# Screening tests for active pulmonary tuberculosis in children

Vonasek, Bryan; Ness, Tara; Takwoingi, Yemisi; Kay, Alexander W; van Wyk, Susanna S; Ouellette, Lara; Marais, Ben J; Steingart, Karen R; Mandalakas, Anna M

*DOI:*
10.1002/14651858.CD013693.pub2



*Document Version*
Publisher's PDF, also known as Version of record

*Citation for published version (Harvard):*
Vonasek, B, Ness, T, Takwoingi, Y, Kay, AW, van Wyk, SS, Ouellette, L, Marais, BJ, Steingart, KR & Mandalakas, AM 2021, 'Screening tests for active pulmonary tuberculosis in children', *The Cochrane Library*, vol. 2021, no. 6, pp. CD013693. https://doi.org/10.1002/14651858.CD013693.pub2

Cochrane Library

**Cochrane** Database of Systematic Reviews

# Screening tests for active pulmonary tuberculosis in children (Review)

Vonasek B, Ness T, Takwoingi Y, Kay AW, van Wyk SS, Ouellette L, Marais BJ, Steingart KR, Mandalakas AM

Vonasek B, Ness T, Takwoingi Y, Kay AW, van Wyk SS, Ouellette L, Marais BJ, Steingart KR, Mandalakas AM.
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.

**www.cochranelibrary.com**




WILEY

# TABLE OF CONTENTS

[Diagnostic Test Accuracy Review]

# Screening tests for active pulmonary tuberculosis in children

Bryan Vonasek[1], Tara Ness[1], Yemisi Takwoingi[2], Alexander W Kay[1], Susanna S van Wyk[3], Lara Ouellette[4], Ben J Marais[5], Karen R Steingart[6], Anna M Mandalakas[1]

[1]The Global Tuberculosis Program, Texas Children's Hospital, Section of Global and Immigrant Health, Department of Pediatrics, Baylor College of Medicine, Houston, Texas, USA. [2]Test Evaluation Research Group, Institute of Applied Health Research, University of Birmingham, Birmingham, UK. [3]Centre for Evidence-based Health Care, Epidemiology and Biostatistics, Department of Global Health, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa. [4]Texas Medical Center Library, Houston, Texas, USA. [5]Marie Bashir Institute for Infectious Diseases and Biosecurity, University of Sydney, Sydney, Australia. [6]Cochrane Infectious Diseases Group, Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, UK

**Contact address:** Anna M Mandalakas, anna.mandalakas@bcm.edu.

**Editorial group:** Cochrane Infectious Diseases Group.
**Publication status and date:** Edited (no change to conclusions), published in Issue 10, 2021.

**Citation:** Vonasek B, Ness T, Takwoingi Y, Kay AW, van Wyk SS, Ouellette L, Marais BJ, Steingart KR, Mandalakas AM. 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.



## ABSTRACT

### Background

Globally, children under 15 years represent approximately 12% of new tuberculosis cases, but 16% of the estimated 1.4 million deaths. This higher share of mortality highlights the urgent need to develop strategies to improve case detection in this age group and identify children without tuberculosis disease who should be considered for tuberculosis preventive treatment. One such strategy is systematic screening for tuberculosis in high-risk groups.

### Objectives

To estimate the sensitivity and specificity of the presence of one or more tuberculosis symptoms, or symptom combinations; chest radiography (CXR); Xpert MTB/RIF; Xpert Ultra; and combinations of these as screening tests for detecting active pulmonary childhood tuberculosis in the following groups.

– Tuberculosis contacts, including household contacts, school contacts, and other close contacts of a person with infectious tuberculosis.

– Children living with HIV.

– Children with pneumonia.

– Other risk groups (e.g. children with a history of previous tuberculosis, malnourished children).

– Children in the general population in high tuberculosis burden settings.

### Search methods

We searched six databases, including the Cochrane Central Register of Controlled Trials, MEDLINE, and Embase, on 14 February 2020 without language restrictions and contacted researchers in the field.

**Selection criteria**

Cross-sectional and cohort studies where at least 75% of children were aged under 15 years. Studies were eligible if conducted for screening rather than diagnosing tuberculosis. Reference standards were microbiological (MRS) and composite reference standard (CRS), which may incorporate symptoms and CXR.

**Data collection and analysis**

Two review authors independently extracted data and assessed study quality using QUADAS-2. We consolidated symptom screens across included studies into groups that used similar combinations of symptoms as follows: one or more of cough, fever, or poor weight gain and one or more of cough, fever, or decreased playfulness. For combination of symptoms, a positive screen was the presence of one or more than one symptom.

We used a bivariate model to estimate pooled sensitivity and specificity with 95% confidence intervals (CIs) and performed analyses separately by reference standard. We assessed certainty of evidence using GRADE.

**Main results**

Nineteen studies assessed the following screens: one symptom (15 studies, 10,097 participants); combinations of symptoms (12 studies, 29,889 participants); CXR (10 studies, 7146 participants); and Xpert MTB/RIF (2 studies, 787 participants). Several studies assessed more than one screening test. No studies assessed Xpert Ultra. For 16 studies (84%), risk of bias for the reference standard domain was unclear owing to concern about incorporation bias. Across other quality domains, risk of bias was generally low.

**Symptom screen (verified by CRS)**

*One or more of cough, fever, or poor weight gain in tuberculosis contacts* (4 studies, tuberculosis prevalence 2% to 13%): pooled sensitivity was 89% (95% CI 52% to 98%; 113 participants; low-certainty evidence) and pooled specificity was 69% (95% CI 51% to 83%; 2582 participants; low-certainty evidence). Of 1000 children where 50 have pulmonary tuberculosis, 339 would be screen-positive, of whom 294 (87%) would not have pulmonary tuberculosis (false positives); 661 would be screen-negative, of whom five (1%) would have pulmonary tuberculosis (false negatives).

*One or more of cough, fever, or decreased playfulness in children aged under five years, inpatient or outpatient* (3 studies, tuberculosis prevalence 3% to 13%): sensitivity ranged from 64% to 76% (106 participants; moderate-certainty evidence) and specificity from 37% to 77% (2339 participants; low-certainty evidence). Of 1000 children where 50 have pulmonary tuberculosis, 251 to 636 would be screen-positive, of whom 219 to 598 (87% to 94%) would not have pulmonary tuberculosis; 364 to 749 would be screen-negative, of whom 12 to 18 (2% to 3%) would have pulmonary tuberculosis.

*One or more of cough, fever, poor weight gain, or tuberculosis close contact (World Health Organization four-symptom screen) in children living with HIV, outpatient* (2 studies, tuberculosis prevalence 3% and 8%): pooled sensitivity was 61% (95% CI 58% to 64%; 1219 screens; moderate-certainty evidence) and pooled specificity was 94% (95% CI 86% to 98%; 201,916 screens; low-certainty evidence). Of 1000 symptom screens where 50 of the screens are on children with pulmonary tuberculosis, 88 would be screen-positive, of which 57 (65%) would be on children who do not have pulmonary tuberculosis; 912 would be screen-negative, of which 19 (2%) would be on children who have pulmonary tuberculosis.

**CXR (verified by CRS)**

*CXR with any abnormality in tuberculosis contacts* (8 studies, tuberculosis prevalence 2% to 25%): pooled sensitivity was 87% (95% CI 75% to 93%; 232 participants; low-certainty evidence) and pooled specificity was 99% (95% CI 68% to 100%; 3281 participants; low-certainty evidence). Of 1000 children, where 50 have pulmonary tuberculosis, 63 would be screen-positive, of whom 19 (30%) would not have pulmonary tuberculosis; 937 would be screen-negative, of whom 6 (1%) would have pulmonary tuberculosis.

**Xpert MTB/RIF (verified by MRS)**

*Xpert MTB/RIF, inpatient or outpatient* (2 studies, tuberculosis prevalence 1% and 4%): sensitivity was 43% and 100% (16 participants; very low-certainty evidence) and specificity was 99% and 100% (771 participants; moderate-certainty evidence). Of 1000 children, where 50 have pulmonary tuberculosis, 31 to 69 would be Xpert MTB/RIF-positive, of whom 9 to 19 (28% to 29%) would not have pulmonary tuberculosis; 931 to 969 would be Xpert MTB/RIF-negative, of whom 0 to 28 (0% to 3%) would have tuberculosis.

Studies often assessed more symptoms than those included in the index test and symptom definitions varied. These differences complicated data aggregation and may have influenced accuracy estimates. Both symptoms and CXR formed part of the CRS (incorporation bias), which may have led to overestimation of sensitivity and specificity.

**Authors' conclusions**

We found that in children who are tuberculosis contacts or living with HIV, screening tests using symptoms or CXR may be useful, but our review is limited by design issues with the index test and incorporation bias in the reference standard.

For Xpert MTB/RIF, we found insufficient evidence regarding screening accuracy.

Prospective evaluations of screening tests for tuberculosis in children will help clarify their use. In the meantime, screening strategies need to be pragmatic to address the persistent gaps in prevention and case detection that exist in resource-limited settings.

## PLAIN LANGUAGE SUMMARY

### Screening tests for active pulmonary tuberculosis in children

#### Why is improving screening for pulmonary tuberculosis in children important?

Tuberculosis is one of the leading causes of death worldwide. Most children who die from tuberculosis are never diagnosed or treated. Screening may be useful to identify children with possible tuberculosis and refer them for further testing. As well, screening could be used to identify children without tuberculosis, who should be considered for preventive treatment. A false-positive result means that children may undergo unnecessary testing and treatment and may not receive preventive treatment promptly. A false-negative result means that children have tuberculosis, but may miss further testing to confirm the diagnosis.

#### What is the aim of this review?

To determine the accuracy of screening tests for active pulmonary tuberculosis in children in high-risk groups, such as children with HIV and close contacts of people with tuberculosis.

#### What was studied in this review?

Screening tests were: one tuberculosis symptom; one or more of a combination of tuberculosis symptoms; the World Health Organization (WHO) four-symptom screen (one or more of cough, fever, poor weight gain, or tuberculosis contact) in children with HIV, recommended at each healthcare visit; chest radiography (CXR); and Xpert MTB/RIF.

#### What are the main results in this review?

Nineteen studies assessed the following screening tests: one symptom (15 studies, 10,097 participants); more than one symptom (12 studies, 29,889 participants); CXR (10 studies, 7146 participants); and Xpert MTB/RIF (two studies, 787 participants).

#### Symptom screening

For every 1000 children screened, if 50 had tuberculosis according to the reference standard:

*One or more of cough, fever, or poor weight gain in tuberculosis contacts (composite reference standard (CRS) (4 studies)*

– 339 would screen positive, of whom 294 (87%) would not have tuberculosis (false positive).

– 661 would screen negative, of whom 5 (1%) would have tuberculosis (false negative).

*One or more of cough, fever, or decreased playfulness in children under five, inpatient or outpatient (CRS) (3 studies)*

– 251 to 636 would screen positive, of whom 219 to 598 (87% to 94%) would not have tuberculosis (false positive).

– 364 to 749 would screen negative, of whom 12 to 18 (2% to 3%) would have tuberculosis (false negative).

*One or more of cough, fever, poor weight gain, or tuberculosis close contact (WHO four-symptom screen) in children with HIV, outpatient (CRS) (2 studies)*

– 88 would screen positive, of which 57 (65%) would not have tuberculosis (false positive).

– 912 would screen negative, of which 19 (2%) would have tuberculosis (false negative).

*Abnormal CXR in tuberculosis contacts (CRS) (8 studies)*

– 63 would screen positive, of whom 19 (30%) would not have tuberculosis (false positive).

– 937 would screen negative, of whom 6 (1%) would have tuberculosis (false negative).

*Xpert MTB/RIF in children, inpatient or outpatient microbiologic reference standard (MRS) (2 studies)*

– 31 to 69 would be Xpert MTB/RIF-positive, of whom 9 to 19 (28% to 29%) would not have tuberculosis (false positive).

– 931 to 969 would be Xpert MTB/RIF-negative, of whom 0 to 28 (0% to 3%) would have tuberculosis (false negative).

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

Diagnosing tuberculosis in children is difficult. This may lead to screening tests appearing more or less accurate than they actually are. For Xpert MTB/RIF, there were few studies and children tested to be confident about results.

**Who do the results of this review apply to?**

Children at risk for pulmonary tuberculosis. Results likely do not apply to children in the general population. Studies mainly took place in countries with a high burden of tuberculosis.

**What are the implications of this review?**

In children who are tuberculosis contacts or living with HIV, screening tests using symptoms or CXR may be useful. However, symptoms and CXR formed part of the reference standard, which may falsely elevate the accuracy of the results. We urgently need better screening tests for tuberculosis in children to better identify children who should be considered for tuberculosis preventive treatment and to increase the timeliness of treatment in those with tuberculosis disease.

**How up-to-date is this review?**

To 14 February 2020.

## SUMMARY OF FINDINGS

### Summary of findings 1. Symptoms for screening of pulmonary tuberculosis

**Review question:** what is the accuracy of symptom groups to screen for pulmonary tuberculosis?

**Studies:** cross-sectional and cohort studies

**Setting:** inpatient and outpatient

**Patients/population:** children with close tuberculosis contacts

**Index tests:** groups of multiple symptoms

**Role:** an initial test

**Threshold for index tests:** any 1 of multiple symptoms

**Reference standards:** composite

| Index test | Population and Setting | Estimation (95% CI) | Number of participants (studies); % with pulmonary TB | Test result | Number of results per 1000 participants tested (95% CI) | | | Certainty of the evidence (GRADE) |
|---|---|---|---|---|---|---|---|---|
| | | | | | Prevalence 0.5% | Prevalence 5% | Prevalence 10% | |
| **≥ 1 of cough, fever, or poor weight gain** | Close TB contacts | **Pooled sensitivity 89%** (52% to 98%) | 113 (4); 2% to 13% | True positives | 4 (3 to 5) | 45 (26 to 49) | 89 (52 to 98) | ⊕⊕⊖⊖ **Low** [a,b] |
| | | | | False negatives | 1 (0 to 2) | 5 (1 to 24) | 11 (2 to 48) | |
| | | **Pooled specificity 69%** (51% to 83%) | 2582 (4) | True negatives | 687 (507 to 826) | 656 (485 to 789) | 621 (459 to 747) | ⊕⊕⊖⊖ **Low** [c,d] |
| | | | | False positives | 308 (169 to 488) | 294 (161 to 465) | 279 (153 to 441) | |
| **≥ 1 of cough, fever, or decreased playfulness** | Children < 5 years old in inpatient and outpatient settings | **Sensitivity range 64% to 76%** [e] | 106 (3); 3% to 13% | True positives | 3 to 4 | 32 to 38 | 64 to 76 | ⊕⊕⊕⊖ **Moderate** [f] |
| | | | | False negatives | 1 to 2 | 12 to 18 | 24 to 36 | |
| | | **Specificity range 37% to 77%** [e] | 2339 (3) | True negatives | 368 to 766 | 352 to 731 | 333 to 693 | ⊕⊕⊖⊖ **Low** [g,h] |
| | | | | False positives | 229 to 627 | 219 to 598 | 207 to 567 | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **≥ 1 of cough, fever, poor weight gain, or tuberculosis close contact (WHO 4-symptom symptom screen)** | Children with HIV in outpatient settings | **Pooled sensitivity 61%** (58 to 64) | 1219[i] (2); 3% and 8% | True positives | 3 (3 to 3) | 31 (29 to 32) | 61 (58 to 64) | ⊕⊕⊕⊖ **Moderate** [j] |
| | | | | False negatives | 2 (2 to 2) | 19 (18 to 21) | 39 (36 to 42) | |
| | | **Pooled specificity 94%** (86 to 98) | 201,916[i] (2 studies) | True negatives | 935 (856 to 975) | 893 (817 to 931) | 846 (774 to 882) | ⊕⊕⊖⊖ **Low** [j,k] |
| | | | | False positives | 60 (20 to 139) | 57 (19 to 133) | 54 (18 to 126) | |

**CI:** confidence interval; **TB:** tuberculosis; **WHO:** World Health Organization.

**GRADE certainty of the evidence**

**High:** further research is very unlikely to change our confidence in the estimate of effect.
**Moderate:** further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
**Low:** further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
**Very low:** any estimate of effect is very uncertain.

We included plausible prevalence estimates for the target condition suggested by the WHO Global Tuberculosis Programme. The upper limit for the prevalence of tuberculosis in children in a high-risk group in a health facility in a high tuberculosis burden country was estimated to be 10% (100/1000 children); the lower limit for the prevalence of tuberculosis in children in the general population in a high tuberculosis burden country was estimated to be 0.5% (5/1000 children).
Confidence intervals were estimated based on those around the point estimates for pooled sensitivity and specificity.
[a]The two studies with relatively lower sensitivity estimates only included children younger than five years of age, which may explain in part the lower sensitivity. We downgraded one level for inconsistency.
[b]There was a low number of children with pulmonary tuberculosis contributing to this analysis for the observed sensitivity. We considered the 95% CI around false negatives and true positives would likely lead to different decisions depending on which confidence limits are assumed. As we had already downgraded for inconsistency, we downgraded one level for imprecision.
[c]The single study with notably lower specificity used a symptom screen that assessed the presence of symptoms over the past month, while the symptom screens of other studies were composed of more recent symptoms. This may explain differences in specificity. We downgraded one level for inconsistency.
[d]We considered the 95% CI around false positives and true negatives would likely lead to different decisions depending on which confidence limits are assumed. We downgraded one level for imprecision.
[e]Reported as range from studies as meta-analysis did not converge and pooled estimates could not be obtained.
[f]There were few participants contributing to the estimation of sensitivity. We downgraded one level for imprecision.
[g]The study with notably higher specificity did not have any obvious characteristics to explain this. We downgraded one level for inconsistency.
[h]The wide range around true negatives and false positives may lead to different decisions depending on which limits are assumed. We downgraded one level for imprecision.
[i]Reported as number of screens rather than participants.
[j]As assessed by QUADAS-2, both studies had high risk of bias in the flow and timing domain. We downgraded one level for risk of bias.
[k]For individual studies, specificity estimates ranged from 89% to 97%. We thought that differences in threshold for clinical diagnosis could explain in part the heterogeneity. We downgraded one level for inconsistency.
The results presented in this table should not be interpreted in isolation from results of the individual included studies contributing to each summary test accuracy measure. These are reported in the main body of the text of the review.

## Summary of findings 2. Chest radiography for screening of pulmonary tuberculosis

**Review question:** what is the accuracy of chest radiography to screen for pulmonary tuberculosis?

**Studies:** cross-sectional and cohort studies

**Setting:** inpatient and outpatient

**Patients/population:** children with close tuberculosis contacts

**Index test:** abnormal chest radiography

**Role:** an initial test

**Threshold for index tests:** author defined and implicit as utilized by the chest radiography reader

**Reference standard:** composite

| Estimation (95% CI) | Number of participants (studies); % with pulmonary TB | Test result | Number of results per 1000 participants tested (95% CI) | | | Certainty of the evidence (GRADE) |
|---|---|---|---|---|---|---|
| | | | Prevalence 0.5% | Prevalence 5% | Prevalence 10% | |
| **Pooled sensitivity 87%** (75% to 93%) | 232 (8); 2% to 25% | True positives | 4 (4 to 5) | 44 (38 to 47) | 87 (75 to 93) | ⊕⊕⊖⊖ **Low** [a,b,c] |
| | | False negatives | 1 (0 to 1) | 6 (3 to 12) | 13 (7 to 25) | |
| **Pooled specificity 99%** (68% to 100%) | 3281 (8) | True negatives | 975 (677 to 985) | 931 (646 to 941) | 882 (612 to 891) | ⊕⊕⊖⊖ **Low** [a,d,e] |
| | | False positives | 20 (10 to 318) | 19 (9 to 304) | 18 (9 to 288) | |

**CI:** confidence interval; **TB:** tuberculosis.

**GRADE certainty of the evidence**

**High:** further research is very unlikely to change our confidence in the estimate of effect.
**Moderate:** further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
**Low:** further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
**Very low:** any estimate of effect is very uncertain.

The results presented in this table should not be interpreted in isolation from results of the individual included studies contributing to each summary test accuracy measure. These are reported in the main body of the text of the review.

Prevalence estimates were suggested by the Child and Adolescent TB Working Group. The upper limit for the prevalence of tuberculosis in children in a high-risk group in a health facility in a high tuberculosis-burden country was estimated to be 10% (100/1000 children); the lower limit for the prevalence of tuberculosis in children in the general population in a high tuberculosis-burden country was estimated to be 0.5% (5/1000 children).
Confidence intervals were estimated based on those around the point estimates for pooled sensitivity and specificity.
[a]As assessed by QUADAS-2, all three studies had high risk of bias because the index test was a component of the reference standard. We downgraded one level for risk of bias.

[b]One study had a low sensitivity (52%), but the other seven had sensitivity of 78% or above. The reason for the difference in sensitivity was unclear. We did not downgrade for inconsistency.
[c]There were relatively few children contributing to the analysis of sensitivity. We downgraded one level for imprecision.
[d]For individual studies, specificity estimates ranged from 28% to 100%. Seven studies had a specificity of 73% or higher. Inter-reader variability in the interpretation of paediatric chest radiographs could in part explain the heterogeneity. We downgraded one level for inconsistency.
[e]The 95% CI around true negatives and false positives would likely lead to different decisions depending on which confidence limits are assumed. However, these are also attributable to inconsistency and have already been downgraded in that domain so we did not downgrade further for imprecision.

### Summary of findings 3. Xpert MTB/RIF for screening of pulmonary tuberculosis

**Review question:** what is the accuracy of Xpert MTB/RIF to screen for pulmonary tuberculosis?

**Studies:** cross-sectional and cohort studies

**Setting:** inpatient and outpatient

**Patients/population:** children evaluated in inpatient or outpatient settings

**Index tests:** Xpert MTB/RIF

**Role:** an initial test

**Threshold for index tests:** an automated result is provided

**Reference standard:** microbiological

| Estimations | Number of participants (studies); prevalence of tuberculosis | Test result | Number of results per 1000 participants tested | | | Certainty of the evidence (GRADE) |
|---|---|---|---|---|---|---|
| | | | Prevalence 0.5% | Prevalence 5% | Prevalence 10% | |
| **Sensitivities 43% and 100%** | 16 (2); 1% and 4% | True positives | 2 to 5 | 22 to 50 | 43 to 100 | ⊕⊖⊖⊖ **Very low** [a,b,c] |
| | | False negatives | 0 to 3 | 0 to 28 | 0 to 57 | |
| **Specificities 99% and 100%** | 771 (2) | True negatives | 975 to 985 | 931 to 941 | 882 to 891 | ⊕⊕⊕⊖ **Moderate** [b] |
| | | False positives | 10 to 20 | 9 to 19 | 9 to 18 | |

**GRADE certainty of the evidence**

**High:** further research is very unlikely to change our confidence in the estimate of effect.
**Moderate:** further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
**Low:** further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
**Very low:** any estimate of effect is very uncertain.

The results presented in this table should not be interpreted in isolation from results of the individual included studies contributing to each summary test accuracy measure. These are reported in the main body of the text of the review.

We included plausible prevalence estimates for the target condition suggested by the World Health Organization Global Tuberculosis Programme. The upper limit for the prevalence of tuberculosis in children in a high-risk group in a health facility in a high tuberculosis-burden country was estimated to be 10% (100/1000 children); the lower limit for the prevalence of tuberculosis in children in the general population in a high tuberculosis-burden country was estimated to be 0.5% (5/1000 children).

[a]The study with the higher sensitivity had only two cases included in the estimation of sensitivity. This study was also conducted in an inpatient setting evaluating children with severe malnutrition, while the other was in an outpatient setting evaluating child tuberculosis contacts. These differences may have explained in part the variability in sensitivity estimates. We downgraded one level for inconsistency.

[b]There were only two studies, both conducted in Africa. Neither was a high tuberculosis-burden country. The applicability to other settings comes with some uncertainty. We downgraded one level for indirectness.

[c]There were few participants contributing to the analysis of sensitivity. We downgraded two levels for imprecision.

## BACKGROUND

Tuberculosis continues to elude traditional control strategies. According to the WHO Global Tuberculosis Report 2020, an estimated 10 million people in 2019 were ill with tuberculosis worldwide. Of these, over 25% were not diagnosed or reported to the World Health Organization (WHO). Children less than 15 years old represented approximately 12% of incident cases, but 16% of the estimated 1.4 million deaths from tuberculosis in 2019. This relatively higher share of mortality in children highlights urgent needs of improved case detection and subsequent access to treatment in this age group (WHO Global Tuberculosis Report 2020).

Case finding is a crucial step in the cascade of care for people with tuberculosis; however, for most deaths from childhood tuberculosis, the disease is never diagnosed (Jenkins 2017). In the "Roadmap towards ending TB in children and adolescents," the WHO identifies case finding for childhood tuberculosis as a key activity (WHO 2018). Major factors that lead to underdiagnosis of childhood tuberculosis include the following: 1. symptoms tend to be less specific in children and overlap with those of other common childhood diseases; 2. existing tests for children are invasive and have suboptimal sensitivity; ideally, tests need to be inexpensive, accessible, and usable at the point of care, allowing for actionable information for patient care; and 3. reliance on a clinical diagnosis of tuberculosis, without microbiological evidence of disease, requires expertise, which is often not available in areas where the burden of disease is greatest. Given these factors, national and international guidelines for child health generally lack systematic screening strategies for tuberculosis (WHO 2018).

For adults, systematic screening for tuberculosis in high-risk groups and vulnerable populations is a more established strategy to improve case detection in high-burden settings. In 2013, the WHO published "Systematic screening for active tuberculosis: principles and recommendations." This document provided guidance for the development of screening approaches for adults (WHO 2013a). One Cochrane protocol (van't Hoog 2014) and an ensuing non-Cochrane systematic review (van't Hoog 2013) contributed to the WHO recommendations (WHO 2013a). Participants included in the systematic review were adults aged 15 years and older. The review excluded studies of children aged zero to five years or studies of childhood tuberculosis only. Since 2013, estimation of the true burden of childhood tuberculosis has improved and several promising strategies for case finding are being either newly implemented or developed (Schumacher 2019; Stop TB Partnership 2019). With this, there is a new call to push forward systematic screening for childhood tuberculosis (Reuter 2019; WHO 2018). This review addressed tuberculosis screening strategies in children under 15 years of age.

### Screening

Tuberculosis screening is a term that has been used differently in the literature depending on the context. We have adopted the definition of tuberculosis screening from the WHO as "the systematic identification of people with suspected active TB [tuberculosis], in a predetermined target group, using tests, examinations or other procedures that can be applied rapidly" (WHO 2013a; WHO 2015). The WHO's more recent End-TB strategy emphasizes early diagnosis of tuberculosis and systematic screening of contacts and high-risk groups (WHO 2018), which is in line with the above definition of tuberculosis screening.

### Target condition being diagnosed

Tuberculosis is a communicable disease caused by the bacterium *Mycobacterium tuberculosis* (*M tuberculosis*). A small fraction of people with tuberculosis infection initially develops active tuberculosis (tuberculosis disease). More commonly, initial infection leads to latent tuberculosis infection, which has the potential to become active tuberculosis throughout a person's lifetime, especially during states of immunosuppression such as HIV infection and malnutrition. *M tuberculosis* is transmitted from person to person through the air and, therefore, most commonly causes disease in the lungs, referred to as pulmonary tuberculosis. Tuberculosis can, however, occur in any organ or tissue outside of the lungs (referred to as extrapulmonary tuberculosis), with lymph node tuberculosis as the most common form and tuberculous meningitis as the most severe form of extrapulmonary disease. As the most common form of active tuberculosis is lung disease, most screening studies in adults and children evaluate tests and strategies for pulmonary tuberculosis and verify tuberculosis using respiratory specimens. In this review, the target condition is pulmonary tuberculosis.

Signs and symptoms of pulmonary tuberculosis include fever, cough, night sweats, weight loss or poor weight gain, visible neck mass, and decreased activity. However, pulmonary tuberculosis symptoms in children, especially those under five years of age, tend to be less specific because they often overlap with other common paediatric conditions such as pneumonia, HIV-associated lung disease, and malnutrition (Jaganath 2012; Oliwa 2015). Compared to adults, children are much more likely to progress from latent tuberculosis infection to tuberculosis disease. Further, among those progressing to disease, younger children are more likely to experience severe manifestations (Marais 2004; Perez-Velez 2012).

Microbiological confirmation of pulmonary tuberculosis in children is complicated by two main factors. First, younger children are not able to voluntarily expectorate sputum, which is the standard specimen used for microbiological detection of pulmonary tuberculosis in adults. Therefore, specimens from young children traditionally are collected from more invasive methods such as gastric aspiration and sputum induction (Graham 2015). Second, lung cavities with high bacillary load as seen in pulmonary tuberculosis in adults are uncommon in children, especially in young children under 10 years of age. The number of bacilli causing disease in children tends to be low and the 'paucibacillary' nature of their disease compromises diagnostic yield (Dunn 2016).

### Index test(s)

This review included the following index tests used in screening for pulmonary childhood tuberculosis: symptoms, chest radiography (CXR), Xpert MTB/RIF and Xpert Ultra, and various combinations of these tests.

With symptom-based screening, individuals or their caregivers are interviewed about symptoms suggestive of pulmonary tuberculosis such as cough or fever of varying duration, weight loss, poor weight gain or reduced appetite, and decreased physical activity. Though not a true symptom, recent contact with an infectious person with tuberculosis is another important factor when interviewing for tuberculosis risk (Graham 2015).

CXR may involve posterior-anterior, anterior-posterior, or lateral recording, or a combination of these. Commonly used types of CXR include conventional CXR (producing 36 cm × 43 cm film), digital radiography, and computed radiography. The most common radiographic finding of pulmonary childhood tuberculosis is hilar lymphadenopathy (Leung 1992), though CXR has limitations identifying this finding (Swingler 2005). Accurate interpretation of CXR findings for pulmonary childhood tuberculosis is dependent on the ability of the healthcare professional interpreting the CXR, and wide interobserver variation has been reported (Du Toit 2002; Kaguthi 2014). Computer-aided interpretation of CXR for pulmonary tuberculosis diagnosis or screening is a promising new technology (Qin 2019; Sodhi 2017) that has been recommended by the WHO as an alternative to human reader interpretation of CXR screening and triage for tuberculosis in people aged 15 years and above (WHO Consolidated Guidelines (Module 2) 2021). However, it has not been adequately assessed in children and may be complicated by the wide variety of intra-thoracic disease manifestations observed in children compared to adults (Reuter 2019).

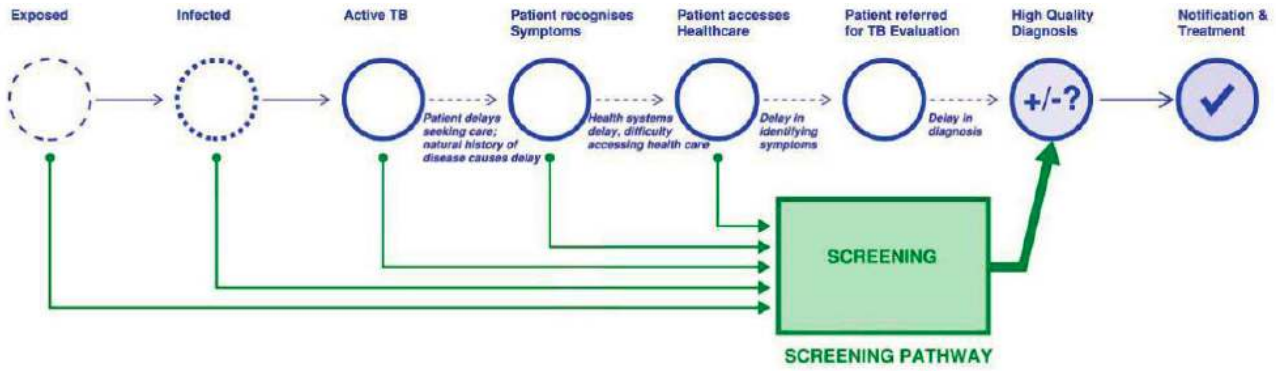
Xpert MTB/RIF and Xpert Ultra, the newest version (Cepheid Inc, CA, USA) are nucleic acid amplification tests (NAATs) that can detect both *M tuberculosis* DNA and rifampicin resistance. We did not assess rifampicin resistance in this review. These two assays are completely automated and self-contained once the sample is loaded into the cartridge. Specimen processing is similar for both Xpert MTB/RIF and Xpert Ultra using Xpert Sample Reagent and requires 15 minutes of incubation. Within two hours, results are available. A consistent supply of electricity, temperature control, and annual calibration of the cartridge modules are needed (Global Laboratory Initiative 2019). Xpert Ultra has approximately 1-log improvement in the lower limit of detection of bacterial load compared to Xpert MTB/RIF (Chakravorty 2017). Xpert Ultra also has a new result category, 'trace call,' that represents minimally detectable bacillary load. According to the WHO, a 'trace call' result is adequate to prompt initiation of tuberculosis treatment in children or people living with HIV (WHO 2017b). The WHO recommends the use of Xpert MTB/RIF and Xpert Ultra as initial diagnostic tests for pulmonary tuberculosis in adults and children. Specifically in children, the guidelines recommend a variety of specimen types for diagnosis of pulmonary tuberculosis, including gastric aspirates, nasopharyngeal aspirates, and stool specimens, in addition to sputum (WHO Consolidated Guidelines (Module 3) 2020). We included Xpert MTB/RIF (all versions) and Xpert Ultra in this review.

Another WHO-recommended NAAT for detection of tuberculosis is Truenat MTB and Truenat MTB Plus (Molbio Diagnostics/ Bigtec Labs, Goa/Bengaluru, India) (WHO Consolidated Guidelines (Module 3) 2020). However, to our knowledge, there are currently no published studies assessing this test in children.

### Clinical pathway

As shown in Figure 1, there are two complementary approaches to detection of tuberculosis disease. The first is the patient-initiated pathway, also known as passive case finding. The second is the provider-initiated screening or active case finding pathway (WHO 2015), which is the analytic framework for this review. One major challenge with either pathway is that 'high-quality diagnosis' is elusive for childhood tuberculosis, especially for younger children and children in resource-limited settings. This diagram also demonstrates the wide range of potential target populations for childhood tuberculosis screening, ranging from contacts of those with tuberculosis ('exposed') to symptomatic children in inpatient or outpatient settings (e.g. children living with HIV, as described below). This review included evidence from all these systematic screening strategies.

**Figure 1. There are two complementary approaches to detection of tuberculosis (TB) disease. The first is the patient-initiated pathway, also known as passive case finding. The second is the provider-initiated screening pathway (WHO 2015), which is the analytic framework for this review. One major challenge with either pathway is that 'high-quality diagnosis' is elusive for child tuberculosis, especially for younger children and in resource-limited settings. This diagram also demonstrates the wide range of potential target populations for tuberculosis screening, ranging from contacts of those with tuberculosis ('exposed') to symptomatic patients accessing healthcare, such as children living with HIV. Copyright © [2015] [World Health Organization]: reproduced with permission.**

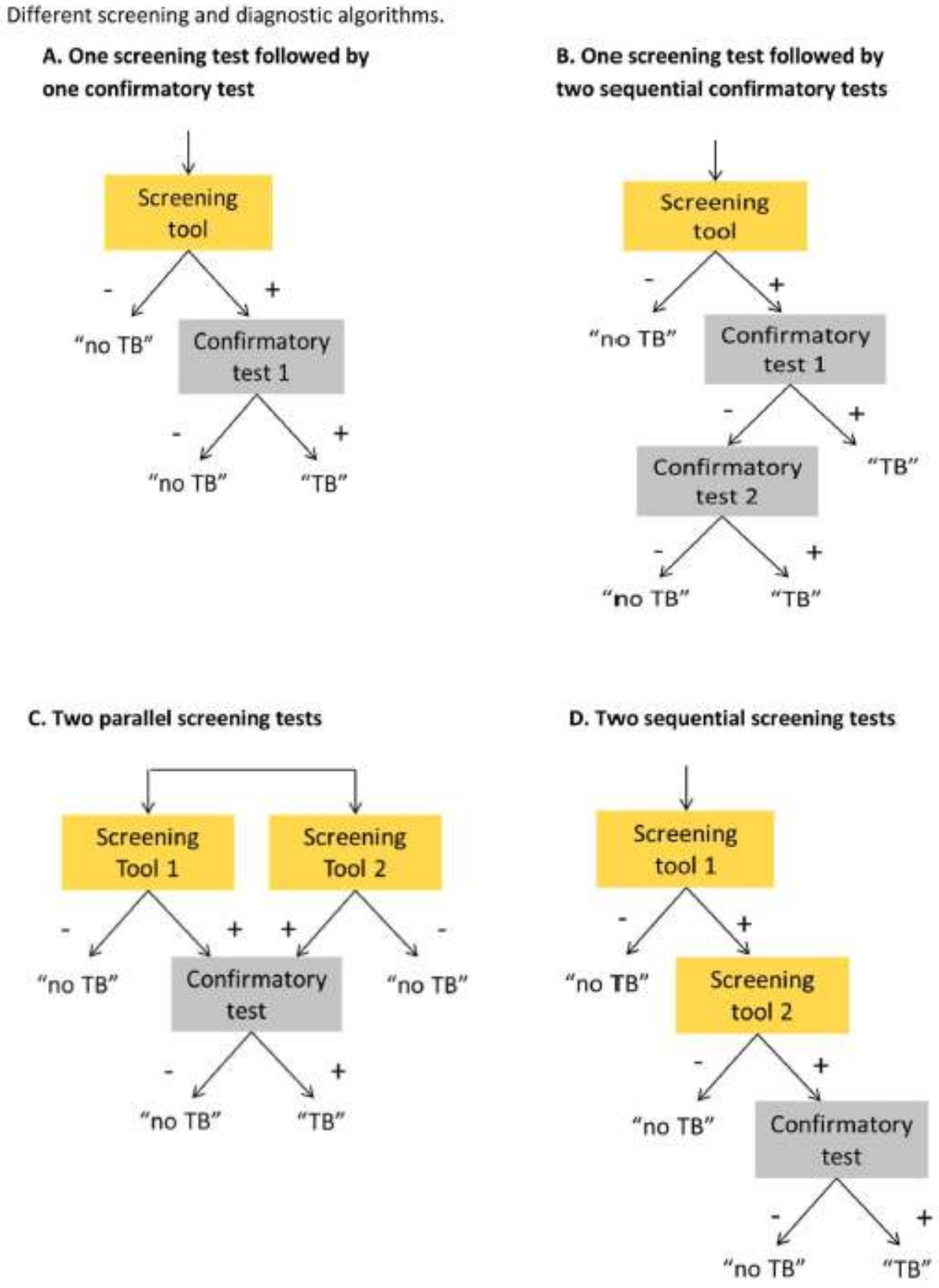



There is no standard screening approach for children, but for the subgroup of children living with HIV, since 2011 the WHO has recommended routine symptom-based screening for all children living with HIV presenting to healthcare facilities as part of the intensified case-finding strategy. Under this guideline, children living with HIV over 12 months of age who report any cough, fever, weight loss or poor weight gain, or history of recent contact with someone with tuberculosis should be further investigated

for tuberculosis. If no symptoms or recent tuberculosis contact are reported they are considered "unlikely to have active TB." Although this 'strong recommendation' was based upon 'low-quality evidence' (WHO 2011), it exemplifies a standardized screening approach for tuberculosis. A similar symptom-based approach has been suggested for household contacts of infectious tuberculosis cases, focusing on any current symptoms (WHO 2014). The main aim here is to allow tuberculosis contacts or children living with HIV, who are completely asymptomatic, prompt access to tuberculosis preventive treatment. For tuberculosis contacts, the WHO Consolidated Guidelines (Module 1) 2020 make a distinction in the strength of recommendation for provision of tuberculosis preventive treatment in children aged under five years (strong recommendation) and in children aged five years and older (conditional recommendation).

Screening may use sequential or parallel strategies (Figure 2). With sequential strategies, only those with a positive result in the first step are screened in the second step. With parallel screening strategies, multiple different screens are done initially, and any positive screen or combinations of positive screens prompts further investigation (i.e. confirmatory test) for the target condition. We included results from various screening strategies in this review. We considered individuals' results to be 'true screen positives' if they were rightfully referred for confirmatory testing; in contrast, we considered individuals' results to be 'false screen positives' if the individuals were referred for confirmatory testing but not diagnosed with tuberculosis. Although individuals with negative screens should not undergo confirmatory testing during routine clinical practice, individuals with negative screens may complete confirmatory testing in a research context to establish true screen negatives and false screen negatives. As described in Types of studies, studies that only conducted confirmatory testing on those with positive screens were excluded in this review. In the context of this review, the intended use of the index tests is considered to be 'screening,' and their role is considered to be triage tests. With triage tests, the index test is used prior to an existing test or strategy, and only those with a specific result on the triage test continue along the clinical pathway (Bossuyt 2006).

**Figure 2. Different screening and diagnostic algorithms.**




The downstream consequences of screening include the following.

- True positive: children would benefit from rapid diagnosis and initiation of appropriate treatment.
- True negative: children would be spared unnecessary treatment and would benefit from reassurance, pursuit of an alternative diagnosis if they have symptoms, and prompt initiation of tuberculosis preventive treatment if eligible.
- False positive: children would probably experience anxiety and morbidity caused by additional testing, unnecessary treatment, and possible adverse events; strain on healthcare resources with unnecessary additional testing and treatment; possible stigma associated with a tuberculosis diagnosis; the chance that a false-positive result may halt further diagnostic evaluation of the true underlying condition; and missed or delayed initiation of tuberculosis preventive treatment if eligible.
- False negative: children would experience an increased risk of morbidity and mortality, and delayed or inappropriate treatment initiation; there would be risk of ongoing tuberculosis transmission particularly in older children; and they may be inappropriately initiated on tuberculosis preventive treatment.

### Alternative test(s)

Two types of immunological tests excluded from this review are the tuberculin skin test (TST) and the interferon gamma release assay (IGRA). Both methods are dependent on the cellular immune response to *M tuberculosis* antigens in individuals previously exposed to the organism, and neither can distinguish between latent tuberculosis infection and active tuberculosis disease (Pai 2014). Further, neither method is sensitive enough to serve as a rule out test for tuberculosis disease in children, but is mainly used to confirm tuberculosis infection and to support clinical decision making; with full consideration of all the stated caveats. The TST has been in clinical use for over a century and involves intradermal injection of *M tuberculosis* purified protein derivative. Drawbacks to the TST include the need for a second clinical encounter 48 to 72 hours after placement for result interpretation, inter-reader variability, a tendency for previous bacillus Calmette-Guerin vaccination to result in false-positive results, and a tendency for false-negative results in immunosuppressed individuals or due to anergy in individuals with active disease (Pai 2014).

Commercially available IGRAs include QuantiFERON-TB Gold In-tube (QFT-GIT; Qiagen, Germantown, MD), QuantiFERON-TB Gold Plus (QFT-Plus; Qiagen), and T-SPOT.TB (Oxford Immunotec Ltd, Oxford, UK). To improve upon the TST, IGRAs were developed to measure release of interferon gamma from T cells stimulated by antigens specific to *M tuberculosis*. The QFT-GIT assay stimulates interferon gamma release from CD4+ T cells, while the QFT-Plus assay can stimulate both CD4+ and CD8+ T-cell responses. CD8+ cytotoxic T cells have been shown to have higher responses in people with active pulmonary tuberculosis compared to those with latent tuberculosis infection (Day 2011; Rozot 2013). Individuals with low CD4+ T-cell counts (e.g. those with advanced HIV) have been shown to maintain CD8+ T-cell antigen responses to *M tuberculosis* (Sutherland 2010). For these reasons, it is theorized that the QFT-Plus assay may be more sensitive for people living with HIV and those with active tuberculosis (Theel 2018), although this has not been demonstrated in clinical practice. The T-SPOT.TB is an enzyme-linked immunoassay that involves incubation of peripheral blood mononuclear cells with antigens specific to *M tuberculosis*. If the number of interferon gamma-producing T cells (spot-forming cells) exceeds a specific threshold relative to negative control wells, the result is positive. All IGRAs utilize positive and negative controls, and they can have indeterminate results if there is a low interferon gamma response in the positive control or if there is a high response in the negative control (Pai 2014).

Beyond the index tests described above, there are several alternative approaches that could be used for screening or diagnosis. This includes examination of sputum smears for acid-fast bacilli under a light microscope using the classical Ziehl-Neelsen staining technique, or fluorescence microscopy with newer light-emitting diode (LED) microscopy. One review found that in children, the sensitivity of smear microscopy was around 22% in gastric aspirates and around 29% in expectorated and induced sputum specimens (WHO 2013b). Microscopy is unable to differentiate *M tuberculosis* from nontuberculous mycobacteria, which may also cause lung disease.

New assays detect lipoarabinomannan (LAM) antigen in the urine of people with tuberculosis disease. LAM is a lipopolysaccharide present in the lipid rich mycobacterial cell wall. Urinary lateral flow LAM assays have the advantages of being rapid and non-invasive. Currently, the only commercially available lateral flow LAM assay is the Alere Determine TB LAM Ag (AlereLAM, Abbott, Chicago, IL, USA). Based on evidence from randomized trials and a Cochrane Review (Bjerrum 2019), the WHO recommends that lateral flow LAM should be used to assist in the diagnosis of active tuberculosis in HIV-positive adults, adolescents, and children. The full recommendations, which differ for inpatients and outpatients, are described in WHO Consolidated Guidelines (Module 3) 2020. Another LAM assay expected to become commercially available is the Fujifilm SILVAMP TB-LAM (Fujifilm, Tokyo, Japan). Early evidence for this assay demonstrates superior sensitivity compared to AlereLAM for adults living with HIV (Bjerrum 2020; Broger 2020). However, accuracy comparisons between these two LAM assays have varied in children (Nicol 2021; Nkereuwem 2021).

The development of novel tools for detection of tuberculosis disease is an active field. Noteworthy tests with emerging evidence include C-reactive protein (Albuquerque 2019), IP-10 (Alsleben 2011; Holm 2014; Jenum 2016; Sudbury 2019; Tebruegge 2015), and C-Tb (Statens Serum Institut, Copenhagen) (Aggerbeck 2019; Ruhwald 2017). During the 2020s, more efficient technologies are anticipated with the hope that these will advance screening strategies and reduce the burden of childhood tuberculosis worldwide (Schumacher 2019; Stop TB Partnership 2019; WHO 2017a).

## Rationale

Effective screening for childhood tuberculosis supports timely and reliable diagnosis, which is essential for reducing tuberculosis-attributable morbidity and mortality. Effective screening also supports disease rule out, thereby guiding treatment for latent tuberculosis infection and consideration for preventive treatment for exposed children or other high-risk groups such as children living with HIV. Historically, screening children for active tuberculosis has been limited by the lack of accurate screening and diagnostic tools. Therefore, systematic screening in children has only been performed within specific populations with increased risk of disease to limit the risk of false-positive test results and consequent overtreatment of tuberculosis. Guidance from the WHO states that "only children who are close contacts of someone with pulmonary tuberculosis and HIV-positive children should

be systematically screened for TB [tuberculosis]" (WHO 2015). Optimal screening strategies for these two high-risk groups are lacking (Szkwarko 2017), although a symptom-based approach has been supported in resource-limited settings (WHO 2014). Limiting systematic screening to child contacts and HIV-positive children may propagate missed opportunities as evidence has identified other high-risk groups of children in certain settings and with health conditions, such as malnutrition or pneumonia, who are also at risk of tuberculosis (Arscott-Mills 2014; Chisti 2014; LaCourse 2014; Munthali 2017; Oliwa 2015). Evidence also demonstrates that children in tuberculosis-endemic settings have considerable risk of tuberculosis exposure outside of their homes (Martinez 2019). However, the unfortunate reality is that systematic screening is rarely implemented in resource-limited settings, even in highly vulnerable young children who are household contacts of infectious tuberculosis cases and at high risk of tuberculosis infection.

This Cochrane Review informed a WHO guideline Development Group meeting convened to update recommendations for systematic screening for active tuberculosis (WHO Consolidated Guidelines (Module 2) 2021). To our knowledge, this is the first systemic review on this topic in children. There have been several systematic reviews evaluating the accuracy of the index tests described above for the diagnosis of active tuberculosis, including a recent Cochrane Review evaluating Xpert MTB/RIF and Xpert Ultra in children (Kay 2020). The lack of knowledge regarding the performance of these tests to complete childhood tuberculosis screening reflects the difficulty of tuberculosis research in children and the predominance of research focused on diagnosis rather than screening. The current review elucidates the potential of these tools for systematic screening for active pulmonary childhood tuberculosis in specific high-risk populations.

## OBJECTIVES

To estimate the sensitivity and specificity of the presence of one or more tuberculosis symptoms, or symptom combinations; chest radiography (CXR); Xpert MTB/RIF; Xpert Ultra; and combinations of these as screening tests for detecting active pulmonary childhood tuberculosis in the following groups.

- Tuberculosis contacts, including household contacts, school contacts, and other close contacts of a person with infectious tuberculosis.
- Children living with HIV.
- Children with pneumonia.
- Other risk groups (e.g. children with a history of previous tuberculosis, malnourished children).
- Children in the general population in high tuberculosis burden settings.

### Secondary objectives

To compare the accuracy of the different index tests and different thresholds (e.g. CXR with any abnormality versus, more specifically, CXR with abnormality suggestive of tuberculosis).

To investigate potential sources of heterogeneity in accuracy estimates in relation to age group, HIV status, whether the study was conducted in a high tuberculosis burden country, whether the child received a single screening or more than one screening, and type and number of CXR interpreters.

We were interested in the accuracy of the index tests in any setting (i.e. community, outpatient, and inpatient).

## METHODS

### Criteria for considering studies for this review

#### Types of studies

We included cross-sectional studies and cohort studies that assessed the accuracy of at least one of the index tests for pulmonary tuberculosis. We also planned to include randomized controlled trials, but none were identified for inclusion. We included studies from all settings and time periods. Data on the results of index test(s) against the reference standard(s) must have been available so that we could construct 2×2 contingency tables containing the number of true positives, false positives, true negatives, and false negatives. We excluded studies in which children with negative screening test results were not verified by the reference standard because true-negative and false-negative test results cannot be obtained. Studies applying index tests multiple times to an individual within a short timeframe (e.g. within a single hospital admission) were considered diagnostic rather than using a screening approach, and we excluded these studies.

We included cohort studies with children with active tuberculosis identified after the time point that the screening test was applied. Especially with studies performed in settings of intended use, the collection of specimens and conduct of the reference standard may occur sometime after the screening test was done. In low-resource settings, this process may take weeks. However, a longer time between the index test and the reference standard would make us less confident that the target condition did not change between the two tests. We addressed this issue in the QUADAS-2 flow and timing domain and in a sensitivity analysis (see Sensitivity analyses).

We included studies that assessed more than one screening test. We excluded case reports and case-control studies, the latter because of the high risk of bias in diagnostic accuracy studies (Rutjes 2006).

#### Participants

We included studies enrolling HIV-positive and HIV-negative children not known to have active tuberculosis prior to screening. We excluded studies if they did not provide data exclusive to participants under 20 years of age with at least 75% participants under 15 years of age. We included children in the general population in high-burden settings and high-risk groups, including children younger than five years old; children living with HIV; children with recent exposure to a person with active tuberculosis; and household, school, or other contacts of a person with active tuberculosis. We included studies in which children were screened only once and studies that reported longitudinal screening with repeated screening tests at predetermined intervals.

#### Index tests

For symptom-based screening, we included studies that assessed any symptom or combinations of symptoms suggestive of possible tuberculosis, as described by the primary study authors. Symptoms of childhood tuberculosis may include cough, fever, night sweats,

decreased appetite, weight loss or failure to thrive, and fatigue or reduced playfulness. Children over 10 years of age experience symptoms similar to those recorded in adults, which may also include haemoptysis. The threshold was presence or absence of symptoms, as defined by the primary study authors. In addition, we included the WHO-recommended intensified case finding (ICF) symptom screen (current cough, fever, poor weight gain, or tuberculosis contact for children; current cough, weight loss, night sweats, or fever for adolescents) for HIV-infected children, applied at each healthcare visit (WHO 2011).

For CXR screening, we included studies that utilized conventional radiography, digital radiography, and computed radiography. We included all classification systems for identification of CXR abnormalities. We categorized all CXR screening results as follows. We used an author defined threshold for CXR results. Essentially this is an implicit threshold utilized by the CXR reader.

- Normal.
- Any CXR abnormality (i.e. abnormalities suggestive of tuberculosis and other abnormalities).
- Abnormalities suggestive of tuberculosis.

For Xpert MTB/RIF and Xpert Ultra, we included studies in which the index tests were evaluated in expectorated or induced sputum, gastric aspirate specimens, nasopharyngeal aspirate specimens, and bronchoalveolar lavage specimens. Tuberculosis bacilli in sputum can be swallowed and detected in stool so we also included studies assessing stool specimens. We included studies assessing more than one type of respiratory specimen collected at the same time and extracted 2×2 data separately for each specimen type.

Xpert MTB/RIF and Xpert Ultra provide the following printed test results:

- MTB (*M tuberculosis*) DETECTED; RIF (rifampicin) resistance DETECTED;
- MTB DETECTED; RIF resistance NOT DETECTED;
- MTB DETECTED; RIF resistance INDETERMINATE;
- MTB NOT DETECTED;
- INVALID (the presence or absence of MTB cannot be determined);
- ERROR (the presence or absence of MTB cannot be determined);
- NO RESULT (the presence or absence of MTB cannot be determined).

Xpert Ultra also gives the following semi-quantitative classifications of *M tuberculosis* bacterial burden from the sample: trace, very low, low, moderate, and high. For this review, Xpert MTB/RIF and Xpert Ultra results were categorized as:

- positive: 'MTB DETECTED,' including 'trace' results from Xpert Ultra;
- negative: 'MTB NOT DETECTED;'
- inconclusive: 'INVALID,' 'ERROR,' or 'NO RESULT.'

We did not evaluate detection of rifampicin resistance in this review.

As shown in Figure 2, with two parallel screening tests, the parallel strategy will entail any of the individual components of the strategy being positive resulting in a positive parallel strategy screen and all individual components being negative resulting in a negative parallel strategy screen. For studies assessing parallel screening tests, if data for the individual components of the parallel strategy against the reference standard were also available, these data were also extracted for analysis.

### Target conditions

The target condition was active pulmonary tuberculosis.

We anticipated that some studies may have evaluated the index tests for active tuberculosis and not explicitly stated 'pulmonary tuberculosis,' the target condition in this review. We included these studies because the most common type of active tuberculosis in children is pulmonary disease; hence, most screening studies in children evaluate tests for pulmonary tuberculosis and diagnose tuberculosis using respiratory specimens.

### Reference standards

We used two reference standards, a microbiological and a composite reference standard.

#### *Microbiological reference standard*

Confirmed pulmonary tuberculosis was defined as a positive culture (on solid or liquid medium) or a positive Xpert MTB/RIF or Xpert Ultra test from a respiratory specimen. When Xpert MTB/RIF was the index test, we excluded it from the reference standard to avoid incorporation bias. We did not include studies where sputum smear microscopy was the reference standard.

Collection of multiple respiratory specimens may improve the diagnostic yield of testing for childhood tuberculosis (Cruz 2012; Zar 2012). With respect to the microbiological reference standard, we included studies that involved multiple specimens collected over time. In these studies, we used the classification of the reference standard as defined by the primary study authors (most commonly at least one positive result representing a positive reference test).

#### *Composite reference standard*

Confirmed pulmonary tuberculosis was defined as microbiological confirmation (as above in 'Microbiological reference standard') or author-defined clinical pulmonary tuberculosis. Clinical pulmonary tuberculosis must have included a component of follow-up to help verify or rule out the diagnosis of active tuberculosis. Hence, the composite reference standard was used to verify disease-positive results and disease-negative results. The consensus research definition for clinical childhood tuberculosis for diagnostic studies was considered too restrictive for the purpose of this review (Graham 2015).

'Not tuberculosis' was defined as negative microbiological test results and establishment of alternative diagnosis during the evaluation for tuberculosis, resolution of symptoms without tuberculosis treatment, or no progression of symptoms for at least one month without tuberculosis treatment.

Two of our index tests, symptoms and CXR, are typically components of case definitions used to support the clinical diagnosis of tuberculosis (i.e. not microbiologically confirmed). This raised the potential for incorporation bias with the composite reference standard, that is, where the result of the index test is used to help determine the reference standard result. We assessed

the composite reference standard for incorporation bias using the QUADAS-2 signalling question: "Were the reference standard results interpreted without knowledge of the results of the index test?" In addition, we discussed incorporation bias as a limitation of the review.

## Search methods for identification of studies

We attempted to identify all relevant published studies regardless of language. Although they were not assessed as index tests in this review, we included immunological tests (TST and IGRA) in the search strategy. This will allow for archiving of relevant studies for a future systematic review assessing immunological tests as index tests.

### Electronic searches

We searched the following databases without language restriction up to 14 February 2020, using the search terms and strategy described in Appendix 1.

- Cochrane Central Register of Controlled Trials (CENTRAL), published in the Cochrane Library.
- MEDLINE and MEDLINE in Process (Ovid), from 1946.
- Embase (Ovid), from 1947.
- Scopus (Elsevier) from 1970.

We also searched ClinicalTrials.gov, the WHO International Clinical Trials Registry Platform (ICTRP; www.who.int/trialsearch), and the International Standard Randomized Controlled Trials Number (ISRCTN) registry (www.isrctn.com/) for trials in progress.

### Searching other resources

To identify any relevant published data not identified with our electronic search, we contacted experts in the field of childhood tuberculosis and checked the references of relevant reviews from the past 10 years. With the studies selected for inclusion in this review, we performed forward and backward reference checking to identify any additional eligible studies.

## Data collection and analysis

### Selection of studies

We used Covidence to manage the selection of studies (Covidence). Two review authors (BV and TN) independently screened all titles and abstracts from the electronic searches to identify potentially eligible studies. We obtained full-text articles of potentially eligible studies, and the two review authors (BV and TN) independently assessed them for study eligibility using the predefined inclusion and exclusion criteria. We resolved any disagreements by discussion or with a third review author (AMM or KRS). As needed, we contacted study authors to clarify the study methods and other information. Studies excluded during the full-text review are listed in Characteristics of excluded studies with reasons for exclusion. We illustrated the study selection process in a PRISMA flow diagram (Moher 2009).

### Data extraction and management

We designed a data extraction form and piloted it on two included studies. After reviewing the piloted forms with the other review authors, we finalized the form. Two review authors independently used the data extraction form to extract data from the included studies (BV, TN, AMM, or KRS). We discussed any inconsistencies with a third review author. We entered the extracted data into an Excel database on password-protected computers (Excel 2013). Data will be secured to the Cochrane Infectious Diseases Group's 'Archive' drives for future access and review updates.

We extracted the following information from each included study.

#### *Study details*

- First author, title, year of publication, journal, language.
- Study design, sampling method, prospective/retrospective, and inclusion criteria for presumptive tuberculosis (if any).
- Number of participants after screening for exclusion and inclusion criteria.
- Number of children included in the primary study analysis.
- Single or initial screening versus more than one screening in the population.
- Any sequential or parallel screening strategies.

#### *Participant characteristics and setting*

- Description of study population.
- Age: median, mean, range, and disaggregation into categories (0 to 4 years, 5 to 14 years).
- Gender.
- HIV status.
- Proportion with severe wasting or severe acute malnutrition.
- Screening location: community, outpatient facility, or inpatient facility.
- Children with prior tuberculosis included, yes/no? If yes, what proportion?
- Country/countries where study was conducted.
- Country WHO classification for tuberculosis high-burden country (WHO Global Tuberculosis Report 2020).
- Years of data collection.

#### *Index test*

- Definition of positive symptom screen.
- Symptoms assessed.
- Details of timing of contact history (i.e. current, within past year, beyond one year).
- Types of CXR used.
- Description of radiographic findings classification.
- Type of CXR reader: radiologist, pulmonologist, general medical officer, clinical officer, nurse, other.
- Types of respiratory specimens used.
- Types of NAATs used.
- For each index test, number of results that were true positive, false positive, true negative, false negative, inconclusive, and missing.

#### *Reference standard*

- Microbiological reference standard used: solid culture, liquid culture, Xpert MTB/RIF, or Xpert Ultra.
- Criteria used for composite reference standard.
- Number of microbiological tests used to exclude tuberculosis.

- Number of contaminated cultures and total number of cultures performed.
- Time between the index test and the reference standard.

We followed Cochrane policy, which states that "authors of primary studies will not extract data from their own study or studies. Instead, another author will extract these data, and check the interpretation against the study report and any available study registration details or protocol."

### Assessment of methodological quality

Two review authors (of BV, TN, AMM, or KRS) independently assessed the methodological quality of the included studies using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool, which we adapted for this review (Whiting 2011). The tool with signalling questions tailored to this review is in Appendix 2. As recommended, we assessed each of the four domains (patient selection, index test, reference standard, and flow and timing) for risk of bias and the first three domains for concerns regarding applicability.

We judged each item as 'yes' (adequately addressed), 'no' (inadequately addressed), or 'unclear' when there was insufficient information reported to make an assessment. One review author piloted the tool on two included studies. We then made revisions to finalize the QUADAS-2 tool, with specific revisions as described in the Differences between protocol and review section. We resolved disagreements between the two review authors' independent assessments through discussion or additional input from a third review author. We presented results of the quality assessment in text, tables, and graphs.

### Statistical analysis and data synthesis

We presented individual study estimates of sensitivity and specificity graphically on forest plots and in receiver operating characteristics (ROC) space using Review Manager 5 (Review Manager 2020).

We considered one index test result per child per time point. However, for studies assessing serial screening over time for individuals, separate screens were assessed if they were also compared against serial confirmatory tests over time (i.e. multiple screens for one individual). In other words, in situations where serial screening of children at each healthcare visit was recommended, screening results (typically multiple per individual) were used as the unit of analysis rather than single results per participant, as with the other analyses here. Within each group listed in Objectives, we performed analyses by index test and reference standard. For symptom screening as the index test, we performed analyses for single and multiple symptoms where data were available. We consolidated symptom screens across included studies into groups that used similar combinations of symptoms as follows: one or more of cough, fever, or poor weight gain and one or more of cough, fever, or decreased playfulness. For combination of symptoms, a positive screen was the presence of one or more than one symptom.

We combined categories depending on the number of studies and screening definitions found in each category. We also stratified the analyses by the type of reference standard used, microbiological or composite.

When there were sufficient data, we performed meta-analyses to estimate summary values of sensitivity and specificity using a bivariate model (Chu 2006; Reitsma 2005). We chose the bivariate model because test results were binary (present/absent), studies used the same threshold or thresholds recommended by the test manufacturer. When we were unable to fit a bivariate model due to sparse data or few studies, we simplified the models to univariate random-effects or fixed-effect logistic regression models (depending on whether or not heterogeneity was observed on forest and summary ROC (SROC) plots) to pool sensitivity and specificity separately (Takwoingi 2015). If there were only two or three studies available for an analysis and there was substantial heterogeneity, we did not perform a meta-analysis. We performed meta-analyses using the meqrlogit command in Stata version 16 (Stata).

Owing to limited data, we did not perform test comparisons.

#### *Approach to inconclusive index test results*

As described above in Index tests, the NAAT assays assessed in this review as index tests may have inconclusive results. We planned to report the proportion of inconclusive index test results as available, but none of the included studies reported inconclusive results.

### Investigations of heterogeneity

We visually inspected forest plots and SROC plots for heterogeneity. We summarized descriptively the type and number of CXR interpreters. We had planned to assess potential sources of heterogeneity using subgroup analyses and bivariate meta-regression. However, owing to limited data, we did not perform subgroup analyses.

### Sensitivity analyses

Owing to limited data we were unable to perform sensitivity analyses to explore the effect of potential sources of bias and study design characteristics on the accuracy of the index tests.

### Assessment of reporting bias

We did not formally assess reporting bias using funnel plots or regression tests as these have not been reported as helpful for diagnostic test accuracy studies (Macaskill 2010).

### Assessment of certainty of the evidence

We assessed the certainty of evidence using the GRADE approach for diagnostic studies (Balshem 2011; Schünemann 2008). As recommended, we rated the certainty of evidence as high (not downgraded), moderate (downgraded by one level), low (downgraded by two levels), or very low (downgraded by more than two levels) based on five domains: risk of bias, indirectness, inconsistency, imprecision, and publication bias. For each outcome, the certainty of evidence started as high when there were high-quality observational studies (cross-sectional or cohort studies) that enrolled participants with diagnostic uncertainty. If we found a reason for downgrading, we used our judgement to classify the reason as either serious (downgraded by one level) or very serious (downgraded by two levels).

Four review authors (BV, TN, AMM, and KRS) discussed judgements and applied GRADE in the following way (Schünemann 2020a; Schünemann 2020b).

#### *Assessment of risk of bias*

We used QUADAS-2 to assess risk of bias.

#### *Indirectness*

We assessed indirectness in relation to the population (including disease spectrum), setting, interventions, and outcomes (accuracy measures). We also used tuberculosis prevalence as a guide to whether there was indirectness in the population.

#### *Inconsistency*

GRADE recommends downgrading for unexplained inconsistency in sensitivity and specificity estimates. We prespecified analyses to investigate potential sources of heterogeneity; however, owing to limited data, we did not perform these. We downgraded when we could not explain inconsistency in the accuracy estimates based on whether the individual point estimates were similar and if the confidence intervals overlapped in the forest plots.

#### *Imprecision*

We considered a precise estimate to be one that would allow a clinically meaningful decision. We considered the width of the confidence interval (CI), and asked, "Would we make a different decision if the lower or upper boundary of the CI represented the truth?" In addition, we worked out projected ranges for true positive, false negative, true negative, and false positive for a given prevalence of tuberculosis and made judgements on imprecision from these calculations.

#### *Publication bias*

We rated publication bias as undetected (not serious) for several reasons, including the comprehensiveness of the literature search and extensive outreach to tuberculosis researchers to identify studies.

## RESULTS

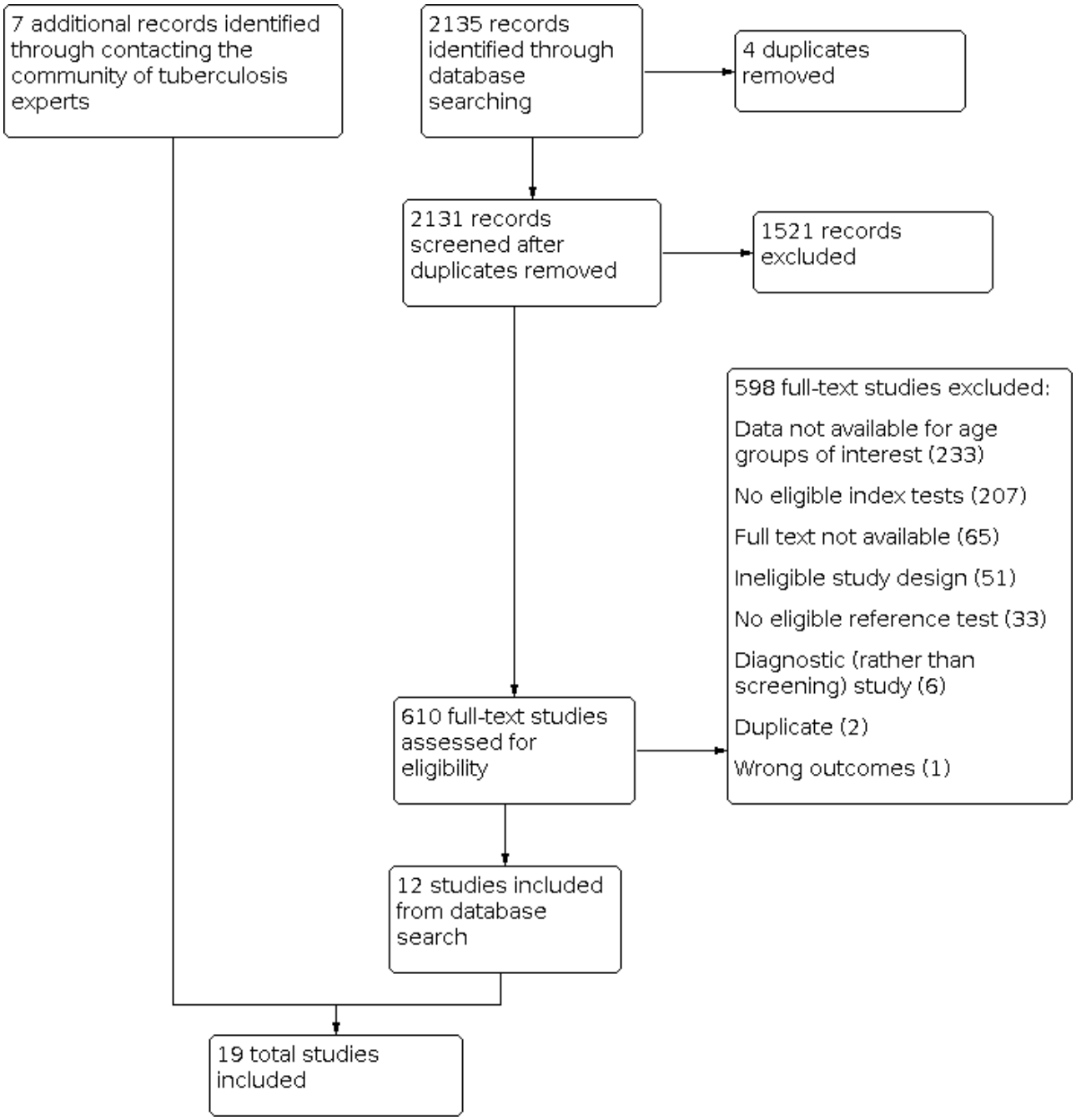
### Results of the search

We identified and screened 2135 records for inclusion in this review. Of these, we assessed 610 full-text papers against our inclusion criteria. We excluded 598 papers for the following reasons: data not available for age groups of interest (233 papers), no eligible index tests (207 papers), full text not available (65 papers), ineligible study design (51 papers), no eligible reference test (33 papers), diagnostic (rather than screening) study (six papers), duplicate (two papers), and wrong outcomes (one paper).

We identified 19 unique studies that met the inclusion criteria of this review, 12 from the database search and seven that were recommended from a community of paediatric tuberculosis experts that we contacted (Aggerbeck 2018; Birungi 2018; Clemente 2017; Dreesman 2017; Jaganath 2013; Kruk 2008; LaCourse 2014; PERCH 2019; Portevin 2014; Rose 2012; Sawry 2018; Schwoebel 2020; Tieu 2014; Togun 2015; Togun 2016; Triasih 2015a; Triasih 2015b; Ustero 2017; Vonasek 2021). All included studies were written in English. Togun 2015 and Togun 2016 assessed different index tests in the same children, and we considered these to be two different studies. Similarly, Triasih 2015a and Triasih 2015b assessed different index tests in the same children, and we designated these as two different studies. We performed descriptive analyses of the included studies and presented their key characteristics in the Characteristics of included studies table and Table 1.

Figure 3 shows the flow of studies through the review process. We listed selected excluded studies and the reasons for their exclusion in the Characteristics of excluded studies table. These studies were selected based upon their relevance to screening for childhood tuberculosis despite not fulfilling inclusion criteria for this review. The full list of excluded studies and the reasons for ineligibility is available from the first author.

**Figure 3. Study flow diagram.**




### Methodological quality of included studies

Figure 4 and Figure 5 show risk of bias and applicability concerns for 19 studies evaluating symptoms, CXR, and Xpert MTB/RIF to screen for pulmonary tuberculosis.

**Figure 4. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies.**

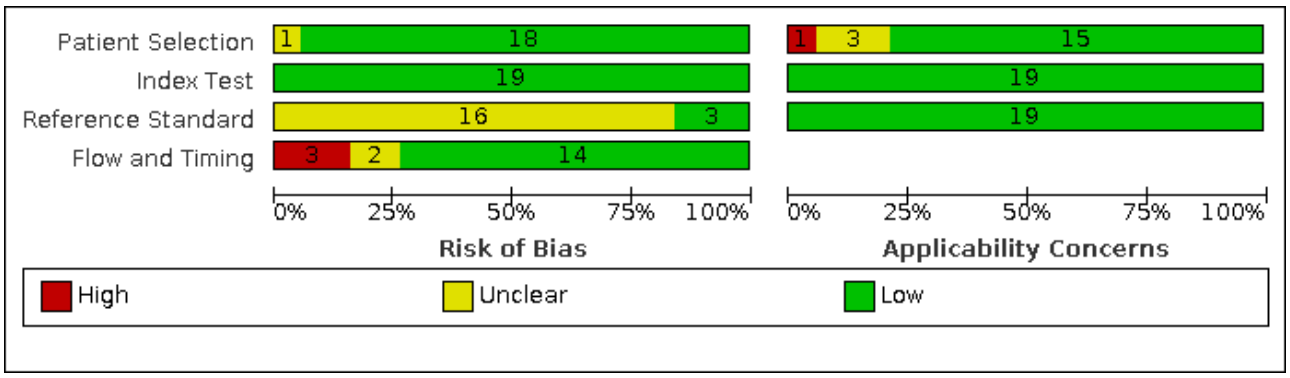


**Figure 5. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.**

In the patient selection domain, we considered 18 studies (95%) at low risk of bias because the studies enrolled a consecutive or random sample of eligible participants and avoided inappropriate exclusions. We considered one study at unclear risk of bias because it was unclear if there was a consecutive or random sample of eligible participants in the study (Jaganath 2013). With respect to applicability, we considered 15 studies at low concern because participants in these studies resembled a population that would typically be considered for screening for tuberculosis. We considered one study to have high concern because enrolment criteria were stricter than is typical for selecting individuals to be screened for tuberculosis (Portevin 2014). We considered three studies (16%) to have unclear concern because we could not determine concerns (Clemente 2017; Rose 2012; Togun 2016).

In the index test domain, we considered all studies at low risk of bias because the results of the index tests were interpreted without knowledge of the results of the reference standard and prespecified thresholds were used, as relevant. Regarding applicability, with respect to the index tests, we considered all studies to have low concern.

In the reference standard domain, we considered three studies (16%) to have low risk of bias because the results of the reference standard were likely to correctly classify the target condition and the results were interpreted without knowledge of the results of the index test (PERCH 2019; Rose 2012; Ustero 2017). We considered 16 studies (84%) at unclear risk of bias because reference standard results may have been influenced by results of the index test. This was particularly a concern for studies assessing CXR against a composite reference standard (Birungi 2018; Clemente 2017; Dreesman 2017; Kruk 2008; LaCourse 2014; Schwoebel 2020; Tieu 2014; Togun 2016; Triasih 2015b), and, to a lesser extent, for studies assessing symptoms against a composite reference standard (Aggerbeck 2018; Birungi 2018; Dreesman 2017; Jaganath 2013; Kruk 2008; LaCourse 2014; Portevin 2014; Rose 2012; Sawry 2018; Schwoebel 2020; Tieu 2014; Togun 2015; Togun 2016; Triasih 2015a; Vonasek 2021 – several studies evaluated more than one index test). Regarding applicability, with respect to the reference standards, we considered all studies to have low concern.

In the flow and timing domain, we considered 14 studies (74%) at low risk of bias because there was an appropriate interval between the index test and reference standard, all children received the same reference standard, and all children were included in the analysis. We considered three studies (16%) at high risk of bias: for one study there was not an appropriate interval between the index test and reference standard, not all children received the same reference standard, and not all children were included in the analysis (Sawry 2018); for one study it was unclear if there was an appropriate interval between the index test and reference standard and not all children received the same reference standard (Ustero 2017); and for one study it was unclear if there was an appropriate interval between the index test and reference standard and not all children were included in the analysis (Vonasek 2021). We considered two studies (10%) at unclear risk of bias: for one study not all children received the same reference standard (PERCH 2019), and for one study it was unclear if there was an appropriate interval between the index test and reference standard (Triasih 2015b).

## Findings

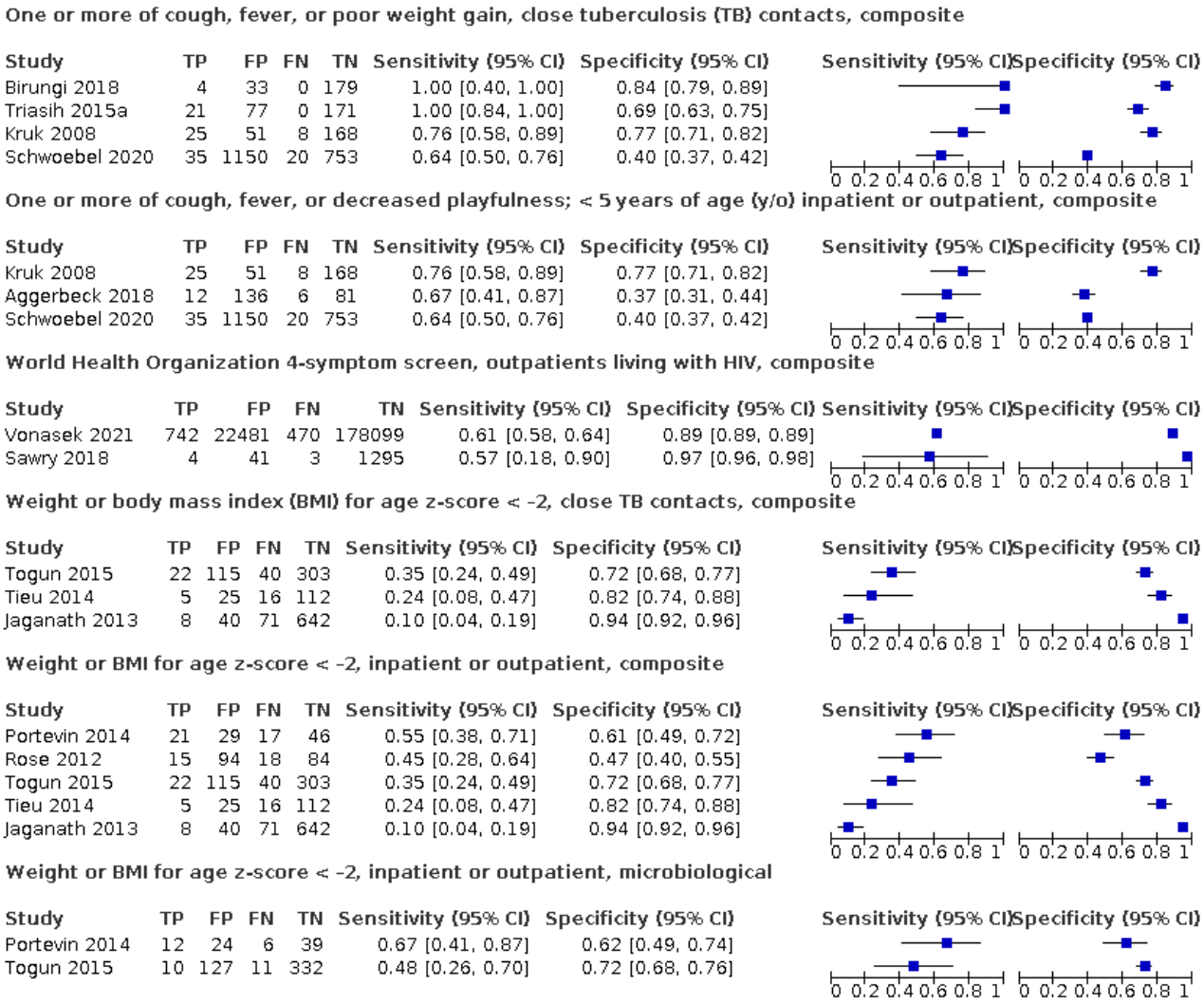
Of the 19 studies, 17 (89%) were conducted mainly or exclusively in low- or middle-income countries and two (11%) were conducted exclusively in high-income countries (Clemente 2017; Dreesman 2017). Two studies only assessed participants living with HIV (Sawry 2018; Vonasek 2021). Six studies did not report the HIV status of participants. One study excluded participants living with HIV (PERCH 2019). HIV prevalence in the remaining 10 studies ranged from 0% (Togun 2015) to 37% (Rose 2012). Fourteen studies were at least partially conducted in sub-Saharan African, four in Asia (PERCH 2019; Tieu 2014; Triasih 2015a; Triasih 2015b), and two in Europe (Clemente 2017; Dreesman 2017). Twelve studies were conducted at least partially in tuberculosis high-burden countries. Fifteen studies evaluated the accuracy of individual symptoms for tuberculosis screening. Twelve studies evaluated the accuracy of combinations of symptoms. Ten studies evaluated CXR. Two studies evaluated Xpert MTB/RIF in a screening context (LaCourse 2014; Togun 2015). Several studies assessed more than one screening test. Six studies (32%) reported results against a microbiological reference standard. Seventeen studies (89%) reported results against a composite reference standard. Table 1 presents a summary of key characteristics of the included studies. We presented details in the Characteristics of included studies table. Table 2 presents summary values of sensitivity and specificity for the following analyses.

### 1. Symptom screening for detection of pulmonary tuberculosis

#### *One or more of cough, fever, or poor weight gain in close tuberculosis contacts, against a composite reference standard*

We identified four studies that used a composite reference standard to estimate the accuracy of the symptom group cough, fever, or poor weight gain to screen for pulmonary tuberculosis in close tuberculosis contacts. Sensitivity estimates ranged from 64% to 100%. The two studies with the lowest sensitivity (64% and 76%) only included children under five years of age (Kruk 2008; Schwoebel 2020), possibly explaining differences in sensitivity given the frequency with which 'asymptomatic hilar adenopathy' may occur in this age group. Specificity estimates ranged from 40% to 84%. Three studies had specificity of 69% or higher (Birungi 2018; Kruk 2008; Triasih 2015a). The single study with notably lower specificity (40%) used a symptom screen that assessed the presence of symptoms over the past month (Schwoebel 2020), while the symptom screens of other studies were composed of more focused symptoms present during a shorter time period. This may explain differences in specificity. Pooled sensitivity was 89% (95% CI 52% to 98%) and pooled specificity was 69% (95% CI 51% to 83%) (4 studies, 2695 participants, 113 (4.2%) with tuberculosis) (Figure 6).

**Figure 6. Forest plots of symptom groups, the WHO four-symptom screen for people living with HIV, and nutrition status to screen for pulmonary tuberculosis by composite reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. The individual studies are ordered by decreasing sensitivity. BMI: body mass index; FN: false negative; FP: false positive; TN: true negative; TP: true positive.**




***One or more of cough, fever, or decreased playfulness in children under five years of age in inpatient or outpatient settings, against a composite reference standard***

We identified three studies that used a composite reference standard to estimate the accuracy of the symptom group cough, fever, or decreased playfulness to screen for pulmonary tuberculosis in children under five years of age in inpatient or outpatient settings (Aggerbeck 2018; Kruk 2008; Schwoebel 2020). Sensitivity estimates ranged from 64% to 76%. Specificity estimates were 37% and 77% (3 studies, 2445 participants, 106 (4.3%) with tuberculosis; Figure 6).

***One or more of cough, fever, poor weight gain, or close tuberculosis contact (WHO four-symptom screen) in children living with HIV in outpatient settings, against a composite reference standard***

We identified two studies that used a composite reference standard to estimate the accuracy of the WHO-recommended four-symptom screen (current cough, fever, poor weight gain, or close tuberculosis contact for children; current cough, weight loss, night sweats, or fever for adolescents) to screen for pulmonary tuberculosis in outpatients living with HIV at every clinical encounter (Sawry 2018; Vonasek 2021). Sensitivity estimates were 57% and 61%. Specificity estimates were 89% and 97%. The WHO four-symptom screen pooled sensitivity was 61% (95% CI 58% to 64%) and pooled specificity was 94% (95% CI 86% to 98%) (2 studies; 20,926 participants, 1219 (5.8%) with tuberculosis; 203,135 screens) (Figure 6).

***Undernutrition in close tuberculosis contacts, against a composite reference standard***

We identified three studies that used a composite reference standard to estimate the accuracy of undernutrition (cutoff of body mass index z-score or weight-for-age z-score of –2) to screen for pulmonary tuberculosis in close tuberculosis contacts (Jaganath 2013; Tieu 2014; Togun 2015). Sensitivity estimates ranged from 10% to 35%. Specificity estimates ranged from 72% to 94%. Undernutrition pooled sensitivity was 21% (95% CI 11% to 38%) and pooled specificity was 85% (95% CI 71% to 93%) (3 studies, 1399 participants, 162 (11.6%) with tuberculosis (Figure 6).

***Undernutrition in children in inpatient or outpatient settings, against a composite reference standard***

We identified five studies that used a composite reference standard to estimate the accuracy of undernutrition (cutoff of body mass index z-score or weight-for-age z-score of –2) to screen for pulmonary tuberculosis in children in inpatient or outpatient settings. Sensitivity estimates ranged from 10% to 55%. The two studies with the highest sensitivities included inpatients likely to have more severe disease (Portevin 2014; Rose 2012), while the other three studies were exclusively conducted in outpatient settings (Jaganath 2013; Tieu 2014; Togun 2015). This could partially explain differences in sensitivity (range 10% to 55%). Specificity estimates ranged from 47% to 94%. Undernutrition pooled sensitivity was 32% (95% CI 18% to 50%) and pooled specificity was 75% (95% CI 56% to 88%) (5 studies, 1723 participants, 233 (13.5%) with tuberculosis) (Figure 6).

***Undernutrition in children in inpatient or outpatient settings, against a microbiological reference standard***

We identified two studies that used a microbiological reference standard to estimate the accuracy of undernutrition (cutoff of body mass index z-score or weight-for-age z-score of –2) to screen for pulmonary tuberculosis in children in inpatient or outpatient settings (Portevin 2014; Togun 2015). Sensitivity estimates were 48% and 67%. Specificity estimates were 62% and 72% (2 studies, 561 participants, 39 (7.0%) with tuberculosis) (Figure 6).

We identified no studies that evaluated symptom screening for detection of pulmonary tuberculosis in children with pneumonia and children in the general population in high-tuberculosis burden settings.

**2. Chest radiography for screening of pulmonary tuberculosis**

Ten studies involving 7146 participants evaluated the accuracy of CXR to screen for pulmonary tuberculosis and included 260 (3.5%) participants with tuberculosis. The median number of participants in the studies was 249 (interquartile range 158 to 300). Table 3 presents details of how CXR was obtained, how results were interpreted, and threshold for positivity for these various studies.

***Abnormal chest radiography in close tuberculosis contacts, against a composite reference standard***

We identified eight studies that used a composite reference standard to estimate the accuracy of abnormal CXR to screen for pulmonary tuberculosis in close tuberculosis contacts (Birungi 2018; Clemente 2017; Dreesman 2017; Kruk 2008; Schwoebel 2020; Tieu 2014; Togun 2016; Triasih 2015b). Sensitivity estimates ranged from 52% to 100%, with only one study (Triasih 2015b) having a sensitivity below 78%. Specificity estimates ranged from 28% to 100%. In the meta-analysis, abnormal CXR pooled sensitivity was 87% (95% CI 75% to 93%) and pooled specificity was 99% (95% CI 68% to 100%) (8 studies, 3513 participants, 232 (6.6%) with tuberculosis) (Figure 7).

**Figure 7. Forest plots of chest radiography (CXR) to screen for pulmonary tuberculosis. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. The individual studies are ordered by decreasing sensitivity. FN: false negative; FP: false positive; TN: true negative; TP: true positive.**

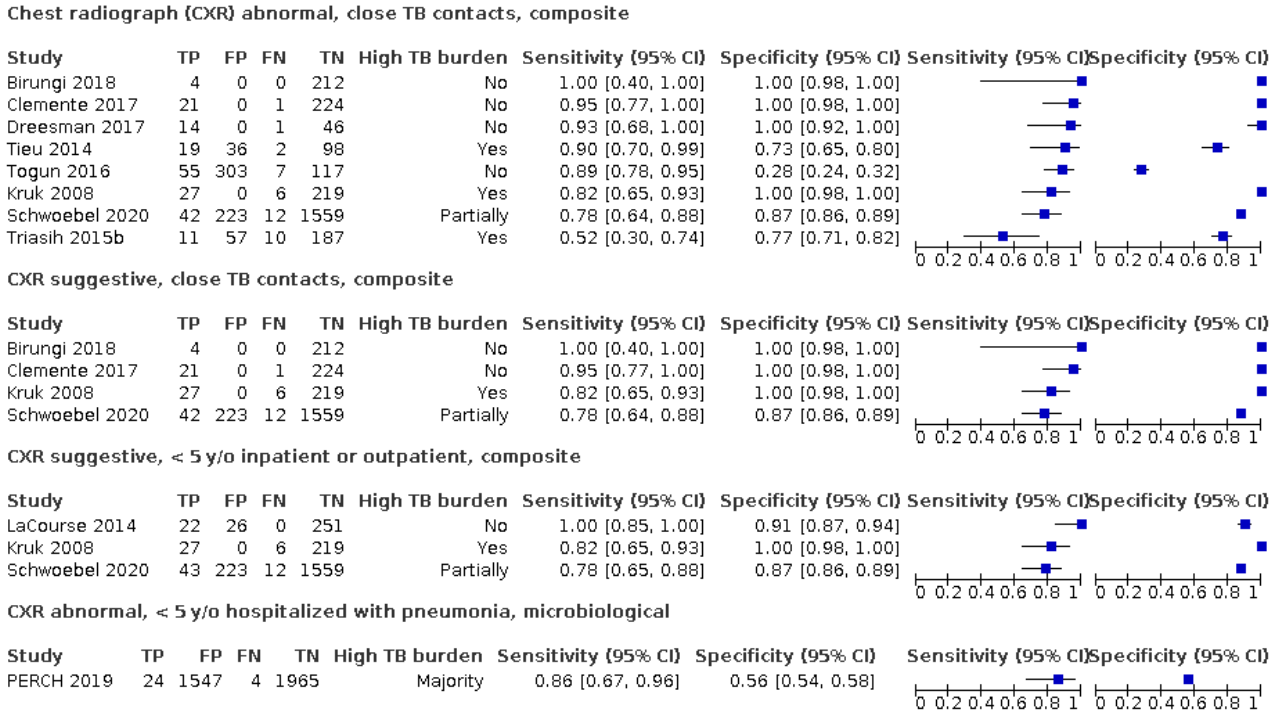
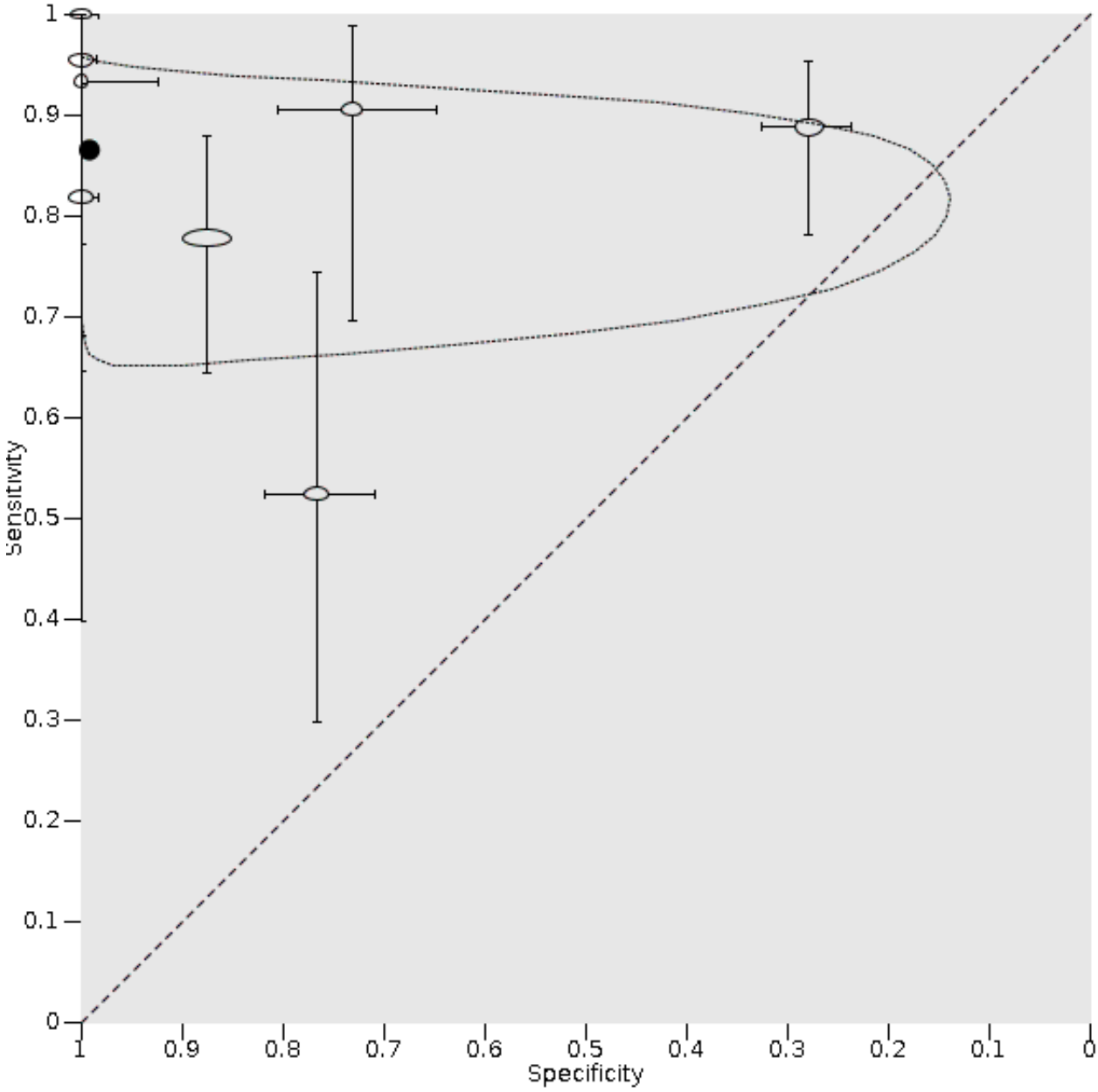



Figure 8 presents a summary plot of abnormal CXR sensitivity and specificity to screen for pulmonary tuberculosis in close tuberculosis contacts. The summary point (pooled value) appears close to the upper left-hand corner of the plot, suggesting high accuracy of this screening test. The 95% prediction region is relatively wide, displaying uncertainty as to where the likely values of sensitivity and specificity might occur in a future study.

**Figure 8. Summary plot of abnormal chest radiography sensitivity and specificity to screen for pulmonary tuberculosis in close tuberculosis contacts. Each individual study is represented by an empty oval. The size of the oval is proportional to the sample size of the study such that larger studies are represented by larger ovals. The dashed curves represent the 95% confidence region.**




***Suggestive chest radiography in close tuberculosis contacts, against a composite reference standard***

Four of the studies in the previous analysis used a composite reference standard to estimate the accuracy of CXR findings more specifically suggestive of tuberculosis (rather than an abnormal CXR more generally) to screen for pulmonary tuberculosis in close tuberculosis contacts. Sensitivity estimates ranged from 78% to 100%, though the estimate of 100% was from a study with only four cases of tuberculosis (Birungi 2018). Three studies had specificity estimates of 100%, though these studies together contributed less than 30% of these data (Birungi 2018; Clemente 2017; Kruk 2008). The largest study had a specificity estimate of 87% (Schwoebel 2020). For CXR suggestive of tuberculosis, pooled sensitivity was 84% (95% CI 70% to 92%) and pooled specificity was 91% (95% CI 90% to 92%) (4 studies, 2550 participants, 113 (4.4%) with tuberculosis) (Figure 7).

***Suggestive chest radiography in children under five years of age in inpatient or outpatient settings, against a composite reference standard***

We identified three studies that used a composite reference standard to estimate the accuracy of suggestive CXR findings to screen for pulmonary tuberculosis in children under five years of age in inpatient or outpatient settings (Kruk 2008; LaCourse 2014; Schwoebel 2020). Two of these studies were also included in the previous analysis as they were conducted with populations and in settings relevant to both analyses (Kruk 2008; Schwoebel 2020). Sensitivity estimates ranged from 78% to 100%. Specificity estimates ranged from 87% to 100%. The largest of these studies, contributing 77% of these data, notably had the lowest sensitivity and specificity estimates (Schwoebel 2020). In the meta-analysis, CXR suggestive of tuberculosis pooled sensitivity was 87% (95% CI 66% to 96%) and pooled specificity was 89% (95% CI 88% to 90%) (3 studies, 2388 participants, 110 (4.6%) with tuberculosis) (Figure 7).

***Abnormal chest radiography in children under five years of age with pneumonia in inpatient settings, against a microbiological reference standard***

We identified one study with participants from seven countries (3540 children in total, 28 (0.8%) with tuberculosis) that used a microbiological reference standard to estimate the accuracy of abnormal CXR findings to screen for pulmonary tuberculosis in children under five years of age hospitalized with pneumonia (PERCH 2019). Sensitivity was 86% (95% CI 67% to 96%) and specificity was 56% (95% CI 54% to 58%) (Figure 7).

We identified no studies that evaluated CXR for screening of pulmonary tuberculosis in children living with HIV and children in the general population in high-tuberculosis burden settings.

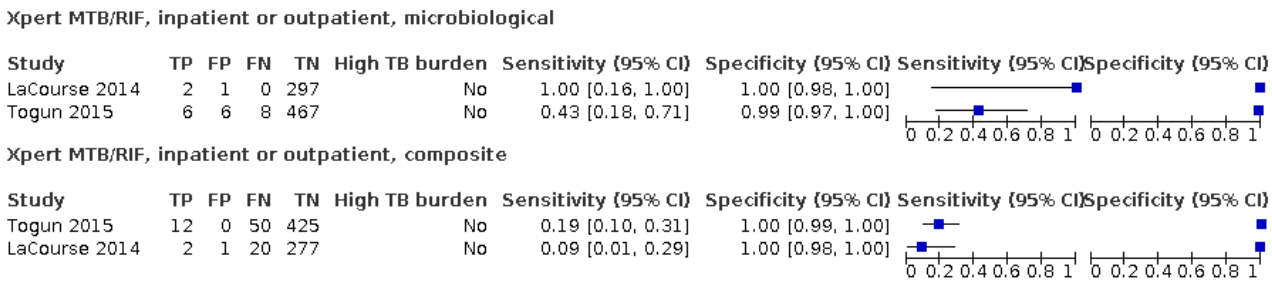
**3. Xpert MTB/RIF for screening of pulmonary tuberculosis**

Two studies involving 787 participants (300 from LaCourse 2014, 487 from Togun 2015) evaluated the accuracy of Xpert MTB/RIF to screen for pulmonary tuberculosis.

***Children in inpatient or outpatient settings, against a microbiological reference standard***

For the two studies, against a microbiological reference standard, sensitivity estimates were 43% and 100%. Of note, these estimates were derived from only two and 14 tuberculosis cases in each study. Specificity estimates were 99% and 100%. These two studies notably selected participants from different populations, with LaCourse 2014 enrolling children under five years of age hospitalized with severe acute malnutrition and Togun 2015 enrolling tuberculosis household contacts under 15 years of age in an outpatient setting (Figure 9).

**Figure 9. Forest plots of Xpert MTB/RIF sensitivity and specificity to screen for pulmonary tuberculosis by reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false negative; FP: false positive; TN: true negative; TP: true positive.**




***Children in inpatient or outpatient settings, against a composite reference standard***

These two studies also used a composite reference standard to estimate the accuracy of Xpert MTB/RIF to screen for pulmonary tuberculosis in children in inpatient or outpatient settings (LaCourse 2014; Togun 2015). Sensitivity estimates were 9% and 19%. Specificity estimates were both 100% (Figure 9).

We did not identify any studies that evaluated Xpert MTB/RIF for screening of pulmonary tuberculosis in children living with HIV, children with pneumonia, and children in the general population in high tuberculosis burden settings.

## DISCUSSION

This systematic review summarized the current literature and included 19 unique studies that estimated the accuracy of symptoms, CXR, and Xpert MTB/RIF to screen for active pulmonary tuberculosis in children.

### Summary of main results

**Symptom-based screening for pulmonary tuberculosis**

- In close tuberculosis contacts, against a composite reference standard, one or more of cough, fever, or poor weight gain pooled sensitivity was 89% (95% CI 52% to 98%) and pooled specificity was 69% (95% CI 51% to 83%) (4 studies, 2695 participants).
- In children under five years of age in inpatient or outpatient settings, against a composite reference standard, one or more of cough, fever, or decreased playfulness sensitivity range was 64% to 76% and specificity range was 37% to 77% (3 studies, 2445 participants).
- In children living with HIV in outpatient settings, against a composite reference standard, one or more of cough, fever, poor

weight gain, or close tuberculosis contact (WHO four-symptom screen), done at each healthcare visit, pooled sensitivity was 61% (95% CI 58% to 64%) and pooled specificity was 94% (95% CI 86% to 98%) (2 studies; 20,926 participants).

- For any setting or population, against a composite reference standard, undernutrition pooled sensitivity was 32% (95% CI 18% to 50%) and pooled specificity was 75% (95% CI 56% to 88%) (5 studies, 1723 participants).
- In close tuberculosis contacts, against a composite reference standard, undernutrition pooled sensitivity was 21% (95% CI 11% to 38%) and pooled specificity was 85% (95% CI 71% to 93%) (3 studies, 1399 participants).
- In children in inpatient or outpatient settings, against a microbiological reference standard, undernutrition sensitivities were 48% and 67% and specificities were 62% and 72% (2 studies, 561 participants).

**Chest radiography screening for pulmonary tuberculosis**

- In close tuberculosis contacts, against a composite reference standard, abnormal CXR pooled sensitivity was 87% (95% CI 75% to 93%) and pooled specificity was 99% (95% CI 68% to 100%) (8 studies, 3513 participants).
- In close tuberculosis contacts, against a composite reference standard, CXR suggestive of tuberculosis pooled sensitivity was 84% (95% CI 70% to 92%) and pooled specificity was 91% (95% CI 90% to 92%) (4 studies, 2550 participants).
- In children under five years of age in inpatient or outpatient settings, against a composite reference standard, CXR suggestive of tuberculosis pooled sensitivity was 87% (95% CI 66% to 96%) and pooled specificity was 89% (95% CI 88% to 90%) (3 studies, 2388 participants).
- In children under five years of age hospitalized with pneumonia, against a microbiological reference standard, abnormal CXR sensitivity was 86% (95% CI 67% to 96%) and specificity was 56% (95% CI 54% to 58%) (1 study, 3540 participants).

**Xpert MTB/RIF screening for pulmonary tuberculosis**

- In children in inpatient or outpatient settings, against a microbiological reference standard, Xpert MTB/RIF sensitivities were 43% and 100% and specificities were 99% and 100% (2 studies, 787 participants).
- In children in inpatient or outpatient settings, against a composite reference standard, Xpert MTB/RIF sensitivities were 9% and 19% and specificities were both 100% (2 studies, 787 participants).

**Illustration of findings in a hypothetical population of 1000 children with 5% prevalence of tuberculosis**

***One or more of cough, fever, or poor weight gain for screening of pulmonary tuberculosis in tuberculosis close contacts***

If 50 of the 1000 children have pulmonary tuberculosis by a composite reference standard, 339 would have cough, fever, or poor weight gain, 294 (87%) of whom would not have tuberculosis (false positives); 661 would not have cough, fever, or poor weight gain, 5 (1%) of whom would have tuberculosis (false negatives) (Summary of findings 1).

***One or more of cough, fever, or decreased playfulness for screening of pulmonary tuberculosis in children under five years of age in inpatient or outpatient settings***

If 50 of the 1000 children have pulmonary tuberculosis by a composite reference standard, 251 to 636 would have cough, fever, or decreased playfulness, 219 to 598 (87% to 94%) of whom would not have tuberculosis (false positives); 364 to 749 would not have cough, fever, or decreased playfulness, 12 to 18 (2% to 3%) of whom would have tuberculosis (false negatives) (Summary of findings 1).

***One or more of cough, fever, poor weight gain, or tuberculosis close contact (WHO four-symptom screen) for pulmonary tuberculosis in outpatients living with HIV at every healthcare visit***

If 50 of 1000 WHO four-symptom screens are on children with pulmonary tuberculosis by a composite reference standard, 88 symptom screens would be positive, 57 (65%) of which would be on children who do not have tuberculosis (false positives); 912 symptom screens would be negative, 19 (2%) of which would be on children who have tuberculosis (false negatives) (Summary of findings 1).

***Abnormal chest radiography for screening of pulmonary tuberculosis in tuberculosis close contacts***

If 50 of the 1000 children have pulmonary tuberculosis by a composite reference standard, 63 would have abnormal CXR, 19 (30%) of whom would not have tuberculosis (false positives); 937 would not have abnormal CXR, 6 (1%) of whom would have tuberculosis (false negatives) (Summary of findings 2).

***Xpert MTB/RIF for screening of pulmonary tuberculosis in children in inpatient or outpatient settings***

If 50 of the 1000 children have pulmonary tuberculosis by a microbiological reference standard, 31 to 69 would be Xpert MTB/RIF-positive, 9 to 19 (28 to 29%) of whom would not have tuberculosis (false positives); 969 to 931 would be Xpert MTB/RIF-negative, 0 to 28 (0 to 3%) of whom would have tuberculosis (false negatives) (Summary of findings 3).

**Symptom-based screening for pulmonary tuberculosis**

Symptom-based screening for tuberculosis has the obvious advantages of not requiring any materials other than a careful interviewer and providing instant results. However, symptoms of childhood tuberculosis, particularly in young children, tend to overlap with symptoms of common childhood conditions and to be non-specific, especially if poorly defined (Marais 2005a; Marais 2005b). Therefore, symptom-based screening is most likely to be beneficial when targeted to high-risk groups.

We reported a meta-analysis of the symptom group 'cough, fever, or poor weight gain' in close tuberculosis contacts. While pooled sensitivity was 89% (95% CI 52% to 98%), specificity was lower (69%, 95% CI 51% to 83%) with this approach tending to have more false-positive screens due to multiple symptoms lowering the threshold for positive screening. Composed symptom screens such as 'cough, fever, or poor weight gain' may lack specificity, but this may be tolerable in contexts where the consequences of false-positive screening are less of a concern. For example, in settings where resources are less constrained and the costs of unnecessary diagnostic work-up are relatively tolerable. Different combinations

of symptoms for composed symptom screens may offer better accuracy for high-risk groups; this is an area in need of research.

For people living with HIV or children in close contact with a tuberculosis case, tuberculosis preventive treatment is highly effective at reducing the risk of developing tuberculosis disease. Tuberculosis preventive treatment is recommended for children in these high-risk groups after tuberculosis disease has been excluded. Screening strategies can dictate who is eligible for preventive treatment. Those screening negative for tuberculosis disease can be considered for preventive treatment, while those screening positive must complete additional diagnostic work-up for tuberculosis disease while preventive treatment is withheld (WHO Consolidated Guidelines (Module 1) 2020). These screening strategies must maximize sensitivity so that false-negative results, with consequent provision of tuberculosis preventive treatment to someone who has tuberculosis disease is prevented. Having a rapid and easy-to-perform screening test should assist preventive treatment reaching more children who could potentially benefit.

Another consideration relevant to screening children eligible for tuberculosis preventive treatment is the severity of tuberculosis disease. Young children with tuberculosis tend to have paucibacillary disease, (often manifesting as asymptomatic hilar adenopathy), which would result in a false-negative screen in this review given the positive CXR. However, concerns regarding risk of transmission and creating drug resistance with one- or two-drug preventive treatment regimens are minimal for those with asymptomatic disease who initiate preventive treatment in this age group (Marais 2009a). Therefore, lower sensitivity may be tolerable if those with false-negative screens mostly have mild, paucibacillary disease and start preventive treatment while those with severe disease are appropriately captured by the screen. Some consideration also needs to be made for screening specificity when it dictates eligibility for preventive treatment. Lower specificity screens would result in more missed or delayed opportunities to initiate preventive treatment due to false-positive screening. Hence, although the symptom group cough, fever, or poor weight gain in close tuberculosis contacts had high sensitivity at 89%, the limited specificity at 69% is an important consideration.

It should be noted that we evaluated these symptom groups in studies that also considered additional symptoms within study-specific symptom screening strategies. Therefore, these additional symptoms may have captured cases of tuberculosis disease when the three symptoms from the group evaluated were not present; in turn, this would lead to enumeration of a true-positive result and possibly inflating sensitivity estimates. Conversely, study-specific symptom screening strategies may have utilized different duration of symptoms for positivity (e.g. 'current cough' versus 'cough greater than two weeks'), and these differing screening definitions may limit applicability of the findings to other settings. Hence, we advise interpreting the symptom group results with caution and emphasize that the 'cough, fever, or poor weight gain' symptom group screen in close tuberculosis contacts requires further investigation as a potentially high sensitivity screening strategy.

We assessed the WHO-recommended symptom screen for people living with HIV (where it is essentially a three-symptom screen accompanied by a question on recent tuberculosis exposure) in children and adolescents presenting to outpatient settings. Against a composite reference standard, as described above, the specificity of 94% is important for this population eligible for tuberculosis preventive treatment. The limited sensitivity (61%) is a concern given that children living with HIV tend to have more rapidly progressive tuberculosis disease (similar to very young children) and tuberculosis preventive treatment could potentially be given to someone with tuberculosis disease. However, synthesis of these data should be considered within the context of this screening strategy that is recommended to be performed serially at every clinical encounter. Therefore, the deleterious effects of an inaccurate screen at a single clinical encounter may be minimized by accurate screening results of the same child in the near future. This highlights the importance of ongoing screening for tuberculosis disease while preventive treatment is being administered.

**Chest radiography screening for pulmonary tuberculosis**

In the absence of a microbiological diagnosis, the diagnosis of pulmonary childhood tuberculosis is heavily influenced by chest imaging, when locally available. This reliance on chest imaging remains the reality in the clinical setting, despite evidence showing limited accuracy of CXR for detecting lesions suggestive of childhood tuberculosis, such as mediastinal lymphadenopathy, compared to computed tomography (Swingler 2005), and poor inter-reader agreement for CXR findings suggestive of tuberculosis (Du Toit 2002; Kaguthi 2014; Swingler 2005). Of the 10 studies assessing CXR, only one study reported inter-reader agreement (Triasih 2015b). Six studies required agreement between at least two interpreters to define positive CXR (Birungi 2018; Kruk 2008; LaCourse 2014; PERCH 2019; Togun 2016; Triasih 2015b), and of the five studies that reported level of training of interpreters, all were trained physicians at minimum (Birungi 2018; PERCH 2019; Schwoebel 2020; Togun 2016; Triasih 2015b). Characteristics of how CXR was obtained and interpreted were not reported in enough included studies to allow for analysis of how these factors influence test accuracy.

Although our data suggest that a lower threshold for positivity ('any abnormality' rather than 'abnormality suggestive of tuberculosis') may give more accurate screening results when applied to close tuberculosis contacts, we noted that these estimates were imprecise (with greatly overlapping 95% CIs) and direct comparisons of the accuracy of these two thresholds were invalid given that they were reported from different studies. Also, study-specific threshold for positivity varied between studies categorized as having thresholds of 'any abnormality' or 'abnormality suggestive of tuberculosis,' and this further complicated comparisons between these thresholds.

Against a composite reference standard, we found that CXR with 'any abnormality' in close tuberculosis contacts had a sensitivity of 87% and a specificity of 99%. We found similarly high accuracy for CXR suggestive of tuberculosis against a composite reference standard in children under five years of age in inpatient or outpatient settings. One systematic review of CXR with any abnormality in the general population of adults, against a microbiological reference standard, reported pooled estimates of sensitivity from three studies of 98% (95% CI 95% to 100%) and specificity of 75% (95% CI 72% to 79%) (van't Hoog 2013). Since it is theorized that comparison against a microbiological reference standard overestimates sensitivity and underestimates specificity of the index test (Drain 2019), we considered estimates of CXR in children against a composite reference standard to be fairly

consistent with these estimates for adults. Indeed, we reported one large, multi-country study with data for abnormal CXR in children under the age of five years against a microbiological reference standard; of note, sensitivity was similarly high at 86% while specificity was much lower at 56% (PERCH 2019). PERCH 2019 evaluated children with severe pneumonia, a population very likely to have CXR abnormalities due to pathology other than tuberculosis, and this also explains the low specificity in this study. As detailed above, in situations where screening would dictate who should be considered for tuberculosis preventive treatment, high sensitivity is the key criterion to guide preventive treatment while low specificity, to a degree, can be tolerable if supplemented by additional testing. Nevertheless, we interpreted the accuracy estimates reported here for CXR against a composite reference standard with caution given the concerns of bias and imprecision (see Strengths and weaknesses of the review, Accuracy of the reference standards used). These findings of high accuracy for the high-risk group of close tuberculosis contacts suggest that this is a promising screening strategy requiring further investigation. We identified limited CXR data for high-risk groups such as malnourished children and the general population in high tuberculosis burden settings.

### Xpert MTB/RIF screening for pulmonary tuberculosis

We found that Xpert MTB/RIF had screening sensitivity of 43% and 100% and specificity of 99% and 100%. Although the sensitivity estimates are based on very few tuberculosis cases, these findings are similar to reported accuracy estimates for Xpert MTB/RIF in paediatric diagnostic studies. Diagnostic studies apply this test to children with presumptive tuberculosis, a context in which the accuracy of Xpert MTB/RIF has been much more robustly evaluated as opposed to a screening context as in this review. One Cochrane Review of the diagnostic accuracy of Xpert MTB/RIF for childhood tuberculosis against a microbiological reference standard reported sensitivities ranging from 45.7% to 73.0% for various specimen types and specificity of over 98% for all specimen types considered (Kay 2020). Given the high specificity but likely limited sensitivity of this assay, Xpert MTB/RIF may have an important role as an early 'rule-in' strategy for tuberculosis case finding in high-risk groups. Although the evidence suggests that Xpert MTB/RIF should not be implemented as a stand-alone screening strategy, the strengths of this test (high specificity and relatively fast results in contexts with adequate resources) may be leveraged with it as one component of a larger screening approach. As resources allow, Xpert MTB/RIF could be used broadly to screen high-risk groups of children so that those with positive results are quickly started on treatment and those with negative results are more carefully evaluated with more sensitive strategies. However, in many high tuberculosis-burden settings, resource limitations require more judicious use of Xpert MTB/RIF testing so it is more appropriately used later in case finding algorithms, after less resource-intensive strategies, such as symptom screening, have been employed. We did not assess combination screening strategies in this review, but high-quality studies evaluating different combinations and sequences of screening tests for high-risk populations are urgently needed to improve childhood tuberculosis case finding.

## Strengths and weaknesses of the review

### Completeness of evidence

We performed comprehensive searches of numerous databases, handsearching references of included studies, and contacting experts in the field of paediatric tuberculosis for additional evidence. Non-English studies were included in the search and assessed for inclusion in this review. Despite the exhaustive approach, we acknowledge that some relevant studies may have been missed. There was a relatively high number of missing full texts (65 of 610 full-text studies sought; 11%) of studies that were, therefore, not fully assessed for inclusion. This was mostly an issue for older studies as 71% of those with missing full texts were published before 1980. Given that knowledge and techniques for diagnosis of tuberculosis in children has changed substantially over the past few decades, this issue of missing pre-1980 full texts is less of a concern.

### Accuracy of the reference standards used

We used two reference standards in this review, microbiological and composite. We do not consider either of these to be superior, as each has its respective limitations for detecting pulmonary tuberculosis in children. Due to the paucibacillary nature of childhood tuberculosis, the lower detection limit of existing microbiological reference standards may be too high to capture a significant proportion of cases; thus, comparison against a microbiological reference standard may potentially overestimate the sensitivity and underestimate the specificity of the index test (Drain 2019). Another consideration for comparisons against the microbiological reference standard is variation between number of specimens tested for a particular individual, with multiple specimen testing likely increasing the yield of the reference standard (Cruz 2012; Zar 2012), and thereby influencing accuracy estimates of the index test. Accuracy of microbiological testing for pulmonary childhood tuberculosis also varies by the type of specimen collected, with invasively collected specimens, such as gastric aspirates, typically more accurate in younger children (Dunn 2016; Kay 2020). Given the limited data available for this review against a microbiological reference standard, we did not investigate the number of specimens tested or type of specimen.

The composite reference standard may overdiagnose tuberculosis; in turn, this may underestimate sensitivity and overestimate specificity of the index test (Drain 2019). We defined the composite reference standard as microbiological confirmation or author-defined clinical pulmonary tuberculosis, with a requirement that any clinical diagnosis have a follow-up visit to help verify the diagnosis. Hence, clinical characteristics and component tests in the composite reference standard differed across studies; these differences may have contributed to variation in accuracy estimates. Incorporation bias was a particular concern when symptoms or CXR were the index tests compared against the composite reference standard. Symptoms and CXR are inherent components of a clinical diagnosis of tuberculosis; if agreement between the index test and the reference standard increases, accuracy will be overestimated due to incorporation bias. Although there is limited evidence that incorporation bias significantly alters accuracy estimates in diagnostic accuracy studies (Rutjes 2006; Whiting 2013), this is potentially a much larger problem in paediatric research where an independent reference standard is much more difficult to achieve.

### Methodological and reporting quality of the included studies

Using QUADAS-2, we considered risk of bias to be low for the patient selection and index test domains. Risk of bias was unclear for the reference standard domain largely due to concerns for incorporation bias with the composite reference standard, but with respect to the microbiological reference standard, risk of bias was low. Risk of bias for the flow and timing domain was low for 14 (74%) studies but high for six studies because of unclear timing between the index test and reference standard. The included studies were generally well reported. For a few of the studies where extraction of the data was not clear, we corresponded with the primary study authors to ensure appropriate data extraction. Overall, the studies had low risk of bias and were well reported.

### Comparison with other systematic reviews

We are not aware of other systematic reviews assessing the accuracy of symptom screening, CXR, or Xpert MTB/RIF for screening of childhood tuberculosis. One systematic review of symptom screening and CXR for tuberculosis in adults was discussed above (see 'Chest radiography screening for pulmonary tuberculosis;' van't Hoog 2013). A Cochrane Review of Xpert MTB/RIF and Xpert Ultra for pulmonary tuberculosis in adults irrespective of signs or symptoms of pulmonary tuberculosis is similar to our assessment of Xpert MTB/RIF as a screening strategy (Shapiro 2021). However, especially for younger children, comparisons between paediatric and adult pulmonary tuberculosis are challenging because the diseases are so different. One Cochrane Review of Xpert MTB/RIF diagnostic accuracy for childhood tuberculosis reported accuracy estimates that were similar to the estimates in this review for Xpert MTB/RIF used in a screening context (Kay 2020). However, our findings for Xpert MTB/RIF accuracy were limited as the numbers of studies and participants enrolled were small.

## Applicability of findings to the review question

To assess the applicability of findings to the review question, we considered QUADAS-2 domains for patient selection, index test, and reference standard. With respect to the patient selection domain, we considered most studies to have low concern about applicability. A few smaller studies had unclear or high concern about applicability in the patient selection domain due to enrolment criteria which implied a diagnostic, rather than screening, application. It should be acknowledged that there is a spectrum between screening and diagnosis rather than a clear distinction. Many studies were excluded from this review because there was consensus among the review authors that they were diagnostic studies. All included studies were determined by the review authors to have applied the index tests in a 'screening context,' as defined in the Background under 'Screening,' although many studies included here may not be considered 'screening' under stricter definitions of the term. With respect to the index test and reference standard domains, all studies had low concern about applicability.

# AUTHORS' CONCLUSIONS

## Implications for practice

We found that in children who are tuberculosis contacts or living with HIV, screening tests using symptoms or chest radiography may be useful; however, both sensitivity and specificity estimates are likely to be overestimated owing to incorporation bias. In close tuberculosis contacts, the symptom screen including one or more of cough, fever, or poor weight gain misses around 10% of children who have tuberculosis at the initial screen; however, these asymptomatic 'cases' are likely to have paucibacillary disease for which tuberculosis preventive treatment may be curative and the risk of inducing drug resistance is minimal. Single use of the World Health Organization (WHO)-recommended 'four-symptom' screen in children living with HIV had limited sensitivity, which is concerning given their risk of rapid disease progression. Repeated use of symptom screening at regular clinical encounters should improve 'cumulative sensitivity' among children living with HIV, but this was not assessed in this review. Chest radiography (any abnormality) seems to be the most accurate screening test for pulmonary tuberculosis in children but is influenced by radiograph quality and inter-reader variability, as well as potential overestimation of both sensitivity and specificity given inclusion bias. Xpert MTB/RIF demonstrates high specificity, though evaluation of sensitivity is limited by few studies and few children with tuberculosis.

## Implications for research

Research to identify accurate and practical screening tests for pulmonary tuberculosis in children remains an urgent need. A major limitation of most studies to date has been the absence of a consistent and objective reference standard. Further, studies assessing the accuracy of screening tests should use both microbiological and composite reference standards and avoid incorporation bias. Although these reference standards have limitations, their combined use provides added value. In addition, to foster robust accuracy estimates, study participants should be tested with a reference standard for tuberculosis regardless of their screening test result being positive or negative. Comparison of studies that assess different screening tests and strategies in the same population should be conducted.

Studies assessing symptom screening tests need to consider the intended use of the test with prioritization of high sensitivity, particularly in high-risk groups. In close tuberculosis contacts, additional studies assessing the utility of simple symptom screening strategies are needed; these studies must use clear and consistent symptom definitions. For the WHO-recommended 'four-symptom' screen for children living with HIV, future studies should assess the added value of serial screening, ideally completed once every one to three months as part of routine clinical care. As accuracy may differ, more data are also needed in children living with HIV who are naive to antiretroviral therapy or those with advanced HIV.

Studies assessing chest radiography screening should consider microbiological testing of multiple and different specimens and clinical follow-up to strengthen the definition of a reference standard without concerns about incorporation bias. Given the promising results for chest radiography screening in tuberculosis close contacts, evaluation of this screening test in other high-risk groups (e.g. malnourished children, children living with HIV) should be a priority.

Xpert MTB/RIF and newer rapid molecular diagnostics (e.g. Xpert Ultra) that have lower limits of bacilli detection are potentially powerful screening tools for high-risk groups. Large, prospective, well-designed studies are needed to assess their screening

accuracy in children living with HIV; close tuberculosis contacts; or children with malnutrition, unremitting cough, and pneumonia. These studies should additionally compare the accuracy of various sampling techniques, such as stool or oral swabs, given the need for optimal feasibility and acceptability. Finally, studies should ideally assess fresh specimens obtained within routine clinical settings.

Assessment of feasibility and cost effectiveness are important to inform implementation strategies, especially in resource-limited settings where chest radiography or rapid molecular tests may not be readily available. Improved screening is paramount if we are to increase tuberculosis preventive treatment in children in high-risk groups without disease and decrease treatment delays in children with disease. New strategies should ideally be rapid, inexpensive, feasible, and acceptable to children and their caregivers.

## ACKNOWLEDGEMENTS

The Cochrane Infectious Diseases Group (CIDG) Academic Editor is Dr Michael Eisenhut, and the Diagnostic Test Accurary (DTA) Contact Editor is Dr Clare Davenport.

We are grateful to Vittoria Lutje, CIDG Information Specialist, for help with the search strategy. The CIDG editorial base 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.

Bryan Vonasek is supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under Award Number T32AI055397. The content is solely the responsibility of the review authors and does not necessarily represent the official views of the National Institutes of Health.

Susanna S van Wyk is supported by the Research, Evidence and Development Initiative (READ-It). READ-It (project number 300342-104) is funded by UK aid from the UK government; however, the views expressed do not necessarily reflect the UK government's official policies.

# REFERENCES

## References to studies included in this review

**Aggerbeck 2018** *{published data only}*

Aggerbeck H, Ruhwald M, Hoff ST, Borregaard B, Hellstrom E, Malahleha M, et al. C-Tb skin test to diagnose Mycobacterium tuberculosis infection in children and HIV-infected adults: a phase 3 trial. *PLOS One* 2018;**13**(9):e0204554.

**Birungi 2018** *{published data only}*

Birungi FM, van Wyk B, Uwimana J, Ntaganira J, Graham SM. Xpert MTB/RIF assay did not improve diagnosis of pulmonary tuberculosis among child contacts in Rwanda. *Pan African Medical Journal* 2018;**30**:39.

**Clemente 2017** *{published data only}*

Clemente MG, Dore E, Abis L, Molicotti P, Zanetti S, Olmeo P, et al. Pediatric tuberculosis in northern Sardinia. *Mediterranean Journal of Hematology and Infectious Diseases* 2017;**9**(1):e2017027.

**Dreesman 2017** *{published data only}*

Dreesman A, Corbiere V, Dirix V, Smits K, Debulpaep S, De Schutter I, et al. Age-stratified T cell responses in children infected with Mycobacterium tuberculosis. *Frontiers in Immunology* 2017;**8**:1059.

**Jaganath 2013** *{published data only}*

Jaganath D, Zalwango S, Okware B, Nsereko M, Kisingo H, Malone L, et al. Contact investigation for active tuberculosis among child contacts in Uganda. *Clinical Infectious Diseases* 2013;**57**(12):1685-92.

**Kruk 2008** *{published data only}*

Kruk A, Gie RP, Schaaf HS, Marais BJ. Symptom-based screening of child tuberculosis contacts: improved feasibility in resource-limited settings. *Pediatrics* 2008;**121**(6):e1646-52.

**LaCourse 2014** *{published data only}*

LaCourse SM, Chester FM, Preidis G, McCrary LM, Arscott-Mills T, Maliwichi M, et al. Use of Xpert for the diagnosis of pulmonary tuberculosis in severely malnourished hospitalized Malawian children. *Pediatric Infectious Disease Journal* 2014;**33**(11):1200-2.

**PERCH 2019** *{published data only}*

Pneumonia Etiology Research for Child Health (PERCH) Study Group. Causes of severe pneumonia requiring hospital admission in children without HIV infection from Africa and Asia: the PERCH multi-country case-control study. *Lancet* 2019;**394**(10200):757-79.

**Portevin 2014** *{published data only}*

Portevin D, Moukambi F, Clowes P, Bauer A, Chachage M, Ntinginya NE, et al. Assessment of the novel T-cell activation marker-tuberculosis assay for diagnosis of active tuberculosis in children: a prospective proof-of-concept study. *Lancet Infectious Diseases* 2014;**14**(10):931-8.

**Rose 2012** *{published data only}*

Rose MV, Kimaro G, Nissen TN, Kroidl I, Hoelscher M, Bygbjerg IC, et al. QuantiFERON-TB gold in-tube performance for diagnosing active tuberculosis in children and adults in a high burden setting. *PLOS One* 2012;**7**(7):e37851.

**Sawry 2018** *{published data only}*

Sawry S, Moultrie H, Van Rie A. Evaluation of the intensified tuberculosis case finding guidelines for children living with HIV. *International Journal of Tuberculosis and Lung Disease* 2018;**22**(11):1322-8.

**Schwoebel 2020** *{published data only}*

Schwoebel V, Koura KG, Adjobimey M, Gnanou S, Wandji AG, Gody J-C, et al. Tuberculosis contact investigation and short-course preventive therapy among young children in Africa. *International Journal of Tuberculosis and Lung Disease* 2020;**24**(4):454-62.

**Tieu 2014** *{published data only}*

Tieu HV, Suntarattiwong P, Puthanakit T, Chotpitayasunondh T, Chokephaibulkit K, Sirivichayakul S, et al. Comparing interferon-gamma release assays to tuberculin skin test in Thai children with tuberculosis exposure. *PLOS One* 2014;**9**(8):e105003.

**Togun 2015** *{published data only}*

Togun TO, Egere U, Sillah AK, Ayorinde A, Mendy F, Tientcheu L, et al. Contribution of Xpert(R) MTB/RIF to the diagnosis of pulmonary tuberculosis among TB-exposed children in The Gambia. *International Journal of Tuberculosis and Lung Disease* 2015;**19**(9):1091-7, i-ii12-18.

**Togun 2016** *{published data only}*

Togun TO, Egere U, Gomez MP, Sillah AK, Daramy M, Tientcheu LD, et al. No added value of interferon-gamma release to a prediction model for childhood tuberculosis. *European Respiratory Journal* 2016;**47**(1):223-32.

**Triasih 2015a** *{published data only}*

Triasih R, Robertson CF, Duke T, Graham SM. A prospective evaluation of the symptom-based screening approach to the management of children who are contacts of tuberculosis cases. *Clinical Infectious Diseases* 2015;**60**(1):12-8.

**Triasih 2015b** *{published data only}*

Triasih R, de Campo J, Duke T, Choridah L, Graham SM. An evaluation of chest X-ray in the context of community-based screening of child tuberculosis contacts. *International Journal of Tuberculosis and Lung Disease* 2015;**19**(12):1428-34.

**Ustero 2017** *{published data only}*

Ustero PA, Kay AW, Ngo K, Golin R, Tsabedze B, Mzileni B, et al. School and household tuberculosis contact investigations in Swaziland: active TB case finding in a high HIV/TB burden setting. *PLOS One* 2017;**12**(6):e0178873.

**Vonasek 2021** *{published data only}*

Vonasek B, Kay A, Devezin T, Bacha JM, Kazembe P, Dhillon D, et al. Tuberculosis symptom screening for children and adolescents living with HIV in six high HIV/TB burden countries in Africa. *AIDS* 2021;**35**(1):73-9.

## References to studies excluded from this review

**Ackerman 2010** *{published data only}*

Ackerman A, Whalen C, Zalwango S, Shwartz JI. Clinical manifestations of tuberculosis among pediatric household contacts with active culture confirmed disease. *International Journal of Infectious Diseases* 2010;**14**:e146-7.

**Aldridge 2016** *{published data only}*

Aldridge RW, Zenner D, White PJ, Williamson EJ, Muzyamba MC, Dhavan P, et al. Tuberculosis in migrants moving from high-incidence to low-incidence countries: a population-based cohort study of 519 955 migrants screened before entry to England, Wales, and Northern Ireland. *Lancet* 2016;**388**(10059):2510-8.

**Alekseev 2018** *{published data only}*

Alekseev A, Fatykhova R. The efficacy of screening for tuberculosis infection in paediatric population in the Republic of Tatarstan. *European Respiratory Journal* 2018;**52**:PA2735.

**Armstrong-Hough 2017** *{published data only}*

Armstrong-Hough M, Turimumahoro P, Meyer AJ, Ochom E, Babirye D, Ayakaka I, et al. Drop-out from the tuberculosis contact investigation cascade in a routine public health setting in urban Uganda: a prospective, multi-center study. *PLOS One* 2017;**12**(11):e0187145.

**Auld 2013** *{published data only}*

Auld SC, Click ES, Heilig CM, Miramontes R, Cain KP, Bisson GP, et al. Association between tuberculin skin test result and clinical presentation of tuberculosis disease. *BMC Infectious Diseases* 2013;**13**:460.

**Azit 2019** *{published data only}*

Azit NA, Ismail A, Ahmad N, Ismail R, Ishak S. Factors associated with tuberculosis disease among children who are household contacts of tuberculosis cases in an urban setting in Malaysia. *BMC Public Health* 2019;**19**(1):1432.

**Bamford 2010** *{published data only}*

Bamford AR, Crook AM, Clark JE, Nademi Z, Dixon G, Paton JY, et al. Comparison of interferon-gamma release assays and tuberculin skin test in predicting active tuberculosis (TB) in children in the UK: a paediatric TB network study. *Archives of Disease in Childhood* 2010;**95**(3):180-6.

**Basta 2010** *{published data only}*

Basta PC, Rios DP, Alves LC, Sant' Anna CC, Coimbra Junior CE. Clinical and radiological study of Surui indigenous children and adolescents, Amazon Region, Brazil. *Revista da Sociedade Brasileira de Medicina Tropical* 2010;**43**(6):719-22.

**Bennet 2017** *{published data only}*

Bennet R, Eriksson M. Tuberculosis infection and disease in the 2015 cohort of unaccompanied minors seeking asylum in Northern Stockholm, Sweden. *Infectious Diseases (London)* 2017;**49**(7):501-6.

**Bonnet 2017** *{published data only}*

Bonnet M, Kyakwera C, Kyomugasho N, Atwine D, Mugabe F, Nansumba M, et al. Prospective cohort study of the feasibility and yield of household child tuberculosis contact screening in Uganda. *International Journal of Tuberculosis and Lung Disease* 2017;**21**(8):862-8.

**Bosa 2017** *{published data only}*

Bosa L, Da Silva L, Mendes DV, Sifna A, Sargento Mendes M, Riccardi F, et al. Feasibility and effectiveness of tuberculosis active case-finding among children living with tuberculosis relatives: a cross-sectional study in Guinea-Bissau. *Mediterranean Journal of Hematology and Infectious Diseases* 2017;**9**(1):e2017059.

**Boullier 2017** *{published data only}*

Boullier M, Saatci D, Keane D, Williams A, Williams B. Choice of screening tool when conducting tuberculosis contact tracing in schools: does it change management outcome? *Archives of Disease in Childhood* 2017;**102**:A57-8.

**Chiappini 2019** *{published data only}*

Chiappini E, Storelli F, Tersigni C, Venturini E, de Martino M, Galli L. QuantiFERON-TB Gold In-Tube test performance in a large pediatric population investigated for suspected tuberculosis infection. *Paediatric Respiratory Reviews* 2019;**32**:36-47.

**Chisti 2014** *{published data only}*

Chisti MJ, Graham SM, Duke T, Ahmed T, Ashraf H, Faruque ASG, et al. A prospective study of the prevalence of tuberculosis and bacteraemia in Bangladeshi children with severe malnutrition and pneumonia including an evaluation of Xpert MTB/RIF assay. *PLOS One* 2014;**9**(4):e93776.

**Curtis 1999** *{published data only}*

Curtis AB, Ridzon R, Vogel R, McDonough S, Hargreaves J, Ferry J, et al. Extensive transmission of Mycobacterium tuberculosis from a child. *New England Journal of Medicine* 1999;**341**(20):1491-5.

**David 2017** *{published data only}*

David SG, Lovero KL, Pombo March MF, Abreu TG, Netto AR, Kritski AL, et al. A comparison of tuberculosis diagnostic systems in a retrospective cohort of HIV-infected children in Rio de Janeiro, Brazil. *International Journal of Infectious Diseases* 2017;**59**:150-5.

**de Lima 2013** *{published data only}*

de Lima LM, Schwartz E, Gonzales RI, Harter J, de Lima JF. The tuberculosis control program in Pelotas/RS, Brazil: home contact investigations. *Revista Gaúcha de Enfermagem* 2013;**34**(2):102-10.

**Do Nascimento Maia 2016** *{published data only}*

Do Nascimento Maia P, Sant'Anna CC, De Fátima Bazhuni Pombo March M, Gonzalez PA, De Paula Santana B, De Rademaker Itagiba PR. Follow-up of children contacts of tuberculosis in a primary health center in Rio De Janeiro, Brazil. *Pediatric Pulmonology* 2016;**51**:S47-8.

**Dorjee 2019** *{published data only}*

Dorjee K, Topgyal S, Dorjee C, Tsundue T, Namdol T, Tsewang T, et al. High prevalence of active and latent tuberculosis in children and adolescents in Tibetan schools in India: the Zero TB Kids Initiative in Tibetan refugee children. *Clinical Infectious Diseases* 2019;**69**(5):760-8.

**Driver 2002** *{published data only}*

Driver CR, Cordova IM, Munsiff SS. Targeting tuberculosis testing: the yield of source case investigations for young children with reactive tuberculin skin tests. *Public Health Reports* 2002;**117**(4):366-72.

**Egere 2017** *{published data only}*

Egere U, Togun T, Sillah A, Mendy F, Otu J, Hoelscher M, et al. Identifying children with tuberculosis among household contacts in The Gambia. *International Journal of Tuberculosis and Lung Disease* 2017;**21**(1):46-52.

**Faccini 2013** *{published data only}*

Faccini M, Codecasa LR, Ciconali G, Cammarata S, Borriello CR, De Gioia C, et al. Tuberculosis outbreak in a primary school, Milan, Italy. *Emerging Infectious Diseases* 2013;**19**(3):485-7.

**Fortunato 2011** *{published data only}*

Fortunato I, Sant'Anna C. Screening and follow-up of children exposed to tuberculosis cases, Luanda, Angola. *International Journal of Tuberculosis and Lung Disease* 2011;**15**(10):1359-61.

**Francis 2002** *{published data only}*

Francis J, Reed A, Yohannes F, Dodard M, Fournier AM. Screening for tuberculosis among orphans in a developing country. *American College of Preventive Medicine* 2002;**22**(2):117-9.

**Galli 2016** *{published data only}*

Galli L, Lancella L, Tersigni C, Venturini E, Chiappini E, Bergamini B, et al. Pediatric tuberculosis in Italian children: epidemiological and clinical data from the Italian register of pediatric tuberculosis. *International Journal of Molecular Sciences* 2016;**17**(6):960.

**Gashu 2016** *{published data only}*

Gashu Z, Jerene D, Ensermu M, Habte D, Melese M, Hiruy N, et al. The yield of community-based "retrospective" tuberculosis contact investigation in a high burden setting in Ethiopia. *PLOS One* 2016;**11**(8):e0160514.

**Girardi 2007** *{published data only}*

Girardi E, Loffredo M, Alessandrini A, Anzidei G, Goletti D. A two-step approach for screening contacts of active tuberculosis. *Infection* 2007;**35**(2):122-3.

**Gomez-Pastrana 1999** *{published data only}*

Gomez-Pastrana D, Torronteras R, Caro P, Anguita ML, Barrio AM, Andres A, et al. Diagnosis of tuberculosis in children using a polymerase chain reaction. *Pediatric Pulmonology* 1999;**28**(5):344-51.

**Gwee 2013** *{published data only}*

Gwee A, Pantazidou A, Ritz N, Tebruegge M, Connell TG, Cain T, et al. To x-ray or not to x-ray? Screening asymptomatic children for pulmonary TB: a retrospective audit. *Archives of Disease in Childhood* 2013;**98**(6):401-4.

**Hanrahan 2019** *{published data only}*

Hanrahan CF, Nonyane BA, Mmolawa L, West NS, Siwelana T, Lebina L, et al. Contact tracing versus facility-based screening for active TB case finding in rural South Africa: a pragmatic cluster-randomized trial (Kharitode TB). *PLOS Medicine* 2019;**16**(4):e1002796.

**Hoffman 1996** *{published data only}*

Hoffman ND, Kelly C, Futterman D. Tuberculosis infection in human immunodeficiency virus-positive adolescents and young adults: a New York City cohort. *Pediatrics* 1996;**97**(2):198-203.

**Huang 2016** *{published data only}*

Huang Y, Zhong J, Wu Q, Liu Z, Pan A, Zhu L, et al. Investigation of a large school-based outbreak of tuberculosis infection in Eastern China. *Pediatria Polska* 2016;**91**(6):541-6.

**Izumi 2017** *{published data only}*

Izumi K, Ohkado A, Uchimura K, Kawatsu L, Suenaga M, Urakawa M, et al. Evaluation of tuberculosis contact investigations in Japan. *International Journal of Tuberculosis and Lung Disease* 2017;**21**(2):188-95.

**Karki 2017** *{published data only}*

Karki B, Kittel G, Bolokon I Jr, Duke T. Active community-based case finding for tuberculosis with limited resources. *Asia Pacific Journal of Public Health* 2017;**29**(1):17-27.

**Kemigisha 2015** *{published data only}*

Kemigisha E, Katawera V, Mwanga AJ. Prevalence and clinical predictors of tuberculosis in severely malnourished Ugandan children. *Annals of Global Health* 2015;**81**(1):7.

**Kim 2017** *{published data only}*

Kim Y, Kim BK, Choi HJ, Ryu SW, Kim ES, Chang YS, et al. Lessons learned from continued TB outbreaks in a high school. *PLOS One* 2017;**12**(11):e0188076.

**Kondo 2003** *{published data only}*

Kondo S, Ito M. Efficacy of tuberculosis contacts investigation and treatment, especially of preventive therapy in infants and young children. *Kekkaku* 2003;**78**(11):677-82.

**Lee 2008** *{published data only}*

Lee MS, Leung CC, Kam KM, Wong MY, Leung MC, Tam CM, et al. Early and late tuberculosis risks among close contacts in Hong Kong. *International Journal of Tuberculosis and Lung Disease* 2008;**12**(3):281-7.

**Leung 2006** *{published data only}*

Leung CC, Yew WW, Chang KC, Tam CM, Chan CK, Law WS, et al. Risk of active tuberculosis among schoolchildren in Hong Kong. *Archives of Pediatric Adolescent Medicine* 2006;**160**(3):247-51.

**Li 2015** *{published data only}*

Li T, Bao L, Diao N, Sun F, Gao Y, Wong K, et al. Influential factors of the performance of interferon-gamma release assays in the diagnosis of childhood tuberculosis. *Clinical and Experimental Medicine* 2015;**15**:303-9.

**Ling 2013** *{published data only}*

Ling DI, Nicol MP, Pai M, Pienaar S, Dendukuri N, Zar HJ. Incremental value of T-SPOT.TB for diagnosis of active pulmonary tuberculosis in children in a high-burden setting: a multivariable analysis. *Thorax* 2013;**86**:860-6.

**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 (Edinburgh, Scotland)* 2013;**93**(3):357-62.

**Malik 2018** *{published data only}*

Malik AA, Amanullah F, Codlin AJ, Siddiqui S, Jaswal M, Ahmed JF, et al. Improving childhood tuberculosis detection and treatment through facility-based screening in rural Pakistan. *International Journal of Tuberculosis and Lung Disease* 2018;**22**(8):851-7.

**Marais 2006** *{published data only}*

Marais BJ, Gie RP, Hesseling AC, Schaaf HS, Lombard C, Enarson DA, et al. A refined symptom-based approach to diagnose pulmonary tuberculosis in children. *Pediatrics* 2006;**118**(5):e1350-9.

**Marais 2009b** *{published data only}*

Marais BJ, Hesseling AC, Schaaf HS, Gie RP, van Helden PD, Warren RM. Mycobacterium tuberculosis transmission is not related to household genotype in a setting of high endemicity. *Journal of Clinical Microbiology* 2009;**47**(5):1338-43.

**Marcy 2019** *{published data only}*

Marcy O, Borand L, Ung V, Msellati P, Teijiokem M, Huu KT, et al. A treatment-decision score for HIV-infected children with suspected tuberculosis. *Pediatrics* 2019;**144**(3):e20182065.

**Masur 2017** *{published data only}*

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.

**Minhas 2017** *{published data only}*

Minhas KS, Andronikou S, Bernatoniene J, Grier D, Sarra A, Roderick M. Role of chest radiography in screening of childhood TB contacts in the UK – A migrant context. *Pediatric Radiology* 2017;**47**:S380-1.

**Moran-Mendoza 2010** *{published data only}*

Moran-Mendoza O, FitzGerald JM, Marion S, Elwood K, Patrick D. Tuberculin skin test size and risk of TB among non-treated contacts of TB cases. *Canadian Respiratory Journal* 2010;**17**:5B.

**Mueller-Hermelink 2018** *{published data only}*

Mueller-Hermelink M, Kobbe R, Methling B, Rau C, Schulze-Sturm U, Auer I, et al. Universal screening for latent and active tuberculosis (TB) in asylum seeking children, Bochum and Hamburg, Germany, September 2015 to November 2016. *Eurosurveillance* 2018;**23**(12):3.

**Murray 2019** *{published data only}*

Murray KO, Castillo-Carandang NT, Mandalakas AM, Cruz AT, Leining LM, Gatchalian SR, et al. Prevalence of tuberculosis in children after natural disasters, Bohol, Philippines. *Emerging Infectious Diseases* 2019;**25**(10):1884-92.

**Nduba 2018** *{published data only}*

Nduba V, Van't Hoog AH, Mitchell EM, Borgdorff M, Laserson KF. Incidence of active tuberculosis and cohort retention among adolescents in western Kenya. *Pediatric Infectious Disease Journal* 2018;**37**(1):10-5.

**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.

**Oh 2018** *{published data only}*

Oh CE, Kwon GY, Kwon YH, Lee EJ, Park MS, Kim SH, et al. High tuberculosis transmission rate in children with nursery exposure to undetected pulmonary tuberculosis. *International Journal of Tuberculosis and Lung Disease* 2018;**22**(9):1031-6.

**Padmapriyadarsini 2016** *{published data only}*

Padmapriyadarsini C, Bhavani PK, Sekar L, Anandhachitra, Selvaraj M, Poornagangadevi N, Mothi SN, et al. Effectiveness of symptom screening and incidence of tuberculosis among adults and children living with HIV infection in India. *National Medical Journal of India* 2016;**29**(6):321-5.

**Pan 2019** *{published data only}*

Pan D, Lan R, Graviss EA, Lin D, Liang D, McNeil E, et al. Adolescent tuberculosis associated with tuberculosis exposure in classrooms and dorm rooms in Guangxi, China. *International Journal of Infectious Diseases* 2019;**78**:8-14.

**Penin 2007** *{published data only}*

Penin Anton M, Gomez Carrasco JA, Lopez Lois G, Merino Villeneuve I, Leal Beckouche M, Garcia de Frias E. Tuberculosis outbreak in a school. *Anales de Pediatria (Barcelona, Spain : 2003)* 2007;**67**(1):18-21.

**Penn-Nicholson 2019** *{published data only}*

Penn-Nicholson A, Hraha T, Thompson EG, Sterling D, Mbandi SK, Wall KM, et al. Discovery and validation of a prognostic proteomic signature for tuberculosis progression: a prospective cohort study. *PLOS Medicine* 2019;**16**(4):e1002781.

**Puryear 2013** *{published data only}*

Puryear S, Seropola G, Ho-Foster A, Arscott-Mills T, Mazhani L, Firth J, et al. Yield of contact tracing from pediatric tuberculosis index cases in Gaborone, Botswana. *International Journal of Tuberculosis and Lung Disease* 2013;**17**(8):1049-55.

**Rachow 2012** *{published data only}*

Rachow A, Clowes P, Saathoff E, Mtafya B, Michael E, Ntinginya EN, et al. Increased and expedited case detection by Xpert MTB/RIF assay in childhood tuberculosis: a prospective cohort study. *Clinical Infectious Diseases* 2012;**54**(10):1388-96.

**Ramirez 2006** *{published data only}*

Ramírez-Cardich ME, Kawai V, Oberhelman RA, Bautista CT, Castillo ME, Gilman RH. Clinical correlates of tuberculosis co-infection in HIV-infected children hospitalized in Peru. *International Journal of Infectious Diseases* 2006;**10**(4):278-81.

**Rossoni 2020** *{published data only}*

Rossoni AM, Lovero KL, Tahan TT, Netto AR, Rossoni MD, Almeida IN, et al. Evaluation of pulmonary tuberculosis diagnostic tests in children and adolescents at a pediatric reference center. *Pulmonology* 2020;**S2531-0437**(20):30001-5.

**Salinas 2002** *{published data only}*

Salinas Solano C, Altube Urrengoetxea L, Españ Yandiola PP, Capelastegui Saiz A, Quintana López JM. Tuberculosis among immigrants in Bilbao (Spain). *Archivos de Bronconeumologia* 2002;**38**(11):506-10.

**Saunders 2014** *{published data only}*

Saunders MJ, Koh GC, Small AD, Dedicoat M. Predictors of contact tracing completion and outcomes in tuberculosis: a 21-year retrospective cohort study. *International Journal of Tuberculosis and Lung Disease* 2014;**18**(6):640-6.

**Shah 2008** *{published data only}*

Shah NS, Anh MH, Thuy TT, Duong Thom BS, Linh T, Nghia DT, et al. Population-based chest X-ray screening for pulmonary tuberculosis in people living with HIV/AIDS, An Giang, Vietnam. *International Journal of Tuberculosis and Lung Disease* 2008;**12**(4):404-10.

**Shaikh 2017** *{published data only}*

Shaikh N, Gupte A, Dharmshale S, Pokkali S, Thakar M, Upadhye VJ, et al. Novel interferon-gamma assays for diagnosing tuberculosis in young children in India. *International Journal of Tuberculosis and Lung Disease* 2017;**21**(4):412-9.

**Sollai 2017** *{published data only}*

Sollai S, Ghetti F, Bianchi L, de Martino M, Galli L, Chiappini E. Infectious diseases prevalence, vaccination coverage, and diagnostic challenges in a population of internationally adopted children referred to a tertiary care children's hospital from 2009 to 2015. *Medicine (Baltimore)* 2017;**96**(12):e6300.

**Spyridis 2003** *{published data only}*

Spyridis P, Tsolia M, Gelesme A, Moustaki M, Spyridis N, Sinaniotis C, et al. The impact of Greece's childhood tuberculosis screening programme on the epidemiological indexes in the greater Athens area. *International Journal of Tuberculosis and Lung Disease* 2003;**7**(3):248-53.

**Swingler 2000** *{published data only}*

Swingler GH. Chest radiography in ambulatory children with acute lower respiratory infections: effective tuberculosis case finding? *Annals of Tropical Paediatrics* 2000;**20**(1):11-5.

**Szkwarko 2018** *{published data only}*

Szkwarko D, Owiti P, Buziba N, Bigelow C, Eaton CB, Carter EJ. Implementation of an active, clinic-based child tuberculosis contact management strategy in western Kenya. *Public Health Action* 2018;**8**(2):91-4.

**Thee 2019** *{published data only}*

Thee S, Kruger R, von Bernuth H, Meisel C, Kolsch U, Kirchberger V, et al. Screening and treatment for tuberculosis in a cohort of unaccompanied minor refugees in Berlin, Germany. *PLOS One* 2019;**14**(5):e0216234.

**van Schalkwyk 2014** *{published data only}*

van Schalkwyk C, Variava E, Shapiro AE, Rakgokong M, Masonoke K, Lebina L, et al. Incidence of TB and HIV in prospectively followed household contacts of TB index patients in South Africa. *PLOS One* 2014;**9**(4):e95372.

**Verver 2005** *{published data only}*

Verver S, van Loenhout-Rooyackers JH, Bwire R, Annee-van Bavel JA, de Lange HJ, van Gerven PJ, et al. Tuberculosis infection in children who are contacts of immigrant tuberculosis patients. *European Respiratory Journal* 2005;**26**(1):126-32.

**Williams 2016** *{published data only}*

Williams B, Pickard L, Grandjean L, Pope S, Anderson SR, Morgan G, et al. The need to implement effective new entrant tuberculosis screening in children: evidence from school 'outbreak'. *Journal of Public Health (Oxford, England)* 2016;**38**(4):e511-5.

**Williams 2019** *{published data only}*

Williams B, Eisen S, Williams A, Cricks Z, Robinson K, Boullier M, et al. Infection screening in unaccompanied asylum-seeking children. *Archives of Disease in Childhood* 2019;**104**:A111-2.

**Yang 2018** *{published data only}*

Yang C, Yasseen AS 3rd, Stimec J, Rea E, Waters V, Lam R, et al. Prevalence of tuberculosis infection and disease in children referred for tuberculosis medical surveillance in Ontario: a single-cohort study. *CMAJ Open* 2018;**6**(3):E365-71.

**Yuan 1995** *{published data only}*

Yuan L, Richardson E, Kendall PR. Evaluation of a tuberculosis screening program for high-risk students in Toronto schools. *Canadian Medical Association Journal* 1995;**153**(7):925-32.

**Zachariah 2003** *{published data only}*

Zachariah R, Spielmann MP, Harries AD, Gomani P, Graham SM, Bakali E, et al. Passive versus active tuberculosis case finding and isoniazid preventive therapy among household contacts in a rural district of Malawi. *International Journal of Tuberculosis and Lung Disease* 2003;**7**(11):1033-9.

## Additional references

**Aggerbeck 2019**

Aggerbeck H, Ruhwald M, Hoff ST, Tingskov PN, Hellstrom E, Malahleha M, et al. Interaction between C-Tb and PPD given concomitantly in a split-body randomised controlled trial. *International Journal of Tuberculosis and Lung Disease* 2019;**23**(1):38-44.

**Albuquerque 2019**

Albuquerque VV, Kumar NP, Fukutani KF, Vasconcelos B, Arriaga MB, Silveira-Mattos PS, et al. Plasma levels of C-reactive protein, matrix metalloproteinase-7 and lipopolysaccharide-binding protein distinguish active pulmonary or extrapulmonary tuberculosis from uninfected controls in children. *Cytokine* 2019;**123**:154773.

**Alsleben 2011**

Alsleben N, Ruhwald M, Rüssmann H, Marx FM, Wahn U, Magdorf K. Interferon-gamma inducible protein 10 as a biomarker for active tuberculosis and latent tuberculosis infection in children: a case-control study. *Scandinavian Journal of Infectious Diseases* 2011;**44**(4):256-62.

**Arscott-Mills 2014**

Arscott-Mills T, Ho-Foster A, Lowenstein M, Jibril H, Masunge J, Mweemba, et al. Yield of screening for TB and HIV among children failing to thrive in Botswana. *Journal of Tropical Pediatrics* 2014;**60**(1):27-32.

**Balshem 2011**

Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. *Journal of Clinical Epidemiology* 2011;**64**(4):401-6.

**Bjerrum 2019**

Bjerrum S, Schiller I, Dendukuri N, Kohli M, Nathavitharana RR, Zwerling AA, et al. Lateral flow urine lipoarabinomannan assay for detecting active tuberculosis in people living with HIV. *Cochrane Database of Systematic Reviews* 2019, Issue 10. Art. No: CD011420. [DOI: 10.1002/14651858.CD011420.pub3]

**Bjerrum 2020**

Bjerrum S, Broger T, Székely R, Mitarai S, Opintan JA, Kenu E, et al. Diagnostic accuracy of a novel and rapid lipoarabinomannan test for diagnosing tuberculosis among people with Human Immunodeficiency Virus. *Open Forum Infectious Diseases* 2020;**7**(1):ofz530.

**Bossuyt 2006**

Bossuyt PM, Irwig L, Craig J, Glasziou P. Comparative accuracy: assessing new tests against existing diagnostic pathways. *BMJ* 2006;**332**(7549):1089-92.

**Broger 2020**

Broger T, Nicol MP, Székely R, Bjerrum S, Sossen B, Schutz C, et al. Diagnostic accuracy of a novel tuberculosis point-of-care urine lipoarabinomannan assay for people living with HIV: a meta-analysis of individual in- and outpatient data. *PLOS Medicine* 2020;**17**(5):e1003113.

**Chakravorty 2017**

Chakravorty S, Simmons AM, Rowneki M, Parmar H, Cao Y, Ryan J, et al. The new Xpert MTB/RIF Ultra: improving detection of Mycobacterium tuberculosis and resistance to rifampin in an assay suitable for point-of-care testing. *mBio* 2017;**8**(4):1-12.

**Chu 2006**

Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *Journal of Clinical Epidemiology* 2006;**59**(12):1331-2.

**Covidence [Computer program]**

Veritas Health Innovation Covidence. Version accessed 10 January 2020. Melbourne, Australia: Veritas Health Innovation, 2017. Available at covidence.org.

**Cruz 2012**

Cruz AT, Revell PA, Starke JR. Gastric aspirate yield or children with suspected pulmonary tuberculosis. *Journal of the Pediatric Infectious Diseases Society* 2012;**2**(2):171-4.

**Day 2011**

Day CL, Abrahams DA, Lerumo L, Janse van Rensburg E, Stone L, O'Rie T, et al. Functional capacity of Mycobacterium tuberculosis-specific T cell responses in humans is associated with mycobacterial load. *Journal of Immunology* 2011;**187**:2222-32.

**Drain 2019**

Drain PK, Gardiner J, Hannah H, Broger T, Dheda K, Fielding K, et al. Guidance for studies evaluating the accuracy of biomarker-based nonsputum tests to diagnose tuberculosis. *Journal of Infectious Diseases* 2019;**S3**:S108-15. [DOI: 10.1093/infdis/jiz356]

**Dunn 2016**

Dunn JJ, Starke JR, Revell PA. Laboratory diagnosis of Mycobacterium tuberculosis infection and disease in children. *Journal of Clinical Microbiology* 2016;**54**(6):1434-41.

**Du Toit 2002**

Du Toit G, Swingler G, Iloni K. Observer variation in detecting lymphadenopathy on chest radiography. *International Journal of Tuberculosis and Lung Disease* 2002;**6**(9):814-7.

**Excel 2013 [Computer program]**

Microsoft Excel Software. Version 16.16.4. Seattle (WA): Microsoft Corporation, 2013.

**Global Laboratory Initiative 2019**

Global Laboratory Initiative. Practical guide to implementing a quality assurance system for Xpert MTB/RIF testing. www.stoptb.org/wg/gli/assets/documents/Xpert-QA-guide-2019.pdf (accessed prior to 11 June 2020).

**Graham 2015**

Graham SM, Cuevas LE, Jean-Philippe P, Browning R, Casenghi M, Detjen AK. Clinical case definitions for classification of intrathoracic tuberculosis in children: an update. *Clinical Infectious Diseases* 2015;**61**(Suppl 3):S179-87.

**Holm 2014**

Holm LL, Rose MV, Kimaro G, Bygbjerg IC, Mfinanga SG, Ravn P, et al. A comparison of interferon-gamma and IP-10 for the diagnosis of tuberculosis. *Pediatrics* 2014;**134**(6):e1568.

**Jaganath 2012**

Jaganath D, Mupere E. Childhood tuberculosis and malnutrition. *Journal of Infectious Diseases* 2012;**206**:1809-15.

**Jenkins 2017**

Jenkins HE, Yuen CM, Rodriguez CA, Nathavitharana RR, McLaughlin MM, Donald P, et al. Mortality in children diagnosed with tuberculosis: a systematic review and meta-analysis. *Lancet Infectious Diseases* 2017;**17**(3):285-95.

**Jenum 2016**

Jenum S, Dhanasekaran S, Ritz C, Macaden R, Doherty TM, Grewal HM, et al. Added value of IP-10 as a read-out of Mycobacterium tuberculosis: specific immunity in young children. *Pediatric Infectious Disease Journal* 2016;**35**(12):1336-8.

**Kaguthi 2014**

Kaguthi G, Nduba V, Nyokabi J, Onchiri F, Gie R, Borgdorff M. Chest radiographs for pediatric TB diagnosis: interrater agreement and utility. *Interdisciplinary Perspectives on Infectious Diseases* 2014;**2014**:291841.

**Kay 2020**

Kay AW, González FL, Takwoingi Y, Eisenhut M, Detjen AK, Steingart KR, et al. Xpert MTB/RIF and Xpert MTB/RIF Ultra assays for active tuberculosis and rifampicin resistance in children. *Cochrane Database of Systematic Reviews* 2020, Issue 8. Art. No: CD013359. [DOI: 10.1002/14651858.CD013359.pub2]

**Leung 1992**

Leung AN, Müller NL, Pineda PR, FitzGerald JM. Primary tuberculosis in childhood: radiographic manifestations. *Radiology* 1992;**182**(1):87-91.

**Macaskill 2010**

Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and presenting results. In: Deeks JJ, Bossuyt PM, Gatsonis C, editor(s). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. The Cochrane Collaboration, 2010. Available from srdta.cochrane.org.

**Marais 2004**

Marais BJ, Gie RP, Schaaf HS, Hesseling AC, Obihara CC, Starke JJ, et al. The natural history of childhood intrathoracic tuberculosis: a critical review of literature from the pre-chemotherapy era. *International Journal of Tuberculosis and Lung Disease* 2004;**8**(4):392-402.

**Marais 2005a**

Marais BJ, Gie RP, Obihara CC, Hesseling AC, Schaaf HS, Beyers N. Well-defined symptoms are of value in the diagnosis of childhood pulmonary tuberculosis. *Archives of Disease in Childhood* 2005;**90**(11):1162-5.

**Marais 2005b**

Marais BJ, Obihara CC, Gie RP, Schaaf HS, Hesseling AC, Lombard C, et al. The prevalence of symptoms associated with pulmonary tuberculosis in randomly selected children from a high-burden community. *Archives of Disease in Childhood* 2005;**90**(11):1166-70.

**Marais 2009a**

Marais BJ, Ayles H, Graham SM, Godfrey-Faussett P. Screening and preventive therapy for tuberculosis. *Clinics in Chest Medicine* 2009;**30**(4):827-46.

**Martinez 2019**

Martinez L, Lo NC, Cords O, Hill PC, Khan P, Hatherill M, et al. Paediatric tuberculosis transmission outside the household: challenging historical paradigms to inform future public health strategies. *Lancet Respiratory Medicine* 2019;**7**(6):544-52.

**Moher 2009**

Moher D, Liberati A, Tetzlaff J, Altman DG, the PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *PLOS Medicine* 2009;**6**(7):e1000097. [DOI: 10.1371/journal.pmed1000097]

**Munthali 2017**

Munthali T, Chabala C, Chama E, Mugode R, Kapata N, Musonda P, et al. Tuberculosis caseload in children with severe acute malnutrition related with high hospital-based mortality in Lusaka, Zambia. *BMC Research Notes* 2017;**10**:206.

**Nicol 2021**

Nicol MP, Schumacher SG, Workman L, Broger T, Baard C, Prins M, et al. Accuracy of a novel urine test, Fujifilm SILVAMP TB LAM, for the diagnosis of pulmonary tuberculosis in children. *Clinical Infectious Diseases* 2021;**72**(9):e280-8.

**Nkereuwem 2021**

Nkereuwem E, Togun T, Gomez MP, Székely R, Macé A, Jobe D, et al. Comparing accuracy of lipoarabinomannan urine tests for diagnosis of pulmonary tuberculosis in children from four African countries: a cross-sectional study. *Lancet Infectious Diseases* 2021;**21**(3):376-84.

**Oliwa 2015**

Oliwa J, Karumbi J, Marais BJ, Madhi SA, Graham SM. Tuberculosis as a cause or comorbidity of childhood pneumonia in tuberculosis-endemic areas: a systematic review. *Lancet Respiratory Medicine* 2015;**3**:235-43.

**Pai 2014**

Pai M, Denkinger CM, Kik SV, Rangaka MX, Zwerling A, Oxlade O, et al. Gamma interferon release assays for detection of Mycobacterium tuberculosis infection. *Clinical Microbiology Reviews* 2014;**27**:3-20.

**Perez-Velez 2012**

Perez-Velez CM, Marais BJ. Tuberculosis in children. *New England Journal of Medicine* 2012;**367**:348-61.

**Qin 2019**

Qin ZZ, Sander MS, Rai B, Titahong CN, Sudrungrot S, Laah SN, et al. Using artificial intelligence to read chest radiographs for tuberculosis detection: a multi-site evaluation of the diagnostic accuracy of three deep learning systems. *Scientific Reports* 2019;**9**(1):15000.

**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.

**Reuter 2019**

Reuter A, Hughes J, Furin J. Challenges and controversies in childhood tuberculosis. *Lancet* 2019;**394**:967-78.

**Review Manager 2020 [Computer program]**

Nordic Cochrane Centre, The Cochrane Collaboration Review Manager 5 (RevMan 5). Version 5.4. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2020.

**Rozot 2013**

Rozot V, Vigano S, Mazza-Stalder J, Idrizi E, Day CL, Perreau M, et al. Mycobacterium tuberculosis-specific CD8 T cells are functionally and phenotypically different between latent infection and active disease. *European Journal of Immunology* 2013;**43**:1568-77.

**Ruhwald 2017**

Ruhwald M, Aggerbeck H, Gallardo RV, Hoff ST, Villate JI, Borregaard B, et al. Safety and efficacy of the C-Tb skin test to diagnose Mycobacterium tuberculosis infection, compared with an interferon γ release assay and the tuberculin skin test: a phase 3, double-blind, randomised, controlled trial. *Lancet Respiratory Medicine* 2017;**5**(4):259-68.

**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. *Canadian Medical Association Journal* 2006;**174**(4):469-76.

**Schumacher 2019**

Schumacher SG, Wells WA, Nicol MP, Steingart KR, Theron G, Dorman SE, et al. Guidance for studies evaluating the accuracy of sputum-based tests to diagnose tuberculosis. *Journal of Infectious Diseases* 2019;**220**(S3):S99-107.

**Schünemann 2008**

Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008;**336**(7653):1106-11.

**Schünemann 2020a**

Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, GRADE Working Group. 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.

**Schünemann 2020b**

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

**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]

**Sodhi 2017**

Sodhi KS, Bhalla AS, Mahomed N, Laya BF. Imaging of thoracic tuberculosis in children: current and future directions. *Pediatric Radiology* 2017;**47**(10):1260-8.

**Stata [Computer program]**

Stata Statistical Software. StataCorp, Version 16. College Station, TX, USA: StataCorp, 2019. Available at www.stata.com.

**Stop TB Partnership 2019**

Global TB Partnership. The paradigm shift 2018-2022. www.stoptb.org/assets/documents/global/plan/GPR_2018-2022_Digital.pdf (accessed 5 April 2020).

**Sudbury 2019**

Sudbury EL, Otero L, Tebruegge M, Messina NL, Seas C, Montes M. Mycobacterium tuberculosis-specific cytokine biomarkers for the diagnosis of childhood TB in a TB-endemic setting. *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases* 2019;**16**:100102.

**Sutherland 2010**

Sutherland JS, Young JM, Peterson KL, Sanneh B, Whittle HC, Rowland-Jones SL, et al. Polyfunctional CD4 and CD8 T cell responses to tuberculosis antigens in HIV-1-infected patients before and after anti-retroviral treatment. *Journal of Immunology* 2010;**184**:6537-44.

**Swingler 2005**

Swingler GH, Du Toit G, Andronikou S, van der Merwe L, Zar HJ. Diagnostic accuracy of chest radiography in detecting mediastinal lymphadenopathy in suspected pulmonary tuberculosis. *Archives of Disease in Childhood* 2005;**90**:1153-6.

**Szkwarko 2017**

Szkwarko D, Hirsch-Moverman Y, Du Plessis L, Du Preez K, Carr C, Mandalakas AM. Child contact management in high tuberculosis burden countries: a mixed-methods systematic review. *PLOS One* 2017;**12**(8):e0182185.

**Takwoingi 2015**

Takwoingi Y, Guo B, Riley RD, Deeks JJ. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. *Statistical Methods in Medical Research* 2015;**0**:1-19.

**Tebruegge 2015**

Tebruegge M, Dutta B, Donath S, Ritz N, Forbes B, Camacho-Badilla K, et al. Mycobacteria-specific cytokine responses detect tuberculosis infection and distinguish latent from active tuberculosis. *American Journal of Respiratory and Critical Care Medicine* 2015;**192**(4):485-99.

**Theel 2018**

Theel ES, Hilgart H, Breen-Lyles M, McCoy K, Flury R, Breeher LE, et al. Comparison of the QuantiFERON-TB Gold Plus and QuantiFERON-TB Gold In-Tube interferon gamma release assays in patients at risk for tuberculosis and in health care workers. *Journal of Clinical Microbiology* 2018;**56**:e00614-18.

**van't Hoog 2013**

van't Hoog AH, Langendam M, Mitchell E, Cobelens FG, Sinclair D, Leeflang MM, 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. www.who.int/tb/Review2Accuracyofscreeningtests.pdf?ua=1 (accessed 5 April 2020).

**van't Hoog 2014**

van't Hoog AH, Langendam M, Mitchell E, Cobelens FG, Sinclair D, Leeflang MM, 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]

**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.

**Whiting 2013**

Whiting PF, Rutjes AW, Westwood ME, Mallett S. A systematic review classifies sources of bias and variation in diagnostic test accuracy studies. *Journal of Clinical Epidemiology* 2013;**66**(10):1093-104.

**WHO 2011**

World Health Organization. Guidelines for intensified case-finding and isoniazid preventive therapy for people living with HIV in resource-constrained settings. apps.who.int/iris/bitstream/handle/10665/44472/9789241500708_eng.pdf?sequence=1 (accessed prior to 28 November 2020).

**WHO 2013a**

World Health Organization. Systematic screening for active tuberculosis: principles and recommendations. www.who.int/tb/tbscreening/en/ (accessed 26 November 2020).

**WHO 2013b**

World Health Organization. Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF assay for the diagnosis of pulmonary and extra-pulmonary TB in adults and children. Policy update. www.who.int/tb/publications/systematic_screening/en/ (accessed prior to 26 November 2020).

**WHO 2014**

World Health Organization. Guidance for national tuberculosis programmes on the management of tuberculosis in children. www.who.int/tb/publications/childtb_guidelines/en/ (accessed 20 January 2021).

**WHO 2015**

World Health Organization. Systematic screening for active tuberculosis: an operational guide. apps.who.int/iris/bitstream/handle/10665/181164/9789241549172_eng.pdf;jsessionid=A852BCDF07D5D7FAC40F5BD7E06498B8?sequence=1 (accessed prior to 26 November 2020).

**WHO 2017a**

World Health Organization. Tuberculosis diagnostics technology landscape. unitaid.org/assets/2017-Unitaid-TB-Diagnostics-Technology-Landscape.pdf (accessed 26 November 2020).

**WHO 2017b**

World Health Organization. WHO meeting report of a technical expert consultation: non-inferiority analysis of Xpert MTB/RIF Ultra compared to Xpert MTB/RIF. www.who.int/tb/publications/2017/XpertUltra/en/ (accessed prior to 26 November 2020).

**WHO 2018**

World Health Organization. Roadmap towards ending TB in children and adolescents. apps.who.int/iris/bitstream/handle/10665/275422/9789241514798-eng.pdf?ua=1 (accessed 5 April 2020).

**WHO Consolidated Guidelines (Module 1) 2020**

World Health Organization. WHO consolidated guidelines on tuberculosis. Module 1, Prevention: tuberculosis preventive treatment. who.int/publications/i/item/who-consolidated-guidelines-on-tuberculosis-module-1-prevention-tuberculosis-preventive-treatment (accessed 26 November 2020).

**WHO Consolidated Guidelines (Module 2) 2021**

World Health Organization. WHO consolidated guidelines on tuberculosis. Module 2: screening – systematic screening for tuberculosis disease. who.int/publications/i/item/9789240022676 (accessed 25 March 2021).

**WHO Consolidated Guidelines (Module 3) 2020**

World Health Organization. WHO consolidated guidelines on tuberculosis. Module 3: diagnosis – rapid diagnostics for tuberculosis detection; June 2020. who.int/publications/i/item/who-consolidated-guidelines-on-tuberculosis-module-3-diagnosis---rapid-diagnostics-for-tuberculosis-detection (accessed 1 July 2020).

**WHO Global Tuberculosis Report 2020**

World Health Organization. Global tuberculosis report 2020. www.who.int/tb/publications/global_report/en/ (accessed 20 October 2020).

**Zar 2012**

Zar HJ, Workman L, Isaacs W, Munro J, Black F, Eley B, et al. Rapid molecular diagnosis of pulmonary tuberculosis in children using nasopharyngeal specimens. *Clinical Infectious Diseases* 2012;**55**(8):1088-95.

## References to other published versions of this review

**Vonasek 2020**

Vonasek B, Ness T, Takwoingi Y, Kay AW, Wyk SS, Ouellette L, et al. Screening tests for active pulmonary tuberculosis in children. *Cochrane Database of Systematic Reviews* 2020, Issue 7. Art. No: CD013693. [DOI: 10.1002/14651858.CD013693]

## CHARACTERISTICS OF STUDIES

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

**Aggerbeck 2018**

| ***Study characteristics*** | |
|---|---|
| Patient Sampling | Prospective, cohort, consecutive |
| Patient characteristics and setting | Enrolment criteria: children with signs of TB, symptoms of TB, or close contact to a sputum smear TB-positive case<br>Age: < 5 years<br>Sex: 51% female overall (not reported for the < 5-year subgroup)<br>HIV infection: 25% overall (not reported for the < 5-year subgroup)<br>Sample size included for analysis: 235<br>Setting: outpatient<br>Country: South Africa<br>World Bank Income Classification: upper middle<br>High TB burden country: yes<br>High TB/HIV burden country: yes<br>Prevalence of TB cases in the study: composite reference standard 8.1%, microbiological reference standard 1.4% |
| Index tests | Children with 1 of following symptoms concerning for TB: fever, cough, decreased playfulness, or night sweats |
| Target condition and reference standard(s) | Active TB not specified as pulmonary<br>Microbiological reference standard and composite reference standard (includes those diagnosed by clinical symptoms) |
| Flow and timing | Timing between index test and reference standard not reported.<br>No missing data reported for the index tests or reference standards. |
| Comparative | |
| Notes | |
| ***Methodological quality*** | |

**Aggerbeck 2018** *(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 (All tests)** | | | |
| 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 reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | Unclear | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |
| **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 receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | Yes | | |
| **Could the patient flow have introduced bias?** | | Low risk | |

**Birungi 2018**

| ***Study characteristics*** | |
|---|---|
| Patient Sampling | Prospective, cross-sectional, consecutive |
| Patient characteristics and setting | Enrolment criteria: child contacts of sputum smear TB-positive cases<br>Age: < 15 years, median 6 years (IQR 2–13 years)<br>Sex: 49% female<br>HIV infection: 6%<br>Sample size included for analysis: 216<br>Setting: outpatient<br>Country: Rwanda<br>World Bank Income Classification: low<br>High TB burden country: no<br>High TB/HIV burden country: no<br>Prevalence of TB cases in the study: 1.9% |
| Index tests | TB contact; CXR; 1 of multiple symptoms – cough > 1 week, haemoptysis, fever, failure to gain weight, absence of appetite, fatigue, or presence of lymphadenopathy |
| Target condition and reference standard(s) | Pulmonary TB<br>Composite reference standard: defined as microbiologically confirmed or unconfirmed TB (symptoms suggestive of TB and CXR consistent with active TB) |
| Flow and timing | Timing between index test and reference standard not reported.<br>No missing data reported for the index tests or reference standards. |
| 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 | |

**Birungi 2018** *(Continued)*

| | | | |
|---|---|---|---|
| **Are there concerns that the included patients and setting do not match the review question?** | | | Low concern |
| **DOMAIN 2: Index Test (All tests)** | | | |
| 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 reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | No | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |
| **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 receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | Yes | | |
| **Could the patient flow have introduced bias?** | | Low risk | |

**Clemente 2017**

| | |
|---|---|
| ***Study characteristics*** | |
| Patient Sampling | Retrospective, cohort, consecutive |
| Patient characteristics and setting | Enrolment criteria: children referred mostly for TB contact and less commonly for concerning symptoms<br>Age: < 15 years, mean 5.8 years (SD 3.9 years)<br>Sex: 50% female<br>HIV infection: not reported |

**Clemente 2017** *(Continued)*

| | |
|---|---|
| | Sample size included for analysis: 246<br>Setting: outpatient<br>Country: Italy<br>World Bank Income Classification: high<br>High TB burden country: no<br>High TB/HIV burden country: no<br>Prevalence of TB cases in the study: 8.9% |
| Index tests | CXR |
| Target condition and reference standard(s) | Active TB not specified as pulmonary<br>Composite reference standard: based upon symptoms, CXR, TST, and microbiological testing |
| Flow and timing | Timing between index test and reference standard not reported.<br>No missing data reported for the index tests or reference standards. |
| 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 (All tests)** | | | |
| 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 | |

**Clemente 2017** *(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 reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | No | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |
| **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 receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | Yes | | |
| **Could the patient flow have introduced bias?** | | Low risk | |

**Dreesman 2017**

| | |
|---|---|
| ***Study characteristics*** | |
| Patient Sampling | Prospective, cohort, consecutive |
| Patient characteristics and setting | Enrolment criteria: children recently exposed to a sputum smear TB-positive adult<br>Age: < 15 years, mean 4.25 years (range 0–14 years)<br>Sex: 44% female<br>HIV infection: not reported<br>Sample size included for analysis: 61<br>Setting: inpatient<br>Country: Belgium<br>World Bank Income Classification: high<br>High TB burden country: no<br>High TB/HIV burden country: no<br>Prevalence of TB cases in the study: 24.6% |
| Index tests | TB contact, CXR |

**Dreesman 2017** *(Continued)*

| | |
|---|---|
| Target condition and reference standard(s) | Active TB not specified as pulmonary.<br><br>Composite reference standard: specific criteria not defined but involved assessment of signs and symptoms, TST, CXR, and microbiological testing |
| Flow and timing | Timing between index test and reference standard not reported.<br><br>No missing data reported for the index tests or reference standards |
| 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 (All tests)** | | | |
| 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 reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | No | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |

**Dreesman 2017** *(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 receive the same reference standard? | Yes | |
| Were all patients included in the analysis? | Yes | |
| **Could the patient flow have introduced bias?** | | Low risk |

**Jaganath 2013**

| | |
|---|---|
| ***Study characteristics*** | |
| Patient Sampling | Prospective, cohort, unclear sampling strategy |
| Patient characteristics and setting | Enrolment criteria: child household contacts of adults with confirmed TB, with ≥ 1 week of contact in the last 3 months<br>Age: < 15 years, median 6 years (IQR 0–12 years)<br>Sex: 47% female<br>HIV infection: 3%<br>Sample size included for analysis: 761<br>Setting: outpatient<br>Country: Uganda<br>World Bank Income Classification: low<br>High TB burden country: no<br>High TB/HIV burden country: yes<br>Prevalence of TB cases in the study: 10.4% |
| Index tests | TB contact, weight for BMI for age z-score < –2 |
| Target condition and reference standard(s) | Active TB not specified as pulmonary.<br>Composite reference standard: defined as culture confirmation or positive response to TB therapy with ≥ 2 of fever, cough > 2 weeks, weight loss, positive TST, CXR consistent with active TB, or failure to respond to empiric antibiotics over 2 weeks. |
| Flow and timing | Timing between Index test and reference standard not reported.<br>No missing data reported for the index tests or reference standards. |
| Comparative | |

**Jaganath 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 (All tests)** | | | |
| 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 reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | No | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |
| **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 receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | Yes | | |
| **Could the patient flow have introduced bias?** | | Low risk | |

**Kruk 2008**

| ***Study characteristics*** | |
|---|---|
| Patient Sampling | Prospective, cohort, purposive |
| Patient characteristics and setting | Enrolment criteria: child household contacts of adults with confirmed TB<br>Age: < 5 years, median 30 months (range 1–60 months)<br>Sex: 44% female<br>HIV infection: not reported<br>Sample size included for analysis: 252<br>Setting: outpatient<br>Country: South Africa<br>World Bank Income Classification: upper middle<br>High TB burden country: yes<br>High TB/HIV burden country: yes<br>Prevalence of TB cases in the study: microbiological reference standard 0.7%, composite reference standard 13.1% |
| Index tests | TB contact, cough, fever, weight loss, fatigue or lethargy, CXR<br>≥ 1 of cough, fever, weight loss, or fatigue |
| Target condition and reference standard(s) | Active TB not specified as pulmonary.<br>Composite reference standard: defined as decision to treat. |
| Flow and timing | Index test and reference standard reported as occurring on the same day.<br>No missing data reported for the index tests or reference standards. |
| 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 | | |

**Kruk 2008** *(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 (All tests)** | | | |
| 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 reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | No | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |
| **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 receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | Yes | | |
| **Could the patient flow have introduced bias?** | | Low risk | |

**LaCourse 2014**

| | |
|---|---|
| ***Study characteristics*** | |
| Patient Sampling | Prospective, cross-sectional, consecutive |
| Patient characteristics and setting | Enrolment criteria: severe acute malnutrition<br>Age: < 5 years, median 1.5 years (IQR 1.0–2.1 years)<br>Sex: 51% female<br>HIV infection: 18% |

**LaCourse 2014** *(Continued)*

| | |
|---|---|
| | Sample size included for analysis: 300<br>Setting: inpatient<br>Country: Malawi<br>World Bank Income Classification: low<br>High TB burden country: no<br>High TB/HIV burden country: yes<br>Prevalence of TB cases in the study: microbiological reference standard 0.7%, composite reference standard 7.3% |
| Index tests | TB contact, cough > 1 week, fever > 1 week, fatigue or lethargy, MUAC < 11.5 cm, weight for height z-score < –3, CXR, Xpert MTB/RIF |
| Target condition and reference standard(s) | Active TB not specified as pulmonary.<br>Microbiological reference standard<br>Composite reference standard: defined as confirmed or probable TB. |
| Flow and timing | Index test and reference standard reported as occurring during a single hospital admission.<br>Minimal missing data for most index tests except for weight for height z-scores, which were only reported for those without oedema in accordance with WHO guidelines. |
| 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 (All tests)** | | | |
| Were the index test results interpreted without knowledge of the results of the reference standard? | Yes | | |

**LaCourse 2014** *(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 3: Reference Standard** | | | |
| Is the reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | No | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |
| **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 receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | Yes | | |
| **Could the patient flow have introduced bias?** | | Low risk | |

**PERCH 2019**

| | |
|---|---|
| ***Study characteristics*** | |
| Patient Sampling | Prospective, case-control with respect to pneumonia (only cases analyzed for this review), consecutive |
| Patient characteristics and setting | Enrolment criteria: WHO-defined severe or very severe pneumonia<br><br>Age: 41% 28 days to 5 months, 23% 6–11 months, 23% 12–23 months, 14% 24–59 months<br><br>Sex: 42% female<br><br>HIV infection: 0%<br><br>Sample size included for analysis: 3540<br><br>Setting: inpatient<br><br>Country: Bangladesh, The Gambia, Kenya, Mali, South Africa, Thailand, and Zambia |

**PERCH 2019** *(Continued)*

| | |
|---|---|
| | World Bank Income Classification: 'low' for The Gambia and Mali; 'low-middle' for Bangladesh, Kenya, and Zambia; 'upper-middle' for South Africa and Thailand<br><br>High TB burden country: Bangladesh, Kenya, South Africa, Thailand, and Zambia<br><br>High TB/HIV burden country: Kenya, South Africa, Thailand, and Zambia<br><br>Prevalence of TB cases in the study: 0.8% |
| Index tests | CXR |
| Target condition and reference standard(s) | Active TB not specified as pulmonary.<br><br>Microbiological reference standard: culture (unspecified type) positive |
| Flow and timing | Index test conducted and reference standard collected both upon enrolment<br><br>All reported participants had index and reference standards.<br><br>11% did not have microbiological reference standard, but they were not included in the analysis (2×2 table). |
| Comparative | |
| Notes | 11% did not have microbiological reference standard, but they were not included in the analysis (2×2 table). |

***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 (All tests)** | | | |
| 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 | |

**PERCH 2019** *(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 reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | Yes | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Low risk | |
| **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 receive the same reference standard? | No | | |
| Were all patients included in the analysis? | Yes | | |
| **Could the patient flow have introduced bias?** | | Unclear risk | |

**Portevin 2014**

| | |
|---|---|
| ***Study characteristics*** | |
| Patient Sampling | Prospective, cohort, consecutive |
| Patient characteristics and setting | Enrolment criteria: ≥ 1 of cough > 14 days, repeated fever, weight loss or poor weight gain; and signs and symptoms that suggested extrapulmonary TB<br>Age: < 15 years; median 6.1 years (IQR 2.1–10.3 years)<br>Sex: 46% female<br>HIV infection: 29%<br>Sample size included for analysis: 113<br>Setting: outpatient and inpatient<br>Country: Tanzania<br>World Bank Income Classification: low<br>High TB burden country: yes<br>High TB/HIV burden country: yes<br>Prevalence of TB cases in the study: microbiological 15.9%, composite 33.6% |

**Portevin 2014** *(Continued)*

| | |
|---|---|
| Index tests | Cough, fever, fatigue or lethargy, weight loss, weight or BMI for age z-score < –2 |
| Target condition and reference standard(s) | Active TB not specified as pulmonary.<br>Microbiological reference standard: MGIT liquid culture or LJ solid culture positive<br>Composite: microbiological diagnosis, highly probable TB, or probable TB |
| Flow and timing | Timing between index tests and reference standard not reported; all participants had reported index and reference standards. |
| Comparative | |
| Notes | Strict selection criteria raised concern that included participants did not match the review question. |

***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?** | | | High |
| **DOMAIN 2: Index Test (All tests)** | | | |
| 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 reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | Unclear | | |

**Portevin 2014** *(Continued)*

| | | | |
|---|---|---|---|
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |
| **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 receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | Yes | | |
| **Could the patient flow have introduced bias?** | | Low risk | |

**Rose 2012**

| ***Study characteristics*** | |
|---|---|
| Patient Sampling | Prospective, cohort, consecutive |
| Patient characteristics and setting | Enrolment criteria: ≥ 1 of the following symptoms – fever, cough, weight loss, or poor weight gain over ≥ 2 weeks; exposure to TB case in the last 2 years; seeking health care multiple times over the last 3 months; or weight-for-age z-score < –2<br>Age: < 15 years; mean 4.4 years (SD 3.8 years)<br>Sex: 41% female<br>HIV infection: 37%<br>Sample size included for analysis: 211<br>Setting: outpatient and inpatient<br>Country: Tanzania<br>World Bank Income Classification: low<br>High TB burden country: yes<br>High TB/HIV burden country: yes<br>Prevalence of TB cases in the study: 15.6% |
| Index tests | Weight-for-age z-score < -2; ≥ 1 of fever, cough, weight loss, or poor weight gain over ≥ 2 weeks; exposure to TB case in the last 2 years; seeking health care multiple times over the last 3 months, or weight-for-age z-score < –2 (only positives) |
| Target condition and reference standard(s) | Active TB not specified as pulmonary.<br>Composite reference standard: microbiological diagnosis (LJ solid culture positive) or highly probable TB |

**Rose 2012** *(Continued)*

| | |
|---|---|
| Flow and timing | Timing between index tests and reference standard not reported; all participants had reported index and reference standards. |
| Comparative | |
| Notes | Strict selection criteria raised concern that included participants did not match the review question.<br><br>The composite reference standard was stricter and was relatively objective, and the main index test here (weight-for-age z-score) did not obviously influence the reference standard. In this way, although it is not clearly stated that the reference standard was determined blinded to the index tests, this reference standard was less susceptible to incorporation 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?** | | | Unclear |
| **DOMAIN 2: Index Test (All tests)** | | | |
| 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 reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | Unclear | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Low risk | |

**Rose 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? | Unclear | |
| Did all patients receive the same reference standard? | Yes | |
| Were all patients included in the analysis? | Yes | |
| **Could the patient flow have introduced bias?** | | Low risk |

**Sawry 2018**

| ***Study characteristics*** | |
|---|---|
| Patient Sampling | Prospective, cohort, consecutive |
| Patient characteristics and setting | Enrolment criteria: HIV positive, initiating antiretroviral therapy<br>Age: < 9 years; median 2.1 years (IQR 0.8–4.7 years)<br>Sex: 47% female<br>HIV infection: 100%<br>Sample size included for analysis: 1346 screens on 220 children<br>Setting: outpatient<br>Country: South Africa<br>World Bank Income Classification: upper-middle<br>High TB burden country: yes<br>High TB/HIV burden country: yes<br>Prevalence of TB cases in the study: 3.2% |
| Index tests | Intensified case finding symptom screen: any 1 of current cough or fever, poor weight gain, or contact with a person with TB |
| Target condition and reference standard(s) | Active TB not specified as pulmonary.<br>Composite reference standard: microbiological diagnosis, probable TB, or possible TB |
| Flow and timing | Timing between index tests and reference standard not reported.<br>Among the 220 children serially screened, receipt of the reference standard was inconsistent because 20 opted out of the study, 17 were lost to follow-up, 13 transferred out, and 2 died. |
| Comparative | |

**Sawry 2018** *(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? | 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 (All tests)** | | | |
| 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 reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | No | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |
| **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? | No | | |
| Did all patients receive the same reference standard? | No | | |
| Were all patients included in the analysis? | No | | |
| **Could the patient flow have introduced bias?** | | High risk | |

### Schwoebel 2020

| ***Study characteristics*** | |
|---|---|
| Patient Sampling | Prospective, cohort, consecutive |
| Patient characteristics and setting | Enrolment criteria: household contact with adult smear-positive TB case, plus either any symptom, clinical sign, or radiographic finding suggestive of TB<br><br>Age: < 5 years; mean 2.6 years<br><br>Sex: 50% female<br><br>HIV infection: 1.8%<br><br>Sample size included for analysis: 1958<br><br>Setting: outpatient<br><br>Country: Benin, Burkina Faso, Cameroon, and Central African Republic<br><br>World Bank Income Classification: Cameroon is 'low-middle', the other 3 are 'low'<br><br>High TB burden country: only Central African Republic<br><br>High TB/HIV burden country: only Central African Republic<br><br>Prevalence of TB cases in the study: 2.3% |
| Index tests | Any 1 of the following within the past 4 weeks: cough, fever, weight loss, reduced appetite, or reduced playfulness; weight-for-height z-score < –3; CXR |
| Target condition and reference standard(s) | Active TB not specified as pulmonary.<br><br>Composite reference standard: microbiological diagnosis, probable TB (based on both clinical and radiological abnormalities, or possible TB (based on signs and symptoms only) |
| Flow and timing | Timing between index tests and reference standard not reported; minor missing data for index tests (highest at 6.5% for CXR). |
| Comparative | |
| Notes | TB close contacts without signs, symptoms, or radiographic findings not suggestive of TB notably did not receive microbiological testing as part of the composite reference. |

| ***Methodological quality*** | | | |
|---|---|---|---|
| **Item** | **Authors' judgement** | **Risk of bias** | **Applicability concerns** |
| **DOMAIN 1: Patient Selection** | | | |
| Was a consecutive or random sample of patients enrolled? | Yes | | |

**Schwoebel 2020** *(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 (All tests)** | | | |
| 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 reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | No | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |
| **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 receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | Yes | | |
| **Could the patient flow have introduced bias?** | | Low risk | |

**Tieu 2014**

| | |
|---|---|
| ***Study characteristics*** | |
| Patient Sampling | Prospective, cohort, consecutive |
| Patient characteristics and setting | Enrolment criteria: close contact with adult TB cases within past year |

**Tieu 2014** *(Continued)*

| | |
|---|---|
| | Age: < 15 years; mean 7.2 years (SD 4.2 years)<br>Sex: 48% female<br>HIV infection: 1.9%<br>Sample size included for analysis: 158<br>Setting: outpatient<br>Country: Thailand<br>World Bank Income Classification: upper-middle<br>High TB burden country: yes<br>High TB/HIV burden country: yes<br>Prevalence of TB cases in the study: microbiological 4.4%, composite 13.3% |
| Index tests | TB contact (only positives); current fever; weight-for-age or BMI-for-age z-score < –2; weight loss or poor weight gain; CXR |
| Target condition and reference standard(s) | Active TB not specified as pulmonary.<br>Composite reference standard: definite, probable, and possible TB |
| Flow and timing | Index test and reference standard reported as occurring on the same day; all participants had reported index and reference standards. |
| 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 (All tests)** | | | |
| Were the index test results interpreted without knowledge of the results of the reference standard? | Yes | | |

**Tieu 2014** *(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 3: Reference Standard** | | | |
| Is the reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | No | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |
| **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 receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | Yes | | |
| **Could the patient flow have introduced bias?** | | Low risk | |

**Togun 2015**

| | |
|---|---|
| ***Study characteristics*** | |
| Patient Sampling | Prospective, cross-sectional, consecutive |
| Patient characteristics and setting | Enrolment criteria: household contact of adult smear-positive TB cases<br>Age: < 15 years; median 6 years (IQR 3–9 years)<br>Sex: 47% female<br>HIV infection: 0%<br>Sample size included for analysis: 487<br>Setting: outpatient<br>Country: The Gambia<br>World Bank Income Classification: low<br>High TB burden country: no |

**Togun 2015** *(Continued)*

| | |
|---|---|
| | High TB/HIV burden country: no<br><br>Prevalence of TB cases in the study: microbiological 4.4%, composite 12.9% |
| Index tests | Cough or fever > 1 week, or both; BMI-for-age z-score < –2; CXR; Xpert MTB/RIF |
| Target condition and reference standard(s) | Active TB not specified as pulmonary.<br><br>Microbiological: both MGIT liquid culture system and LJ solid culture<br><br>Composite reference standard: microbiological diagnosis or clinical diagnosis, defined as: suggestive appearance on chest radiograph; and either favourable response to specific anti-TB therapy, positive TST, or suggestive histological appearances on biopsy material. |
| Flow and timing | Index test and reference standard reported as occurring on the same day; all participants had reported index and reference standards. |
| 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 (All tests)** | | | |
| 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 |

**Togun 2015** *(Continued)*

| | | | |
|---|---|---|---|
| **DOMAIN 3: Reference Standard** | | | |
| Is the reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | No | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |
| **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 receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | Yes | | |
| **Could the patient flow have introduced bias?** | | Low risk | |

**Togun 2016**

| | |
|---|---|
| ***Study characteristics*** | |
| Patient Sampling | Prospective, cohort, consecutive |
| Patient characteristics and setting | Enrolment criteria: household contact of adult smear-positive TB cases<br>Age: < 15 years; median 6 years (IQR 3–9 years)<br>Sex: 49% female<br>HIV infection: 0%<br>Sample size included for analysis: 150<br>Setting: outpatient<br>Country: The Gambia<br>World Bank Income Classification: low<br>High TB burden country: no<br>High TB/HIV burden country: no<br>Prevalence of TB cases in the study: 23.3% |
| Index tests | Cough > 2 weeks; weight loss; BMI-for-age z-score < –2; fatigue; night sweats; fever; CXR |
| Target condition and reference standard(s) | Active TB not specified as pulmonary. |

**Togun 2016** *(Continued)*

| | |
|---|---|
| | Composite reference standard: microbiological diagnosis or clinical diagnosis, defined as: suggestive appearance on chest radiograph; and favourable response to specific antituberculosis therapy or positive TST (or both) or suggestive histological appearances on biopsy material |
| Flow and timing | Index test and reference standard reported as occurring on the same day; all participants had reported index and reference standards. |
| 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 (All tests)** | | | |
| 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 reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | No | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |

**Togun 2016** *(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 receive the same reference standard? | Yes | |
| Were all patients included in the analysis? | Yes | |
| **Could the patient flow have introduced bias?** | | Low risk |

**Triasih 2015a**

| | |
|---|---|
| ***Study characteristics*** | |
| Patient Sampling | Cohort, prospective, consecutive |
| Patient characteristics and setting | Enrolment criteria: close contact of adult pulmonary TB cases<br>Age: < 15 years; median 6 years (IQR 3–10 years)<br>Sex: NR<br>HIV infection: NR<br>Sample size included for analysis: 265<br>Setting: outpatient<br>Country: Indonesia<br>World Bank Income Classification: lower middle<br>High TB burden country: yes<br>High TB/HIV burden country: yes<br>Prevalence of TB cases in the study: 7.9% |
| Index tests | Any 1 of: persistent cough, fever, weight loss, or failure to thrive |
| Target condition and reference standard(s) | Active TB not specified as pulmonary.<br>Composite reference standard: defined as certain, probable, or possible TB ('possible' defined as "at least one of the well-defined symptoms and either of the following: a positive clinical response to anti-TB treatment OR chest radiography was consistent with intrathoracic TB") |
| Flow and timing | Index test and reference standard reported as occurring on the same day; all participants have reported index and reference standards |
| Comparative | |

**Triasih 2015a** *(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? | 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 (All tests)** | | | |
| 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 reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | No | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |
| **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 receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | Yes | | |
| **Could the patient flow have introduced bias?** | | Low risk | |

### Triasih 2015b

| ***Study characteristics*** | |
|---|---|
| Patient Sampling | Cohort, prospective, consecutive |
| Patient characteristics and setting | Enrolment criteria: close contact of pulmonary TB cases<br>Age: < 15 years; median 6 years (IQR 3–10 years)<br>Sex: NR<br>HIV infection: NR<br>Sample size included for analysis: 265<br>Setting: community (household)<br>Country: Indonesia<br>World Bank Income Classification: lower middle<br>High TB burden country: yes<br>High TB/HIV burden country: yes<br>Prevalence of TB cases in the study: 7.9% |
| Index tests | CXR |
| Target condition and reference standard(s) | Active TB: not defined as pulmonary a priori but all diagnosed cases at least had pulmonary disease.<br>Composite reference standard: defined as certain, probable, or possible TB (possible defined as "at least one of the well-defined symptoms and either of the following: a positive clinical response to anti-TB treatment OR chest radiography was consistent with intrathoracic TB") |
| Flow and timing | Time between index test and reference standard not reported; of 269 eligible participants, 265 received the index and reference standards. |
| Comparative | |
| Notes | None of the 21 children diagnosed with active TB had microbiological confirmation. Chest radiographs were reviewed by 4 readers, and data were only extracted for the reader with the highest agreement with the other 3. |

| ***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 | | |

**Triasih 2015b** *(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 (All tests)** | | | |
| 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 reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | No | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |
| **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 receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | Yes | | |
| **Could the patient flow have introduced bias?** | | Unclear risk | |

**Ustero 2017**

| | |
|---|---|
| ***Study characteristics*** | |
| Patient Sampling | Cross-sectional, prospective, consecutive |
| Patient characteristics and setting | Enrolment criteria: household contacts of microbiologically confirmed paediatric TB cases<br><br>Age: < 20 years; mean 11.9 years (SD 7.9 years) |

**Ustero 2017** *(Continued)*

| | |
|---|---|
| | Sex: 62.5% female<br>HIV infection: 16.7%<br>Sample size included for analysis: 24<br>Setting: community (household)<br>Country: Eswatini<br>World Bank Income Classification: lower middle<br>High TB burden country: no<br>High TB/HIV burden country: yes<br>Prevalence of TB cases in the study: 4.5% |
| Index tests | Any cough, fever, night sweats, or weight loss |
| Target condition and reference standard(s) | Pulmonary TB, microbiological reference standard |
| Flow and timing | Time between index test and reference standard not reported, portion of participants with negative index test (< 10%) did not receive the reference standard. |
| 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 (All tests)** | | | |
| 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 | |

**Ustero 2017** *(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 reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | Unclear | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Low risk | |
| **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 receive the same reference standard? | No | | |
| Were all patients included in the analysis? | Yes | | |
| **Could the patient flow have introduced bias?** | | High risk | |

**Vonasek 2021**

| | |
|---|---|
| ***Study characteristics*** | |
| Patient Sampling | Cohort, consecutive, retrospective |
| Patient characteristics and setting | Enrolment criteria: children and adolescents living with HIV receiving routine outpatient HIV care<br><br>Age: < 20 years; median 11.2 years (IQR 6.9–15.0 years)<br><br>Sex: 50% female<br><br>HIV infection: 100%<br><br>Sample size included for analysis: 240,161 screens on 20,706 participants<br><br>Setting: outpatient<br><br>Country: Botswana, Eswatini, Lesotho, Malawi, Tanzania (2 sites), Uganda<br><br>World Bank Income Classification: 'low' for Malawi, Tanzania and Uganda; 'low-middle' for Eswatini and Lesotho; 'upper-middle' for Botswana<br><br>High TB burden country: Lesotho, Tanzania |

**Vonasek 2021** *(Continued)*

| | |
|---|---|
| | High TB/HIV burden country: Botswana, Eswatini, Lesotho, Malawi, Tanzania, Uganda<br><br>Prevalence of TB cases in the study: 7.7% |
| Index tests | Children: any current fever, cough, poor weight gain, or recent TB contact; adolescents: any current fever, cough, night sweats, or weight loss |
| Target condition and reference standard(s) | Target condition: active TB not specified as pulmonary.<br><br>Reference standard: composite, clinician decision based on clinical signs and symptoms, radiograph imaging, and Xpert MTB/RIF testing |
| Flow and timing | Timing between index tests and reference standard not reported.<br><br>For those diagnosis of TB disease, only the single symptom screen nearest the date of diagnosis was analyzed, and all other future and previous screens were excluded. |
| 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 (All tests)** | | | |
| 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** | | | |

**Vonasek 2021** *(Continued)*

| | | | |
|---|---|---|---|
| Is the reference standards likely to correctly classify the target condition? | Yes | | |
| Were the reference standard results interpreted without knowledge of the results of the index tests? | Unclear | | |
| **Could the reference standard, its conduct, or its interpretation have introduced bias?** | | Unclear risk | |
| **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 receive the same reference standard? | Yes | | |
| Were all patients included in the analysis? | No | | |
| **Could the patient flow have introduced bias?** | | High risk | |

BMI: body mass index; CXR: chest radiography; IQR: interquartile range; LJ: Löwenstein-Jensen; MGIT: Mycobacterium Growth Indicator Tube; MUAC: mid-upper arm circumference; SD: standard deviation; TB: tuberculosis; TST: tuberculin skin test; WHO: World Health Organization.

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

| Study | Reason for exclusion |
|---|---|
| Ackerman 2010 | Ineligible study design. |
| Aldridge 2016 | Data not available for age group of interest. |
| Alekseev 2018 | Data not available for age group of interest. |
| Armstrong-Hough 2017 | Diagnostic study. |
| Auld 2013 | Ineligible study design. |
| Azit 2019 | No eligible reference test(s). |
| Bamford 2010 | Ineligible study design. |
| Basta 2010 | No eligible reference test(s). |
| Bennet 2017 | No eligible index test(s). |
| Bonnet 2017 | No eligible reference test(s). |
| Bosa 2017 | No eligible reference test(s). |
| Boullier 2017 | No eligible index test(s). |

| Study | Reason for exclusion |
|---|---|
| Chiappini 2019 | Data not available for age group of interest. |
| Chisti 2014 | Diagnostic study. |
| Curtis 1999 | No eligible index test(s). |
| David 2017 | Diagnostic study. |
| de Lima 2013 | No eligible index test(s). |
| Do Nascimento Maia 2016 | No eligible index test(s). |
| Dorjee 2019 | Data not available for age group of interest. |
| Driver 2002 | No eligible index test(s). |
| Egere 2017 | No eligible index test(s). |
| Faccini 2013 | No eligible index test(s). |
| Fortunato 2011 | No eligible reference test(s). |
| Francis 2002 | No eligible index test(s). |
| Galli 2016 | No eligible index test(s). |
| Gashu 2016 | No eligible reference test(s). |
| Girardi 2007 | No eligible index test(s). |
| Gomez-Pastrana 1999 | No eligible index test(s). |
| Gwee 2013 | Ineligible study design. |
| Hanrahan 2019 | Ineligible study design. |
| Hoffman 1996 | No eligible index test(s). |
| Huang 2016 | No eligible index test(s). |
| Izumi 2017 | No eligible index test(s). |
| Karki 2017 | No eligible index test(s). |
| Kemigisha 2015 | No eligible index test(s). |
| Kim 2017 | Ineligible study design. |
| Kondo 2003 | No eligible index test(s). |
| Lee 2008 | No eligible index test(s). |
| Leung 2006 | No eligible index test(s). |
| Li 2015 | Diagnostic study. |

| Study | Reason for exclusion |
|---|---|
| Ling 2013 | Diagnostic study. |
| Mahomed 2013 | Data not available for age group of interest. |
| Malik 2018 | Reference standard not applied to those with negative screens. |
| Marais 2006 | Diagnostic study. |
| Marais 2009b | No eligible index test(s). |
| Marcy 2019 | Diagnostic study. |
| Masur 2017 | No eligible reference test(s). |
| Minhas 2017 | No eligible reference test(s). |
| Moran-Mendoza 2010 | No eligible index test(s). |
| Mueller-Hermelink 2018 | Ineligible study design. |
| Murray 2019 | No eligible index test(s). |
| Nduba 2018 | No eligible index test(s). |
| Ntinginya 2012 | No eligible index test(s). |
| Oh 2018 | No eligible reference test(s). |
| Padmapriyadarsini 2016 | No eligible reference test(s). |
| Pan 2019 | No eligible index test(s). |
| Penin 2007 | No eligible index test(s). |
| Penn-Nicholson 2019 | Ineligible study design. |
| Puryear 2013 | No eligible reference test(s). |
| Rachow 2012 | Diagnostic study. |
| Ramirez 2006 | Diagnostic study. |
| Rossoni 2020 | Diagnostic study. |
| Salinas 2002 | No eligible index test(s). |
| Saunders 2014 | No eligible index test(s). |
| Shah 2008 | No eligible reference test(s). |
| Shaikh 2017 | Diagnostic study. |
| Sollai 2017 | No eligible index test(s). |
| Spyridis 2003 | No eligible index test(s). |

| Study | Reason for exclusion |
|---|---|
| Swingler 2000 | No eligible reference test(s). |
| Szkwarko 2018 | No eligible index test(s). |
| Thee 2019 | No eligible index test(s). |
| van Schalkwyk 2014 | No eligible index test(s). |
| Verver 2005 | No included index test(s). |
| Williams 2016 | No eligible index test(s). |
| Williams 2019 | No eligible index test(s). |
| Yang 2018 | No eligible index test(s). |
| Yuan 1995 | No eligible index test(s). |
| Zachariah 2003 | No eligible reference test(s). |

## 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 One or more of cough, fever, or poor weight gain, close tuberculosis (TB) contacts, composite | 4 | 2695 |
| 2 One or more of cough, fever, or decreased playfulness; < 5 years of age (y/o) inpatient or outpatient, composite | 3 | 2445 |
| 3 World Health Organization 4-symptom screen, outpatients living with HIV, composite | 2 | 203135 |
| 4 Chest radiograph (CXR) abnormal, close TB contacts, composite | 8 | 3513 |
| 5 CXR suggestive, close TB contacts, composite | 4 | 2550 |
| 6 CXR suggestive, < 5 y/o inpatient or outpatient, composite | 3 | 2388 |
| 7 CXR abnormal, < 5 y/o hospitalized with pneumonia, microbiological | 1 | 3540 |
| 8 Weight or body mass index (BMI) for age z-score < –2, close TB contacts, composite | 3 | 1399 |
| 9 Weight or BMI for age z-score < –2, inpatient or outpatient, composite | 5 | 1723 |
| 10 Weight or BMI for age z-score < –2, inpatient or outpatient, microbiological | 2 | 561 |

| Test | No. of studies | No. of participants |
|---|---|---|
| 11 Xpert MTB/RIF, inpatient or outpatient, microbiological | 2 | 787 |
| 12 Xpert MTB/RIF, inpatient or outpatient, composite | 2 | 787 |
| 13 Current cough, < 15 y/o, microbiological | 1 | 113 |
| 14 Cough > 1 week, < 5 y/o, microbiological | 1 | 300 |
| 15 Cough > 3 weeks, < 5 y/o, microbiological | 0 | 0 |
| 16 Cough > 3 weeks, < 15 y/o, microbiological | 0 | 0 |
| 17 Cough > 4 weeks, < 5 y/o, microbiological | 0 | 0 |
| 18 Cough > 4 weeks, < 15 y/o, microbiological | 0 | 0 |
| 19 Cough > 3 weeks, < 15 y/o, and HIV+, microbiological | 0 | 0 |
| 20 Cough > 4 weeks, < 15 y/o, and HIV+, microbiological | 0 | 0 |
| 21 Any cough, < 15 y/o, microbiological | 2 | 413 |
| 22 Current cough, < 5 y/o, composite | 1 | 252 |
| 23 Current cough, < 15 y/o, composite | 1 | 113 |
| 24 Cough > 1 week, < 5 y/o, composite | 1 | 300 |
| 25 Cough > 2 weeks, < 15 y/o, composite | 1 | 150 |
| 26 Any cough, < 15 y/o, composite | 4 | 815 |
| 27 TB contact, < 5 y/o, microbiological | 1 | 300 |
| 28 TB contact, < 15 y/o, microbiological | 1 | 300 |
| 29 TB contact, < 20 y/o, microbiological | 1 | 300 |
| 30 TB contact, < 15 y/o, and HIV+, microbiological | 0 | 0 |
| 31 TB contact, < 5 y/o, composite | 1 | 300 |
| 32 TB contact, < 20 y/o, composite | 1 | 300 |
| 33 Current fever, < 5 y/o, microbiological | 0 | 0 |
| 34 Current fever, < 15 y/o, microbiological | 2 | 413 |
| 35 Fever > 1 week, < 5 y/o, microbiological | 1 | 300 |
| 36 Fever, < 15 y/o, and HIV+, microbiological | 0 | 0 |
| 37 Current fever, < 5 y/o, composite | 2 | 552 |
| 38 Current fever, < 15 y/o, composite | 5 | 973 |

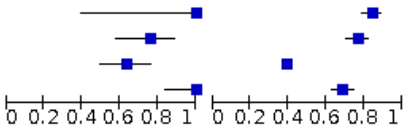
| Test | No. of studies | No. of participants |
|---|---|---|
| 39 Fever > 1 week, < 5 y/o, composite | 1 | 300 |
| 40 Weight for height z-score < –3, < 5 y/o, microbiological | 1 | 127 |
| 41 Weight for height z-score < –3, < 5 y/o, composite | 2 | 1985 |
| 42 Severe malnutrition, < 5 y/o, composite | 1 | 300 |
| 43 Severe malnutrition, < 5 y/o, microbiological | 1 | 300 |
| 44 Weight loss or poor weight gain, < 5 y/o, microbiological | 0 | 0 |
| 45 Weight loss or poor weight gain, < 15 y/o, microbiological | 1 | 113 |
| 46 Weight loss or poor weight gain, < 20 y/o, microbiological | 1 | 113 |
| 47 Weight loss or poor weight gain, < 15 y/o, and HIV+, microbiological | 0 | 0 |
| 48 Weight loss or poor weight gain, < 5 y/o, composite | 1 | 252 |
| 49 Weight loss or poor weight gain, < 15 y/o, composite | 4 | 673 |
| 50 Weight loss or poor weight gain, < 20 y/o, composite | 4 | 673 |
| 51 Fatigue or lethargy, < 5 y/o, microbiological | 1 | 299 |
| 52 Fatigue or lethargy, < 15 y/o, microbiological | 2 | 412 |
| 53 Fatigue or lethargy, < 15 y/o, and HIV+, microbiological | 0 | 0 |
| 54 Fatigue or lethargy, < 5 y/o, composite | 2 | 551 |
| 55 Fatigue or lethargy, < 15 y/o, composite | 4 | 814 |
| 56 Fatigue or lethargy, < 20 y/o, composite | 4 | 814 |
| 57 Night sweats, < 5 y/o, microbiological | 0 | 0 |
| 58 Night sweats, < 15 y/o, microbiological | 0 | 0 |
| 59 Night sweats, < 15 y/o, and HIV+, microbiological | 0 | 0 |
| 60 Night sweats, < 15 y/o, composite | 1 | 150 |
| 61 CXR abnormal, < 15 y/o, microbiological | 1 | 482 |
| 62 CXR suggestive, < 5 y/o, microbiological | 1 | 299 |
| 63 CXR suggestive, < 15 y/o, microbiological | 1 | 299 |
| 64 CXR abnormal, < 15 y/o, composite | 5 | 1113 |
| 65 CXR suggestive, < 15 y/o, composite | 5 | 2850 |
| 66 Xpert MTB/Rif, < 5 y/o, microbiological | 1 | 300 |

| Test | No. of studies | No. of participants |
|---|---|---|
| 67 Xpert MTB/Rif, < 5 y/o, composite | 1 | 300 |
| 68 One of multiple symptoms, < 5 y/o, microbiological | 1 | 235 |
| 69 One of multiple symptoms, < 15 y/o, microbiological | 3 | 740 |
| 70 One of multiple symptoms, < 20 y/o, microbiological | 3 | 744 |
| 71 One of multiple symptoms, < 5 y/o, composite | 4 | 2553 |
| 72 One of multiple symptoms, < 15 y/o, composite | 7 | 4760 |
| 73 One of multiple symptoms, < 20 y/o, composite | 7 | 4760 |
| 74 Any cough, < 15 y/o, contact tracing, composite | 2 | 402 |
| 75 Current fever, < 15 y/o, contact tracing, composite | 3 | 560 |
| 76 Weight loss or poor weight gain, < 20 y/o, contact tracing, composite | 3 | 560 |
| 77 CXR abnormal, < 15 y/o, contact tracing, composite | 4 | 963 |
| 78 CXR suggestive, < 5 y/o, contact tracing, composite | 2 | 2089 |
| 79 One of multiple symptoms, < 15 y/o, contact tracing, composite | 5 | 3182 |
| 80 TB contact, < 20 y/o in inpatient or outpatient settings, microbiological | 1 | 300 |
| 81 Weight loss or poor weight gain, < 20 y/o in inpatient or outpatient settings, microbiological | 1 | 113 |
| 82 One of multiple symptoms, < 20 y/o in inpatient or outpatient settings, microbiological | 2 | 722 |
| 83 Weight loss or poor weight gain, < 20 y/o in inpatient or outpatient settings, composite | 4 | 673 |
| 84 Fatigue or lethargy, < 20 y/o in inpatient or outpatient settings, composite | 4 | 814 |
| 85 CXR abnormal, < 15 y/o, contact tracing, composite | 4 | 963 |
| 86 One of multiple symptoms, < 15 y/o in inpatient or outpatient settings, composite | 7 | 4760 |
| 87 Mid-upper arm circumference (MUAC) < 11.5 cm, < 5 y/o, microbiological | 1 | 300 |
| 88 MUAC < 11.5 cm, < 5 y/o, composite | 1 | 300 |
| 89 CXR abnormal, < 15 y/o in community, composite | 1 | 265 |
| 90 One of cough, fever, or decreased playfulness; < 15 y/o in inpatient or outpatient settings, composite | 4 | 2661 |
| 91 One of cough, fever, or decreased playfulness; < 15 y/o, contact tracing, composite | 3 | 2426 |

### Test 1. One or more of cough, fever, or poor weight gain, close tuberculosis (TB) contacts, composite

One or more of cough, fever, or poor weight gain, close tuberculosis (TB) contacts, composite

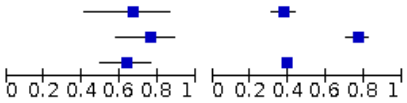
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Birungi 2018 | 4 | 33 | 0 | 179 | 1.00 [0.40, 1.00] | 0.84 [0.79, 0.89] | | |
| Kruk 2008 | 25 | 51 | 8 | 168 | 0.76 [0.58, 0.89] | 0.77 [0.71, 0.82] | | |
| Schwoebel 2020 | 35 | 1150 | 20 | 753 | 0.64 [0.50, 0.76] | 0.40 [0.37, 0.42] | | |
| Triasih 2015a | 21 | 77 | 0 | 171 | 1.00 [0.84, 1.00] | 0.69 [0.63, 0.75] | | |




### Test 2. One or more of cough, fever, or decreased playfulness; < 5 years of age (y/o) inpatient or outpatient, composite

One or more of cough, fever, or decreased playfulness; < 5 years of age (y/o) inpatient or outpatient, composite

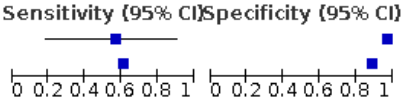
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Aggerbeck 2018 | 12 | 136 | 6 | 81 | 0.67 [0.41, 0.87] | 0.37 [0.31, 0.44] | | |
| Kruk 2008 | 25 | 51 | 8 | 168 | 0.76 [0.58, 0.89] | 0.77 [0.71, 0.82] | | |
| Schwoebel 2020 | 35 | 1150 | 20 | 753 | 0.64 [0.50, 0.76] | 0.40 [0.37, 0.42] | | |




### Test 3. World Health Organization 4-symptom screen, outpatients living with HIV, composite

World Health Organization 4-symptom screen, outpatients living with HIV, composite

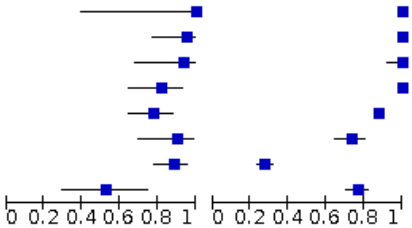
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Sawry 2018 | 4 | 41 | 3 | 1295 | 0.57 [0.18, 0.90] | 0.97 [0.96, 0.98] | | |
| Vonasek 2021 | 742 | 22481 | 470 | 178099 | 0.61 [0.58, 0.64] | 0.89 [0.89, 0.89] | | |




### Test 4. Chest radiograph (CXR) abnormal, close TB contacts, composite

Chest radiograph (CXR) abnormal, close TB contacts, composite

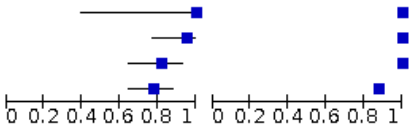
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Birungi 2018 | 4 | 0 | 0 | 212 | 1.00 [0.40, 1.00] | 1.00 [0.98, 1.00] | | |
| Clemente 2017 | 21 | 0 | 1 | 224 | 0.95 [0.77, 1.00] | 1.00 [0.98, 1.00] | | |
| Dreesman 2017 | 14 | 0 | 1 | 46 | 0.93 [0.68, 1.00] | 1.00 [0.92, 1.00] | | |
| Kruk 2008 | 27 | 0 | 6 | 219 | 0.82 [0.65, 0.93] | 1.00 [0.98, 1.00] | | |
| Schwoebel 2020 | 42 | 223 | 12 | 1559 | 0.78 [0.64, 0.88] | 0.87 [0.86, 0.89] | | |
| Tieu 2014 | 19 | 36 | 2 | 98 | 0.90 [0.70, 0.99] | 0.73 [0.65, 0.80] | | |
| Togun 2016 | 55 | 303 | 7 | 117 | 0.89 [0.78, 0.95] | 0.28 [0.24, 0.32] | | |
| Triasih 2015b | 11 | 57 | 10 | 187 | 0.52 [0.30, 0.74] | 0.77 [0.71, 0.82] | | |



### Test 5. CXR suggestive, close TB contacts, composite

CXR suggestive, close TB contacts, composite

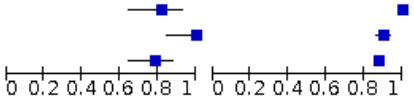
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Birungi 2018 | 4 | 0 | 0 | 212 | 1.00 [0.40, 1.00] | 1.00 [0.98, 1.00] | | |
| Clemente 2017 | 21 | 0 | 1 | 224 | 0.95 [0.77, 1.00] | 1.00 [0.98, 1.00] | | |
| Kruk 2008 | 27 | 0 | 6 | 219 | 0.82 [0.65, 0.93] | 1.00 [0.98, 1.00] | | |
| Schwoebel 2020 | 42 | 223 | 12 | 1559 | 0.78 [0.64, 0.88] | 0.87 [0.86, 0.89] | | |




### Test 6. CXR suggestive, < 5 y/o inpatient or outpatient, composite

CXR suggestive, < 5 y/o inpatient or outpatient, composite

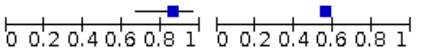
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Kruk 2008 | 27 | 0 | 6 | 219 | 0.82 [0.65, 0.93] | 1.00 [0.98, 1.00] | | |
| LaCourse 2014 | 22 | 26 | 0 | 251 | 1.00 [0.85, 1.00] | 0.91 [0.87, 0.94] | | |
| Schwoebel 2020 | 43 | 223 | 12 | 1559 | 0.78 [0.65, 0.88] | 0.87 [0.86, 0.89] | | |




### Test 7. CXR abnormal, < 5 y/o hospitalized with pneumonia, microbiological

CXR abnormal, < 5 y/o hospitalized with pneumonia, microbiological

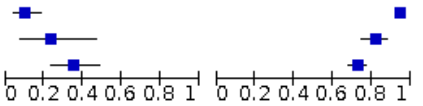
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| PERCH 2019 | 24 | 1547 | 4 | 1965 | 0.86 [0.67, 0.96] | 0.56 [0.54, 0.58] | | |




### Test 8. Weight or body mass index (BMI) for age z-score < –2, close TB contacts, composite

Weight or body mass index (BMI) for age z-score < -2, close TB contacts, composite

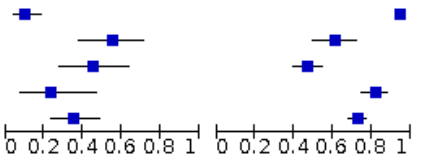
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Jaganath 2013 | 8 | 40 | 71 | 642 | 0.10 [0.04, 0.19] | 0.94 [0.92, 0.96] | | |
| Tieu 2014 | 5 | 25 | 16 | 112 | 0.24 [0.08, 0.47] | 0.82 [0.74, 0.88] | | |
| Togun 2015 | 22 | 115 | 40 | 303 | 0.35 [0.24, 0.49] | 0.72 [0.68, 0.77] | | |




### Test 9. Weight or BMI for age z-score < –2, inpatient or outpatient, composite

Weight or BMI for age z-score < -2, inpatient or outpatient, composite

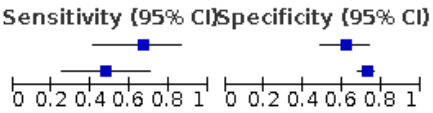
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Jaganath 2013 | 8 | 40 | 71 | 642 | 0.10 [0.04, 0.19] | 0.94 [0.92, 0.96] | | |
| Portevin 2014 | 21 | 29 | 17 | 46 | 0.55 [0.38, 0.71] | 0.61 [0.49, 0.72] | | |
| Rose 2012 | 15 | 94 | 18 | 84 | 0.45 [0.28, 0.64] | 0.47 [0.40, 0.55] | | |
| Tieu 2014 | 5 | 25 | 16 | 112 | 0.24 [0.08, 0.47] | 0.82 [0.74, 0.88] | | |
| Togun 2015 | 22 | 115 | 40 | 303 | 0.35 [0.24, 0.49] | 0.72 [0.68, 0.77] | | |



### Test 10. Weight or BMI for age z-score < –2, inpatient or outpatient, microbiological

Weight or BMI for age z-score < -2, inpatient or outpatient, microbiological

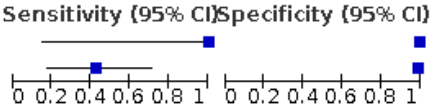
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Portevin 2014 | 12 | 24 | 6 | 39 | 0.67 [0.41, 0.87] | 0.62 [0.49, 0.74] |
| Togun 2015 | 10 | 127 | 11 | 332 | 0.48 [0.26, 0.70] | 0.72 [0.68, 0.76] |




### Test 11. Xpert MTB/RIF, inpatient or outpatient, microbiological

Xpert MTB/RIF, inpatient or outpatient, microbiological

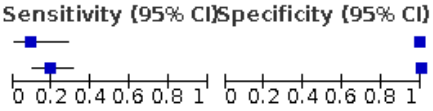
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| LaCourse 2014 | 2 | 1 | 0 | 297 | 1.00 [0.16, 1.00] | 1.00 [0.98, 1.00] |
| Togun 2015 | 6 | 6 | 8 | 467 | 0.43 [0.18, 0.71] | 0.99 [0.97, 1.00] |




### Test 12. Xpert MTB/RIF, inpatient or outpatient, composite

Xpert MTB/RIF, inpatient or outpatient, composite

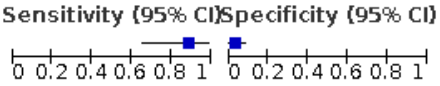
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| LaCourse 2014 | 2 | 1 | 20 | 277 | 0.09 [0.01, 0.29] | 1.00 [0.98, 1.00] |
| Togun 2015 | 12 | 0 | 50 | 425 | 0.19 [0.10, 0.31] | 1.00 [0.99, 1.00] |




### Test 13. Current cough, < 15 y/o, microbiological

Current cough, < 15 y/o, microbiological

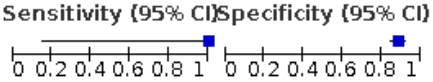
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Portevin 2014 | 16 | 92 | 2 | 3 | 0.89 [0.65, 0.99] | 0.03 [0.01, 0.09] |




### Test 14. Cough > 1 week, < 5 y/o, microbiological

Cough > 1 week, < 5 y/o, microbiological

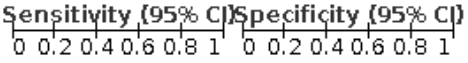
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| LaCourse 2014 | 2 | 34 | 0 | 264 | 1.00 [0.16, 1.00] | 0.89 [0.84, 0.92] |




### Test 15. Cough > 3 weeks, < 5 y/o, microbiological

Cough > 3 weeks, < 5 y/o, microbiological

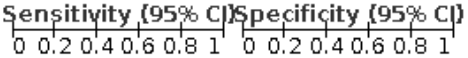
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|



### Test 16. Cough > 3 weeks, < 15 y/o, microbiological

Cough > 3 weeks, < 15 y/o, microbiological

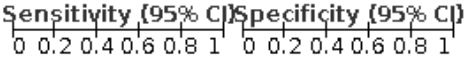
Study TP FP FN TN Sensitivity (95% CI) Specificity (95% CI) Sensitivity (95% CI) Specificity (95% CI)




### Test 17. Cough > 4 weeks, < 5 y/o, microbiological

Cough > 4 weeks, < 5 y/o, microbiological

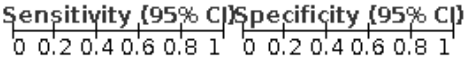
Study TP FP FN TN Sensitivity (95% CI) Specificity (95% CI) Sensitivity (95% CI) Specificity (95% CI)




### Test 18. Cough > 4 weeks, < 15 y/o, microbiological

Cough > 4 weeks, < 15 y/o, microbiological

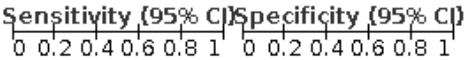
Study TP FP FN TN Sensitivity (95% CI) Specificity (95% CI) Sensitivity (95% CI) Specificity (95% CI)




### Test 19. Cough > 3 weeks, < 15 y/o, and HIV+, microbiological

Cough > 3 weeks, < 15 y/o, and HIV+, microbiological

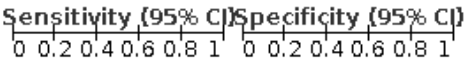
Study TP FP FN TN Sensitivity (95% CI) Specificity (95% CI) Sensitivity (95% CI) Specificity (95% CI)




### Test 20. Cough > 4 weeks, < 15 y/o, and HIV+, microbiological

Cough > 4 weeks, < 15 y/o, and HIV+, microbiological

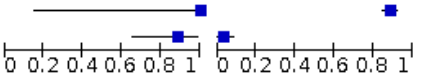
Study TP FP FN TN Sensitivity (95% CI) Specificity (95% CI) Sensitivity (95% CI) Specificity (95% CI)




### Test 21. Any cough, < 15 y/o, microbiological

Any cough, < 15 y/o, microbiological

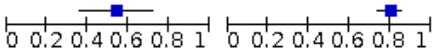
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 2 | 34 | 0 | 264 | 1.00 [0.16, 1.00] | 0.89 [0.84, 0.92] | | |
| Portevin 2014 | 16 | 92 | 2 | 3 | 0.89 [0.65, 0.99] | 0.03 [0.01, 0.09] | | |



### Test 22. Current cough, < 5 y/o, composite

Current cough, < 5 y/o, composite

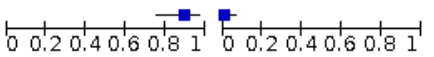
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Kruk 2008 | 18 | 44 | 15 | 175 | 0.55 [0.36, 0.72] | 0.80 [0.74, 0.85] | | |




### Test 23. Current cough, < 15 y/o, composite

Current cough, < 15 y/o, composite

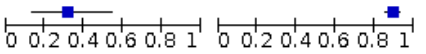
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Portevin 2014 | 34 | 74 | 4 | 1 | 0.89 [0.75, 0.97] | 0.01 [0.00, 0.07] | | |




### Test 24. Cough > 1 week, < 5 y/o, composite

Cough > 1 week, < 5 y/o, composite

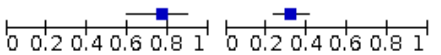
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 7 | 29 | 15 | 249 | 0.32 [0.14, 0.55] | 0.90 [0.85, 0.93] | | |




### Test 25. Cough > 2 weeks, < 15 y/o, composite

Cough > 2 weeks, < 15 y/o, composite

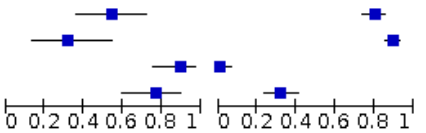
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Togun 2016 | 27 | 78 | 8 | 37 | 0.77 [0.60, 0.90] | 0.32 [0.24, 0.42] | | |




### Test 26. Any cough, < 15 y/o, composite

Any cough, < 15 y/o, composite

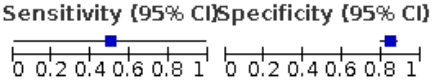
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Kruk 2008 | 18 | 44 | 15 | 175 | 0.55 [0.36, 0.72] | 0.80 [0.74, 0.85] | | |
| LaCourse 2014 | 7 | 29 | 15 | 249 | 0.32 [0.14, 0.55] | 0.90 [0.85, 0.93] | | |
| Portevin 2014 | 34 | 74 | 4 | 1 | 0.89 [0.75, 0.97] | 0.01 [0.00, 0.07] | | |
| Togun 2016 | 27 | 78 | 8 | 37 | 0.77 [0.60, 0.90] | 0.32 [0.24, 0.42] | | |




### Test 27. TB contact, < 5 y/o, microbiological

TB contact, < 5 y/o, microbiological

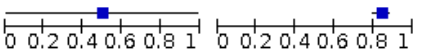
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 1 | 47 | 1 | 251 | 0.50 [0.01, 0.99] | 0.84 [0.80, 0.88] | | |



### Test 28. TB contact, < 15 y/o, microbiological

TB contact, < 15 y/o, microbiological

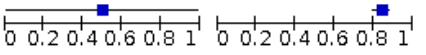
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 1 | 47 | 1 | 251 | 0.50 [0.01, 0.99] | 0.84 [0.80, 0.88] | | |




### Test 29. TB contact, < 20 y/o, microbiological

TB contact, < 20 y/o, microbiological

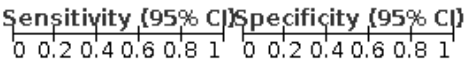
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 1 | 47 | 1 | 251 | 0.50 [0.01, 0.99] | 0.84 [0.80, 0.88] | | |




### Test 30. TB contact, < 15 y/o, and HIV+, microbiological

TB contact, < 15 y/o, and HIV+, microbiological

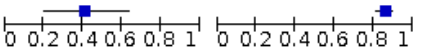
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|




### Test 31. TB contact, < 5 y/o, composite

TB contact, < 5 y/o, composite

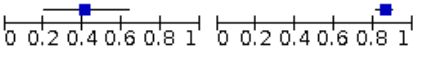
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 9 | 39 | 13 | 239 | 0.41 [0.21, 0.64] | 0.86 [0.81, 0.90] | | |




### Test 32. TB contact, < 20 y/o, composite

TB contact, < 20 y/o, composite

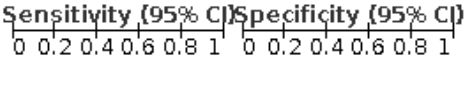
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 9 | 39 | 13 | 239 | 0.41 [0.21, 0.64] | 0.86 [0.81, 0.90] | | |




### Test 33. Current fever, < 5 y/o, microbiological

Current fever, < 5 y/o, microbiological

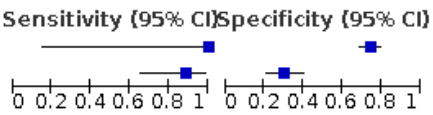
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|



### Test 34. Current fever, < 15 y/o, microbiological

Current fever, < 15 y/o, microbiological

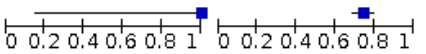
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| LaCourse 2014 | 2 | 76 | 0 | 222 | 1.00 [0.16, 1.00] | 0.74 [0.69, 0.79] |
| Portevin 2014 | 16 | 66 | 2 | 29 | 0.89 [0.65, 0.99] | 0.31 [0.21, 0.41] |




### Test 35. Fever > 1 week, < 5 y/o, microbiological

Fever > 1 week, < 5 y/o, microbiological

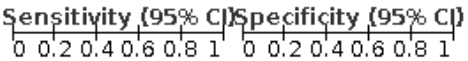
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| LaCourse 2014 | 2 | 76 | 0 | 222 | 1.00 [0.16, 1.00] | 0.74 [0.69, 0.79] |




### Test 36. Fever, < 15 y/o, and HIV+, microbiological

Fever, < 15 y/o, and HIV+, microbiological

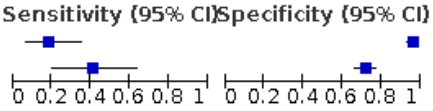
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|




### Test 37. Current fever, < 5 y/o, composite

Current fever, < 5 y/o, composite

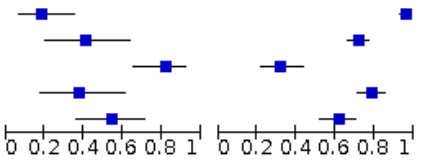
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Kruk 2008 | 6 | 8 | 27 | 211 | 0.18 [0.07, 0.35] | 0.96 [0.93, 0.98] |
| LaCourse 2014 | 9 | 79 | 13 | 199 | 0.41 [0.21, 0.64] | 0.72 [0.66, 0.77] |




### Test 38. Current fever, < 15 y/o, composite

Current fever, < 15 y/o, composite

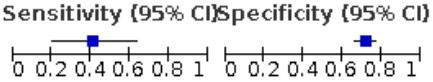
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Kruk 2008 | 6 | 8 | 27 | 211 | 0.18 [0.07, 0.35] | 0.96 [0.93, 0.98] |
| LaCourse 2014 | 9 | 79 | 13 | 199 | 0.41 [0.21, 0.64] | 0.72 [0.66, 0.77] |
| Portevin 2014 | 31 | 51 | 7 | 24 | 0.82 [0.66, 0.92] | 0.32 [0.22, 0.44] |
| Tieu 2014 | 8 | 29 | 13 | 108 | 0.38 [0.18, 0.62] | 0.79 [0.71, 0.85] |
| Togun 2016 | 19 | 44 | 16 | 71 | 0.54 [0.37, 0.71] | 0.62 [0.52, 0.71] |




### Test 39. Fever > 1 week, < 5 y/o, composite

Fever > 1 week, < 5 y/o, composite

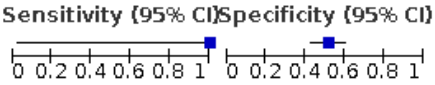
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| LaCourse 2014 | 9 | 79 | 13 | 199 | 0.41 [0.21, 0.64] | 0.72 [0.66, 0.77] |



### Test 40. Weight for height z-score < –3, < 5 y/o, microbiological

Weight for height z-score < -3, < 5 y/o, microbiological

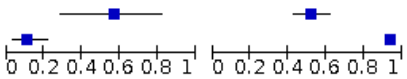
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 1 | 61 | 0 | 65 | 1.00 [0.03, 1.00] | 0.52 [0.43, 0.61] | | |




### Test 41. Weight for height z-score < –3, < 5 y/o, composite

Weight for height z-score < -3, < 5 y/o, composite

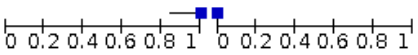
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 8 | 54 | 6 | 59 | 0.57 [0.29, 0.82] | 0.52 [0.43, 0.62] | | |
| Schwoebel 2020 | 6 | 122 | 49 | 1681 | 0.11 [0.04, 0.22] | 0.93 [0.92, 0.94] | | |




### Test 42. Severe malnutrition, < 5 y/o, composite

Severe malnutrition, < 5 y/o, composite

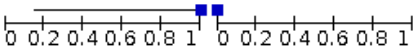
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 22 | 278 | 0 | 0 | 1.00 [0.85, 1.00] | 0.00 [0.00, 0.01] | | |




### Test 43. Severe malnutrition, < 5 y/o, microbiological

Severe malnutrition, < 5 y/o, microbiological

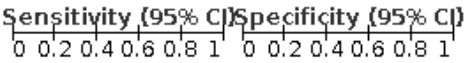
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 2 | 298 | 0 | 0 | 1.00 [0.16, 1.00] | 0.00 [0.00, 0.01] | | |




### Test 44. Weight loss or poor weight gain, < 5 y/o, microbiological

Weight loss or poor weight gain, < 5 y/o, microbiological

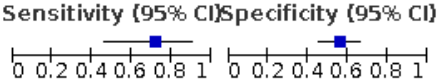
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|




### Test 45. Weight loss or poor weight gain, < 15 y/o, microbiological

Weight loss or poor weight gain, < 15 y/o, microbiological

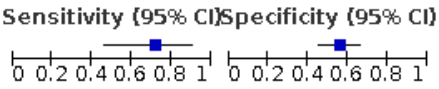
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Portevin 2014 | 13 | 42 | 5 | 53 | 0.72 [0.47, 0.90] | 0.56 [0.45, 0.66] | | |



### Test 46. Weight loss or poor weight gain, < 20 y/o, microbiological

Weight loss or poor weight gain, < 20 y/o, microbiological

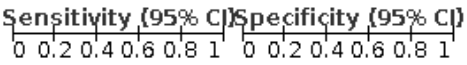
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Portevin 2014 | 13 | 42 | 5 | 53 | 0.72 [0.47, 0.90] | 0.56 [0.45, 0.66] |




### Test 47. Weight loss or poor weight gain, < 15 y/o, and HIV+, microbiological

Weight loss or poor weight gain, < 15 y/o, and HIV+, microbiological

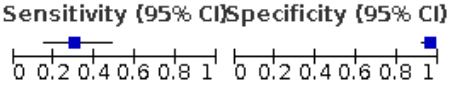
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|




### Test 48. Weight loss or poor weight gain, < 5 y/o, composite

Weight loss or poor weight gain, < 5 y/o, composite

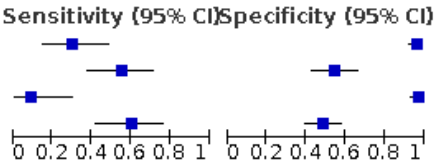
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Kruk 2008 | 10 | 9 | 23 | 210 | 0.30 [0.16, 0.49] | 0.96 [0.92, 0.98] |




### Test 49. Weight loss or poor weight gain, < 15 y/o, composite

Weight loss or poor weight gain, < 15 y/o, composite

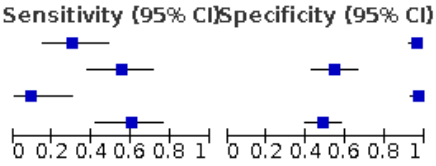
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Kruk 2008 | 10 | 9 | 23 | 210 | 0.30 [0.16, 0.49] | 0.96 [0.92, 0.98] |
| Portevin 2014 | 21 | 34 | 17 | 41 | 0.55 [0.38, 0.71] | 0.55 [0.43, 0.66] |
| Tieu 2014 | 2 | 4 | 19 | 133 | 0.10 [0.01, 0.30] | 0.97 [0.93, 0.99] |
| Togun 2016 | 21 | 59 | 14 | 56 | 0.60 [0.42, 0.76] | 0.49 [0.39, 0.58] |




### Test 50. Weight loss or poor weight gain, < 20 y/o, composite

Weight loss or poor weight gain, < 20 y/o, composite

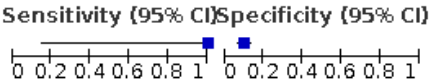
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Kruk 2008 | 10 | 9 | 23 | 210 | 0.30 [0.16, 0.49] | 0.96 [0.92, 0.98] |
| Portevin 2014 | 21 | 34 | 17 | 41 | 0.55 [0.38, 0.71] | 0.55 [0.43, 0.66] |
| Tieu 2014 | 2 | 4 | 19 | 133 | 0.10 [0.01, 0.30] | 0.97 [0.93, 0.99] |
| Togun 2016 | 21 | 59 | 14 | 56 | 0.60 [0.42, 0.76] | 0.49 [0.39, 0.58] |




### Test 51. Fatigue or lethargy, < 5 y/o, microbiological

Fatigue or lethargy, < 5 y/o, microbiological

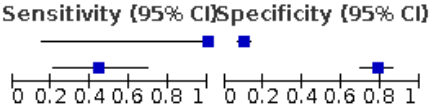
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| LaCourse 2014 | 2 | 267 | 0 | 30 | 1.00 [0.16, 1.00] | 0.10 [0.07, 0.14] |



### Test 52. Fatigue or lethargy, < 15 y/o, microbiological

Fatigue or lethargy, < 15 y/o, microbiological

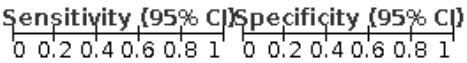
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| LaCourse 2014 | 2 | 267 | 0 | 30 | 1.00 [0.16, 1.00] | 0.10 [0.07, 0.14] |
| Portevin 2014 | 8 | 20 | 10 | 75 | 0.44 [0.22, 0.69] | 0.79 [0.69, 0.87] |




### Test 53. Fatigue or lethargy, < 15 y/o, and HIV+, microbiological

Fatigue or lethargy, < 15 y/o, and HIV+, microbiological

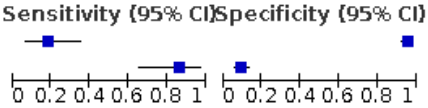
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|




### Test 54. Fatigue or lethargy, < 5 y/o, composite

Fatigue or lethargy, < 5 y/o, composite

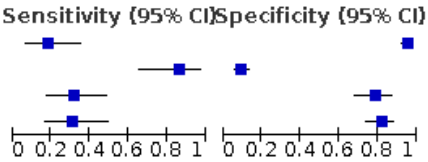
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Kruk 2008 | 6 | 10 | 27 | 209 | 0.18 [0.07, 0.35] | 0.95 [0.92, 0.98] |
| LaCourse 2014 | 19 | 250 | 3 | 27 | 0.86 [0.65, 0.97] | 0.10 [0.07, 0.14] |




### Test 55. Fatigue or lethargy, < 15 y/o, composite

Fatigue or lethargy, < 15 y/o, composite

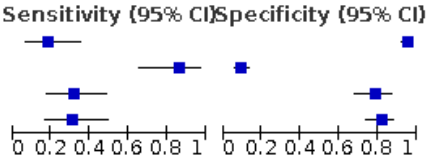
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Kruk 2008 | 6 | 10 | 27 | 209 | 0.18 [0.07, 0.35] | 0.95 [0.92, 0.98] |
| LaCourse 2014 | 19 | 250 | 3 | 27 | 0.86 [0.65, 0.97] | 0.10 [0.07, 0.14] |
| Portevin 2014 | 12 | 16 | 26 | 59 | 0.32 [0.18, 0.49] | 0.79 [0.68, 0.87] |
| Togun 2016 | 11 | 21 | 24 | 94 | 0.31 [0.17, 0.49] | 0.82 [0.73, 0.88] |




### Test 56. Fatigue or lethargy, < 20 y/o, composite

Fatigue or lethargy, < 20 y/o, composite

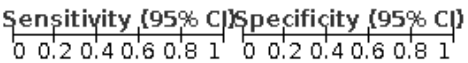
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Kruk 2008 | 6 | 10 | 27 | 209 | 0.18 [0.07, 0.35] | 0.95 [0.92, 0.98] |
| LaCourse 2014 | 19 | 250 | 3 | 27 | 0.86 [0.65, 0.97] | 0.10 [0.07, 0.14] |
| Portevin 2014 | 12 | 16 | 26 | 59 | 0.32 [0.18, 0.49] | 0.79 [0.68, 0.87] |
| Togun 2016 | 11 | 21 | 24 | 94 | 0.31 [0.17, 0.49] | 0.82 [0.73, 0.88] |



### Test 57. Night sweats, < 5 y/o, microbiological

Night sweats, < 5 y/o, microbiological

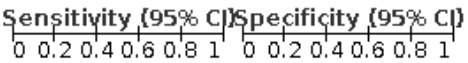
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|




### Test 58. Night sweats, < 15 y/o, microbiological

Night sweats, < 15 y/o, microbiological

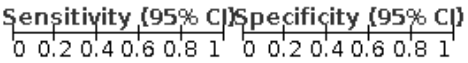
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|




### Test 59. Night sweats, < 15 y/o, and HIV+, microbiological

Night sweats, < 15 y/o, and HIV+, microbiological

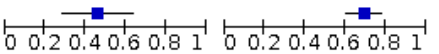
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|




### Test 60. Night sweats, < 15 y/o, composite

Night sweats, < 15 y/o, composite

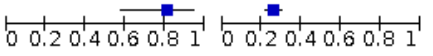
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Togun 2016 | 16 | 35 | 19 | 80 | 0.46 [0.29, 0.63] | 0.70 [0.60, 0.78] | | |




### Test 61. CXR abnormal, < 15 y/o, microbiological

CXR abnormal, < 15 y/o, microbiological

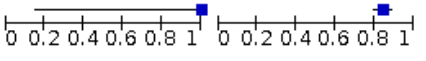
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Togun 2015 | 17 | 341 | 4 | 120 | 0.81 [0.58, 0.95] | 0.26 [0.22, 0.30] | | |




### Test 62. CXR suggestive, < 5 y/o, microbiological

CXR suggestive, < 5 y/o, microbiological

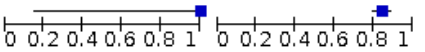
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 2 | 46 | 0 | 251 | 1.00 [0.16, 1.00] | 0.85 [0.80, 0.88] | | |



### Test 63. CXR suggestive, < 15 y/o, microbiological

CXR suggestive, < 15 y/o, microbiological

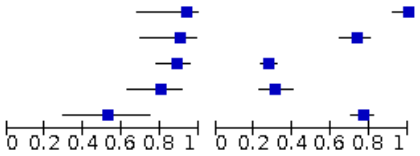
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 2 | 46 | 0 | 251 | 1.00 [0.16, 1.00] | 0.85 [0.80, 0.88] | | |




### Test 64. CXR abnormal, < 15 y/o, composite

CXR abnormal, < 15 y/o, composite

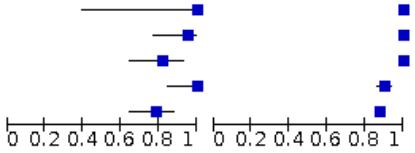
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Dreesman 2017 | 14 | 0 | 1 | 46 | 0.93 [0.68, 1.00] | 1.00 [0.92, 1.00] | | |
| Tieu 2014 | 19 | 36 | 2 | 98 | 0.90 [0.70, 0.99] | 0.73 [0.65, 0.80] | | |
| Togun 2015 | 55 | 303 | 7 | 117 | 0.89 [0.78, 0.95] | 0.28 [0.24, 0.32] | | |
| Togun 2016 | 28 | 79 | 7 | 36 | 0.80 [0.63, 0.92] | 0.31 [0.23, 0.41] | | |
| Triasih 2015b | 11 | 57 | 10 | 187 | 0.52 [0.30, 0.74] | 0.77 [0.71, 0.82] | | |




### Test 65. CXR suggestive, < 15 y/o, composite

CXR suggestive, < 15 y/o, composite

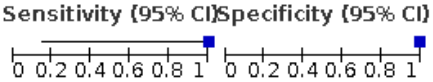
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Birungi 2018 | 4 | 0 | 0 | 212 | 1.00 [0.40, 1.00] | 1.00 [0.98, 1.00] | | |
| Clemente 2017 | 21 | 0 | 1 | 224 | 0.95 [0.77, 1.00] | 1.00 [0.98, 1.00] | | |
| Kruk 2008 | 27 | 0 | 6 | 219 | 0.82 [0.65, 0.93] | 1.00 [0.98, 1.00] | | |
| LaCourse 2014 | 22 | 26 | 0 | 251 | 1.00 [0.85, 1.00] | 0.91 [0.87, 0.94] | | |
| Schwoebel 2020 | 43 | 223 | 12 | 1559 | 0.78 [0.65, 0.88] | 0.87 [0.86, 0.89] | | |




### Test 66. Xpert MTB/Rif, < 5 y/o, microbiological

Xpert MTB/Rif, < 5 y/o, microbiological

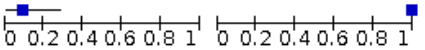
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 2 | 1 | 0 | 297 | 1.00 [0.16, 1.00] | 1.00 [0.98, 1.00] | | |




### Test 67. Xpert MTB/Rif, < 5 y/o, composite

Xpert MTB/Rif, < 5 y/o, composite

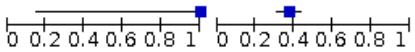
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 2 | 1 | 20 | 277 | 0.09 [0.01, 0.29] | 1.00 [0.98, 1.00] | | |



### Test 68. One of multiple symptoms, < 5 y/o, microbiological

One of multiple symptoms, < 5 y/o, microbiological

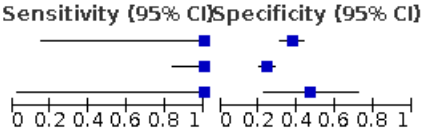
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Aggerbeck 2018 | 2 | 146 | 0 | 87 | 1.00 [0.16, 1.00] | 0.37 [0.31, 0.44] | | |




### Test 69. One of multiple symptoms, < 15 y/o, microbiological

One of multiple symptoms, < 15 y/o, microbiological

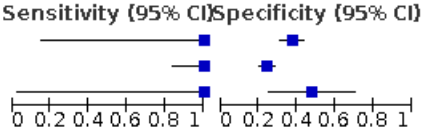
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Aggerbeck 2018 | 2 | 146 | 0 | 87 | 1.00 [0.16, 1.00] | 0.37 [0.31, 0.44] | | |
| Togun 2015 | 21 | 353 | 0 | 113 | 1.00 [0.84, 1.00] | 0.24 [0.20, 0.28] | | |
| Ustero 2017 | 1 | 9 | 0 | 8 | 1.00 [0.03, 1.00] | 0.47 [0.23, 0.72] | | |




### Test 70. One of multiple symptoms, < 20 y/o, microbiological

One of multiple symptoms, < 20 y/o, microbiological

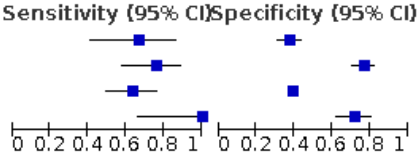
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Aggerbeck 2018 | 2 | 146 | 0 | 87 | 1.00 [0.16, 1.00] | 0.37 [0.31, 0.44] | | |
| Togun 2015 | 21 | 353 | 0 | 113 | 1.00 [0.84, 1.00] | 0.24 [0.20, 0.28] | | |
| Ustero 2017 | 1 | 11 | 0 | 10 | 1.00 [0.03, 1.00] | 0.48 [0.26, 0.70] | | |




### Test 71. One of multiple symptoms, < 5 y/o, composite

One of multiple symptoms, < 5 y/o, composite

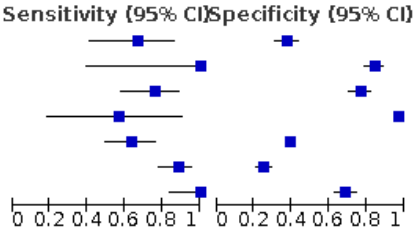
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Aggerbeck 2018 | 12 | 136 | 6 | 81 | 0.67 [0.41, 0.87] | 0.37 [0.31, 0.44] | | |
| Kruk 2008 | 25 | 51 | 8 | 168 | 0.76 [0.58, 0.89] | 0.77 [0.71, 0.82] | | |
| Schwoebel 2020 | 35 | 1150 | 20 | 753 | 0.64 [0.50, 0.76] | 0.40 [0.37, 0.42] | | |
| Triasih 2015a | 9 | 28 | 0 | 71 | 1.00 [0.66, 1.00] | 0.72 [0.62, 0.80] | | |




### Test 72. One of multiple symptoms, < 15 y/o, composite

One of multiple symptoms, < 15 y/o, composite

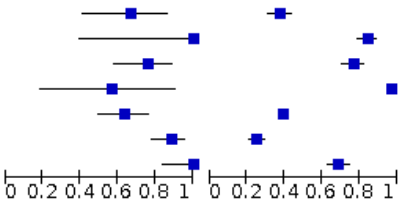
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Aggerbeck 2018 | 12 | 136 | 6 | 81 | 0.67 [0.41, 0.87] | 0.37 [0.31, 0.44] | | |
| Birungi 2018 | 4 | 33 | 0 | 179 | 1.00 [0.40, 1.00] | 0.84 [0.79, 0.89] | | |
| Kruk 2008 | 25 | 51 | 8 | 168 | 0.76 [0.58, 0.89] | 0.77 [0.71, 0.82] | | |
| Sawry 2018 | 4 | 41 | 3 | 1295 | 0.57 [0.18, 0.90] | 0.97 [0.96, 0.98] | | |
| Schwoebel 2020 | 35 | 1150 | 20 | 753 | 0.64 [0.50, 0.76] | 0.40 [0.37, 0.42] | | |
| Togun 2015 | 55 | 319 | 7 | 106 | 0.89 [0.78, 0.95] | 0.25 [0.21, 0.29] | | |
| Triasih 2015a | 21 | 77 | 0 | 171 | 1.00 [0.84, 1.00] | 0.69 [0.63, 0.75] | | |



### Test 73. One of multiple symptoms, < 20 y/o, composite

One of multiple symptoms, < 20 y/o, composite

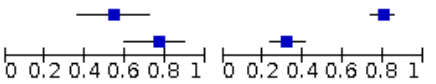
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Aggerbeck 2018 | 12 | 136 | 6 | 81 | 0.67 [0.41, 0.87] | 0.37 [0.31, 0.44] | | |
| Birungi 2018 | 4 | 33 | 0 | 179 | 1.00 [0.40, 1.00] | 0.84 [0.79, 0.89] | | |
| Kruk 2008 | 25 | 51 | 8 | 168 | 0.76 [0.58, 0.89] | 0.77 [0.71, 0.82] | | |
| Sawry 2018 | 4 | 41 | 3 | 1295 | 0.57 [0.18, 0.90] | 0.97 [0.96, 0.98] | | |
| Schwoebel 2020 | 35 | 1150 | 20 | 753 | 0.64 [0.50, 0.76] | 0.40 [0.37, 0.42] | | |
| Togun 2015 | 55 | 319 | 7 | 106 | 0.89 [0.78, 0.95] | 0.25 [0.21, 0.29] | | |
| Triasih 2015a | 21 | 77 | 0 | 171 | 1.00 [0.84, 1.00] | 0.69 [0.63, 0.75] | | |




### Test 74. Any cough, < 15 y/o, contact tracing, composite

Any cough, < 15 y/o, contact tracing, composite

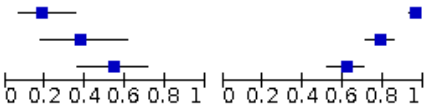
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Kruk 2008 | 18 | 44 | 15 | 175 | 0.55 [0.36, 0.72] | 0.80 [0.74, 0.85] | | |
| Togun 2016 | 27 | 78 | 8 | 37 | 0.77 [0.60, 0.90] | 0.32 [0.24, 0.42] | | |




### Test 75. Current fever, < 15 y/o, contact tracing, composite

Current fever, < 15 y/o, contact tracing, composite

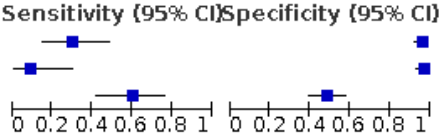
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Kruk 2008 | 6 | 8 | 27 | 211 | 0.18 [0.07, 0.35] | 0.96 [0.93, 0.98] | | |
| Tieu 2014 | 8 | 29 | 13 | 108 | 0.38 [0.18, 0.62] | 0.79 [0.71, 0.85] | | |
| Togun 2016 | 19 | 44 | 16 | 71 | 0.54 [0.37, 0.71] | 0.62 [0.52, 0.71] | | |




### Test 76. Weight loss or poor weight gain, < 20 y/o, contact tracing, composite

Weight loss or poor weight gain, < 20 y/o, contact tracing, composite

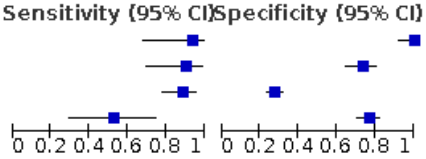
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Kruk 2008 | 10 | 9 | 23 | 210 | 0.30 [0.16, 0.49] | 0.96 [0.92, 0.98] | | |
| Tieu 2014 | 2 | 4 | 19 | 133 | 0.10 [0.01, 0.30] | 0.97 [0.93, 0.99] | | |
| Togun 2016 | 21 | 59 | 14 | 56 | 0.60 [0.42, 0.76] | 0.49 [0.39, 0.58] | | |




### Test 77. CXR abnormal, < 15 y/o, contact tracing, composite

CXR abnormal, < 15 y/o, contact tracing, composite

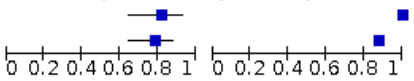
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Dreesman 2017 | 14 | 0 | 1 | 46 | 0.93 [0.68, 1.00] | 1.00 [0.92, 1.00] | | |
| Tieu 2014 | 19 | 36 | 2 | 98 | 0.90 [0.70, 0.99] | 0.73 [0.65, 0.80] | | |
| Togun 2015 | 55 | 303 | 7 | 117 | 0.89 [0.78, 0.95] | 0.28 [0.24, 0.32] | | |
| Triasih 2015b | 11 | 57 | 10 | 187 | 0.52 [0.30, 0.74] | 0.77 [0.71, 0.82] | | |



### Test 78. CXR suggestive, < 5 y/o, contact tracing, composite

CXR suggestive, < 5 y/o, contact tracing, composite

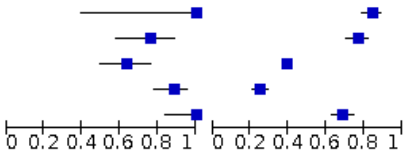
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Kruk 2008 | 27 | 0 | 6 | 219 | 0.82 [0.65, 0.93] | 1.00 [0.98, 1.00] | | |
| Schwoebel 2020 | 43 | 223 | 12 | 1559 | 0.78 [0.65, 0.88] | 0.87 [0.86, 0.89] | | |




### Test 79. One of multiple symptoms, < 15 y/o, contact tracing, composite

One of multiple symptoms, < 15 y/o, contact tracing, composite

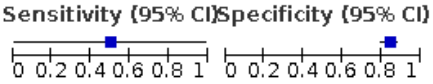
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Birungi 2018 | 4 | 33 | 0 | 179 | 1.00 [0.40, 1.00] | 0.84 [0.79, 0.89] | | |
| Kruk 2008 | 25 | 51 | 8 | 168 | 0.76 [0.58, 0.89] | 0.77 [0.71, 0.82] | | |
| Schwoebel 2020 | 35 | 1150 | 20 | 753 | 0.64 [0.50, 0.76] | 0.40 [0.37, 0.42] | | |
| Togun 2015 | 55 | 319 | 7 | 106 | 0.89 [0.78, 0.95] | 0.25 [0.21, 0.29] | | |
| Triasih 2015a | 21 | 77 | 0 | 171 | 1.00 [0.84, 1.00] | 0.69 [0.63, 0.75] | | |




### Test 80. TB contact, < 20 y/o in inpatient or outpatient settings, microbiological

TB contact, < 20 y/o in inpatient or outpatient settings, microbiological

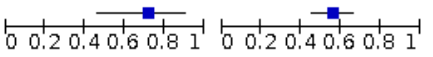
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 1 | 47 | 1 | 251 | 0.50 [0.01, 0.99] | 0.84 [0.80, 0.88] | | |




### Test 81. Weight loss or poor weight gain, < 20 y/o in inpatient or outpatient settings, microbiological

Weight loss or poor weight gain, < 20 y/o in inpatient or outpatient settings, microbiological

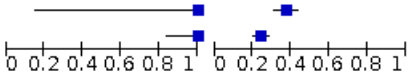
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Portevin 2014 | 13 | 42 | 5 | 53 | 0.72 [0.47, 0.90] | 0.56 [0.45, 0.66] | | |




### Test 82. One of multiple symptoms, < 20 y/o in inpatient or outpatient settings, microbiological

One of multiple symptoms, < 20 y/o in inpatient or outpatient settings, microbiological

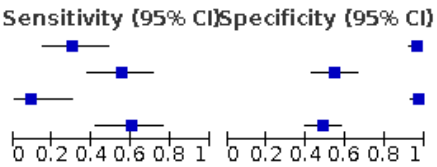
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Aggerbeck 2018 | 2 | 146 | 0 | 87 | 1.00 [0.16, 1.00] | 0.37 [0.31, 0.44] | | |
| Togun 2015 | 21 | 353 | 0 | 113 | 1.00 [0.84, 1.00] | 0.24 [0.20, 0.28] | | |



### Test 83. Weight loss or poor weight gain, < 20 y/o in inpatient or outpatient settings, composite

Weight loss or poor weight gain, < 20 y/o in inpatient or outpatient settings, composite

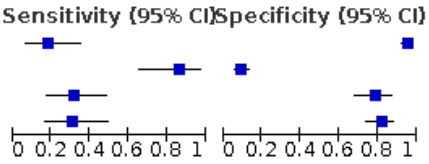
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Kruk 2008 | 10 | 9 | 23 | 210 | 0.30 [0.16, 0.49] | 0.96 [0.92, 0.98] |
| Portevin 2014 | 21 | 34 | 17 | 41 | 0.55 [0.38, 0.71] | 0.55 [0.43, 0.66] |
| Tieu 2014 | 2 | 4 | 19 | 133 | 0.10 [0.01, 0.30] | 0.97 [0.93, 0.99] |
| Togun 2016 | 21 | 59 | 14 | 56 | 0.60 [0.42, 0.76] | 0.49 [0.39, 0.58] |




### Test 84. Fatigue or lethargy, < 20 y/o in inpatient or outpatient settings, composite

Fatigue or lethargy, < 20 y/o in inpatient or outpatient settings, composite

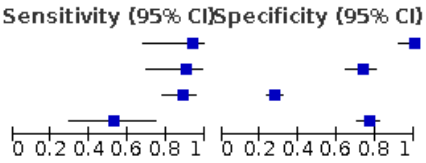
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Kruk 2008 | 6 | 10 | 27 | 209 | 0.18 [0.07, 0.35] | 0.95 [0.92, 0.98] |
| LaCourse 2014 | 19 | 250 | 3 | 27 | 0.86 [0.65, 0.97] | 0.10 [0.07, 0.14] |
| Portevin 2014 | 12 | 16 | 26 | 59 | 0.32 [0.18, 0.49] | 0.79 [0.68, 0.87] |
| Togun 2016 | 11 | 21 | 24 | 94 | 0.31 [0.17, 0.49] | 0.82 [0.73, 0.88] |




### Test 85. CXR abnormal, < 15 y/o, contact tracing, composite

CXR abnormal, < 15 y/o, contact tracing, composite

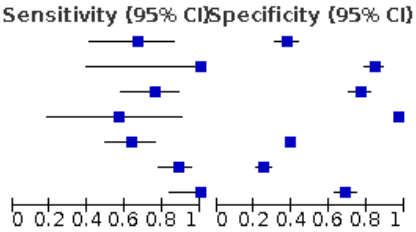
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Dreesman 2017 | 14 | 0 | 1 | 46 | 0.93 [0.68, 1.00] | 1.00 [0.92, 1.00] |
| Tieu 2014 | 19 | 36 | 2 | 98 | 0.90 [0.70, 0.99] | 0.73 [0.65, 0.80] |
| Togun 2015 | 55 | 303 | 7 | 117 | 0.89 [0.78, 0.95] | 0.28 [0.24, 0.32] |
| Triasih 2015b | 11 | 57 | 10 | 187 | 0.52 [0.30, 0.74] | 0.77 [0.71, 0.82] |




### Test 86. One of multiple symptoms, < 15 y/o in inpatient or outpatient settings, composite

One of multiple symptoms, < 15 y/o in inpatient or outpatient settings, composite

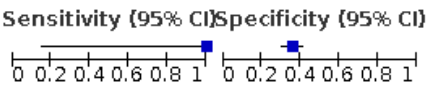
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Aggerbeck 2018 | 12 | 136 | 6 | 81 | 0.67 [0.41, 0.87] | 0.37 [0.31, 0.44] |
| Birungi 2018 | 4 | 33 | 0 | 179 | 1.00 [0.40, 1.00] | 0.84 [0.79, 0.89] |
| Kruk 2008 | 25 | 51 | 8 | 168 | 0.76 [0.58, 0.89] | 0.77 [0.71, 0.82] |
| Sawry 2018 | 4 | 41 | 3 | 1295 | 0.57 [0.18, 0.90] | 0.97 [0.96, 0.98] |
| Schwoebel 2020 | 35 | 1150 | 20 | 753 | 0.64 [0.50, 0.76] | 0.40 [0.37, 0.42] |
| Togun 2015 | 55 | 319 | 7 | 106 | 0.89 [0.78, 0.95] | 0.25 [0.21, 0.29] |
| Triasih 2015a | 21 | 77 | 0 | 171 | 1.00 [0.84, 1.00] | 0.69 [0.63, 0.75] |




### Test 87. Mid-upper arm circumference (MUAC) < 11.5 cm, < 5 y/o, microbiological

Mid-upper arm circumference (MUAC) < 11.5 cm, < 5 y/o, microbiological

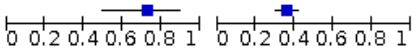
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| LaCourse 2014 | 2 | 191 | 0 | 107 | 1.00 [0.16, 1.00] | 0.36 [0.30, 0.42] |



### Test 88. MUAC < 11.5 cm, < 5 y/o, composite

MUAC < 11.5 cm, < 5 y/o, composite

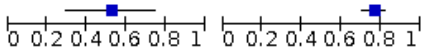
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| LaCourse 2014 | 16 | 177 | 6 | 101 | 0.73 [0.50, 0.89] | 0.36 [0.31, 0.42] | | |




### Test 89. CXR abnormal, < 15 y/o in community, composite

CXR abnormal, < 15 y/o in community, composite

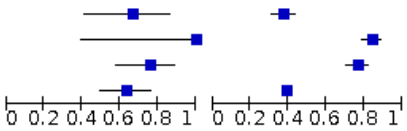
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Triasih 2015b | 11 | 57 | 10 | 187 | 0.52 [0.30, 0.74] | 0.77 [0.71, 0.82] | | |




### Test 90. One of cough, fever, or decreased playfulness; < 15 y/o in inpatient or outpatient settings, composite

One of cough, fever, or decreased playfulness; < 15 y/o in inpatient or outpatient settings, composite

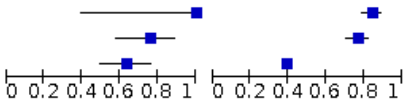
| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Aggerbeck 2018 | 12 | 136 | 6 | 81 | 0.67 [0.41, 0.87] | 0.37 [0.31, 0.44] | | |
| Birungi 2018 | 4 | 33 | 0 | 179 | 1.00 [0.40, 1.00] | 0.84 [0.79, 0.89] | | |
| Kruk 2008 | 25 | 51 | 8 | 168 | 0.76 [0.58, 0.89] | 0.77 [0.71, 0.82] | | |
| Schwoebel 2020 | 35 | 1150 | 20 | 753 | 0.64 [0.50, 0.76] | 0.40 [0.37, 0.42] | | |




### Test 91. One of cough, fever, or decreased playfulness; < 15 y/o, contact tracing, composite

One of cough, fever, or decreased playfulness; < 15 y/o, contact tracing, composite

| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|---|
| Birungi 2018 | 4 | 33 | 0 | 179 | 1.00 [0.40, 1.00] | 0.84 [0.79, 0.89] | | |
| Kruk 2008 | 25 | 51 | 8 | 168 | 0.76 [0.58, 0.89] | 0.77 [0.71, 0.82] | | |
| Schwoebel 2020 | 35 | 1150 | 20 | 753 | 0.64 [0.50, 0.76] | 0.40 [0.37, 0.42] | | |




## ADDITIONAL TABLES

### Table 1. Summary of included studies

| Study | Country or countries of sampling | Sampling in TB high-burden country?[a] |
|---|---|---|
| Aggerbeck 2018 | South Africa | Yes |
| Birungi 2018 | Rwanda | No |
| Clemente 2017 | Italy | No |
| Dreesman 2017 | Belgium | No |
| Jaganath 2013 | Uganda | No |

**Table 1. Summary of included studies** *(Continued)*

| | | |
|---|---|---|
| Kruk 2008 | South Africa | Yes |
| LaCourse 2014 [b] | Malawi | No |
| PERCH 2019 [b] | Bangladesh, The Gambia, Kenya, Mali, South Africa, Thailand, and Zambia | Majority |
| Portevin 2014 | Tanzania | Yes |
| Rose 2012 | Tanzania | Yes |
| Sawry 2018 [b] | South Africa | Yes |
| Schwoebel 2020 [b] | Benin, Burkina Faso, Cameroon, and CAR | Only 1 of 4 countries (CAR) |
| Tieu 2014 | Thailand | Yes |
| Togun 2015 [b] | The Gambia | No |
| Togun 2016 | The Gambia | No |
| Triasih 2015a | Indonesia | Yes |
| Triasih 2015b [b] | Indonesia | Yes |
| Ustero 2017 | Eswatini (Swaziland) | Yes |
| Vonasek 2021 [b] | Botswana, Eswatini, Lesotho, Malawi, Tanzania, and Uganda | 2 of 6 countries |
| **Publication year range:** 2008 to 2021 | **Africa:** 14 studies<br>**Asia:** 4 studies<br>**Europe:** 2 studies | **Sampling at least partially in TB high-burden countries:** 12 studies |

CAR: Central African Republic, TB: tuberculosis.
[a]TB high-burden countries are defined in the WHO Global Tuberculosis Report 2020.
[b]Studies not captured through database searching but identified through contacting the community of TB experts. All other studies identified through database searching.

**Table 2. Screening tests for active pulmonary tuberculosis in children by population and setting**

| Population of children and adolescents | Index test | Reference standard | Studies | TB prevalence | Number of children (TB cases) | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|---|
| Close TB contacts | ≥ 1 of cough, fever, or poor weight gain | CRS | 4 | 2% to 13% | 2695 (113) | 89% (52% to 98%) | 69% (51% to 83%) |
| Inpatient or outpatient settings, < 5 years | ≥ 1 of cough, fever, or decreased playfulness | CRS | 3 | (3)% to 13% | 2445 (106) | 64% to 76%[a] | 37% to 77%[a] |
| Outpatients living with HIV | ≥ 1 of cough, fever, poor weight gain, or TB close contact (WHO 4-symptom screen) performed at each healthcare visit | CRS | 2 | 3% and 8% | 203,135 (1219)[b] | 61% (58% to 64%) | 94% (86% to 98%) |
| Close TB contacts | Undernutrition | CRS | 3 | 10% to 13% | 1399 (162) | 21% (11% to 38%) | 85% (71% to 93%) |
| Inpatient or outpatient settings | Undernutrition | CRS | 5 | 10% to 34% | 1723 (233) | 32% (18% to 50%) | 75% (56% to 88%) |
| Inpatient or outpatient settings | Undernutrition | MRS | 2 | 4% and 16% | 561 (39) | 48% (26% to 70%) and 67% (41% to 87%) | 62% (49% to 74%) and 72% (68% to 76%) |
| Close TB contacts | Abnormal CXR | CRS | 8 | 2% to 25% | 3513 (232) | 87% (75% to 93%) | 99% (68% to 100%) |
| Close TB contacts | Suggestive CXR | CRS | 4 | 2% to 13% | 2550 (113) | 84% (70% to 92%) | 91% (90% to 92%) |
| Inpatient or outpatient settings, < 5 years | Suggestive CXR | CRS | 3 | 2% to 13% | 2388 (110) | 87% (66% to 96%) | 89% (88% to 90%) |
| Inpatients with pneumonia, < 5 years | Abnormal CXR | MRS | 1 | 1% | 3540 (28) | 86% (67% to 96%) | 56% (54% to 58%) |
| Inpatient or outpatient settings | Xpert MTB/RIF | MRS | 2 | 1% and 4% | 787 (16) | 43% (18% to 71%) and 100% (16% to 100%) | 99% (97% to 100%) and 100% (98% to 100%) |
| Inpatient or outpatient settings | Xpert MTB/RIF | CRS | 2 | 7% and 13% | 787 (84) | 9% (1% to 29%) and 19% (10% to 31%) | 100% (98% to 100%) and 100% (99% to 100%) |

CI: confidence interval; CRS: composite reference standard; CXR: chest radiography; MRS: microbiological reference standard; TB: tuberculosis; WHO: World Health Organization.
[a]Reported as range from studies as meta-analysis did not converge and pooled estimates could not be obtained.
[b]Reported as: number of screens (cases).

**Table 3. Chest radiography details by study**

| Study | Views obtained | Level of training of interpreters | Number of interpreters per radiograph | Interobserver variability | Threshold for a positive test |
|---|---|---|---|---|---|
| Birungi 2018 | AP and lateral | Experienced general practitioners with findings proofread by an experienced radiologist | 2 | NR | Any of these features, at the same location, detected by both interpreters: air compression or tracheal displacement, soft tissue density suggestive of lymphadenopathy, air space opacification, bilateral nodule picture (military or larger widespread), pleural effusion, cavities, calcified parenchyma, and vertebral spondylitis |
| Clemente 2017 | NR | NR | NR | NR | Hilar lymphadenopathy, pleurisy, pneumonia with calcifications, miliary pattern |
| Dreesman 2017 | NR | NR | NR | NR | Suggestive of active tuberculosis |
| Kruk 2008 | AP and lateral | NR | 2 | NR | Lymph node disease, airway compression, lung cavitation, pleural effusion, or miliary pattern |
| LaCourse 2014 | AP and lateral | NR | 2 | NR | Chest radiography consistent with tuberculosis |
| PERCH 2019 | NR | Trained radiologists and paediatricians | 2 | NR | Presence of lung consolidation, other infiltrate, or both |
| Schwoebel 2020 | AP only | Medical doctor | 1 | NA | Suggestive of tuberculosis |
| Tieu 2014 | NR | NR | NR | NR | Hilar, interstitial, or other types of lung infiltrates, other infiltrates; and lymph node disease |
| Togun 2016 | NR | Study physicians | 2 | NR | Abnormality consistent with active tuberculosis disease |
| Triasih 2015b | AP and lateral | 2 paediatricians and 2 radiologists | 4 | k = 0.25–0.46 | Hilar lymphadenopathy, parenchymal infiltrate or consolidation, pleural effusion, miliary pattern, Gohn focus, calcification |

AP: anteroposterior view; k: kappa statistic; NA: not applicable; NR: not reported.

# APPENDICES

## Appendix 1. Search strategy

### MEDLINE (OVID)

1 exp child/ or exp infant/

2 (newborn* or new-born* or neonat* or neo-nat* or infancy* or infant* or baby* or babies* or toddler*).ti,ab,kw.

3 (child* or children* or boy or boys or girl* or youth* or pediatric* or paediatric* or kid or kids or "school-age*" or juvenile* or preteen* or tween*).ti,ab,kw.

4 (preteen* or pre-teen* or fifteen* or fourteen* or thirteen* or teen* or adolescen* or preadolescen* or "pre-adolescen*" or pubescen* or prepubescen* or "pre-pubescen*").ti,ab,kw.

5 1 or 2 or 3 or 4

6 exp Mycobacterium tuberculosis/

7 exp Tuberculosis/

8 (tuberculos* or tb*).ti,ab,kw.

9 6 or 7 or 8

10 ((active* or symptomatic*) adj3 (tuberculosis* or tb*)).ti,ab,kw.

11 ("active tuberculos*" or "active tb*").kw.

12 ("symptomatic* tuberculos*" or "symptomatic* tb*").kw.

13 10 or 11 or 12

14 9 and 13

15 exp Symptom Assessment/ or exp symptom flare up/

16 (symptom* or manifest*).ti,ab,kw.

17 15 or 16

18 Cough/

19 Cough*.ti,ab,kw.

20 Hemoptysis/

21 (hemoptysis* or "hemo-ptysis*").ti,ab,kw.

22 (cough* adj3 blood*).ti,ab,kw.

23 ("blood* cough*" or "cough* blood*").kw.

24 Fever/

25 (fever* or "high* temp*").ti,ab,kw.

26 Weight Loss/

27 ("weight loss*" or weightloss*).ti,ab,kw.

28 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27

29 17 and 28

**Embase**

1 juvenile'/de OR 'child'/exp

2 newborn*:ti,ab,kw OR 'new born*':ti,ab,kw OR neonat*:ti,ab,kw OR 'neo nat*':ti,ab,kw OR infancy*:ti,ab,kw OR infant*:ti,ab,kw OR baby*:ti,ab,kw OR babies*:ti,ab,kw OR toddler*:ti,ab,kw

3 child*:ti,ab,kw OR children*:ti,ab,kw OR boy:ti,ab,kw OR boys:ti,ab,kw OR girl*:ti,ab,kw OR youth*:ti,ab,kw OR pediatric*:ti,ab,kw OR paediatric*:ti,ab,kw OR kid:ti,ab,kw OR kids:ti,ab,kw OR 'school-age*':ti,ab,kw OR juvenile*:ti,ab,kw OR preteen*:ti,ab,kw OR tween*:ti,ab,kw

4 preteen*:ti,ab,kw OR 'pre teen*':ti,ab,kw OR fifteen*:ti,ab,kw OR fourteen*:ti,ab,kw OR thirteen*:ti,ab,kw OR teen*:ti,ab,kw OR adolescen*:ti,ab,kw OR preadolescen*:ti,ab,kw OR 'pre-adolescen*':ti,ab,kw OR pubescen*:ti,ab,kw OR prepubescen*:ti,ab,kw OR 'pre-pubescen*':ti,ab,kw

5 #1 OR #2 OR #3 OR #4

6 mycobacterium tuberculosis'/exp

7 tuberculosis'/exp

8 tuberculos*:ti,ab,kw OR tb:ti,ab,kw

9 #6 OR #7 OR #8

10 ((active* OR symptomatic*) NEAR/3 (tuberculosis* OR tb)):ti,ab,kw

11 #9 AND #10

12 symptom'/exp

13 symptom*:ti,ab,kw OR manifest*:ti,ab,kw

14 #12 OR #13

15 coughing'/de

16 cough*:ti,ab,kw

17 hemoptysis'/de

18 hemoptysis*:ti,ab,kw OR 'hemo-ptysis*':ti,ab,kw

19 (cough* NEAR/3 blood*):ti,ab,kw

20 fever'/exp

21 fever*:ti,ab,kw OR 'high* temp*':ti,ab,kw

22 body weight loss'/de

23 weight loss*':ti,ab,kw OR weightloss*:ti,ab,kw

24 #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23

25 #14 AND #24

26 thorax radiography'/exp

27 ((chest* OR lung* OR thoracic*) NEAR/3 ('x-ray*' OR xray* OR radiogra* OR imag*)):ti,ab,kw

28 tuberculin test'/exp

29 tubercul* skin test*':ti,ab,kw OR tst:ti,ab,kw OR 'tubercul* test*':ti,ab,kw OR 'tb skin test*':ti,ab,kw OR 'tb test*':ti,ab,kw

**SCOPUS**

1 TITLE-ABS-KEY ( newborn* OR new-born* OR neonat* OR neo-nat* OR infancy* OR infant* OR baby* OR babies* OR toddler* )

2 TITLE-ABS-KEY ( child* OR children* OR boy OR boys OR girl* OR youth* OR pediatric* OR paediatric* OR kid OR kids OR "school-age*" OR juvenile* OR preteen* OR tween* )

3 "

TITLE-ABS-KEY ( preteen* OR pre-teen* OR fifteen* OR fourteen* OR thirteen* OR teen* OR adolescen* OR preadolescen* OR ""pre-adolescen*"" OR pubescen* OR prepubescen* OR ""pre-pubescen*"" ) "

4 ( TITLE-ABS-KEY ( newborn* OR new-born* OR neonat* OR neo-nat* OR infancy* OR infant* OR baby* OR babies* OR toddler* ) ) OR ( TITLE-ABS-KEY ( child* OR children* OR boy OR boys OR girl* OR youth* OR pediatric* OR paediatric* OR kid OR kids OR "school-age*"

OR juvenile* OR preteen* OR tween* ) ) OR ( TITLE-ABS-KEY ( preteen* OR pre-teen* OR fifteen* OR fourteen* OR thirteen* OR teen* OR adolescen* OR preadolescen* OR "pre-adolescen*" OR pubescen* OR prepubescen* OR "pre-pubescen*" ) )

5 TITLE-ABS-KEY ( tuberculos* OR tb* )

6 TITLE-ABS-KEY ( ( active* OR symptomatic* ) W/3 ( tuberculosis* OR tb* ) )

7 ( TITLE-ABS-KEY ( tuberculos* OR tb* ) ) AND ( TITLE-ABS-KEY ( ( active* OR symptomatic* ) W/3 ( tuberculosis* OR tb* ) ) )

8 TITLE-ABS-KEY ( symptom* OR manifest* )

9 TITLE-ABS-KEY ( cough* )

10 TITLE-ABS-KEY ( hemoptysis* OR "hemo-ptysis*" )

11 TITLE-ABS-KEY ( cough* W/3 blood* )

12 TITLE-ABS-KEY ( fever* OR "high* temp*" )

13 "

TITLE-ABS-KEY ( ""weight loss*"" OR weightloss* ) "

14 ( TITLE-ABS-KEY ( cough* ) ) OR ( TITLE-ABS-KEY ( hemoptysis* OR "hemo-ptysis*" ) ) OR ( TITLE-ABS-KEY ( cough* W/3 blood* ) ) OR ( TITLE-ABS-KEY ( fever* OR "high* temp*" ) ) OR ( TITLE-ABS-KEY ( "weight loss*" OR weightloss* ) )

15 ( TITLE-ABS-KEY ( symptom* OR manifest* ) ) AND ( ( TITLE-ABS-KEY ( cough* ) ) OR ( TITLE-ABS-KEY ( hemoptysis* OR "hemo-ptysis*" ) ) OR ( TITLE-ABS-KEY ( cough* W/3 blood* ) ) OR ( TITLE-ABS-KEY ( fever* OR "high* temp*" ) ) OR ( TITLE-ABS-KEY ( "weight loss*" OR weightloss* ) ) )

16 TITLE-ABS-KEY ( ( chest* OR lung* OR thoracic* ) W/3 ( "x-ray*" OR xray* OR radiogra* OR imag* ) )

17 TITLE-ABS-KEY ( "tubercul* skin test*" OR tst OR "tb skin test*" )

18 TITLE-ABS-KEY ( ( "interferon-gamma releas*" OR "IFN-gamma releas*" ) W/3 ( test* OR assay* ) )

19 TITLE-ABS-KEY ( igra )

20 TITLE-ABS-KEY ( "QuantiFERON-TB*" OR quantiferontb* OR qft* OR "T-Spot*" OR tspot* )

21 TITLE-ABS-KEY ( "immunologic* test*" OR "immuno-logic test*" )

22 TITLE-ABS-KEY ( "microbiologic* confirm*" OR "micro-biologic* confirm*" )

23 TITLE-ABS-KEY ( ( "Mycobacter* tubercul*" OR mtb ) W/3 ( culture* OR test* OR assay* ) )

24 TITLE-ABS-KEY ( ( tubercul* OR tb ) W/3 ( test* OR assay* ) )

25 TITLE-ABS-KEY ( ( xpert* OR genexpert* ) W/3 ( mtb OR rif OR rifampicin* OR ultra ) )

26 TITLE-ABS-KEY ( genexpert* OR xpert* )

27 TITLE-ABS-KEY ( truenat OR "True-Nat" OR trunat OR "Tru-Nat" )

28 TITLE-ABS-KEY ( "nucleic acid amplification test*" OR naat )

29 ( TITLE-ABS-KEY ( symptom* OR manifest* ) ) OR ( ( TITLE-ABS-KEY ( symptom* OR manifest* ) ) AND ( ( TITLE-ABS-KEY ( cough* ) ) OR ( TITLE-ABS-KEY ( hemoptysis* OR "hemo-ptysis*" ) ) OR ( TITLE-ABS-KEY ( cough* W/3 blood* ) ) OR ( TITLE-ABS-KEY ( fever* OR "high* temp*" ) ) OR ( TITLE-ABS-KEY ( "weight loss*" OR weightloss* ) ) ) ) OR ( TITLE-ABS-KEY ( ( chest* OR lung* OR thoracic* ) W/3 ( "x-ray*" OR xray* OR radiogra* OR imag* ) ) ) OR ( TITLE-ABS-KEY ( "tubercul* skin test*" OR tst OR "tb skin test*" ) ) OR ( TITLE-ABS-KEY ( ( "interferon-gamma releas*" OR "IFN-gamma releas*" ) W/3 ( test* OR assay* ) ) ) OR ( TITLE-ABS-KEY ( igra ) ) OR ( TITLE-ABS-KEY ( "QuantiFERON-TB*" OR quantiferontb* OR qft* OR "T-Spot*" OR tspot* ) ) OR ( TITLE-ABS-KEY ( "immunologic* test*" OR "immuno-logic test*" ) ) OR # 23 OR ( TITLE-ABS-KEY ( ( "Mycobacter* tubercul*" OR mtb ) W/3 ( culture* OR test* OR assay* ) ) ) OR ( TITLE-ABS-KEY ( ( tubercul* OR tb ) W/3 ( test* OR assay* ) ) ) OR ( TITLE-ABS-KEY ( ( xpert* OR genexpert* ) W/3 ( mtb OR rif OR rifampicin* OR ultra ) ) ) OR ( TITLE-ABS-KEY ( genexpert* OR xpert* ) ) OR ( TITLE-ABS-KEY ( truenat OR "True-Nat" OR trunat OR "Tru-Nat" ) ) OR ( TITLE-ABS-KEY ( "nucleic acid amplification test*" OR naat ) )

30 ( ( TITLE-ABS-KEY ( newborn* OR new-born* OR neonat* OR neo-nat* OR infancy* OR infant* OR baby* OR babies* OR toddler* ) ) OR ( TITLE-ABS-KEY ( child* OR children* OR boy OR boys OR girl* OR youth* OR pediatric* OR paediatric* OR kid OR kids OR "school-age*" OR juvenile* OR preteen* OR tween* ) ) OR ( TITLE-ABS-KEY ( preteen* OR pre-teen* OR fifteen* OR fourteen* OR thirteen* OR teen* OR adolescen* OR preadolescen* OR "pre-adolescen*" OR pubescen* OR prepubescen* OR "pre-pubescen*" ) ) ) AND ( ( TITLE-ABS-KEY ( tuberculos* OR tb* ) ) AND ( TITLE-ABS-KEY ( ( active* OR symptomatic* ) W/3 ( tuberculosis* OR tb* ) ) ) ) AND ( ( TITLE-ABS-KEY ( symptom* OR manifest* ) ) OR ( ( TITLE-ABS-KEY ( symptom* OR manifest* ) ) AND ( ( TITLE-ABS-KEY ( cough* ) ) OR ( TITLE-ABS-KEY ( hemoptysis* OR "hemo-ptysis*" ) ) OR ( TITLE-ABS-KEY ( cough* W/3 blood* ) ) OR ( TITLE-ABS-KEY ( fever* OR "high* temp*" ) ) OR ( TITLE-ABS-KEY ( "weight loss*" OR weightloss* ) ) ) ) ) OR ( TITLE-ABS-KEY ( ( chest* OR lung* OR thoracic* ) W/3 ( "x-ray*" OR xray* OR radiogra* OR imag* ) ) ) OR ( TITLE-ABS-KEY ( "tubercul* skin test*" OR tst OR "tb skin test*" ) ) OR ( TITLE-ABS-KEY ( ( "interferon-gamma releas*" OR "IFN-gamma releas*" ) W/3 ( test* OR assay* ) ) ) OR ( TITLE-ABS-KEY ( igra ) ) OR ( TITLE-ABS-KEY ( "QuantiFERON-TB*" OR quantiferontb* OR qft* OR "T-Spot*" OR tspot* ) ) OR ( TITLE-ABS-KEY ( "immunologic* test*" OR "immuno-logic test*" ) ) OR # 23 OR ( TITLE-ABS-KEY ( ( "Mycobacter* tubercul*" OR mtb ) W/3 ( culture* OR test* OR assay* ) ) ) OR ( TITLE-ABS-KEY ( ( tubercul* OR tb ) W/3 ( test* OR assay* ) ) ) OR ( TITLE-ABS-KEY ( ( xpert* OR genexpert* ) W/3 ( mtb OR rif OR rifampicin* OR ultra ) ) ) OR ( TITLE-ABS-KEY ( genexpert* OR xpert* ) ) OR ( TITLE-ABS-KEY ( truenat OR "True-Nat" OR trunat OR "Tru-Nat" ) ) OR ( TITLE-ABS-KEY ( "nucleic acid amplification test*" OR naat ) ) )

**Cochrane Library issue 2 of 12, February 2021**

1 MeSH descriptor: [Child] explode all trees

2 MeSH descriptor: [Infant] explode all trees

3 (newborn* or new-born* or neonat* or neo-nat* or infancy* or infant* or baby* or babies* or toddler*):ti,ab,kw

4 (child* or children* or boy or boys or girl* or youth* or pediatric* or paediatric* or kid or kids or "school-age*" or juvenile* or preteen* or tween*):ti,ab,kw

5 (preteen* or pre-teen* or fifteen* or fourteen* or thirteen* or teen* or adolescen* or preadolescen* or "pre-adolescen*" or pubescen* or prepubescen* or "pre-pubescen*"):ti,ab,kw

6 #1 or #2 or #3 or #4 or #5

7 MeSH descriptor: [Mycobacterium tuberculosis] explode all trees

8 MeSH descriptor: [Tuberculosis] explode all trees

9 (tuberculos* or tb*):ti,ab,kw

10 #7 or #8 or #9

11 (active* or symptomatic*) NEAR/3 (tuberculosis* or tb*)):ti,ab,kw

12 #10 and #11

13 MeSH descriptor: [Symptom Assessment] explode all trees

14 MeSH descriptor: [Symptom Flare Up] explode all trees

15 (symptom* or manifest*):ti,ab,kw

16 #13 or #14 or #15

17 MeSH descriptor: [Cough] explode all trees

18 (cough*):ti,ab,kw

19 MeSH descriptor: [Hemoptysis] explode all trees

20 (hemoptysis* or "hemo-ptysis*"):ti,ab,kw

21 (cough* NEAR/3 blood*):ti,ab,kw

22 MeSH descriptor: [Fever] explode all trees

23 (fever* or "high* temp*"):ti,ab,kw

24 MeSH descriptor: [Weight Loss] explode all trees

25 ("weight loss*" or weightloss*):ti,ab,kw"

26 #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25

27 #16 and #26

28 MeSH descriptor: [Radiography, Thoracic] explode all trees

29 ((chest* or lung* or thoracic*) NEAR/3 ("x-ray*" or xray* or radiogra* or imag*)):ti,ab,kw

29 tubercul* skin test*':ti,ab,kw OR tst:ti,ab,kw OR 'tubercul* test*':ti,ab,kw OR 'tb skin test*':ti,ab,kw OR 'tb test*':ti,ab,kw

30 interferon gamma release assay'/exp

31 (('interferon-gamma releas*' OR 'ifn-gamma releas*') NEAR/3 (test* OR assay)):ti,ab,kw

32 igra:ti,ab,kw

33 mycobacterium tuberculosis test kit'/exp

34 (('mycobacter* tubercul*' OR mtb) NEAR/3 (culture* OR test* OR assay*)):ti,ab,kw

35 ((tubercul* OR tb) NEAR/3 (test* OR assay*)):ti,ab,kw

36 quantiferon-tb*':ti,ab,kw OR quantiferontb*:ti,ab,kw OR qft*:ti,ab,kw OR 't-spot*':ti,ab,kw OR tspot*:ti,ab,kw

37 immunologic* test*':ti,ab,kw OR 'immuno-logic test*':ti,ab,kw

38 microbiologic* confirm*':ti,ab,kw OR 'micro-biologic* confirm*':ti,ab,kw

39 ((xpert* OR genexpert*) NEAR/3 (mtb OR rif OR rifampicin* OR ultra)):ti,ab,kw

40 genexpert*:ti,ab,kw OR xpert*:ti,ab,kw

41 truenat:ti,ab,kw OR 'true-nat':ti,ab,kw OR trunat:ti,ab,kw OR 'tru-nat':ti,ab,kw

42 nucleic acid amplification test*':ti,ab,kw OR naat:ti,ab,kw

43 #14 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40 OR #41 OR #42

44 #5 AND #11 AND #43

## Appendix 2. QUADAS-2 review-specific guidance

**Domain 1: patient selection**

Risk of bias: could the selection of patients have introduced bias?

*Signalling question 1: was a consecutive or random sample of patients enrolled?*

- Yes: if the study enrolled a consecutive or random sample of eligible participants.
- No: if the study selected participants by convenience.
- Unclear: if the study did not report the manner of participant selection or we could not determine.

*Signalling question 2: did the study avoid inappropriate exclusions?* Examples of inappropriate exclusions may have included children with distant history of tuberculosis, children experiencing severe signs and symptoms of tuberculosis, or children with negative screening test.

- Yes: if no study participants were excluded after inclusion.
- No: if study participants were excluded.
- Unclear: if we could not determine.

Applicability: are there concerns that the included participants and setting do not match the review question?

Based upon the inclusion criteria, included studies focused primarily on pulmonary tuberculosis. Therefore, all included studies assessed as 'low concern.'

**Domain 2: index test**

Risk of bias: could the conduct or interpretation of the index test have introduced bias?

***Symptom screen, chest radiography, and Xpert MTB/RIF and Xpert Ultra***

*Signalling question 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 active tuberculosis. Also, with respect to Xpert MTB/RIF and Xpert Ultra, the test results are automatically generated and the user is provided with printable test results. Thus, there is no room for subjective interpretation of test results.
- No: if symptom questions were asked after the results of the reference test were known, or the chest radiograph was interpreted with knowledge of the results of the reference test.
- Unclear: if we could not determine. For example, if it was unclear whether the chest radiograph reader was blinded to the results of the reference standard.

*Signalling question 2: if a threshold was used, was it prespecified?*

**For tuberculosis symptoms**

This question was not applicable.

**For chest radiography**

- Yes: if the study clearly reported positivity criteria for abnormalities suggestive of tuberculosis or other abnormalities.
- No: if the study did not report the positivity criteria for abnormalities suggestive of tuberculosis or other abnormalities.
- Unclear: if we could not determine.

**For Xpert MTB/RIF and Xpert Ultra**

The threshold is prespecified in all versions of Xpert.

- Yes: for all studies using Xpert MTB/RIF or Xpert Ultra as the index test.

Applicability: are there concerns that the index test, its conduct, or its interpretation differ from the review question?

- High concern: if the index tests were used for diagnosis rather than for screening.
- Low concern: if the index tests were performed with the intention to screen.
- Unclear concern: if we could not determine.

**Domain 3: reference standard**

Risk of bias: could the reference standard, its conduct, or its interpretation have introduced bias?

*Signalling question 1: is the reference standard likely to correctly classify the target condition?*

- Yes: for all studies using either a microbiological reference standard (i.e. culture, Xpert MTB/RIF, or Xpert Ultra) or a composite reference standard as described in Reference standards. These are the acceptable reference tests for inclusion of studies in the review.
- Given the criteria for including studies in this review, all included studies had a 'yes' response.

*Signalling question 2: were the reference standard results interpreted without knowledge of the results of the index test?*

- Yes: if the reference test provided an automated result (e.g. MGIT 960), blinding was explicitly stated, or it was clear that the reference standard was performed at a separate laboratory or performed by different people, or both.
- No: if the study stated that the reference standard result was interpreted with knowledge of the index test result.
- Unclear: if we could not determine. We also answered unclear if the study used a composite reference standard in which the index test was one of the components of the reference standard. In the latter situation, the study may have had incorporation bias where there could not be blinding of the reference standard to the index test. Incorporation of the index test in the reference standard may increase the amount of agreement between the index test results and reference standard thereby overestimating diagnostic accuracy.

Applicability: are there concerns that the target condition as defined by the reference standard does not match the question?

- High concern: if more than 50% of tuberculosis cases identified in the study did not have microbiologically confirmed tuberculosis.
- Low concern: if the children with tuberculosis in the study had signs and symptoms or chest radiograph abnormalities in addition to a positive culture or Xpert result.
- Unclear concern: if we could not determine.

**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 specimens obtained) at the same time or within one week.
- No: if the time between the screening test and reference standard (specimen collection) was more than one week.
- Unclear: if insufficient information was provided to decide.

2. Did all participants receive the same reference standard?

- Yes: if all participants were evaluated with the reference standard, and if all or most participants were evaluated with the same test(s).
- No: if not all participants were evaluated with the reference standard, or participants received different number of reference tests.
- Unclear: if insufficient information was provided to decide.

3. Were all participants included in the analysis?

- Yes: if all participants were included.
- No: if participants who participated were excluded, for example, cultures were lost or because they did not provide sputum for a reference test.
- Unclear: if insufficient information was provided to decide.

Judgements for 'risk of bias' assessments for a given domain.

- If we answered all signalling questions for a domain 'yes,' then we judged risk of bias as 'low.'
- If we answered all or most signalling questions for a domain 'no,' then we judged risk of bias as 'high.'
- If we answered only one signalling question for a domain 'no,' we discussed further the risk of bias judgement.
- If we answered all or most signalling questions for a domain 'unclear,' then we judged risk of bias as 'unclear.'
- If we answered only one signalling question for a domain 'unclear,' we discussed further the risk of bias judgement for the domain.

## WHAT'S NEW

| Date | Event | Description |
|---|---|---|
| 4 October 2021 | Amended | Minor typos corrected in PLS and Abstract |

## HISTORY

Protocol first published: Issue 7, 2020
Review first published: Issue 6, 2021

| Date | Event | Description |
|---|---|---|
| 2 July 2021 | Amended | Minor typos corrected in Abstract and PLS |

## CONTRIBUTIONS OF AUTHORS

LO developed the search strategy.

BV, TN, KRS, and AMM assessed articles for inclusion and extracted data.

BV entered data into Review Manager 5.

BV, YT, KRS, and AMM analyzed the data and interpreted the analyses. In particular, YT performed statistical analyses.

BV drafted the manuscript with contributions on content from TN, YT, AWK, SvW, BJM, KRS, and AMM.

All review authors reviewed and approved the final manuscript.

## DECLARATIONS OF INTEREST

BV: none.

TN: none.

YT: none.

AWK has conducted prior primary research on tuberculosis diagnostics. The Baylor College of Medicine Children's Foundation-Swaziland, where Dr Kay is based, received a discount from Cepheid on Xpert MTB/RIF Ultra cartridges for a tuberculosis case finding programme. The Baylor College of Medicine Children's Foundation-Swaziland is separate from Baylor College of Medicine (AK's employer).

SvW: none.

LO: none.

BJM: none.

KRS has received financial support for the preparation of systematic reviews and educational materials, consultancy fees from FIND (for the preparation of systematic reviews), honoraria, and travel support to attend WHO guideline meetings.

AMM has conducted prior primary research on tuberculosis diagnostics and 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
- World Health Organization (WHO), Switzerland

  Development and publication of this manuscript was in part made possible with financial support from the WHO.
  WHO Registration 2021/1090755-0
  Purchase Order 202638664

## DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We made the following changes from the review protocol (Vonasek 2020).

Assessment of type and number of chest radiography interpreters as a potential source of heterogeneity was added to Secondary objectives and noted in Investigations of heterogeneity.

We clarified in Types of studies that we excluded studies evaluating the index tests for extrapulmonary tuberculosis and studies where more than 25% of participants with active tuberculosis had extrapulmonary disease. Hence, included studies focused primarily on pulmonary tuberculosis.

In QUADAS-2, applicability for the reference standard was no longer determined by the proportion of those diagnosed clinically versus microbiologically as stated in the protocol. Therefore, we assessed all included studies as 'low concern' for reference standard applicability.

In Types of studies, we removed inclusion of studies that did not apply the reference standard to participants screening negative. We attempted to assess these types of studies, including only those with strict design criteria. However, given that only a small number of studies would be included under this criterion and those studies included heterogeneous populations, we decided not to analyse them in this review. Only calculation of positive predictive value would be feasible with these data, and sensitivity and specificity were the focus of this review.

As described in Participants, we modified the inclusion criteria for studies to allow for inclusion of participants between 15 and 19 years of age, but requiring that at least 75% of participants in any single included study were less than 15 years of age.

In Target conditions and Sensitivity analyses, we removed mention of performing sensitivity analysis for those studies that explicitly evaluated the index tests for pulmonary tuberculosis. This sensitivity analysis was not conducted because most included studies did not clearly describe explicit evaluation of only pulmonary tuberculosis.

As described in Statistical analysis and data synthesis, we developed symptom groups for meta-analysis of similar composite symptom screens.

In Statistical analysis and data synthesis, we removed mention of performing test comparisons because these were not done due to limited data.

In 'Assessment of certainty of the evidence', we explained that prespecified analyses were not performed owing to limited data.

## INDEX TERMS

### Medical Subject Headings (MeSH)

Bias; Child Behavior; Cohort Studies; Confidence Intervals; *Contact Tracing; Cough [diagnosis]; Cross-Sectional Studies; False Negative Reactions; False Positive Reactions; Fever [diagnosis]; HIV Infections [epidemiology]; Mass Screening [statistics & numerical data]; Molecular Diagnostic Techniques; Radiography, Thoracic; Reference Standards; Sensitivity and Specificity; Symptom Assessment [*methods] [statistics & numerical data]; Tuberculosis, Pulmonary [*diagnosis] [epidemiology] [prevention & control]; Weight Gain

### MeSH check words

Adolescent; Child; Child, Preschool; Humans