



Cochrane
Library

Cochrane Database of Systematic Reviews

Screening tests for active pulmonary tuberculosis in children (Protocol)

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 (Protocol).
Cochrane Database of Systematic Reviews 2020, Issue 7. Art. No.: CD013693.
DOI: [10.1002/14651858.CD013693](https://doi.org/10.1002/14651858.CD013693).

www.cochranelibrary.com

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
BACKGROUND	2
Figure 1.	3
Figure 2.	5
OBJECTIVES	6
METHODS	7
ACKNOWLEDGEMENTS	11
REFERENCES	12
APPENDICES	15
HISTORY	18
CONTRIBUTIONS OF AUTHORS	18
DECLARATIONS OF INTEREST	18
SOURCES OF SUPPORT	19

[Diagnostic Test Accuracy Protocol]

Screening tests for active pulmonary tuberculosis in children

Bryan Vonasek¹, Tara Ness¹, Yemisi Takwoingi², Alexander W Kay¹, Susanna S van Wyk³, Lara Ouellette⁴, Ben J Marais⁵, Karen R Steingart^{6a}, Anna M Mandalakas^{1b}

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

^aThese authors contributed equally to this work. ^bThese authors contributed equally to this work

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

Editorial group: Cochrane Infectious Diseases Group.

Publication status and date: New, published in Issue 7, 2020.

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 (Protocol). *Cochrane Database of Systematic Reviews* 2020, Issue 7. Art. No.: CD013693. DOI: [10.1002/14651858.CD013693](https://doi.org/10.1002/14651858.CD013693).

Copyright © 2020 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Objectives

This is a protocol for a Cochrane Review (diagnostic). The objectives are as follows:

To determine the sensitivity, specificity, and positive and negative predictive value of 1) the presence of one or more tuberculosis symptoms, or symptom combinations; 2) chest radiography; 3) Xpert MTB/RIF; 4) Xpert Ultra; and 5) combinations of the aforementioned tests as screening tests for detecting active pulmonary tuberculosis in children in the following groups.

- Household contacts of a person with active tuberculosis;
- School contacts of a person with active tuberculosis;
- Other close contacts of a person with active 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 burden settings

Secondary objectives

To compare the accuracy of the different index tests, including different applications of tests (e.g. CXR with any abnormality versus, more specifically, CXR with abnormality suggestive of tuberculosis); we are interested in the accuracy of the index tests in any setting (i.e. community, outpatient, and inpatient).

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, and whether the child received a single screening or more than one screening.

BACKGROUND

Tuberculosis continues to elude traditional control strategies. According to the [WHO Global TB Report 2019](#), an estimated 10 million people in 2018 fell 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 11% of incident cases but 14% of the estimated 1.3 million deaths from tuberculosis in 2018. 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 TB Report 2019](#)).

Case finding is a crucial step in the cascade of care for patients with tuberculosis; however, for most deaths from tuberculosis in children, the disease is never diagnosed ([Jenkins 2017](#)). In the "Roadmap towards ending TB in children and adolescents," the WHO identifies case finding for child tuberculosis as a key activity ([WHO 2018](#)). Major factors lead to underdiagnosis of tuberculosis in children including 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 adult populations, 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 adult populations ([WHO 2013a](#)). A 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 15 years and older. The review excluded studies of children 0 to 5 years of age or studies of child tuberculosis only. Since 2013, estimation of the true burden of child 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 active tuberculosis in children ([Reuter 2019](#); [WHO 2018](#)). This review will address tuberculosis screening strategies in children less than 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, 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 those with tuberculosis infection initially develop active tuberculosis (tuberculosis disease). More commonly, initial infection leads to latent tuberculosis infection, which has the potential to become active tuberculosis throughout an individual'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 (termed 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 will be pulmonary tuberculosis.

Signs and symptoms of pulmonary tuberculosis include fever, cough, night sweats, weight loss or poor weight gain, visible neck mass and lethargy or decreased activity/playfulness. 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](#); [Marais 2014](#)).

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 adult pulmonary tuberculosis is uncommon in children, especially in young children (<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 includes the following index tests used in screening for pulmonary tuberculosis in children: tuberculosis symptoms, chest radiography (CXR), Xpert MTB/RIF and Xpert Ultra, and various combinations of these tests.

With symptom screening, individuals or their caregivers are interviewed about symptoms suggestive of pulmonary tuberculosis such as cough of varying duration, fever of varying duration, weight loss or poor weight gain, night sweats, visible neck mass, and decreased activity. Though not a true symptom, contact with individuals with tuberculosis is another important factor when interviewing for tuberculosis risk ([Graham 2015](#)).

CXR can involve posterior-anterior, anterior-posterior, and/or lateral recording. Commonly used types of CXR include conventional CXR (producing 36 cm x 43 cm film), digital radiography, and computed radiography. The most common radiographic finding of pulmonary tuberculosis in children is hilar lymphadenopathy (Leung 1992), though CXR has limitations identifying this finding (Swingler 2005). Accurate interpretation of CXR findings for pulmonary tuberculosis in children is dependent on the ability of the individual interpreting the CXR, and wide interobserver variation has been reported (Du 2002; Kaguthi 2014). Computer-aided interpretation of CXR for pulmonary tuberculosis is a promising new technology, especially for resource-limited settings where expertise in CXR interpretation is limited (Qin 2019; Sodhi 2017), but this technology has currently not been assessed in children (Reuter 2019).

Xpert MTB/RIF and Xpert Ultra, the newest version, (Cepheid Inc, CA, USA) are nucleic acid amplification tests (NAAT) that can detect both *M tuberculosis* DNA and rifampicin resistance. We will 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. Consistent supply of electricity, temperature control, and annual calibration of the cartridge modules are needed for Xpert MTB/RIF and Xpert Ultra (Global Laboratory Initiative 2019). Xpert Ultra has approximately 1-log improvement in the lower limit of detection of bacterial load compared to previous generations of the test (Chakravorty 2017). Xpert Ultra also has a new result category, 'trace call,' that represents minimally detectable bacillary burden. According to the WHO, a 'trace call' result is adequate to prompt

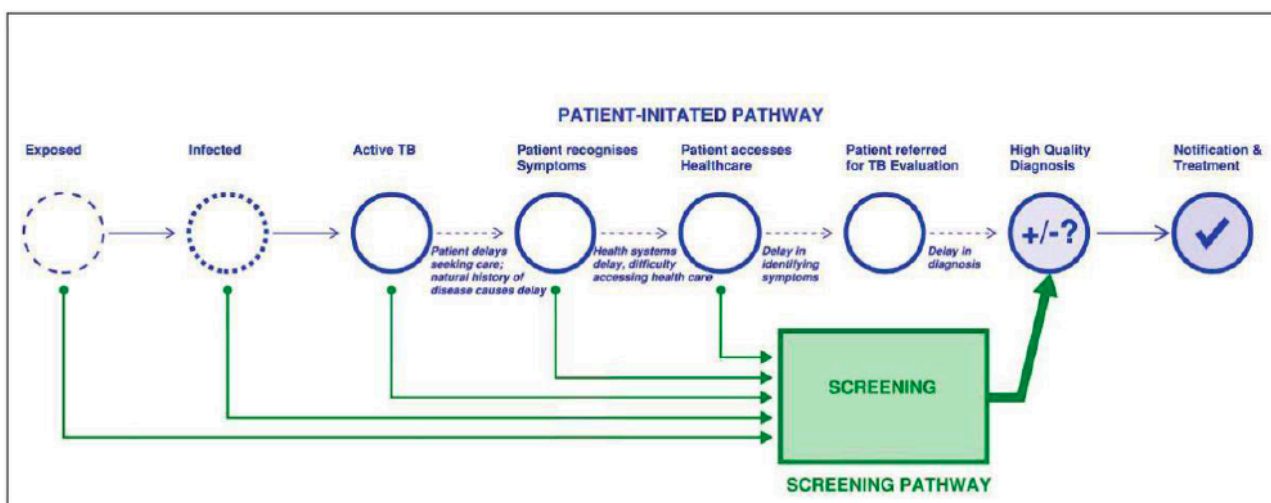
initiation of anti-tuberculosis therapy in children or people living with HIV (WHO 2017b). The WHO recommends the use of Xpert MTB/RIF and Xpert Ultra as initial diagnostic tests for pulmonary tuberculosis in adults and children. Specifically in children, the guidelines recommend use of a variety of specimen types for diagnosis including gastric specimens, nasopharyngeal specimens, and stool specimens, in addition to sputum (WHO 2020). We will include Xpert MTB/RIF (all versions) and Xpert Ultra in this review.

Another WHO-recommended NAAT for detection of tuberculosis is the TrueNat assay (Molbio Diagnostics/Bigtec Labs, Goa/Bengaluru, India) (WHO 2020). However, to our knowledge, there are currently no published studies assessing this test in children. We plan to include TrueNat in this review if data become available while we perform the review.

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 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 in children, ranging from contacts of those with tuberculosis ('exposed') to symptomatic patients accessing healthcare (e.g. children living with HIV, as described above). This review will include evidence from all of these systematic screening strategies.

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 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 less than 15 years old. For the subgroup of children living with HIV, since

2011 the WHO has recommended symptom screening for all children living with HIV presenting to healthcare facilities. Under

this guideline, children living with HIV older than 12 months of age presenting with any cough, fever, weight loss or poor weight gain, or history of contact with someone with tuberculosis should be further investigated for tuberculosis. In the absence of any of these four symptoms, they "are 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 that is otherwise lacking for the paediatric population.

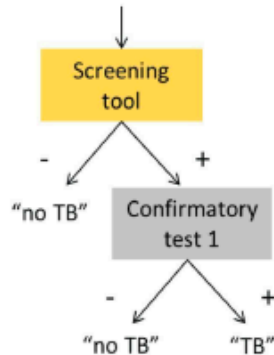
Screening may use sequential or parallel strategies (Figure 2). With sequential strategies, only those with a positive result in the first step are screened in the second step. With parallel screening strategies, multiple different screens are done initially, and any positive screen or combinations of positive screens prompts further investigation (i.e. confirmatory test) for the target condition. Results from various screening strategies will be included in this

review. We will consider individuals' results to be 'true screen positives' if they were rightfully referred for confirmatory testing; in contrast, we will consider 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 conduct confirmatory testing on those with positive screens will be analysed in this review. In the context of this review, the purpose 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).

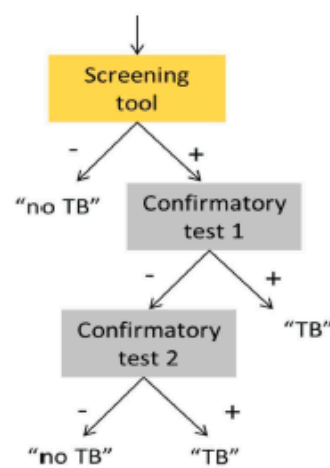
Figure 2. Different screening and diagnostic algorithms

Different screening and diagnostic algorithms.

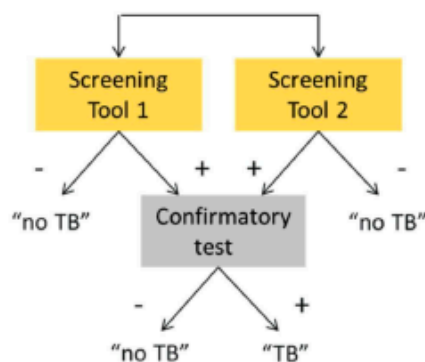
A. One screening test followed by one confirmatory test



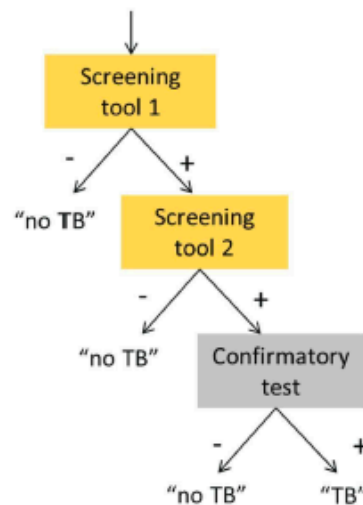
B. One screening test followed by two sequential confirmatory tests



C. Two parallel screening tests



D. Two sequential screening tests



Alternative test(s)

Two types of immunologic tests not included in this review are the tuberculin skin test (TST) and the interferon gamma release assay (IGRA). Both of these 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 rather may be used to support tuberculosis diagnosis. 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, United Kingdom). 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 subjects with active pulmonary tuberculosis compared to 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 those with active tuberculosis and people living with HIV (Theel 2018). The T-SPOT.TB is an enzyme-linked immunoassay that involves incubation of separated and counted 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 a number of alternative approaches and tests for screening (or diagnosis) of tuberculosis in children. Prior to the widespread roll out of Xpert MTB/RIF worldwide over the past decade, testing for tuberculosis mostly relied upon examination of sputum smears for acid-fast bacilli under a light microscope using the classical Ziehl-Neelsen staining technique, fluorescence microscopy, or light-emitting diode (LED)-based fluorescence microscopy. A recent review found that in children, the sensitivity of smear microscopy was about 22% in gastric specimens and about 29% in expectorated and induced sputum specimens (WHO 2013b).

A variety of newer assays detect lipoarabinomannan (LAM) antigen in urine of people with tuberculosis disease. LAM is a lipopolysaccharide present in the mycobacterial cell wall. Urine LAM assays have the advantages of being non-invasive and rapid. Currently, the only rapid commercially available LAM assay is the Determine TB-LAM (Alere, Waltham, MA, USA). Based on evidence from randomized trials and a Cochrane Review (Bjerrum 2019), the WHO recommends that AlereLAM 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 here: WHO Lateral flow LAM 2019. Another LAM assay expected to become commercially available in 2020 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 2019).

It is noteworthy that development of novel tools for detection of tuberculosis disease is a very 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)). Over the next decade, more efficient technologies are anticipated with the hope that these will advance screening strategies and reduce the burden of child tuberculosis worldwide (Schumacher 2019; Stop TB Partnership 2019; WHO 2017a).

Rationale

Effective screening for active tuberculosis in children 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 preventive treatment for exposed or other high-risk groups such as people 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 errors and consequent over-treatment of tuberculosis. Guidance from the WHO states that "only children who are close contacts of someone with tuberculosis and HIV-positive children should be systematically screened for TB" (WHO 2015). Optimal screening strategies for these two high-risk groups are lacking, particularly in resource-limited settings (Szkwardo 2017, WHO 2011). Further, limiting systematic screening to child contacts and HIV-positive children propagates 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). Finally, increasing evidence demonstrates that children have considerable risk of tuberculosis exposure outside of their homes with up to 70% to 90% of children with tuberculosis having no known exposure (Martinez 2019).

This Cochrane Review will inform an upcoming WHO meeting to update guidelines for systematic screening for active tuberculosis. To our knowledge, this is the first systemic review on this topic in children. There have been several systematic reviews evaluating the diagnostic accuracy of the index tests described above for active tuberculosis, including an ongoing Cochrane Review evaluating Xpert MTB/RIF and Xpert Ultra in children (Kay 2019). The lack of knowledge regarding the performance of screening tests in children likely reflects the predominance of paediatric research which has assessed the performance of tuberculosis tests for diagnosis rather than screening. The current review will shed light on the potential of these tools for systematic screening for active pulmonary tuberculosis in children in specific high-risk populations.

OBJECTIVES

To determine the sensitivity, specificity, and positive and negative predictive value of 1) the presence of one or more tuberculosis symptoms, or symptom combinations; 2) chest radiography; 3) Xpert MTB/RIF; 4) Xpert Ultra; and 5) combinations of the aforementioned tests as screening tests for detecting active pulmonary tuberculosis in children in the following groups.

- Household contacts of a person with active tuberculosis;
- School contacts of a person with active tuberculosis;
- Other close contacts of a person with active 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 burden settings

Secondary objectives

To compare the accuracy of the different index tests, including different applications of tests (e.g. CXR with any abnormality versus, more specifically, CXR with abnormality suggestive of tuberculosis); we are interested in the accuracy of the index tests in any setting (i.e. community, outpatient, and inpatient).

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, and whether the child received a single screening or more than one screening.

METHODS

Criteria for considering studies for this review

Types of studies

We will include cross-sectional studies, cohort studies, and randomised controlled trials that assessed the accuracy of at least one of the defined index tests for pulmonary tuberculosis. We will only include studies that used a microbiologic reference standard (defined below). We will include studies from all settings and time periods. Randomized controlled trials will be included because these studies may report sensitivity and specificity in addition to patient health outcomes. For randomised studies that compare different screening strategies, we will evaluate each arm as a separate cohort. Data on the index test(s) must be available to be extracted as true positive, false positive, true negative, and false negative against the reference standard(s) so that we can construct two-by-two contingency tables.

Studies applying index tests multiple times to an individual within a short timeframe (e.g. within a single hospital admission), will be considered diagnostic rather than using a screening approach, and these studies will be excluded. Studies in which children with negative screening test results were not subjected to the reference standard will be included. As shown in [Figure 2](#), this often occurs in tuberculosis prevalence studies when it is assumed that those with a negative screen (e.g. no CXR abnormalities) do not have active tuberculosis. However, this leads to poor enumeration of true negative and false negative test results. Therefore, we will only include such studies that partially verified tuberculosis status if they were conducted prospectively and enrolled a consecutive or random sample of eligible children. Our rationale for specifying these strict design criteria is to enable us to calculate positive predictive values from such studies and include them in a separate set of analyses. Due to the direct relationship between prevalence and predictive values, to reduce potential variation in prevalence between studies, we will only include studies in the analyses of positive predictive value if the studies were done in the same setting (e.g. community or accessing healthcare), target population, and for the same purpose (e.g. contact tracing).

We will include 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 some time 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 sure that the target condition did not change between the two tests. We will address this issue in the QUADAS-2 domain 4: Flow and timing and with a sensitivity analysis (see [Sensitivity analyses](#)).

We will exclude case reports and case-control studies, the latter because of the high risk of bias in diagnostic accuracy studies ([Rutjes 2006](#)).

Participants

We will include studies enrolling HIV-positive and HIV-negative children younger than 15 years old not known to have active tuberculosis prior to screening. We will exclude studies if they include older children and we are unable to extract data for children younger than 15 years old from the publication. We will include children in the general population in high burden settings, children living in areas of high tuberculosis burden, 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 will include studies in which children are screened only once and studies that report longitudinal screening with repeated screening tests at predetermined intervals.

Index tests

For symptom screening, we will include studies that assess symptoms of tuberculosis or combinations of symptoms as described by the primary study authors. Symptoms of active tuberculosis in children may include cough, fever, decreased appetite, weight loss or failure to thrive, and fatigue or reduced playfulness. Older children may experience symptoms similar to those in adults and include persistent cough, haemoptysis, and weight loss, fever, night sweats and fatigue. The threshold was presence or absence of symptoms.

For CXR screening, we will include studies that utilize conventional radiography, digital radiography, and computed radiography. We will include all classification systems for identification of CXR abnormalities, including automated interpretation of radiography using deep learning or artificial intelligence technology. We will categorize all CXR screening results as follows. We will use 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 will include studies in which the index tests are 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 will also include studies assessing stool specimens. We will include studies assessing more than one type of respiratory specimen collected at the same time and extract 2 x 2 data separately for each specimen type.

Xpert MTB/RIF provides 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 will be categorized as:

- Positive: 'MTB DETECTED,' including 'trace' results from Xpert Ultra
- Negative: 'MTB NOT DETECTED'
- Inconclusive: 'INVALID,' 'ERROR,' or 'NO RESULT'

We will not evaluate 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 is also available, these data will also be extracted for analysis.

Target conditions

The target condition is active pulmonary tuberculosis.

We anticipate that some studies may evaluate the index tests for active tuberculosis and not explicitly state 'pulmonary tuberculosis', the target condition in this review. We will include these studies because the most common type of active tuberculosis in children is lung disease; hence, most screening studies in children evaluate tests for pulmonary tuberculosis and diagnose tuberculosis using respiratory specimens. If data are sufficient, we will perform a sensitivity analysis limiting inclusion to those studies that explicitly evaluated the index tests for pulmonary tuberculosis.

Reference standards

We will utilize two reference standards, a microbiological and a composite reference standard.

Microbiological reference standard

Confirmed pulmonary tuberculosis will be 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 either the Xpert MTB/RIF or Xpert Ultra test is the index test, we will not include the test as a reference standard to avoid incorporation bias. We will not include studies where sputum smear microscopy is the reference standard.

Collection of multiple respiratory specimens may improve the diagnostic yield of testing for tuberculosis in children ([Cruz 2012](#); [Zar 2012](#)). With respect to the microbiological reference standard, we will include studies that involve multiple specimens collected over time. In these studies, we will utilize 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

We will define the composite reference standard as microbiological confirmation (as above) or author-defined clinical pulmonary tuberculosis. Clinical pulmonary tuberculosis must include a component of follow-up to verify the diagnosis of active tuberculosis. The consensus research definition of clinical tuberculosis in children ([Graham 2015](#)) is likely too restrictive for the purpose of this review. 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 raises the potential for incorporation bias with the composite reference standard, i.e. where the result of the index test is used to help determine the reference standard diagnosis. Therefore, we will assess the composite reference standard for incorporation bias using QUADAS-2, which will be enhanced with the following additional signalling question: "Was incorporation bias avoided (inclusion of index test as part of the reference standard)?" In addition, we will discuss incorporation bias as a limitation of the review.

We will define 'not tuberculosis' 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.

Search methods for identification of studies

We will attempt to identify all relevant published studies regardless of language. We will describe unpublished studies in the Ongoing studies section of the review. Although they will not be assessed as index tests in this review, we will include immunologic tests in the search strategy. This will allow for archiving of relevant studies for a future systematic review assessing immunologic tests as index tests.

Electronic searches

We will search the following databases without language restriction, using the search terms and strategy described in [Appendix 1](#).

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

We will also search 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 will contact experts in the field, and check the references of relevant reviews from the past ten years. With the studies selected for inclusion in this review, we will perform forward and backward reference checking to identify any additional eligible studies.

Data collection and analysis

Selection of studies

We will use Covidence to manage the selection of studies (Covidence 2017). Two review authors (BV and TN) will independently screen all titles and abstracts from the electronic searches to identify potentially eligible studies. We will obtain full-text articles of potentially eligible studies, and the two review authors (BV and TN) will independently assess the full-text articles for study eligibility using the predefined inclusion and exclusion criteria. We will resolve any disagreements by discussion or with a third review author (AMM). As needed, we will contact study authors to clarify the study methods and other information. Studies excluded during the full-text review will be listed in the 'Characteristics of excluded studies' table with a summary of reasons for exclusion. We will illustrate the study selection process in a PRISMA flow diagram (Moher 2009).

Data extraction and management

We will design a data extraction form and pilot it on at least two included studies. After reviewing the piloted forms with the other review authors, we will finalize the form. Two review authors will use the data extraction form to independently extract data from the included studies. We will discuss any inconsistencies with a third review author. We will enter the extracted data into an Excel database (Excel 2013) on password-protected computers and secured in the cloud storage Dropbox for future review updates.

We will extract the following information for 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 systematic review analysis.
- Single or initial screening versus more than one screening in the population.
- Any sequential or parallel screening strategies.

Patient characteristics and setting

- Description of study population.
- Age: median, mean, range, and disaggregation into categories (0 to 4, 5 to 14)
- 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 high tuberculosis burden country (WHO Global TB Report 2019).
- Years of data collection.

Index test

- Definition of positive symptom screen.
- List 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 are 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.
- Reference standard applied to all children or only those with a positive screening test result?
- 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.

Assessment of methodological quality

Two review authors will independently assess the methodological quality of the included studies using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) instrument, which we will adapt for this review (Whiting 2011). The preliminary tool with signalling questions tailored to this review is in Appendix 2. As recommended, we will assess each of the four domains (patient selection, index tests, reference standards, and flow and timing) for risk of bias and the first three domains for concerns regarding applicability.

We will judge each item as 'yes' (adequately addressed), 'no' (inadequately addressed), or 'unclear' when there is insufficient information reported to make an assessment. One review author will pilot the tool on two included studies. We will then make any necessary revisions to finalize the tool. We will resolve disagreements between the two review authors' independent assessments through discussion or additional input from a third review author. We will present results of the quality assessment in text, tables, and graphs.

Statistical analysis and data synthesis

We will perform descriptive analyses of the included studies and present their key characteristics in the 'Characteristics of included studies' table and a summary table. We will present individual study estimates of sensitivity and specificity graphically on forest plots and in receiver operating characteristics (ROC) space using Review Manager 5 (RevMan 2020).

We will consider one index test result per child per time point. However, for studies assessing serial screening over time for individuals, separate screens may be assessed if they are also

compared against serial confirmatory tests over time (i.e. multiple screens for one individual). Within each group listed in [Objectives](#), we will perform analyses by index tests and reference standards. For symptom screening as the index test, we plan to perform analyses by single symptoms and multiple symptoms (such as the WHO four-symptom screen), where data are available.

We will consider combining categories depending on the number of studies and screening definitions found in each category. We will also stratify the analyses by the type of reference standard used, microbiological or composite. Separate analyses will be performed for studies that verify participants regardless of their index test results (i.e. complete verification) and those that only verify test positives (i.e. partial verification).

When there are sufficient data, we will perform meta-analyses to estimate summary values of sensitivity and specificity using a bivariate model ([Chu 2006](#); [Reitsma 2005](#)). We chose the bivariate model because we anticipate dealing with binary test results or studies that used the same threshold because they applied the threshold recommended by the test manufacturer. Also, we note that the bivariate model is appropriate to use for index tests such as Xpert MTB/RIF and Xpert Ultra, which apply a common positivity criterion ([Macaskill 2010](#)). When we are unable to fit the models due to sparse data or few studies, we will simplify the models to univariate random-effects logistic regression models to pool sensitivity and specificity separately ([Takwoingi 2015](#)). For studies that verify only test positives, we will pool positive predictive values using a univariate random-effects logistic regression model. We will perform meta-analyses using the `meqrlogit` command in Stata version 16 ([Stata 16](#)).

For test comparisons, we will exclude studies where one index test was used as the reference standard for another index test in the comparison e.g. if Xpert MTB/RIF was used as the reference standard for CXR. We will perform comparative meta-analyses by first including all studies with relevant data in indirect comparisons to make use of all available data. We will then perform additional analyses by restricting the analyses to only comparative studies that made direct comparisons between the index tests within the same study population. Comparative meta-analyses will be performed using bivariate meta-regression by adding test-type as a covariate to bivariate models. We will assess model fit using likelihood ratio tests to compare models with and without the covariate terms. We will calculate absolute differences in sensitivity and specificity using the model parameters. We will obtain 95% confidence intervals and P values for the absolute differences using the delta method and Wald tests, respectively.

For subgroups or screening definitions that do not have sufficient data for a meta-analysis, we will summarize findings using descriptive methods.

Approach to inconclusive index test results

As described above, the NAAT assays assessed in this review as index tests may have inconclusive results. We will report the proportion of inconclusive index test results. Depending on the available data, we will reclassify these results as positive or negative and perform additional analyses to determine the impact of including these test results on test accuracy.

Investigations of heterogeneity

We will visually inspect forest plots and summary ROC (SROC) plots for heterogeneity. When data allow, we will evaluate potential sources of heterogeneity using subgroup analyses and bivariate meta-regression.

For subgroup analyses, we will assess the following subgroups: children aged 0 to 4 years, children aged 5 to 14 years, HIV positive children, and HIV negative children.

For meta-regression analyses, we will include high tuberculosis burden country (yes or no) and single or initial screening versus more than one screening and consider each source of heterogeneity as a single covariate in a bivariate model.

Sensitivity analyses

When data allow, we will perform sensitivity analyses and explore the effect of risk of bias and study characteristics on the accuracy of index test results by limiting inclusion in the meta-analyses to the following.

- Studies that only used consecutive or random selection of participants.
- Studies with an appropriate interval between the index test and the reference standard.
- Studies that avoid incorporation bias (inclusion of index test as part of the reference standard).
- Studies that explicitly evaluated the index tests for pulmonary tuberculosis.

Assessment of reporting bias

We will 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 will assess the certainty of evidence using the GRADE approach for diagnostic studies ([Balslem 2011](#); [Schünemann 2008](#)). As recommended, we will rate the certainty of evidence as either 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 starts as high when there are high-quality observational studies (cross-sectional or cohort studies) that enrolled participants with diagnostic uncertainty. If we find a reason for downgrading, we will use our judgement to classify the reason as either serious (downgraded by one level) or very serious (downgraded by two levels).

Three review authors will discuss judgements and apply GRADE in the following way ([Schünemann 2020a](#); [Schünemann 2020b](#)).

Assessment of risk of bias

We will use QUADAS-2 to assess risk of bias.

Indirectness

We will assess indirectness in relation to the population (including disease spectrum), setting, interventions, and outcomes (accuracy

measures). We will also use 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 will carry out prespecified analyses to investigate potential sources of heterogeneity and downgrade when we cannot explain inconsistency in the accuracy estimates.

Imprecision

We will consider a precise estimate to be one that would allow a clinically meaningful decision. We will consider the width of the CI, and ask, "Would we make a different decision if the lower or upper boundary of the CI represented the truth?" In addition, we will work out projected ranges for TP, FN, TN, and FP for a given prevalence of tuberculosis and make judgements on imprecision from these calculations.

Publication bias

We will rate 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.

ACKNOWLEDGEMENTS

We are grateful to Vittoria Lutje, the Cochrane Infectious Diseases Group (CIDG) Information Specialist, for help with the search strategy.

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.

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.

REFERENCES

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 VVS, 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 P, 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, Zwering 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](https://doi.org/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. *British Medical Journal* 2006;**332**(7549):1089-92.

Broger 2019

Broger T, Sossen B, Du Toit E, Kerkhoff AD, Schutz C, Reipold EI, et al. Novel lipoarabinomannan point-of-care tuberculosis test for people living with HIV: a diagnostic accuracy study. *Lancet Infectious Diseases* 2019;**19**(8):707-16.

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.

Chisti 2014

Chisti MJ, Graham SM, Duke T, Ahmed T, Ashraf H, Faruque AS, 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.

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 2017 [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.

Dropbox [Computer program]

Dropbox Inc. Dropbox. San Francisco: Dropbox Inc., 2007.

Du 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.

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.

Excel 2013 [Computer program]

Microsoft Excel Software.. Version 16.16.4. Seattle: 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 5 Feb 2020) (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 HMS, 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 2019

Kay AW, González FL, Takwoingi Y, Eisenhut M, Vu RD, Steingart KR. Xpert MTB/RIF and Xpert MTB/RIF Ultra assays for active tuberculosis and rifampicin resistance in children. *Cochrane Database of Systematic Reviews* 2019, Issue 6. Art. No: CD013359. [DOI: [10.1002/14651858.CD013359](https://doi.org/10.1002/14651858.CD013359)]

LaCourse 2014

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.

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, 2013. Available from srdta.cochrane.org.

Marais 2004

Marais BJ, Gie RP, Schaaf HS, Hesselting 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 2014

Marais BJ, Schaaf HS. Tuberculosis in children. *Cold Spring Harbor Perspectives in Medicine* 2014;**4**:1-12.

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](https://doi.org/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.

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.

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.

RevMan 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.

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 16 [Computer program]

Stata Statistical Software. StataCorp, Version 16. College Station: StataCorp LP, 2019.

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 MMG, et al. A systematic review of the sensitivity and specificity of symptom- and chest-radiography screening for active pulmonary tuberculosis in HIV-negative persons and persons with unknown HIV status. 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 MMG, 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](https://doi.org/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.

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 12 June 2020).

WHO 2013a

World Health Organization. Systematic screening for active tuberculosis: principles and recommendations. <https://www.who.int/tb/tbscreening/en/> (accessed prior to 12 June 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 12 June 2020).

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=A852BCDF07D5D7FAC40F5BD7E064988?sequence=1 (accessed prior to 12 June 2020).

APPENDICES

Appendix 1. MEDLINE search strategy

TB Original Search Strategy Designed for MEDLINE OVID

- 1.exp child/ or exp infant/
- 2.exp pediatrics/
- 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 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 preadolescenc* or "pre-adolescenc*" or pubescen* or prepubescen* or "pre-pubescen*").ti,ab,kw.
- 6.1 or 2 or 3 or 4 or 5
- 7.exp Mycobacterium tuberculosis/
- 8.exp Tuberculosis/
- 9.(tuberculos* or tb* or ptb).ti,ab,kw.
- 10.7 or 8 or 9

WHO 2017a

World Health Organization. Tuberculosis diagnostics technology landscape. unitaid.org/assets/2017-Unitaid-TB-Diagnostics-Technology-Landscape.pdf (accessed prior to 11 June 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/ 2017.

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 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](https://www.who.int/publications/i/item/who-consolidated-guidelines-on-tuberculosis-module-3-diagnosis---rapid-diagnostics-for-tuberculosis-detection) (accessed 1 July 2020).

WHO Global TB Report 2019

World Health Organization. Global tuberculosis report 2019. www.who.int/tb/publications/global_report/en/ (accessed prior to 12 June 2020).

WHO Lateral flow LAM 2019

World Health Organization. Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV. Policy update 2019. https://www.who.int/tb/publications/2019/diagnose_tb_hiv/en/ (accessed prior to 12 June 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**(6):1088-95.

- 11.((active* or symptomatic*) adj3 (tuberculosis* or tb*)).ti,ab,kw.
- 12.("active tuberculosis*" or "active tb*").kw.
- 13.("symptomatic* tuberculosis*" or "symptomatic* tb*").kw.
- 14.11 or 12 or 13
- 15.10 and 14
- 16.exp Symptom Assessment/ or exp symptom flare up/
- 17.(symptom* or manifest*).ti,ab,kw.
- 18.16 or 17
- 19.Cough/
- 20.Cough*.ti,ab,kw.
- 21.Hemoptysis/
- 22.(hemoptysis* or "hemo-ptysis*").ti,ab,kw.
- 23.Fever/
- 24.(fever* or "high* temp*").ti,ab,kw.
- 25.Weight Loss/
- 26.("weight loss*" or weightloss*).ti,ab,kw.
- 27.19 or 20 or 21 or 22 or 23 or 24 or 25 or 26
- 28.18 and 27
- 29.exp Radiography, Thoracic/
- 30.((chest* or lung* or thoracic*) adj3 ("x-ray*" or xray* or radiogra* or imag*)).ti,ab,kw.
- 31.("chest x-ray*" or "chest xray*" or "chest radiogra*" or "chest imag*").kw.
- 32.("lung x-ray*" or "lung xray*" or "lung radiogra*" or "lung imag*").kw.
- 33.("thoracic x-ray*" or "thoracic xray*" or "thoracic radiogra*" or "thoracic imag*").kw.
- 34.exp tuberculin test/
- 35.("tubercul* skin test*" or tst or "tb skin test*").ti,ab,kw.
- 36.exp Interferon-gamma Release Tests/
- 37.((interferon-gamma releas*" or "IFN-gamma releas*") adj3 (test* or assay*)).ti,ab,kw.
- 38.("interferon-gamma releas* test*" or "interferon-gamma releas* assay*").kw.
- 39.("IFN-gamma releas* test*" or "IFN-gamma releas* assay*").kw.
- 40.igra.ti,ab,kw.
- 41.("QuantiFERON-TB*" or QuantiFeronTB* or QFT* or "T-Spot*" or TSpot*).ti,ab,kw.
- 42.("immunologic* test*" or "immuno-logic test*").ti,ab,kw.
- 43.("microbiologic* confirm*" or "micro-biologic* confirm*").ti,ab,kw.
- 44.("Mycobacter* tubercul*" or MTB) adj3 (culture* or test* or assay*).ti,ab,kw.
- 45.("Mycobacter* tubercul* culture*" or "Mycobacter* tubercul* test*" or "Mycobacter* tubercul* assay*").kw.
- 46.("MTB culture*" or "MTB test*" or "MTB culture*").kw.
- 47.((tubercul* or TB) adj3 (test* or assay*)).ti,ab,kw.
- 48.("tubercul* test*" or "tubercul* assay*" or "TB test*" or "TB assay*").kw.
- 49.((Xpert* or GeneXpert*) adj3 (MTB or RIF or rifampicin* or Ultra)).ti,ab,kw.
- 50.("Xpert MTB*" or "Xpert RIF*" or "Xpert rifampicin*" or "Xpert Ultra*").kw.
- 51.("GeneXpert MTB*" or "GeneXpert RIF*" or "GeneXpert rifampicin*" or "GeneXpert Ultra*").kw.
- 52.(GeneXpert* or Xpert*).ti,ab,kw.
- 53.(TrueNat or "True-Nat" or TruNat or "Tru-Nat").ti,ab,kw.
- 54.("nucleic acid amplification test*" or NAAT).ti,ab,kw.
- 55.18 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 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54
56. 6 and 15 and 55

We will use this search strategy for Medline OVID and adapt it for the other electronic databases listed.

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 patients.
- No: if the study selected patients by convenience.
- Unclear: if the study did not report the manner of patient selection or we cannot tell.

Signalling question 2: Did the study avoid inappropriate exclusions? Examples of inappropriate exclusions may include 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 cannot tell.

Applicability: Are there concerns that the included patients and setting do not match the review question?

High concern: if the study population does not resemble a population that would be considered for screening for tuberculosis in a community setting, such as household or schools.

Low concern: if the study population resembles a population that would be considered for screening for tuberculosis in a community setting; screening program in practice.

Unclear concern: if we cannot tell.

Domain 2: Index test

Risk of bias: Could the conduct or interpretation of the index test have introduced bias?

Symptom screen, CXR, 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 CXR was interpreted with knowledge of the results of the reference test.
- Unclear: if we cannot tell. For example, if it was unclear whether the CXR 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 CXR

Yes: if the study clearly reports positivity criteria for abnormalities suggestive of tuberculosis or other abnormalities.

No: if the study does not report the positivity criteria for abnormalities suggestive of tuberculosis or other abnormalities.

Unclear: if we cannot tell.

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 cannot tell.

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 will have 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 (for example, MGIT 960), blinding was explicitly stated, or it was clear that the reference standard was performed at a separate laboratory and/or performed by different people.

- No: if the study stated that the reference standard result was interpreted with knowledge of the index test result.
- Unclear: if we cannot tell. We will also answer 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 suffer from incorporation bias where there cannot 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 > 50% of tuberculosis cases identified in the study did not have microbiologically confirmed tuberculosis.
- Low concern: if the children with tuberculosis in the study have signs and symptoms or CXR abnormalities in addition to a positive culture or Xpert result.
- Unclear concern: if we cannot tell.

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 1 week; and
- Unclear: if insufficient information is provided to decide.

2. Did all patients receive the same reference standard?

- Yes: if all participants were evaluated with the reference standard, and if all or 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 is provided to decide.

3. Were all patients included in the analysis?

- Yes: if all participants were included;
- No: if participants who participated were excluded, for example, cultures were lost or because they did not provide sputum for a reference test;
- Unclear: if insufficient information is 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 will judge risk of bias as 'low'.
- If we answered all or most signalling questions for a domain 'no', then we will judge risk of bias as 'high'.
- If we answered only one signalling question for a domain 'no', we will discuss further the risk of bias judgement.
- If we answered all or most signalling questions for a domain 'unclear', then we will judge risk of bias as 'unclear'.
- If we answered only one signalling question for a domain 'unclear', we will discuss further the risk of bias judgement for the domain.

HISTORY

Protocol first published: Issue 7, 2020

CONTRIBUTIONS OF AUTHORS

BV drafted the protocol with contributions on content from TN, AWK, SvW, BJM, and AMM. LO drafted the search strategy. YT and KRS provided methodological input, and YT provided statistical expertise. All authors reviewed the final protocol.

DECLARATIONS OF INTEREST

BV has no declared interests.

TV has no declared interests.

YT has no declared interests.

AWK has conducted prior primary research on tuberculosis diagnostics. AWK reports in-kind donations of Xpert MTB/RIF (or Ultra cartridges) for an unrelated project from Cepheid.

SvW has no declared interests.

LO has no declared interests.

BJM has no declared interests.

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

- Department for International Development (DFID), UK

Project number: 300342-104

- World Health Organization, Switzerland

Development and publication of this manuscript was in part made possible with financial support from the World Health Organization