examination. Cases were probably extracted from records.
Costantino 2010	No cases of bacteriologically confirmed active tuberculosis identified; insufficient data; mostly screening for Mtb infection
Crampin 2011	Only one screen applied and screen negatives are not further evaluated; case control study; symptoms/CXR data not reported
Cruz-Hervert 2012	Only one screen applied and screen negatives are not further evaluated
Cuevas 2011a	Population ineligible
Cuevas 2011b	Population ineligible

Study	Reason for exclusion
Dara 2013	Review
Datta 2001	Participants were screened with MMR
de Vries 2007	Only one screen applied and screen negatives are not further evaluated
Eckhoff 2000	Relevant data collected but not reported
Forman 2003	No reference test
Fox 2011	Conference abstract
Fox 2012a	Conference abstract, insufficient data. Unclear from abstract whether all participants were examined with all methods (including smear).
Gopi 2003	Participants were screened with MMR
Gopi 2008a	Does not contain original survey data
Gopi 2008b	Duplicates Gopi 2003
Hamid 2012	Population ineligible. Patients attending a tertiary care hospital with cough.
Hanrahan 2013	The tuberculosis suspect definition was stratified by HIV status. For HIV-infected or unknown status: cough, fever, night sweats, weight loss of any duration; for HIV-uninfected: cough or fever for 2 or more weeks, or night sweats or weight loss of any duration. These screening questions were used to determine if the patients would be tested for tuberculosis. Only symptom positives received a reference standard.
Henostroza 2016	A sampling procedure was applied that needs correction in order to create 2 x 2 tables. Authors reply not received. The sampling fraction of CXRs differed for different index test classification categories: 1) all patients with CXRs deemed abnormal by the project clinical officer; 2) all persons with normal CXRs who were diagnosed with TB (based on smear results, clinical criteria, and/or culture confirmation); and 3) a random sample of inmates with normal CXRs and not diagnosed with TB. After evaluating the number of CXRs in categories 1) and 2), we decided to select 80 CXRs from HIV-positive and 80 from HIV-negative inmates in category 3) to strike a balance between feasibility to conduct all CXR readings and ensuring that there were an adequate number of normal CXRs in the sample.
Hoffmann 2013	Population ineligible. HIV-infected pregnant women who had tested HIV seropositive at a prior prenatal clinic visit.
Hong 1993	Only one screen applied and screen negatives are not further evaluated
Hong 1998	Only one screen applied and screen negatives are not further evaluated
Horie 2007	Not enough data on symptoms or CXR abnormalties on non-cases to fill 2x2 tables
Kim 2012	Population ineligible (HIV-infected only)
Leung 2005	Only one screen applied and screen negatives are not further evaluated
Mahomed 2013	In our earlier version of this review, we included studies with sputum smear-positive cases only and discussed those separately. Such studies were excluded from this review as NAATs are now recommended instead of sputum smear microscopy as the primary test to diagnose tuberculosis, and the number of such studies was small.

Study	Reason for exclusion
Masur 2017	Studies evaluating CXR screening in a pre-selected symptomatic population were outside the scope of this review and were excluded. And studies with sputum smear-positive cases were excluded from this review as NAATs are now recommended instead of sputum smear microscopy as the primary test to diagnose tuberculosis, and the number of such studies was small.
Mathez 2007	Only one screen applied and screen negatives are not further evaluated
Mor 2013	Narrative review
Santha 2005	Population ineligible. Outpatients with cough for 2 or more weeks
Sebhatu 2007	In our earlier version of this review, we included studies with sputum smear-positive cases only and discussed those separately. Such studies were excluded from this review as NAATs are now recommended instead of sputum smear microscopy as the primary test to diagnose tuberculosis, and the number of such studies was small.
Senkoro 2016	Discrepancy between different source documents in case definitions. Unable to construct 2 x 2 tables.
Tupasi 2000	Data about symptoms in those with abnormal CXRs. Outside the scope of this review

CXR: chest radiography
HIV: human immunodeficiency virus
Mtb: *Mycobacterium tuberculosis*
TB: tuberculosis
CAD: computer assisted reading of digital radiographs
IGRA: Interferon Gamma Release Assay
MMR: mass miniature radiography
SSM+: sputum smear microscopy positive
TST: tuberculin skin test
NAAT: Nucleic Acid Amplification Test
More information on excluded studies and their characteristics is available from the review authors.

Characteristics of studies awaiting classification *[ordered by study ID]*

Assefa 2019

Patient Sampling	Consecutive: all patients with a non-tuberculosis medical diagnosis who were admitted to the general medical wards during a 3-month period
Patient characteristics and setting	Inpatients of a large referral hospital in Ethiopia, admitted with a medical illness other than tuberculosis disease, in a general medical ward; 19% HIV-infected
Index tests	Data on index tests cough of > 2 weeks duration and any tuberculosis symptom
Target condition and reference standard(s)	Bacteriologically confirmed pulmonary tuberculosis case: a patient whose submitted sputum sample was positive either by smear microscopy and/or Xpert MTB/RIF test, and diagnosed during the study
Flow and timing	15 were not able to submit sputum and excluded from analysis
Comparative	—
Notes	The study objective is to assess if there are missed pulmonary tuberculosis cases among medical inpatients. At least 2 concerns: the population is not representative for the study question: medical inpatients, high HIV prevalence, the data are reported aggregately for HIV+ and HIV- patients. The

Assefa 2019 *(Continued)*

	symptom screening was not systematic but extracted from retrospective chart review. N = 10 for sensitivity and 290 for specificity.

Bekken 2020

Patient Sampling	Consecutive: all contacts in a large prospective household contact (HHC) study were eligible. Contacts were persons living ≥ 75% of the time in the same household as the index case and sharing the same kitchen.
Patient characteristics and setting	Household contacts of smear-positive tuberculosis cases. 40% below 15 years of age. Palamaner Taluk, Andhra Pradesh, India.
Index tests	CXR any abnormality and tuberculosis abnormality
Target condition and reference standard(s)	Subclinical tuberculosis defined by positive Mtb culture in sputum or gastric aspirate (subjects < 5 years) specimen. Of 29 cases 90% had no symptoms and no CXR abnormalities.
Flow and timing	Adequate specimen for Mtb culture were harvested in 493 (93.9%) HHCs. Of these, 488 (99%) (and all 38 children < 5 years) had 2 samples harvested on 2 consecutive days.
Comparative	—
Notes	The study objective is to evaluate a Tuberculosis Contact Score (TCS) and Infectivity Score in identifying subclinical tuberculosis. Concerns: representativeness of the population - household contacts of whom 40% below 15 years of age. Also representativeness of the cases: 90% have no symptoms and no CXR abnormalities. The study provides data on N = 29 for sensitivity and n = 496 for specificity.

Bonsu 2020

Patient Sampling	—
Patient characteristics and setting	National tuberculosis Prevalence Survey Ghana
Index tests	—
Target condition and reference standard(s)	—
Flow and timing	—
Comparative	—
Notes	Already included in current meta-analysis based on Survey Report. See reference 'Ghana NTP 2015'.

Hamda 2020

Patient Sampling	Stratified random sampling proportional to size using a lottery method to select ANC clients. All pregnant women qualifying for inclusion (18 to 49 years, providing consent and able to produce sputum) were then consecutively enrolled until the required sample size was reached.

Hamda 2020 *(Continued)*

Patient characteristics and setting	Pregnant women attending ANC in 7 health facilities in Greater Gaborone and Kweneng East districts in Botswana. 17% HIV-infected.
Index tests	Cough ≥ 2 weeks and any tuberculosis symptoms (cough, fever, night sweats, and weight loss) ≥ 2 weeks' duration
Target condition and reference standard(s)	Two sputum samples (Spot 1 and 2) were collected; 1 was tested using Xpert MTB/RIF, the 2nd by culture using MGIT. A tuberculosis case was defined based on a positive culture.
Flow and timing	Of the 429 clients enrolled, 22 (5.1%) were excluded as their sputum samples did not fulfil specimen criteria or due to incomplete information or mismatched names, leaving 407 (94.9%) in the analysis.
Comparative	—
Notes	The study objectives are to determine 1) the prevalence of tuberculosis (TB) and HIV-TB co-infection in pregnancy, and 2) the sensitivities of symptomatic tuberculosis screening and Xpert MTB/RIF testing against gold standard culture. Concerns: the definition of any tuberculosis symptoms is different from how it is used in the review, which is usually of any one symptom of any duration. The risk of bias in domain 1 is unclear. N = 407 ANC women in the analysis, 2 had culture positive for Mtb. 8 had symptoms, none of those had tuberculosis. Sensitivity of symptoms is 0/2 = 0%. For specificity calculation n = 405.

Lansang 2021

Patient Sampling	—
Patient characteristics and setting	National tuberculosis Prevalence Survey Philippines 2016
Index tests	—
Target condition and reference standard(s)	—
Flow and timing	—
Comparative	—
Notes	Already included in current meta-analysis based on Survey Report. See reference 'NTP Philippines 2018'.

Li 2019

Patient Sampling	Study design in accordance with WHO guidance on tuberculosis prevalence surveys
Patient characteristics and setting	Subnational Tuberculosis Prevalence Survey in The Tibet Autonomous Region of China
Index tests	Persistent cough of 2 weeks' duration or haemoptysis; CXR suggestive of tuberculosis
Target condition and reference standard(s)	Approximately 10% of the study population had a microbacteriological test as the reference standard. Those were expected to submit 3 sputum samples, 1 for smear, 2 for culture. Bacteriologically confirmed tuberculosis cases were defined as those with positive smear (n = 24) or culture sputum (n = 39) results. Other, bacteriologically-negative cases (n = 176) included those with abnormal CXR

Li 2019 *(Continued)*

	suggestive of tuberculosis and those clinically diagnosed with tuberculosis by clinicians and radiologists after ruling out other diseases.
Flow and timing	No report on % who did not give sputum despite eligibility.
Comparative	—
Notes	Study objective: to determine the prevalence of pulmonary tuberculosis (TB) in Tibet. Concerns: substantial incorporation bias in domain 3; approximately 10% of > 30,000 study population had a microbacteriological test as the reference standard. The remaining 90% was classified as TB-negative based on report of absence of a persistent cough of 2 weeks' duration or haemoptysis and absence of CXR abnormalities suggestive of TB. Of 215 active TB cases, less than half would count as a case for the systematic review. The remainder is either bacteriologically negative (= clinically diagnosed) or were already diagnosed withTB before the survey (n = 49; 23%). To determine if this can be disaggregated needs more work. Risk of bias in domain 4 unclear.

Migambi 2020

Patient Sampling	—
Patient characteristics and setting	National TB Prevalence Survey Rwanda 2012
Index tests	—
Target condition and reference standard(s)	—
Flow and timing	—
Comparative	—
Notes	Already included in current meta-analysis based on Survey Report. See reference 'Rwanda MoH 2014'.

Nalunjogi 2021

Patient Sampling	—
Patient characteristics and setting	National Tuberculosis Prevalence Survey Uganda.
Index tests	—
Target condition and reference standard(s)	—
Flow and timing	—
Comparative	—
Notes	Already included in current meta-analysis based on Survey Report. See reference 'Republic of Uganda 2018'. The data as presented in the publication by Nalunjogi is not additional for the review, as their analysis includes only participants with presumptive tuberculosis, and not the full screened population.

Nguyen HV 2020

Patient Sampling	Study design in accordance with WHO guidance on tuberculosis prevalence surveys; n = 61,783
Patient characteristics and setting	National Tuberculosis Prevalence Survey Vietnam conducted in 2018
Index tests	Cough for 2 weeks or more; CXR abnormalities consistent with tuberculosis; combination of these
Target condition and reference standard(s)	Bacteriologically confirmed tuberculosis case. Xpert MTB/RIF or culture positive for Mtb or both. 4738 (7.7%) of participants were eligible for microbacteriological examination
Flow and timing	91% of 4738 submitted sputum
Comparative	In instances when data were available to construct a 2 x 2 table more than once for the same index test (CXR) in (almost) the same study population, we selected the most complete or most recent report
Notes	National Tuberculosis Prevalence Survey. This one (conducted in 2018) is not included in the review, but the previous one (2006) is as reference 'Hoa 2012'. In the review, in instances when data were available to construct a 2 x 2 table more than once for the same index test(s) in (almost) the same study population, we selected the most complete or most recent report, or made a random choice (see Results section). This to avoid further systematic bias. This implies that the Nguyen HV 2020 study would replace the Hoa 2012 study that is currently in the review.

Nguyen TBP 2020

Patient Sampling	96 randomly selected clusters. The clusters were sub-communes, with ± population of 1000 persons aged 15 years or older. Of 72,985 eligible participants, 57,597 (78.9%) participated in the Xpert MTB/RIF screening and 12,752 (17.5%) participated in the CXR screen; 11,235 (15.4% of the eligible study population) took part in both Xpert MTB/RIF screening and CXR, and hence formed the final study population. Women < 45 years were excluded because of the use of CXR.
Patient characteristics and setting	Men aged > 15 years and women > 45 years in Ca Mau province of VietNam, 2017/2018
Index tests	CXR abnormalities consistent with tuberculosis, by one or both of 2 readers
Target condition and reference standard(s)	All individuals provided sputum for Xpert MTB/RIF. Those who had either sputum that was Xpert MTB-positive (n = 33), or CXR that was reported as 'consistent with tuberculosis' by one or both readers (n = 659), were to submit 2 sputum specimens for liquid and solid mycobacterial culture.
Flow and timing	7168 (63.8% of the final study population) provided sputum specimens suitable for Xpert MTB/RIF testing. Among these, valid results were obtained for 7132 (99.5%). 10% of those providing sputum for culture had no results.
Comparative	—
Notes	The study objective is to compare (sensitivity/yield and feasibility/participation) of 2 community screening tests for tuberculosis: sputum examination using Xpert MTB/RIF and chest radiography. Concerns: high risk of bias in domain 1 (participant selection) and domain 4 (flow and timing); representativeness of population excluding women < 45 years; index test is defined based on 2 readers - different from the studies included in the review. N = 51 cases of tuberculosis (for sensitivity). The publication reports sensitivity based on 59 estimated by a capture-recapture model. The remainder provide data for specificity.

Pasipamire 2020

Patient Sampling	Stratified, consecutive. Four groups of women were enrolled: HIV-positive pregnant, HIV-negative pregnant, HIV-positive postpartum, and HIV-negative postpartum. Participants were consecutively enrolled until the sample size of 183 in each group was reached.
Patient characteristics and setting	Pregnant and postpartum women at 3 public health facilities in 3 of the 4 regions of Eswatini, 2015. Of 990 participants half are HIV+ (purposely selected for stratification purposes). 47% HIV+
Index tests	Any cough, any tuberculosis symptom
Target condition and reference standard(s)	All participants were expected to submit 2 samples of sputum (for Xpert MTB/RIF, smear microscopy, and culture using MGIT 960). Case definition: bacteriologically confirmed tuberculosis: culture positive for Mtb. 14/15 cases had no symptoms or signs of disease. Unclear how many had >1 positive microbacteriological test.
Flow and timing	In total, 776 of 990 (78%) participants produced sputum samples and of those 758 (98%) had samples available for culture testing of whom 704 (93%) had valid culture results
Comparative	—
Notes	The study objectives are to determine tuberculosis prevalence stratified by HIV status and identify screening algorithms that maximise detection of active tuberculosis among pregnant and postpartum women in Eswatini. Concerns: Domain1 - representativeness of the population. Of 990 participants half are HIV+ (purposely selected for stratification purposes). Domain 4: 704 (approximately 70%) had culture results available. The remaining did not give sputum. Domain 3: % cases without signs of disease or 2nd confirmed bacteriology unclear. It is possible to calculate accuracy for n = 990 and n = 704, but not for the HIV-negatives only. Data on sensitivity: 15 women had a positive tuberculosis culture (reference standard) of which 11 were HIV+.

Reichler 2020

Patient Sampling	Consecutive
Patient characteristics and setting	Contacts to culture-confirmed adult tuberculosis patients and their close contacts at 9 US and Canadian sites, between 2002 and 2006
Index tests	Any cough
Target condition and reference standard(s)	Cases defined as 'co-prevalent tuberculosis': clinician-defined tuberculosis identified from tuberculosis treatment registry matches if the start date for tuberculosis treatment was before or < 30 days after blood draw/interview. Microbacteriological procedures unclear.
Flow and timing	The date of tuberculosis diagnosis was defined as the start date for tuberculosis treatment. Contacts with tuberculosis diagnosed before or < 30 days after blood draw (for immunology study) were considered co-prevalent cases.
Comparative	—
Notes	The study objective is to examined cytokine immune response profiles among contacts to tuberculosis patients to identify immunologic and epidemiologic correlates of tuberculosis. Data for index test 'any cough' could be identified. The study provides 41 cases for sensitivity and > 1200 for specificity. High risk of bias in at least 3 domains: Domain 1: of 3221 eligible contacts 61% did not enrol. Domain 3: tuberculosis cases were identified from tuberculosis registry matches, implying that the

Reichler 2020 *(Continued)*	case definition is unclear - essentially clinician-defined - bacteriology results are unknown. Domain 2/4: date of tuberculosis diagnosis may have been before date of symptom (index test) interview.

ANC: antenatal clinic
CXR: chest radiography
Mtb: *Mycobacterium tuberculosis*
WHO: World Health Organization

DATA

Presented below are all the data for all of the tests entered into the review.

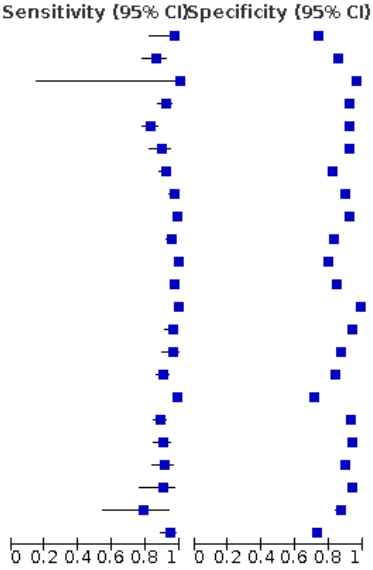
Table Tests. Data tables by test

Test	No. of studies	No. of participants
1 Any CXR abnormality	23	1039057
2 CXR abnormalities suggestive of tuberculosis	19	466970
3 Cough for 2 or more weeks	41	1547358
4 Cough of any duration	21	771025
5 Any tuberculosis symptom	29	510892
6 Parallel CXR and symptom screening	25	1143894
7 CXR for parallel comparison	24	1093540
8 Cough for 3 or more weeks	6	333737
9 Combination of tuberculosis symptoms - out of several	11	444276

Test 1. Any CXR abnormality

Any CXR abnormality

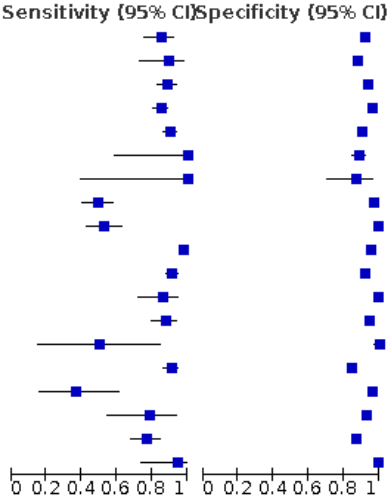
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
den Boon 2006	28	674	1	1905	0.97 [0.82, 1.00]	0.74 [0.72, 0.76]
Federal MoH Sudan 2018	88	11573	14	66255	0.86 [0.78, 0.92]	0.85 [0.85, 0.85]
Fox 2012	2	20	0	523	1.00 [0.16, 1.00]	0.96 [0.94, 0.98]
Ghana NTP 2015	152	5006	13	54547	0.92 [0.87, 0.96]	0.92 [0.91, 0.92]
Kapata 2016	219	3539	46	41829	0.83 [0.78, 0.87]	0.92 [0.92, 0.92]
Kebede 2014	98	3720	12	42718	0.89 [0.82, 0.94]	0.92 [0.92, 0.92]
Kenya MoH 2018	274	11275	24	50911	0.92 [0.88, 0.95]	0.82 [0.82, 0.82]
Law 2015	230	4167	7	34591	0.97 [0.94, 0.99]	0.89 [0.89, 0.90]
MoH Cambodia 2012	306	3103	5	33807	0.98 [0.96, 0.99]	0.92 [0.91, 0.92]
MoH Indonesia 2015	401	10801	21	53115	0.95 [0.92, 0.97]	0.83 [0.83, 0.83]
MoH Myanmar 2012	290	10347	2	39602	0.99 [0.98, 1.00]	0.79 [0.79, 0.80]
Mongolia MoH 2016	238	7643	7	41633	0.97 [0.94, 0.99]	0.84 [0.84, 0.85]
MoPH DPRK 2017	333	695	2	58619	0.99 [0.98, 1.00]	0.99 [0.99, 0.99]
MoPH Thailand 2017	136	4157	6	57644	0.96 [0.91, 0.98]	0.93 [0.93, 0.93]
Morasert 2018	81	509	3	3414	0.96 [0.90, 0.99]	0.87 [0.86, 0.88]
NTP Bangladesh 2017	251	16126	27	82155	0.90 [0.86, 0.94]	0.84 [0.83, 0.84]
NTP Philippines 2018	430	11716	8	29289	0.98 [0.96, 0.99]	0.71 [0.71, 0.72]
Qadeer 2016	299	7562	38	93353	0.89 [0.85, 0.92]	0.93 [0.92, 0.93]
Republic of Uganda 2018	144	2706	16	38137	0.90 [0.84, 0.94]	0.93 [0.93, 0.94]
Republic of Zimbabwe 2015	92	3339	9	29061	0.91 [0.84, 0.96]	0.90 [0.89, 0.90]
Rwanda MoH 2014	36	2616	4	40413	0.90 [0.76, 0.97]	0.94 [0.94, 0.94]
Telisinghe 2014	15	82	4	532	0.79 [0.54, 0.94]	0.87 [0.84, 0.89]
van't Hoog 2012	113	5229	7	13867	0.94 [0.88, 0.98]	0.73 [0.72, 0.73]



Test 2. CXR abnormalities suggestive of tuberculosis

CXR abnormalities suggestive of tuberculosis

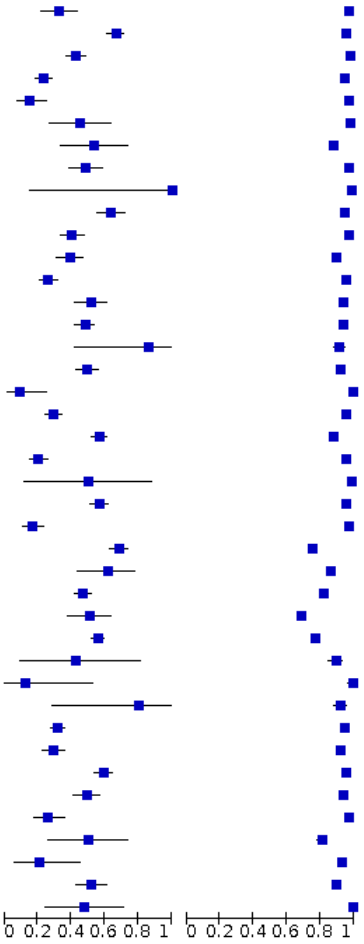
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Adetifa 2016	64	3343	11	39529	0.85 [0.75, 0.92]	0.92 [0.92, 0.92]
den Boon 2006	26	311	3	2268	0.90 [0.73, 0.98]	0.88 [0.87, 0.89]
FRoNigeria 2014	128	2840	16	40215	0.89 [0.83, 0.94]	0.93 [0.93, 0.94]
Hoa 2012	229	3452	40	83921	0.85 [0.80, 0.89]	0.96 [0.96, 0.96]
Kenya MoH 2018	269	6156	29	56030	0.90 [0.86, 0.93]	0.90 [0.90, 0.90]
Koesoemadinata 2018	7	38	0	301	1.00 [0.59, 1.00]	0.89 [0.85, 0.92]
Lu 2016	4	4	0	27	1.00 [0.40, 1.00]	0.87 [0.70, 0.96]
Malawi MoH 2016	65	951	67	30476	0.49 [0.40, 0.58]	0.97 [0.97, 0.97]
Melendez 2017	56	95	50	23637	0.53 [0.43, 0.63]	1.00 [1.00, 1.00]
MoH Cambodia 2012	305	1892	6	35018	0.98 [0.96, 0.99]	0.95 [0.95, 0.95]
MoH Myanmar 2012	267	3954	25	45995	0.91 [0.88, 0.94]	0.92 [0.92, 0.92]
Mor 2012	37	113	6	13223	0.86 [0.72, 0.95]	0.99 [0.99, 0.99]
Morasert 2018	74	208	10	3715	0.88 [0.79, 0.94]	0.95 [0.94, 0.95]
Nair 2016a	4	0	4	143	0.50 [0.16, 0.84]	1.00 [0.97, 1.00]
Pelissari 2018	165	1575	16	8570	0.91 [0.86, 0.95]	0.84 [0.84, 0.85]
Seri 2017	7	36	12	888	0.37 [0.16, 0.62]	0.96 [0.95, 0.97]
Telisinghe 2014	15	44	4	570	0.79 [0.54, 0.94]	0.93 [0.90, 0.95]
van't Hoog 2011	86	133	26	898	0.77 [0.68, 0.84]	0.87 [0.85, 0.89]
Wei 2014	18	326	1	53923	0.95 [0.74, 1.00]	0.99 [0.99, 0.99]



Test 3. Cough for 2 or more weeks

Cough for 2 or more weeks

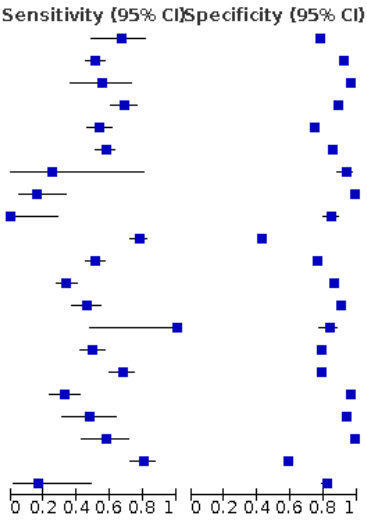
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Adetifa 2016	25	1366	52	41657	0.32 [0.22, 0.44]	0.97 [0.97, 0.97]
Chadha 2018	223	3851	112	83344	0.67 [0.61, 0.72]	0.96 [0.95, 0.96]
Cheng 2015	135	5157	180	247322	0.43 [0.37, 0.49]	0.98 [0.98, 0.98]
Claassens 2017b	72	802	230	14345	0.24 [0.19, 0.29]	0.95 [0.94, 0.95]
Claassens 2017d	11	719	61	22473	0.15 [0.08, 0.26]	0.97 [0.97, 0.97]
Corbett 2010b	14	166	17	6924	0.45 [0.27, 0.64]	0.98 [0.97, 0.98]
den Boon 2006	14	314	12	2171	0.54 [0.33, 0.73]	0.87 [0.86, 0.89]
Federal MoH Sudan 2018	51	2612	54	80293	0.49 [0.39, 0.59]	0.97 [0.97, 0.97]
Fox 2012	2	10	0	533	1.00 [0.16, 1.00]	0.98 [0.97, 0.99]
FRoNigeria 2014	92	2381	52	41661	0.64 [0.55, 0.72]	0.95 [0.94, 0.95]
Ghana NTP 2015	82	1887	120	59637	0.41 [0.34, 0.48]	0.97 [0.97, 0.97]
Ho 2016	66	4710	103	38556	0.39 [0.32, 0.47]	0.89 [0.89, 0.89]
Hoa 2012	71	4129	198	89360	0.26 [0.21, 0.32]	0.96 [0.95, 0.96]
Kebede 2014	57	2968	53	43619	0.52 [0.42, 0.61]	0.94 [0.93, 0.94]
Kenya MoH 2018	147	3990	158	58755	0.48 [0.42, 0.54]	0.94 [0.93, 0.94]
Koesoemadinata 2018	6	29	1	310	0.86 [0.42, 1.00]	0.91 [0.88, 0.94]
Law 2015	118	3079	119	35679	0.50 [0.43, 0.56]	0.92 [0.92, 0.92]
Lewis 2009b	3	9	28	1338	0.10 [0.02, 0.26]	0.99 [0.99, 1.00]
MoH Cambodia 2012	92	1712	222	35391	0.29 [0.24, 0.35]	0.95 [0.95, 0.96]
MoH Indonesia 2015	242	8310	184	59208	0.57 [0.52, 0.62]	0.88 [0.87, 0.88]
Mongolia MoH 2016	51	2495	197	47451	0.21 [0.16, 0.26]	0.95 [0.95, 0.95]
Moosazadeh 2015	3	7	3	655	0.50 [0.12, 0.88]	0.99 [0.98, 1.00]
MoPH DPRK 2017	194	2750	146	57593	0.57 [0.52, 0.62]	0.95 [0.95, 0.96]
MoPH Thailand 2017	24	1751	118	60643	0.17 [0.11, 0.24]	0.97 [0.97, 0.97]
Morishita 2017a	195	3138	89	9485	0.69 [0.63, 0.74]	0.75 [0.74, 0.76]
Morishita 2017b	21	227	13	1364	0.62 [0.44, 0.78]	0.86 [0.84, 0.87]
Morishita 2017c	177	1059	201	4696	0.47 [0.42, 0.52]	0.82 [0.81, 0.83]
Morishita 2017d	32	659	31	1423	0.51 [0.38, 0.64]	0.68 [0.66, 0.70]
Morishita 2017e	425	5083	334	16968	0.56 [0.52, 0.60]	0.77 [0.76, 0.78]
Muyoyeta 2017	3	28	4	228	0.43 [0.10, 0.82]	0.89 [0.85, 0.93]
Nair 2016a	1	1	7	142	0.13 [0.00, 0.53]	0.99 [0.96, 1.00]
Ntinginya 2012	4	17	1	197	0.80 [0.28, 0.99]	0.92 [0.88, 0.95]
NTP Philippines 2018	150	2665	316	43558	0.32 [0.28, 0.37]	0.94 [0.94, 0.94]
Pelissari 2018	53	828	128	9317	0.29 [0.23, 0.36]	0.92 [0.91, 0.92]
Qadeer 2016	197	4866	137	99431	0.59 [0.53, 0.64]	0.95 [0.95, 0.95]
Republic of Uganda 2018	79	2635	81	38359	0.49 [0.41, 0.57]	0.94 [0.93, 0.94]
Republic of Zimbabwe 2015	28	1040	78	32544	0.26 [0.18, 0.36]	0.97 [0.97, 0.97]
Seri 2017	9	153	9	640	0.50 [0.26, 0.74]	0.81 [0.78, 0.83]
Telisinghe 2014	4	44	15	570	0.21 [0.06, 0.46]	0.93 [0.90, 0.95]
van't Hoog 2012	64	2200	59	18243	0.52 [0.43, 0.61]	0.89 [0.89, 0.90]
Wei 2014	9	299	10	53950	0.47 [0.24, 0.71]	0.99 [0.99, 1.00]



Test 4. Cough of any duration

Cough of any duration

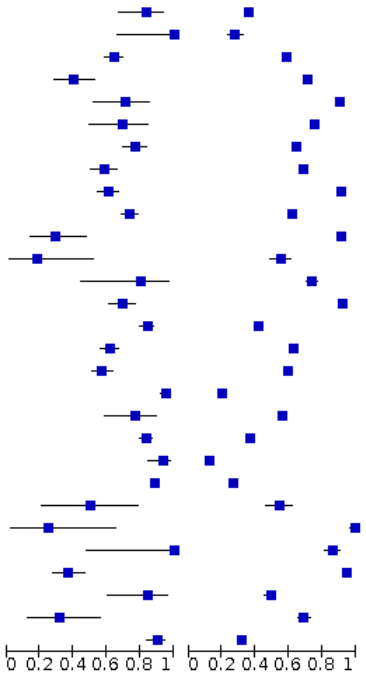
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ayles 2009b	24	1250	12	4380	0.67 [0.49, 0.81]	0.78 [0.77, 0.79]
Cheng 2015	162	20481	153	231998	0.51 [0.46, 0.57]	0.92 [0.92, 0.92]
Corbett 2010b	17	306	14	6784	0.55 [0.36, 0.73]	0.96 [0.95, 0.96]
FRoNigeria 2014	99	5053	45	38989	0.69 [0.61, 0.76]	0.89 [0.88, 0.89]
Ho 2016	91	11209	78	32057	0.54 [0.46, 0.62]	0.74 [0.74, 0.75]
Kenya MoH 2018	175	9130	130	53615	0.57 [0.52, 0.63]	0.85 [0.85, 0.86]
Kimerling 1999	1	8	3	115	0.25 [0.01, 0.81]	0.93 [0.88, 0.97]
Lewis 2009b	5	19	26	1328	0.16 [0.05, 0.34]	0.99 [0.98, 0.99]
Little 2018	0	41	11	227	0.00 [0.00, 0.28]	0.85 [0.80, 0.89]
MoH Cambodia 2012	243	21312	71	15791	0.77 [0.72, 0.82]	0.43 [0.42, 0.43]
MoH Myanmar 2012	155	12113	149	38950	0.51 [0.45, 0.57]	0.76 [0.76, 0.77]
Mongolia MoH 2016	84	7007	164	42939	0.34 [0.28, 0.40]	0.86 [0.86, 0.86]
MoPH Thailand 2017	65	6330	77	56064	0.46 [0.37, 0.54]	0.90 [0.90, 0.90]
Ntinginya 2012	5	36	0	178	1.00 [0.48, 1.00]	0.83 [0.77, 0.88]
Pelissari 2018	89	2137	92	8008	0.49 [0.42, 0.57]	0.79 [0.78, 0.80]
Republic of Uganda 2018	108	8798	52	32196	0.68 [0.60, 0.75]	0.79 [0.78, 0.79]
Republic of Zimbabwe 2015	35	1422	72	32207	0.33 [0.24, 0.42]	0.96 [0.96, 0.96]
Rwanda MoH 2014	19	2618	21	40463	0.47 [0.32, 0.64]	0.94 [0.94, 0.94]
Singh 2013	30	29	22	1527	0.58 [0.43, 0.71]	0.98 [0.97, 0.99]
van't Hoog 2012	99	8455	24	11982	0.80 [0.72, 0.87]	0.59 [0.58, 0.59]
Wood 2007	2	134	10	605	0.17 [0.02, 0.48]	0.82 [0.79, 0.85]



Test 5. Any tuberculosis symptom

Any tuberculosis symptom

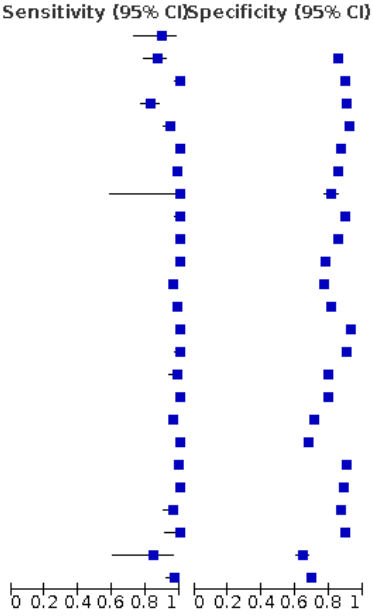
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ayles 2009b	30	3614	6	2016	0.83 [0.67, 0.94]	0.36 [0.35, 0.37]
Chheng 2008b	9	261	0	102	1.00 [0.66, 1.00]	0.28 [0.24, 0.33]
Claassens 2017b	195	6240	107	8907	0.65 [0.59, 0.70]	0.59 [0.58, 0.60]
Claassens 2017d	29	6702	43	16490	0.40 [0.29, 0.53]	0.71 [0.71, 0.72]
Corbett 2010b	22	686	9	6404	0.71 [0.52, 0.86]	0.90 [0.90, 0.91]
den Boon 2006	20	637	9	1941	0.69 [0.49, 0.85]	0.75 [0.74, 0.77]
FRoNigeria 2014	111	15701	33	28341	0.77 [0.69, 0.84]	0.64 [0.64, 0.65]
Ho 2016	99	13545	70	29721	0.59 [0.51, 0.66]	0.69 [0.68, 0.69]
Kapata 2016	161	4292	104	41542	0.61 [0.55, 0.67]	0.91 [0.90, 0.91]
Kenya MoH 2018	225	24109	80	38636	0.74 [0.68, 0.79]	0.62 [0.61, 0.62]
Lewis 2009b	9	117	22	1230	0.29 [0.14, 0.48]	0.91 [0.90, 0.93]
Little 2018	2	120	9	148	0.18 [0.02, 0.52]	0.55 [0.49, 0.61]
Mabuto 2015	8	178	2	498	0.80 [0.44, 0.97]	0.74 [0.70, 0.77]
Malawi MoH 2016	92	2623	40	28824	0.70 [0.61, 0.77]	0.92 [0.91, 0.92]
MoH Cambodia 2005	229	12673	43	9215	0.84 [0.79, 0.88]	0.42 [0.41, 0.43]
MoH Myanmar 2012	188	18922	116	32141	0.62 [0.56, 0.67]	0.63 [0.63, 0.63]
Mongolia MoH 2016	142	20373	106	29573	0.57 [0.51, 0.63]	0.59 [0.59, 0.60]
Morishita 2017a	270	10055	14	2568	0.95 [0.92, 0.97]	0.20 [0.20, 0.21]
Morishita 2017b	26	693	8	898	0.76 [0.59, 0.89]	0.56 [0.54, 0.59]
Morishita 2017c	317	3632	61	2123	0.84 [0.80, 0.87]	0.37 [0.36, 0.38]
Morishita 2017d	59	1820	4	262	0.94 [0.85, 0.98]	0.13 [0.11, 0.14]
Morishita 2017e	672	16200	87	5851	0.89 [0.86, 0.91]	0.27 [0.26, 0.27]
Muyoyeta 2017	6	76	6	90	0.50 [0.21, 0.79]	0.54 [0.46, 0.62]
Nair 2016a	2	1	6	142	0.25 [0.03, 0.65]	0.99 [0.96, 1.00]
Ntinginya 2012	5	29	0	185	1.00 [0.48, 1.00]	0.86 [0.81, 0.91]
Republic of Zimbabwe 2015	39	1794	67	31790	0.37 [0.28, 0.47]	0.95 [0.94, 0.95]
Seri 2017	16	471	3	453	0.84 [0.60, 0.97]	0.49 [0.46, 0.52]
Telisinghe 2014	6	191	13	423	0.32 [0.13, 0.57]	0.69 [0.65, 0.73]
van't Hoog 2012	111	13878	12	6565	0.90 [0.84, 0.95]	0.32 [0.31, 0.33]



Test 6. Parallel CXR and symptom screening

Parallel CXR and symptom screening

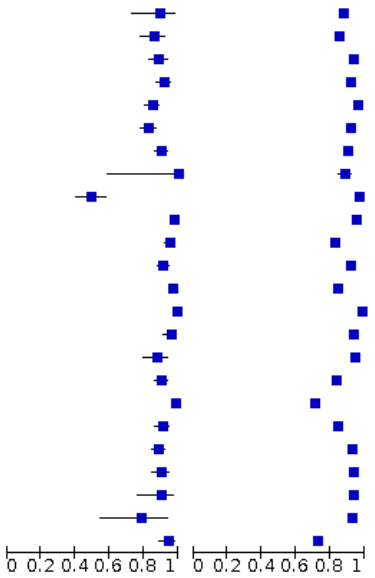
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
den Boon 2006	26	0	3	0	0.90 [0.73, 0.98]	Not estimable
Federal MoH Sudan 2018	96	12553	15	70346	0.86 [0.79, 0.92]	0.85 [0.85, 0.85]
FRoNigeria 2014	144	4544	0	39498	1.00 [0.97, 1.00]	0.90 [0.89, 0.90]
Ghana NTP 2015	167	6189	35	55335	0.83 [0.77, 0.88]	0.90 [0.90, 0.90]
Hoa 2012	253	7245	16	86244	0.94 [0.91, 0.97]	0.92 [0.92, 0.92]
Kapata 2016	265	5883	0	39951	1.00 [0.99, 1.00]	0.87 [0.87, 0.87]
Kenya MoH 2018	301	9020	4	53725	0.99 [0.97, 1.00]	0.86 [0.85, 0.86]
Koesoemadinata 2018	7	63	0	276	1.00 [0.59, 1.00]	0.81 [0.77, 0.85]
Malawi MoH 2016	132	3300	0	28147	1.00 [0.97, 1.00]	0.90 [0.89, 0.90]
MoH Cambodia 2005	272	3301	0	18859	1.00 [0.99, 1.00]	0.85 [0.85, 0.86]
MoH Indonesia 2015	426	15020	0	52498	1.00 [0.99, 1.00]	0.78 [0.77, 0.78]
MoH Myanmar 2012	299	11936	12	39120	0.96 [0.93, 0.98]	0.77 [0.76, 0.77]
Mongolia MoH 2016	245	9365	3	40696	0.99 [0.97, 1.00]	0.81 [0.81, 0.82]
MoPH DPRK 2017	340	4462	0	55881	1.00 [0.99, 1.00]	0.93 [0.92, 0.93]
MoPH Thailand 2017	142	5908	0	56486	1.00 [0.97, 1.00]	0.91 [0.90, 0.91]
Morasert 2018	83	817	1	3106	0.99 [0.94, 1.00]	0.79 [0.78, 0.80]
NTP Bangladesh 2017	278	20316	0	78116	1.00 [0.99, 1.00]	0.79 [0.79, 0.80]
NTP Philippines 2018	448	13517	18	32706	0.96 [0.94, 0.98]	0.71 [0.70, 0.71]
Pelissari 2018	181	3318	0	6827	1.00 [0.98, 1.00]	0.67 [0.66, 0.68]
Qadeer 2016	339	10120	1	95441	1.00 [0.98, 1.00]	0.90 [0.90, 0.91]
Republic of Uganda 2018	160	4852	0	36142	1.00 [0.98, 1.00]	0.88 [0.88, 0.88]
Republic of Zimbabwe 2015	103	4533	4	29096	0.96 [0.91, 0.99]	0.87 [0.86, 0.87]
Rwanda MoH 2014	40	4707	0	38381	1.00 [0.91, 1.00]	0.89 [0.89, 0.89]
Telisinghe 2014	16	219	3	395	0.84 [0.60, 0.97]	0.64 [0.60, 0.68]
van't Hoog 2012	119	6319	4	14114	0.97 [0.92, 0.99]	0.69 [0.68, 0.70]



Test 7. CXR for parallel comparison

CXR for parallel comparison

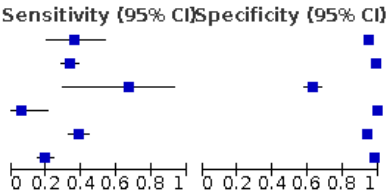
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
den Boon 2006	26	311	3	2268	0.90 [0.73, 0.98]	0.88 [0.87, 0.89]
Federal MoH Sudan 2018	88	11573	14	66255	0.86 [0.78, 0.92]	0.85 [0.85, 0.85]
FRoNigeria 2014	128	2840	16	40215	0.89 [0.83, 0.94]	0.93 [0.93, 0.94]
Ghana NTP 2015	152	5006	13	54547	0.92 [0.87, 0.96]	0.92 [0.91, 0.92]
Hoa 2012	229	3452	40	83921	0.85 [0.80, 0.89]	0.96 [0.96, 0.96]
Kapata 2016	219	3539	46	41829	0.83 [0.78, 0.87]	0.92 [0.92, 0.92]
Kenya MoH 2018	269	6156	29	56030	0.90 [0.86, 0.93]	0.90 [0.90, 0.90]
Koesoemadinata 2018	7	38	0	301	1.00 [0.59, 1.00]	0.89 [0.85, 0.92]
Malawi MoH 2016	65	951	67	30476	0.49 [0.40, 0.58]	0.97 [0.97, 0.97]
MoH Cambodia 2005	305	1892	6	35018	0.98 [0.96, 0.99]	0.95 [0.95, 0.95]
MoH Indonesia 2015	401	10801	21	53115	0.95 [0.92, 0.97]	0.83 [0.83, 0.83]
MoH Myanmar 2012	267	3954	25	45995	0.91 [0.88, 0.94]	0.92 [0.92, 0.92]
Mongolia MoH 2016	238	7643	7	41633	0.97 [0.94, 0.99]	0.84 [0.84, 0.85]
MoPH DPRK 2017	333	695	2	58619	0.99 [0.98, 1.00]	0.99 [0.99, 0.99]
MoPH Thailand 2017	136	4157	6	57644	0.96 [0.91, 0.98]	0.93 [0.93, 0.93]
Morasert 2018	74	208	10	3715	0.88 [0.79, 0.94]	0.95 [0.94, 0.95]
NTP Bangladesh 2017	251	16126	27	82155	0.90 [0.86, 0.94]	0.84 [0.83, 0.84]
NTP Philippines 2018	430	11716	8	29289	0.98 [0.96, 0.99]	0.71 [0.71, 0.72]
Pelissari 2018	165	1575	16	8570	0.91 [0.86, 0.95]	0.84 [0.84, 0.85]
Qadeer 2016	299	7562	38	93353	0.89 [0.85, 0.92]	0.93 [0.92, 0.93]
Republic of Uganda 2018	144	2706	16	38137	0.90 [0.84, 0.94]	0.93 [0.93, 0.94]
Rwanda MoH 2014	36	2616	4	40413	0.90 [0.76, 0.97]	0.94 [0.94, 0.94]
Telisinghe 2014	15	44	4	570	0.79 [0.54, 0.94]	0.93 [0.90, 0.95]
van't Hoog 2012	113	5229	7	13867	0.94 [0.88, 0.98]	0.73 [0.72, 0.73]



Test 8. Cough for 3 or more weeks

Cough for 3 or more weeks

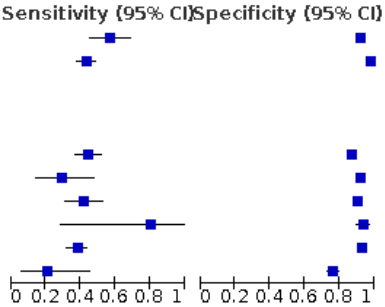
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ayles 2009b	13	338	23	5292	0.36 [0.21, 0.54]	0.94 [0.93, 0.95]
Cheng 2015	106	3285	209	249194	0.34 [0.28, 0.39]	0.99 [0.99, 0.99]
Chheng 2008b	6	134	3	229	0.67 [0.30, 0.93]	0.63 [0.58, 0.68]
Lewis 2009b	2	7	29	1340	0.06 [0.01, 0.21]	0.99 [0.99, 1.00]
MoH Cambodia 2005	105	1396	167	20492	0.39 [0.33, 0.45]	0.94 [0.93, 0.94]
MoH Myanmar 2012	60	1373	244	49690	0.20 [0.15, 0.25]	0.97 [0.97, 0.97]



Test 9. Combination of tuberculosis symptoms - out of several

Combination of tuberculosis symptoms - out of several

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Adetifa 2016	44	3453	33	39570	0.57 [0.45, 0.68]	0.92 [0.92, 0.92]
Cheng 2015	136	5254	179	247225	0.43 [0.38, 0.49]	0.98 [0.98, 0.98]
Claassens 2017b	0	0	0	0	Not estimable	Not estimable
Claassens 2017d	0	0	0	0	Not estimable	Not estimable
Corbett 2010b	0	0	0	0	Not estimable	Not estimable
Ho 2016	75	5681	94	37585	0.44 [0.37, 0.52]	0.87 [0.87, 0.87]
Lewis 2009b	9	111	22	1236	0.29 [0.14, 0.48]	0.92 [0.90, 0.93]
Morasert 2018	35	397	49	3526	0.42 [0.31, 0.53]	0.90 [0.89, 0.91]
Ntinginya 2012	4	13	1	201	0.80 [0.28, 0.99]	0.94 [0.90, 0.97]
NTP Bangladesh 2017	106	7188	172	91244	0.38 [0.32, 0.44]	0.93 [0.93, 0.93]
Telisinghe 2014	4	145	15	469	0.21 [0.06, 0.46]	0.76 [0.73, 0.80]



ADDITIONAL TABLES

Table 1. Investigations of heterogeneity: factors modifying sensitivity or specificity or both of symptom index tests

Variable	Category	Cough ≥ 2 weeks (n = 41)					Cough of any duration (n = 21)					Any tuberculosis symptom (n = 29)				
		# studies	Sensitivity (95% CI)		Specificity (95% CI)		P value[a]	# studies	Sensitivity (95% CI)		Specificity (95% CI)		P value	# studies	Sensitivity (95% CI)		Specificity (95% CI)		P value
Study-level characteristics																			
Country income level	LIC	11	48.5%	38.0% to 58.9%	94.6%	91.9% to 97.3%	0.057	9	64.8%	54.8% to 73.6%	*80.8%*	*69.1% to 88.9%*	0.010[b]	10	79.5%	67.4% to 87.8%	*62.5%*	*42.1% to 79.2%*	0.013[b]
	LMIC	20	43.9%	36.7% to 51.1%	92.5%	89.8% to 95.2%		5	48.6%	35.9% to 61.5%	*91.6%*	*82.2% to 96.2%*		13	70.8%	58.7% to 80.5%	*63.3%*	*45.4% to 78.2%*	
	UMIC	10	31.6%	21.7% to 41.6%	96.8%	95.1% to 98.6%		7[d]	34.4%	23.3% to 47.5%	*90.8%*	*82.7% to 95.4%*		6	50.0%	30.6% to 69.4%	*72.7%*	*47.8% to 88.6%*	
Tuberculosis prevalence	< 0.5%	19	*39.9%*	*32.5% to 47.9%*	96.4%	94.8% to 97.5%	0.003[c]	*10*	*51.0%*	*39.3% to 62.7%*	*89.6%*	*81.8% to 94.3%*	0.59[d]	*8*	*60.8%*	*43.5% to 75.8%*	*79.1%*	*61.1% to 90.1%*	0.20[d]
	≥ 0.5%	22	*43.4%*	*35.8% to 51.4%*	91.7%	88.6% to 94.1%		*11*	*50.8%*	*38.3% to 63.3%*	*85.3%*	*75.7% to 91.6%*		*21*	*74.0%*	*81.9% to 64.2%*	*58.7%*	*45.2% to 71.0%*	
Participant characteristics																			
Proportion with previous tuberculosis	< 5%	13	48.2%	38.0% to 58.5%	*93.8%*	*90.1% to 96.2%*	0.021[b]	11	60.6%	49.4% to 70.9%	*84.9%*	*74.5% to 91.5%*	0.024[b]	*10*	*74.8%*	*59.3% to 85.8%*	*64.3%*	*42.9% to 81.2%*	0.61[d]
	≥ 5%	18	37.4%	29.0% to 45.7%	*93.6%*	*90.9% to 96.2%*		6	34.1%	20.6% to 50.8%	*91.6%*	*81.7% to 96.4%*		*15*	*68.9%*	*54.6% to 80.2%*	*63.8%*	*46.2% to 78.3%*	
	Missing	10						4						4					

Abbreviations: CI: confidence interval; LIC: low-income country; LMIC: lower middle-income country; UMIC: upper middle-income country.
*We examined whether the variable modified sensitivity, specificity, or both.
P values are shown of the most parsimonious model with the best fit:
[a]Only sensitivity was modified by the variable, but specificity was not. Specificity is therefore shown in italic font.

[b]Only specificity was modified by the variable, but sensitivity was not. Sensitivity is therefore shown in italic font.
[c]Neither sensitivity nor specificity was statistically significantly modified by the variable. Both are shown in italic font.
[d]Includes one study from a high-income country.

Table 2. Sensitivity analyses for symptom index tests

	Cough ≥2 weeks					Cough of any duration					Any tuberculosis symptom				
	# studies	Sensitivity (95% CI)		Specificity (95% CI)		# studies	Sensitivity (95% CI)		Specificity (95% CI)		# studies	Sensitivity (95% CI)		Specificity (95% CI)	
Overall meta-analytic summary	41	42.1%	36.6% to 47.7%	94.4%	92.6% to 95.8%	21	51.3%	42.8% to 59.7%	87.6%	81.6% to 91.8%	29	70.6%	61.7% to 78.2%	65.1%	53.3% to 75.4%
Patient Selection domain - low risk of bias only	32	40.3%	34.3% to 46.5%	95.0%	93.3% to 96.3%	19	51.0%	41.7% to 60.2%	86.8%	80.6% to 91.2%	23	65.2%	55.8% to 73.6%	70.5%	58.4% to 80.2%
Reference Standard domain - low risk of bias only	8	29.3%	19.4% to 41.7%	94.7%	89.5% to 97.4%	8	41.7%	18.8% to 56.8%	92.0%	83.5% to 96.3%	12	62.9%	47.4% to 76.1%	66.5%	53.8% to 77.2%
Flow and Timing domain - low risk of bias only	11	38.3%	28.7% to 48.9%	96.4%	93.9% to 97.8%	10	41.4%	29.2% to 54.8%	89.4%	84.6% to 92.9%	9	62.9%	43.5% to 78.9%	79.3%	55.6% to 92.1%
Studies with a microbiological reference standard on 20% or more participants	16	37.8%	29.2% to 47.3%	93.2%	89.4% to 95.7%	10	43.2%	26.7% to 61.4%	89.1%	78.9% to 94.7%	17	67.2%	55.8% to 76.9%	61.5%	50.9% to 71.0%
Studies with a culture-based microbiological reference standard	34	40.4%	34.3% to 46.8%	95.7%	94.3% to 96.7%	20	51.4%	42.3% to 60.4%	87.9%	81.8% to 92.1%	23	65.1%	55.6% to 73.6%	72.7%	61.6% to 81.6%

CI: confidence interval

Table 3. Other symptom combinations

Study ID	Population	Any tuberculosis symptom	Other symptom combination	Cough for 2 or 3 or more weeks

Table 3. Other symptom combinations *(Continued)*

		Sensitivity	Specificity	Definition	Sensitivity	Specificity	Definition	Sensitivity	Specificity
Ayles 2009b	HIV-negative subset of general population	83%	64%	Cough ≥ 3 weeks, haemoptysis, or any other 2 symptoms	75%	59%	Cough for 3 or more weeks (includes haemoptysis)	36%	94%
Adetifa 2016	General population	Not available		Cough ≥ 2 weeks, or 3 or more other symptoms out of cough, fever, chest pain, night sweats, shortness of breath, anorexia, weight loss, haemoptysis	57%	92%	Cough for 2 or more weeks	32%	97%
Chheng 2008a	HIV-negative VCT attendants only	100%	28%	Fever or haemoptysis or weight loss	100%	40%	Cough for 3 weeks or more	67%	63%
Claassens 2017b	HIV-negative subset of general population	65%	59%	Cough ≥ 2 weeks or ≥ 2 symptoms of cough less than 2 weeks, night sweats, weight loss, fever, chest pain, shortness of breath	47%	77%	Cough for 2 or more weeks	24%	95%
Claassens 2017d	HIV-negative subset of general population	40%	71%	Cough ≥ 2 weeks or ≥ 2 symptoms of cough less than 2 weeks, night sweats, weight loss, fever, chest pain, shortness of breath	26%	89%	Cough for 2 or more weeks	15%	97%
Corbett 2010b	HIV-negative subset of general population	71%	90%	Either of: cough or weight loss	65%	93%	Cough for 2 or more weeks	45%	98%
Ho 2016	General population	59%	69%	Any symptom for the past 2 weeks or more	44%	87%	Cough for 2 or more weeks	39%	89%
Lewis 2009b	Occupational (gold miners) - HIV-negative subset	29%	91%	At least one of cough or weight loss; new or worsening sputum for at least 2 weeks	29%	92%	Cough for 2 or more weeks	9.7%	99%
Ntinginya 2012	(Household) contacts	100%	86%	≥ 2 tuberculosis symptoms out of 7	80%	94%	Cough for 2 or more weeks	80%	92%

Table 3. Other symptom combinations *(Continued)*

Telisinghe 2014	Prison inmates	32%	69%	Any one of cough ≥ 2 weeks, night sweats or unintentional weight loss	21%	76%	Cough for 2 or more weeks	21%	93%

Table 4. Investigations of heterogeneity: factors modifying sensitivity or specificity or both of CXR index test(s)

Variable	Category	Any CXR abnormality (n = 23)						CXR abnormalities suggestive of tuberculosis (n = 19)					
		# studies	Sensitivity (95% CI)		Specificity (95% CI)		P value[a]	# studies	Sensitivity (95% CI)		Specificity (95% CI)		P value
Study-level characteristics													
WHO region	Africa	11	90.2%	85.7% to 93.5%	*88.1%*	*82.4% to 92.1%*	0.0006[b]	10	77.3%	65.1% to 86.2%	*95.6%*	*91.2% to 97.8%*	0.013[b]
	Other (mostly Asia and Pacific regions)	12	96.9%	95.2% to 98.0%	*90.0%*	*85.3% to 93.3%*		9	91.3%	83.7% to 95.5%	*95.6%*	*90.7% to 98.0%*	
Tuberculosis prevalence	< 0.5%	12	92.1%	87.8% to 95.0%	*90.7%*	*86.4% to 93.7%*	0.014[b]	8	*81.5%*	*67.3% to 90.5%*	97.5%	95.1% to 98.7%	0.050[c]
	≥ 0.5%	11	96.6%	94.4% to 98.0%	*87.1%*	*81.3% to 91.4%*		11	*86.9%*	*76.5% to 93.1%*	93.0%	87.8% to 96.1%	
Participant-level characteristics													
Proportion smoking	< 10%	6	89.3%	82.0% to 93.8%	*90.0%*	*82.9% to 94.4%*	0.007[b]	*7*	*77.2%*	*59.9% to 88.5%*	*97.6%*	*94.7% to 99.0%*	*0.13[d]*
	≥ 10%	17	96.1%	94.2% to 97.4%	*88.8%*	*84.5% to 92.0%*		*12*	*88.0%*	*79.8% to 93.2%*	*93.8%*	*89.1% to 96.5%*	
Median/mean population age	< 35 years	13	91.1%	87.7% to 93.6%	*89.9%*	*85.4% to 93.1%*	0.0001[b]	*12*	*84.1%*	*73.1% to 91.1%*	*95.3%*	*91.1% to 97.5%*	*0.86[d]*

Table 4. Investigations of heterogeneity: factors modifying sensitivity or specificity or both of CXR index test(s) *(Continued)*

	≥ 35 years	10	97.6%	96.1% to 98.5%	*88.0%*	*82.0% to 92.2%*		*7*	*86.0%*	*71.7% to 93.8%*	*96.0%*	*91.0% to 98.3%*	
HIV prevalence	*Very high > 5%*	*6*	*95.6%*	*92.5% to 97.5%*	*89.9%*	*85.0% to 93.4%*	*0.73[d]*	*5*	*74.2%*	*54.3% to 87.4%*	*95.9%*	*89.3% to 98.5%*	*0.074[b]*
	High > 1%	*11*	*94.6%*	*89.3% to 97.3%*	*89.8%*	*82.7% to 94.2%*		*7*	*85.6%*	*73.8% to 92.6%*	*94.2%*	*87.1% to 97.5%*	
	Low < 1%	*6*	*92.5%*	*85.3% to 96.3%*	*86.6%*	*77.8% to 92.3%*		*7*	*91.3%*	*80.8% to 96.3%*	*96.6%*	*92.0% to 98.6%*	

[a]We examined whether the variable modified sensitivity, or specificity, or both.
P values of the most parsimonious model with the best fit are shown:
[b]Only sensitivity was modified by the variable, but specificity was not. Specificity is therefore shown in italic font.
[c]Only specificity was modified by the variable, but sensitivity was not. Sensitivity is therefore shown in italic font.
[d]Neither sensitivity nor specificity was statistically significantly modified by the variable. Both are shown in italic font.
CI: confidence interval; CXR: chest radiography

Table 5. Sensitivity analyses for CXR index tests

	Any CXR abnormality					CXR abnormalities suggestive of tuberculosis				
	# studies	Sensitivity (95% CI)		Specificity (95% CI)		# studies	Sensitivity (95% CI)		Specificity (95% CI)	
Overall meta-analytic summary	23	94.7%	92.2% to 96.4%	89.1%	85.6% to 91.8%	19	84.8%	76.7% to 90.4%	95.6%	92.6% to 97.4%
Reference Standard domain - low risk of bias only	2	did not converge				3	did not converge			
Reference Standard domain - low applicability concerns only	18	94.4%	91.0% to 96.5%	88.1%	83.5% to 91.5%	16	81.5%	72.9% to 87.9%	95.9%	92.6% to 97.8%
Flow and Timing domain - low risk of bias only	10	95.0%	91.8% to 97.0%	91.7%	87.2% to 94.8%	3	73.5%	42.8% to 91.1%	98.6%	96.8% to 99.4%
Studies with a microbiological reference standard on 20% or more participants	7	95.8%	91.3% to 98.0%	79.2%	74.7% to 83.0%	7	84.2%	67.2% to 93.2%	91.1%	87.9% to 93.5%

CI: confidence interval; CXR: chest radiography

Table 6. Parallel CXR and Symptom Index Tests: description of studies and their populations, index tests definitions, and difference in sensitivity and specificity compared to CXR alone

Study ID	Study population	% with bact ref. test	CXR only			Parallel CXR and symptom question(s)			Difference		Median (range)	
			Classification	Sens	Spec	Symptom	Sens	Spec	Sens	Spec	Sens	Spec
van't Hoog 2012	General population	32%	Any CXR abnormality	94.2%	72.6%	Cough ≥ 2 weeks	96.7%	69.1%	2.6%	-3.5%	+1.6% (-9.4% to +11%)	-3.2% (-0.3% to -6.2%)
MoPH DPRK 2017	General population	8%	Any pulmonary CXR abnormality	99.4%	98.8%	Cough ≥ 2 weeks and/or haemoptysis	100.0%	92.6%	0.6%	-6.2%		
Qadeer 2016	General population	8%	Any pulmonary CXR abnormality	88.7%	92.5%	Cough ≥ 2 weeks	99.7%	90.4%	11.0%	-2.1%		
Ghana NTP 2015	General population	12%	Any pulmonary CXR abnormality	92.1%	91.6%	Cough ≥ 2 weeks	82.7%	89.9%	-9.4%	-1.7%		
NTP Philippines 2018	General population	35%	Any pulmonary CXR abnormality	98.2%	71.4%	Cough ≥ 2 weeks and/or haemoptysis	96.1%	70.8%	-2.0%	-0.7%		
Mongolia MoH 2016	General population	19%	Any pulmonary CXR abnormality	97.1%	84.5%	Cough ≥ 2 weeks	98.8%	81.3%	1.6%	-3.2%		
Federal MoH Sudan 2018	General population	17%	Any pulmonary CXR abnormality	86.3%	85.1%	Cough ≥ 2 weeks	86.5%	84.9%	0.2%	-0.3%		
Republic of Uganda 2018	General population	12%	Any pulmonary CXR abnormality	90.0%	93.4%	Cough ≥ 2 weeks	100.0%	88.2%	10.0%	-5.2%		

Table 6. Parallel CXR and Symptom Index Tests: description of studies and their populations, index tests definitions, and difference in sensitivity and specificity compared to CXR alone *(Continued)*

MoH Indonesia 2015	General population	22%	Any pulmonary CXR abnormality	95.0%	83.1%	Cough ≥ 2 weeks	100.0%	77.8%	5.0%	-5.3%		
Rwanda MoH 2014	General population	11%	Any pulmonary CXR abnormality	90.0%	93.9%	Cough any duration	100.0%	89.1%	10.0%	-4.8%		
MoPH Thailand 2017	General population	9%	Any pulmonary CXR abnormality	95.8%	93.3%	Clinical score of several symptoms[a]	100.0%	90.5%	4.2%	-2.7%	+7.4% (+4.2% to +17%)	-3.7% (-2.7% to -5.0%)
NTP Bangladesh 2017	General population	21%	Any pulmonary CXR abnormality	90.3%	83.6%	Clinical score of several symptoms[a]	100.0%	79.4%	9.7%	-4.2%		
Kapata 2016	General population	13%	Any pulmonary CXR abnormality [b]	82.6%	92.2%	Any one of several symptoms[b]	100.0%	87.2%	17.4%	-5.0%		
Republic of Zimbabwe 2015	General population	17%	Any CXR abnormality	91.1%	89.7%	Any one of several symptoms[c]	96.3%	86.5%	5.2%	-3.2%		
Koesoemadinata 2018	Diabetes patients	20%	CXR abnormality suggestive of tuberculosis	100.0%	88.8%	Cough ≥ 2 weeks	100.0%	81.4%	0.0%	-7.4%	+6.6% (0% to 11%)	-4.8% (-3.7% % to -15.5%)
MoH Myanmar 2012	General population	23%	CXR abnormality suggestive of tuberculosis	91.4%	92.1%	Cough ≥ 3 weeks	96.1%	76.6%	4.7%	-15.5%		
Kenya MoH 2018	General population	14%	CXR abnormality suggestive of tuberculosis	90.3%	90.1%	Cough ≥ 2 weeks	98.7%	85.6%	8.4%	-4.5%		
MoH Cambodia 2005	General population	15%	CXR abnormality suggestive of tuberculosis	97.1%	90.1%	Cough ≥ 3 weeks	100.0%	85.1%	2.9%	-5.0%		

Table 6. Parallel CXR and Symptom Index Tests: description of studies and their populations, index tests definitions, and difference in sensitivity and specificity compared to CXR alone *(Continued)*

Hoa 2012	General population	8%	CXR abnormality suggestive of tuberculosis	85.1%	96.0%	Productive cough ≥ 2 weeks or tuberculosis history	94.1%	92.3%	8.9%	-3.8%		
FRoNigeria 2014	General population	10%	CXR abnormality suggestive of tuberculosis	88.9%	93.4%	Cough ≥ 2 weeks	100.0%	89.7%	11.1%	-3.7%		
Pelissari 2018	Prison inmates	19%	CXR abnormality suggestive of tuberculosis	91.2%	84.5%	Cough any duration	100.0%	67.3%	8.8%	-17.2%		
Malawi MoH 2016	General population	11%	CXR abnormality suggestive of tuberculosis	49.2%	97.0%	Any one of several symptoms[d]	100.0%	89.5%	50.8%	-7.5%	+10.7% (+5.3% to +51%)	-15.5% (-7.5% to -29%)
Telisinghe 2014	Prison inmates	100%	CXR abnormality suggestive of tuberculosis	78.9%	92.8%	Any one of several symptoms[e]	84.2%	64.3%	5.3%	-28.5%		
Morasert 2018	Prison inmates	17%	CXR abnormality suggestive of tuberculosis	88.1%	94.7%	Clinical score of several symptoms[f]	98.8%	79.2%	10.7%	-15.5%		
den Boon 2006	General population	45%	CXR abnormality suggestive of tuberculosis	89.7%	87.9%	Any one of several symptoms[g]	89.7%		0.0%			

Abbreviations: bact ref. test: microbacteriological reference test; CI: confidence interval; CXR: chest radiography; sens: sensitivity; spec: specificity.

[a]Clinical score ≥ 3 out of cough ≥ 2 weeks (3 points), haemoptysis over the past month (3 points), cough < 2 weeks (2 points), weight loss in the past month (1 point), fever ≥ 1 week in the past 2 weeks (1 point), night sweats in the past month (1 point).

[b]CXR any pulmonary abnormality and heart abnormalities. Symptom combination comprises any one of the following 3 symptoms for at least 2 weeks: cough, fever, and/or chest pain.

[c]Any one of the following 3 symptoms: cough of any duration, haemoptysis in the past 12 months, drenching night sweats.

[d]Any of the following symptoms for at least 1 week: cough, sputum production, haemoptysis, chest pain, weight loss, night sweats, fatigue, fever, or shortness of breath.

[e]WHO 4 symptom tool for People Living with HIV.

[f]Cough, haemoptysis, weight loss, night sweats, and fever.

gScore ≥ 3 points out of the prison screening questionnaire composed of 8 questions: 3 points for any history of previous anti-tuberculosis treatment, cough ≥ 2 weeks, haemoptysis in the past month or cervical lymphadenopathy ≥ 2 cm; 2 points for cough < 2 weeks; 1 point for intermittent or persistent fever in the past month, weight loss 5% of body weight in the past month or night sweats in the past month.

Table 7. Subgroup and sensitivity analyses of parallel CXR and symptom index tests

	# studies	Sensitivity (95% CI)		Specificity (95% CI)	
Overall meta-analytic summary of parallel combinations of CXR and symptoms[a]	24	99.6%	98.3% to 99.9%	84.2%	81.1% to 87.0%
CXR any (pulmonary) abnormality and cough ≥ 2 weeks	9	99.2%	94.8% to 99.9%	84.3%	78.6% to 88.7%
CXR abnormalities suggestive of tuberculosis and cough ≥ 2 weeks	6	99.1%	94.3% to 99.9%	86.0%	81.3% to 89.6%
CXR (either of the above definitions) and cough ≥ 2 weeks	15	99.2%	96.8% to 99.8%	84.9%	81.2% to 88.1%
CXR any (pulmonary) abnormality and several symptoms[b]	4	100%	58.2% to 100%	86.4%	82.1% to 89.7%
CXR (either of the above definitions) and several symptoms[a]	7	99.8%	95.4% to 100%	83.7%	77.6% to 88.4%
Reference Standard domain - low risk of bias only	1	too few studies			
Reference Standard domain - low applicability concerns only	20	99.6%	98.0% to 99.9%	83.7%	80.0% to 86.9%
Flow and Timing domain - low risk of bias only	9	100%	96.6% to 100%	86.8%	83.4% to 89.6%
Studies with a microbiological reference standard on at least 20% of participants	7	99.0%	93.5% to 99.8%	74.5%	70.3% to 78.3%

[a]One study lacked data to calculate specificity and was excluded.
[b]See Table 6 for study definitions of 'several symptoms'.

Table 8. Comparison with earlier version of review

	This report									2012 report (van't Hoog 2013a)							
	n	Sensitivity	95% CI		CiE	Specificity	95% CI		CiE	n	Sensitivity	95% CI		CiE	Specificity	95% CI		CiE
Any tuberculosis symptom	29	70.6%	61.7%	78.2%	very low	65.1%	53.3%	75.4%	low	8	77.0%	68.0%	86.0%	very low	67.7%	50.2%	85.1%	very low
Cough any duration	21	51.3%	42.8%	59.7%	very low	87.6%	81.6%	91.8%	low	6	62.7%	49.3%	76.1%	low	77.5%	65.5%	89.5%	low
Cough 2 or more weeks[a]	41	42.1%	36.5%	47.7%	very low	94.4%	92.9%	95.8%	high	8*	35.1%	24.4%	45.7%	very low	94.7%	92.5%	96.8%	low
CXR any (pulmonary) abnormality	23	94.7%	92.2%	96.4%	very low	89.1%	85.6%	91.8%	low	3	97.8%	95.1%	100.0%	moderate	75.4%	72.0%	78.8%	moderate
CXR abnormalities suggestive of tuberculosis	19	84.8%	76.7%	90.4%	low	95.6%	92.6%	97.4%	high	4	86.8%	79.2%	94.5%	low	89.4%	86.7%	92.0%	low
Parallel cough 2 or more weeks and CXR	15	99.2%	96.8%	99.8%	—	84.9%	81.2%	88.1%	—	—	—	—	—	—	—	—	—	—

Abbreviations: CI: confidence interval; CiE: certainty of evidence; CXR: chest radiography.
[a]The 2012 report includes 3 of 8 studies reporting on cough for 3 or more weeks, which has lower sensitivity (Figure 20), and was excluded in the current report.

Table 9. Seven studies probably eligible for inclusion

Study ID	Study objective	Index test(s)	# tuberculosis cases	# concerns	Concerns and comments
Assefa 2019	To assess if there are missed pulmonary tuberculosis (PTB) cases among medical inpatients in a referral hospital in Ethiopia.	Cough ≥ 2 weeks, any tuberculosis symptom	10	≥ 2	Concerns: The population is not representative for the study question: medical inpatients, high HIV prevalence, the data are reported aggregately for HIV+ and HIV- patients. The symptom screening was not systematic but extracted from retrospective chart review. Consensus discussion with several authors would be required to determine eligibility.
Bekken 2020	To evaluate a Tuberculosis Contact Score (TCS) and Infectivity Score in identifying subclinical tuberculosis, in household contacts in India.	Any CXR abnormality, CXR abnormalities suggestive of tuberculosis	29	≥ 2	Concerns: the population are household contacts and is not fully representative as 40% is below 15 years of age. Of 29 cases 90% has subclinical tuberculosis, meaning no symptoms and no CXR abnormalities. The study provides data on 29 cases for sensitivity and n = 496 for specificity.
Hamda 2020	To determine 1) the prevalence of tuberculosis (TB) and HIV-TB co-infection in pregnancy, and 2) the sensitivities of symptomatic TB screening and Xpert MTB/RIF testing against gold standard culture, in Botswana.	Cough ≥ 2 weeks, other symptom definition	2	1	Concerns. The definition of any tuberculosis symptoms is different from how it is used in the review, which is usually of any one symptom of any duration. The risk of bias in domain 1 is unclear. # of tuberculosis cases is very low.
Li 2019	To determine the prevalence of pulmonary tuberculosis in Tibet.	Cough ≥ 2 weeks, CXR abnormalities suggestive of tuberculosis, CXR + symptoms in parallel	63	≥ 2	Standard prevalence survey design. Substantial incorporation bias in domain 3; approximately 10% of > 30,000 study population had a microbacteriological test as the reference standard. Of 215 active tuberculosis cases, less than half would count as a case for the systematic review. To determine if this can be disaggregated needs more work. Risk of bias in domain 4 unclear.
Nguyen TBP 2020	To compare (sensitivity/yield and feasibility/participation) of two community screening tests for tuberculosis: sputum examination using Xpert MTB/RIF and chest radiography, in Vietnam.	CXR abnormalities suggestive of tuberculosis	51	≥ 2	Concerns: High risk of bias in domain 1 (participant selection) and domain 4 (flow and timing); representativeness of population excluding women < 45 years; index test is defined based on 2 readers - different from the studies included in the review. N = 51 cases of tuberculosis (for sensitivity). The publication reports sensitivity based on 59 estimated by a capture-recapture model.

Table 9. Seven studies probably eligible for inclusion *(Continued)*

Pasipamire 2020	To determine tuberculosis prevalence stratified by HIV status and identify screening algorithms that maximise detection of active tuberculosis among pregnant and postpartum women in Eswatini.	Any cough, any tuberculosis symptom	15	≥ 2	Concerns: Domain1 - representativeness of the population Of 990 participants half are HIV+ (purposely selected for stratification purposes). Domain 4: 704 (approximately 70%) had culture results available. The remaining did not give sputum. Domain 3: % cases without signs of disease or 2nd confirmed bacteriology unclear. It is possible to calculate accuracy for n = 990 and n = 704, but not for the HIV-negatives only.
Reichler 2020	To examined cytokine immune response profiles among contacts to tuberculosis patients to identify immunologic and epidemiologic correlates of tuberculosis, in the USA and Canada.	Any cough	41	≥ 2	Concerns: High risk of bias in 3 domains. Domain1: of 3221 eligible contacts 61% did not enrol. D3: tuberculosis cases were identified from tuberculosis registry matches, implying that the case definition is unclear - essentially clinician defined - bacteriology results are unknown. D4: Contacts with tuberculosis diagnosed before or < 30 days after blood draw were considered co-prevalent cases. The date of tuberculosis diagnosis was defined as the start date for tuberculosis treatment. Consensus discussion with several authors would be required to determine eligibility.

APPENDICES

Appendix 1. QUADAS-2 tool

Key questions	Signalling questions
Domain 1: Patient selection	
Risk of bias: Could the selection of patients have introduced bias?	1. Did the study enrol a consecutive or random sample of patients? • Yes: if all eligible patients were enrolled; or if the authors reported that the patients were either a consecutive series or randomly selected; • No: if the authors report that the selection was based on clinical judgement of health workers, or participation of randomly selected people in the study was low; • Unclear: if there is discrepancy between the numbers of eligible people and the number of included people, but no reasons given for that, or the selection procedure is not clearly described. 2. Was a case–control design avoided? • Yes: if a case–control design was avoided; • No: if a case–control design was not avoided; • Unclear: if not reported or insufficient information is provided to decide.

(Continued)

	3. Did the study avoid inappropriate exclusions? • Yes: if no study participants were excluded after inclusion; • No: if study participants were excluded (for example, participants with mild or severe symptoms or signs); • Unclear: if insufficient information is provided to decide.
Applicability: Are there concerns that the included patients and setting do not match the review question?	• High concern: if the study population does not resemble a population that would be considered for a tuberculosis screening program in practice; • Low concern: if the study population does resemble a population that would be considered for a tuberculosis screening program in practice; • Unclear: if not reported or insufficient information is provided to decide.
Domain 2: Index test	
Risk of bias: Could the conduct or interpretation of the index test have introduced bias?	1. Were the index test results interpreted without knowledge of the results of the reference standard? • Yes: if the screening test was performed without knowing whether the person had infectious tuberculosis. • No: if symptom questions were asked after the results of the reference test were known, or the CXR was interpreted with knowledge of the results of the reference test. • Unclear: if insufficient information is provided to decide. For example, if it was unclear whether the CXR reader was blinded to the results of the reference test. 2. If a threshold was used, was it pre-specified? • This question was not applicable for our review question.
Applicability: Are there concerns that the index test, its conduct or its interpretation differ from the review question?	• High concern: if the symptom questions or CXR classification were intended as a diagnostic rather than a screening tool; or if part of the population was screened with MMR; • Low concern: if the symptom questions or CXR assessment were done with the intention to screen; • Unclear: if insufficient information is provided to decide.
Domain 3: Reference standard	
Risk of bias: Could the reference standard, its conduct or its interpretation have introduced bias?	1. Is the reference standard likely to correctly classify the target condition? • Yes: if the reference standard was an author-defined combination of mycobacterial culture (on solid or liquid medium) and possibly sputum smear microscopy, or Xpert MTB/RIF, other NAAT, or both, and cases defined by sputum microscopy only are limited to a small proportion (≦ 10%) in whom culture was contaminated or negative or missing but smears were positive; • No: if the reference standard was not an author-defined combination of mycobacterial culture (on solid or liquid medium) and possibly sputum smear microscopy, or Xpert MTB/RIF, other NAAT, or both. This includes studies where sputum smear microscopy was the only reference test; • Unclear: if insufficient information is provided to decide. 2. Were the reference standard results interpreted without knowledge of the results of the index test? • Yes: if the screening test results were not known to the people interpreting the reference standard results; • No: if the screening test results were known to the people interpreting the reference standard results;

(Continued)

	• Unclear: if insufficient information is provided to decide.
Applicability: Are there concerns that the target condition as defined by the reference standard does not match the question?	• High concern: if there was a high probability that a considerable proportion of the tuberculosis cases identified in the study did not have bacteriologically confirmed tuberculosis or did not have active tuberculosis; • Low concern: (i) if the tuberculosis cases in the study have tuberculosis symptoms or CXR abnormalities in addition to a positive culture, or positive smear microscopy, or both; or (ii) if they have at least two different samples positive on culture, or on smear microscopy, or both. • Moderate concern: Because we perceive a large contrast between “low” and “high” we added a category “moderate” for the applicability sections. We applied the "moderate" category if the tuberculosis cases in the study could include people with one positive sputum culture or Xpert MTB/RIF, NAAT or smear only, without the presence or symptoms or CXR abnormalities; • Unclear: if insufficient information is provided to decide.
Domain 4: Flow and timing	
Risk of bias: Could the patient flow have introduced bias?	1. Was there an appropriate interval between the index test and reference standard? • Yes: if the screening test and reference standard were applied (or samples taken) at the same time or within 1 week; • No: if the time between the screening test and reference standard (sample collection) was more than 1 week; • Unclear: if insufficient information is provided to decide. 2. Did all patients receive the same reference standard? • Yes: if all participants were evaluated with the reference standard, and if all or a large majority of participants were evaluated with the same test(s); • No: if not all participants were evaluated with the reference standard, or participants received different tests (for example, some smear only, some culture, or different numbers of samples were submitted for testing); • Unclear: if insufficient information is provided to decide. 3. Were all patients included in the analysis? • Yes: if all participants were included; • No: if participants who participated were excluded. For instance because they did not provide sputum for a reference test; • Unclear: if insufficient information is provided to decide.

Appendix 2. Search strategy

A. MEDLINE search strategy

Platform: OvidSP

Database: MEDLINE(R) In-Process & Other Non-Indexed Citations

Limits: no limits were used

Methodological filters: none

1	exp Mycobacterium/
2	mycobacterium.ti,ab.
3	tuberculosis/
4	peritonitis, tuberculous/
5	exp tuberculoma/
6	tuberculosis, bovine/
7	exp tuberculosis, cardiovascular/
8	exp tuberculosis, central nervous system/
9	tuberculosis, cutaneous/
10	erythema induratum/
11	tuberculosis, endocrine/
12	tuberculosis, gastrointestinal/
13	tuberculosis, hepatic/
14	exp tuberculosis, lymph node/
15	tuberculosis, miliary/
16	tuberculosis, multidrug-resistant/
17	tuberculosis, ocular/
18	tuberculosis, oral/
19	tuberculosis, osteoarticular/
20	tuberculosis, pleural/
21	tuberculosis, pulmonary/
22	tuberculosis, splenic/
23	tuberculosis, urogenital/
24	(tuberculo* or TB or scrofuloderma).ti,ab.
25	or/1-24
26	(case adj finding).ti,ab.
27	screen*.ti,ab.
28	Mass Screening/ or Mass Chest X-ray/

(Continued)

29	exp Population Surveillance/
30	(disease adj3 surveillance).ti,ab.
31	(case adj detection).ti,ab.
32	Contact Tracing/
33	(contact adj tracing).ti,ab.
34	exp Health Surveys/
35	survey.ti,ab.
36	exp "Sensitivity and Specificity"/
37	(false adj negative).ti,ab.
38	odds.mp.
39	((ROC or HSROC or SROC) adj2 (curve* or analys?s or plot*1)).ti,ab.
40	(predictive adj3 value).ti,ab.
41	specificit*.ti,ab.
42	accuracy.ti,ab.
43	or/36-42
44	prevalence.mp. or Prevalence/
45	Cross-Sectional Studies/ or cross sectional.mp.
46	44 or 45
47	34 or 35 or 46
48	(mycobacteri$ adj2 culture).ti,ab.
49	(microscopy adj2 (sputum smear or ZN or Ziehl-neelsen or FM or fluorescence)).ti,ab.
50	lowenstein-jensen.ti,ab.
51	(LJ adj2 medium).ti,ab.
52	"mycobacteria growth incubator tube".ti,ab.
53	mgit.ti,ab.
54	Xpert.ti,ab.
55	(auramine adj2 staining).ti,ab.
56	((culture or smear) adj positiv*).ti,ab.

(Continued)

57	or/48-56
58	25 and 47 and 57
59	or/26-35
60	25 and 59 and 43
61	25 and 59 and 57
62	58 or 60 or 61
63	Limit 62 to ed=19920101-20130801

B. Embase search strategy

#	Searches
1	exp Mycobacterium/ or exp tuberculosis/ or central nervous system tuberculosis/ or congenital tuberculosis/ or drug resistant tuberculosis/ or extrapulmonary tuberculosis/ or intestine tuberculosis/ or kidney tuberculosis/ or laryngeal tuberculosis/ or latent tuberculosis/ or lung tuberculosis/ or miliary tuberculosis/ or ocular tuberculosis/ or postprimary tuberculosis/ or primary tuberculosis/ or skin tuberculosis/ or tuberculoma/ or (tuberculo* or TB or scrofuloderma).ti,ab. or mycobacterium.ti,ab.
2	case finding/
3	(case adj finding).ti,ab.
4	screen*.ti,ab.
5	exp screening/
6	(disease adj3 surveillance).ti,ab.
7	exp disease surveillance/
8	case-finding.ti,ab.
9	(case adj detection).tw.
10	(contact adj tracing).ti,ab.
11	exp health survey/ or screening/ or mass screening/ or screening test/ or mass radiography/ or thorax radiography/
12	survey.ti,ab.

(Continued)

13	or/2-12
14	"sensitivity and specificity"/
15	(false adj negative).ti,ab.
16	odds.mp.
17	(predictive adj3 value).ti,ab.
18	((ROC or HSROC or SROC) adj2 (curve* or analys?s or plot*1)).ti,ab.
19	specificit*.ti,ab.
20	accuracy.tw.
21	or/14-20
22	exp prevalence/
23	prevalence.ti,ab.
24	cross sectional.ti,ab.
25	cross-sectional study/
26	or/22-25
27	11 or 12 or 26
28	(mycobacteri$ adj2 culture).ti,ab.
29	(microscopy adj2 (sputum smear or ZN or Ziehl-neelsen or FM or fluorescence)).ti,ab.
30	lowenstein-jensen.ti,ab.
31	(LJ adj2 medium).ti,ab.
32	"mycobacteria growth incubator tube".ti,ab.
33	mgit.ti,ab.
34	geneXpert.ti,ab.
35	(auramine adj2 staining).ti,ab.
36	((culture or smear) adj positiv*).ti,ab.
37	or/28-36
38	1 and 27 and 37
39	1 and 13 and 21
40	1 and 13 and 37

(Continued)

41	38 or 39 or 40

C. LILACS search strategy

((tw:(tuberculo*)) AND (tw:(prevalence OR cross-sectional OR survey)) AND (tw:(culture OR "sputum smear" OR ziehl-neelsen OR flurescence OR lowenstein-jensen OR "incubator tube" mgit OR genexpert OR smear)) AND (instance:"regional")) OR ((tw:(tuberculo*)) AND (tw:(screen* OR limpaduras OR cribado OR "case finding" OR "detección de casos" OR "population surveillance" OR "disease surveillance" OR "contact tracing" OR survey)) AND (tw:(culture OR "sputum smear" OR ziehl-neelsen OR flurescence OR lowenstein-jensen OR "incubator tube" mgit OR genexpert OR smear))) OR ((tw:(tuberculo*)) AND (tw:(screen* OR limpaduras OR cribado OR "case finding" OR "detección de casos" OR "population surveillance" OR "disease surveillance" OR "contact tracing" OR survey)) AND (tw:(specificit* OR accuracy OR predict* OR odds OR roc))) AND (db:("LILACS")) AND (instance:"regional")

D. CRD HTA database search strategy

tubercul* in all fields

Limit: HTA (published)

Appendix 3. Certainty of the evidence and search date

Exploration to assess if updating the search increases the certainty of evidence

After completion of the review based on the search of 10 December 2018, we explored if adding an updated search would potentially reach a higher certainty of evidence and therefore more firm conclusions. We conducted an updated search in MEDLINE and Embase on 2 July 2021 and then restricted the search to records that included tuberculosis terms, DTA terms, and terms for our reference standard in the title or/and abstract (TiAb). This set would most likely include studies of high quality. We assessed the restricted set using the review's eligibility criteria and of the studies that were most likely eligible for inclusion we assessed risks of bias and applicability concerns in the QUADAS-2 domains.

The updated search would add approximately 3000 records of which 262 remained after restricting to inclusion in the TiAbs of tuberculosis terms, DTA terms, and terms for our reference standard. Assessment of these 262 by two review authors independently resulted in 40 articles for full-text selection of which 12 studies fulfilled our inclusion criteria. These are listed as Studies awaiting classification in the review and the section on Characteristics of studies awaiting classification provides information on each of these studies. Five are national tuberculosis prevalence studies of which four (Bonsu 2020; Lansang 2021; Migambi 2020; Nalunjogi 2021) are already included in the review based on grey literature reports (Ghana NTP 2015; NTP Philippines 2018; Rwanda MoH 2014; Republic of Uganda 2018). The Vietnam survey (Nguyen HV 2020) would replace the data from the previous version of the Vietnam prevalence survey which is already included in the review (Hoa 2012). These five studies would thus not add data to the review.

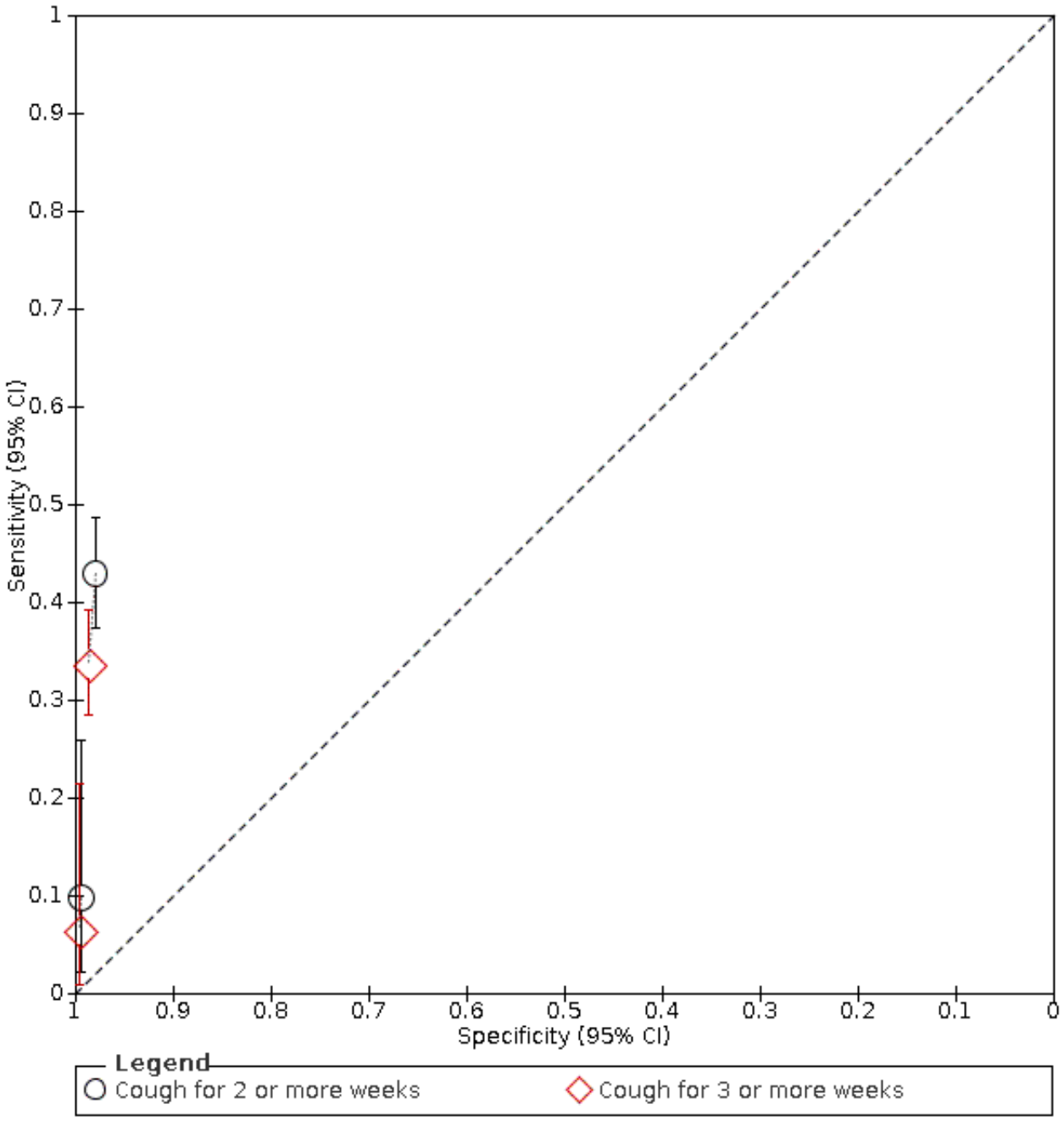
Of the remaining seven studies (see Table 9), one is a subnational tuberculosis prevalence surveys from Tibet/China (Li 2019) which follows the WHO-recommended study design for national tuberculosis prevalence surveys and thus shares the concerns for the review that arise from this design: high risk of bias in the QUADAS-2 reference standard domain due to incorporation bias. The remaining six studies have different study objectives, designs, and populations, and five have concerns about high risk of bias or/and applicability concerns in at least two QUADAS-2 domains (Assefa 2019; Bekken 2020; Nguyen TBP 2020; Pasipamire 2020; Reichler 2020), and one study in one domain (Hamda 2020). Two of these six studies have a very low (10 or fewer) number of tuberculosis cases (Assefa 2019; Hamda 2020). For at least two studies consensus discussion with several authors would be required to determine final eligibility (Assefa 2019; Reichler 2020).

From these data we conclude that an update would not increase the certainty of evidence. We expect the same level of heterogeneity and risk of bias. The large number of prevalence surveys in the current review already provide us with a considerable amount of data to estimate specificity, and more data will not further increase the certainty of evidence. The uncertainty in the estimates for sensitivity will not be solved, as the data from prevalence surveys suffer from incorporation bias (as we show in the Results section of the review), while the additional studies with other designs have other biases or concerns or are small in sample size and number of tuberculosis cases to contribute to sensitivity calculations. Incorporating the update in the review would take a considerable amount of time, with no gain in information value.

Appendix 4. Cough for ≥ 3 versus ≥ 2 weeks; within-study comparison

Figure 20

Figure 20. Summary ROC plot: 'Cough for 2 or more weeks' and 'Cough for 3 or more weeks' in studies reporting on both tests.



WHAT'S NEW

Date	Event	Description
9 May 2022	Amended	Amendment published to resolve an issue with figures not displaying correctly.

HISTORY

Protocol first published: Issue 1, 2014
Review first published: Issue 3, 2022

CONTRIBUTIONS OF AUTHORS

Anja van't Hoog and Miranda Langendam wrote the protocol. Mariska Leeflang provided methodological advice.

Anja van't Hoog, Kerri Viney, Olivia Biermann, Bada Yang, and Miranda Langendam did study selection, data extraction, and QUADAS-2 assessments.

Anja van't Hoog and Mariska Leeflang did the statistical analyses.

Anja van't Hoog wrote the first manuscript draft. All authors provided input on the manuscript and approved the final version.

DECLARATIONS OF INTEREST

Anja van't Hoog: an earlier version of this work was funded by the WHO/TB-CARE 1, through a contract with AvH's employer at the time, Amsterdam Institute for Global Health and Development. This work is included in the current version of the review. At the WHO's request, AvH presented the results to the guideline development group for the updated TB screening guidelines. She received financial compensation for the time spent on the presentations, but not for the actual work on the Cochrane Review itself.

Kerri Viney works for the Global TB Programme, WHO and this review contributed to WHO guidelines. KV started work on the review prior to employment with WHO.

Olivia Biermann has no known conflicts of interest.

Bada Yang has no known conflicts of interest.

Mariska Leeflang has no known conflicts of interest.

Miranda Langendam has no known conflicts of interest.

SOURCES OF SUPPORT

Internal sources

- Liverpool School of Tropical Medicine, UK

External sources

- Foreign, Commonwealth and Development Office (FCDO), UK

  Project number 300342-104
- USAID - TB Care, USA

  The earlier version of this review was funded by TB-CARE, through the WHO. For the update/expansion from 2013 onwards, the review authors did not receive dedicated funding.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We added to the participants section that "Studies evaluating CXR screening in a pre-selected symptomatic population were outside the scope of this review". This is a modification of the protocol in the sense that we mentioned in the initial protocol that we may include such pre-screened populations (van't Hoog 2014b). This would, however, be a very different population, resembling a presumptive tuberculosis (passive case detection) population. The latter is an exclusion criterion. For consistency we restricted the index tests in this review to populations targeted for screening who are not pre-screened with other tests.

The sentence: "In our earlier report we included a study on smear-positive cases only, but excluded such studies from this report as NAATs are increasingly recommended in stead of sputum smear microscopy alone, and the number of such studies was small" is a new addition. These studies are not included in the current results as WHO no longer recommends sputum smear microscopy as the only diagnostic test.

INDEX TERMS

Medical Subject Headings (MeSH)

Cough; *HIV Infections [complications]; Mass Screening; Radiography; Sensitivity and Specificity; *Tuberculosis, Pulmonary [diagnostic imaging] [epidemiology]

MeSH check words

Adult; Humans