Cochrane Library

Cochrane Database of Systematic Reviews

Xpert MTB/RIF and Xpert MTB/RIF Ultra assays for active tuberculosis and rifampicin resistance in children (Protocol)

Kay AW, González Fernández L, Takwoingi Y, Eisenhut M, Vu RD, Steingart KR, Detjen AK, Mandalakas AM

WILEY

T A B L E   O F   C O N T E N T S

[Diagnostic Test Accuracy Protocol]

Xpert MTB/RIF and Xpert MTB/RIF Ultra assays for active tuberculosis and rifampicin resistance in children

Alexander W Kay[1], Lucia González Fernández[2], Yemisi Takwoingi[3], Michael Eisenhut[4], Ryan D Vu[5], Karen R Steingart[6], Anne K Detjen[7], Anna M Mandalakas[5]

[1]Pediatrics, Baylor College of Medicine, Mbabane, Swaziland. [2]Department of Paediatrics, Baylor College of Medicine, Houston, Texas, USA. [3]Institute of Applied Health Research, University of Birmingham, Birmingham, UK. [4]Paediatric Department, Luton & Dunstable University Hospital NHS Foundation Trust, Luton, UK. [5]Pediatrics, Baylor College of Medicine, Houston, Texas, USA. [6]Honorary Research Fellow, Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, UK. [7]Health Section, UNICEF, New York, New York, USA

Contact address: Anna M Mandalakas, Pediatrics, Baylor College of Medicine, 1102 Bates Street - FC630, Houston, Texas, 77030, USA. anna.mandalakas@bcm.edu.

ABSTRACT

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

Primary objectives

To determine the diagnostic accuracy of Xpert MTB/RIF and Xpert Ultra for a) PTB in children presumed to have tuberculosis; b) lymph node tuberculosis in children presumed to have tuberculosis; c) tuberculous meningitis in children presumed to have tuberculosis; and d) rifampicin resistance in children presumed to have tuberculosis.

For tuberculosis detection, the role of the index tests would be a replacement for standard practice.

For rifampicin resistance detection, the role of the index tests would be an initial replacement test for culture-based drug susceptibility testing.

BACKGROUND

Tuberculosis is one of the top 10 causes of death and the leading cause from a single infectious agent (above HIV/AIDS), causing an estimated 1.6 million deaths in 2017. Globally during that year, an estimated 10 million people developed tuberculosis disease, including one million children (WHO 2018a). Recent models that have been accepted and supported by the World Health Organization (WHO) suggest that there is substantial underreporting as well as underdiagnosis of child tuberculosis. Even accounting for underdiagnosis, children represent 10% of the annual 10.4 million incident cases. Furthermore, tuberculosis-associated deaths take a disproportionate toll in children: 253,000 deaths were estimated in 2016 in children below 15 years, accounting for 6.9% of the

total deaths notified in that year; of these deaths, 80% occurred in children under five years of age (Dodd 2017; Jenkins 2017). Estimates suggest that the majority of deaths occurring in children occur in undiagnosed cases and represent a missed opportunity to start adequate treatment (Jenkins 2017). Tuberculosis treatment in children follows the same principles as that in adults, with the same drugs used in most cases. The standard four-drug combination regimen of isoniazid, rifampicin, pyrazinamide, and ethambutol given daily for a period of two months followed by isoniazid and rifampicin given daily for additional four to six months is used for the treatment of drug-susceptible tuberculosis, both the pulmonary and extrapulmonary forms with the exception of central nervous system tuberculosis where the treatment with isoniazid and rifampicin is extended to a total of 12 months. The recent introduction of paediatric fixed-dose combinations with optimized dosing and taste masking has improved the efficiency of treatment. Treatment of drug-resistant tuberculosis in children generally has better outcomes than in adults, and in August 2018, the WHO released a rapid communication containing new recommendations for treatment, including the use of all-oral regimens (WHO 2018b; Furin 2019).

The diagnosis of child tuberculosis relies on a mix of clinical, epidemiological, radiological, and laboratory information. Child tuberculosis disease is typically paucibacillary (tuberculosis disease caused by a smaller number of bacteria), and young children cannot voluntarily produce sputum specimens (Marais 2005; Theart 2005). Hence, even under ideal clinical and laboratory conditions, only 30% to 40% of child tuberculosis cases are microbiologically confirmed (Dunn 2016). The probability of a microbiological confirmation is increased in children with more severe or advanced disease (Marais 2006c; Marais 2006d). However, the diagnostic gap also exists because conventional smear microscopy, which is of little value in diagnosing child tuberculosis, remains the most used and widely available tuberculosis diagnostic method in low- and middle-income countries. Tuberculosis culture methods have shown a greater, yet highly variable, sensitivity in child tuberculosis (Frigati 2015; Chiang 2017); unfortunately, tuberculosis culture is not widely available in high-burden settings to support diagnosis.

Xpert MTB/RIF represents a promising diagnostic modality for child tuberculosis. Since 2010 the WHO has recommended the use of Xpert MTB/RIF as the preferred initial microbiological test for people thought to have multidrug-resistant tuberculosis or HIV-associated tuberculosis (strong recommendation), and has included children with presumed tuberculosis on the strength of evidence reported in adults (WHO 2011). In 2013 this guidance was updated with a recommendation specific to children, that is that Xpert MTB/RIF should be used as the preferred initial diagnostic test in children thought to have multidrug-resistant tuberculosis or HIV-associated tuberculosis (strong recommendation, very low-quality) and as the initial diagnostic test in all children with presumptive tuberculosis (conditional recommendation acknowledging resource implications, very low-quality evidence) (WHO 2013). At present, the WHO supports the use of Xpert MTB/RIF for diagnosis of child tuberculosis in four scenarios:

- the initial diagnostic test of choice, rather than conventional smear microscopy or culture (conditional recommendation acknowledging resource implications, very low-quality evidence (also called certainty of the evidence);

- diagnosis in children suspected of having drug-resistant tuberculosis or HIV-associated tuberculosis (strong recommendation, very low-quality evidence);

- as a replacement test for specific non-respiratory specimens (lymph nodes and other tissues) in children presumed to have extrapulmonary tuberculosis (conditional recommendation, very low-quality evidence); and

- as the preferred initial diagnostic on cerebrospinal fluid (CSF) in children suspected of having tuberculous meningitis (strong recommendation given the urgency of rapid diagnosis, very low-quality evidence) (WHO 2014a).

The WHO does not currently recommend Xpert MTB/RIF for use on other specimen types, and the existing guidelines acknowledge that all current recommendations regarding use of Xpert MTB/RIF in children rely on "very low-certainty evidence" and are currently evolving with the expansion of the use of Xpert MTB/RIF Ultra (Xpert Ultra) (WHO 2017).

A non-inferiority analysis of Xpert Ultra compared to Xpert MTB/RIF found that Xpert Ultra has higher sensitivity than Xpert MTB/RIF, particularly in smear-negative, culture-positive specimens and in specimens from HIV-positive patients. Xpert Ultra was also found to have accuracy that was at least as good as Xpert MTB/RIF for rifampicin resistance detection. However, it was noted that Xpert Ultra may have reduced specificity in high tuberculosis burden settings. The current WHO recommendations for the use of Xpert MTB/RIF now also apply to the use of Xpert Ultra as the initial diagnostic test for all adults and children with signs and symptoms of tuberculosis and in the testing of selected extrapulmonary specimens (cerebrospinal fluid, lymph nodes and tissue specimens). However, a negative test result does not exclude tuberculosis in children (WHO 2017). This systematic review will estimate and compare the diagnostic accuracy of Xpert MTB/RIF and Xpert Ultra in children presumed to have pulmonary tuberculosis (PTB) or specific forms of extrapulmonary tuberculosis.

Target condition being diagnosed

Tuberculosis is an infectious disease caused by bacteria within the *Mycobacterium tuberculosis* complex, most commonly *Mycobacterium tuberculosis* (*M tuberculosis*). Typically disseminated

through the air, *M tuberculosis* predominantly affects the lungs, causing PTB, and less typically can cause disease in other organs of the body in extrapulmonary tuberculosis forms. Lymph node tuberculosis is the most common extrapulmonary tuberculosis presentation in children, while tuberculous meningitis results in the highest morbidity and mortality. In this review we will limit extrapulmonary tuberculosis to lymph node tuberculosis and tuberculous meningitis because other forms of extrapulmonary tuberculosis in children are less common, and literature supporting Xpert (Xpert MTB/RIF and Xpert Ultra) as a diagnostic tool in other forms of child extrapulmonary tuberculosis is sparse.

The natural history of tuberculosis in children is distinct from adults due to the more frequent progression to primary tuberculosis disease (Marais 2004). Children under five years of age are at particularly high risk of progression to tuberculosis disease following infection, but the risk for older children and adolescents is also higher than that in adults. Overall, it is estimated that 90% of tuberculosis disease in young children occurs within one year of infection (Marais 2014). Despite age being a key predictor of disease progression, other factors such as nutritional status, immune-compromising conditions (e.g. HIV infection), BCG (bacillus Calmette-Guérin) vaccination status, and genetic susceptibility also contribute to the risk of disease progression. Immediately following infection with *M tuberculosis* in a child, haematogenous spread (by way of the bloodstream) can occur. The highest risk period for presentation with tuberculous meningitis and miliary tuberculosis is one to three months following primary infection. Children between six months and two years of age are at particularly high risk of these severe forms of tuberculosis disease. Approximately 50% of children in this age range progress to tuberculosis disease following infection, and 20% to 40% of those children will present with disseminated disease (Marais 2004; Marais 2014). Children under five years of age most commonly present with hilar lymph node and bronchial forms of intrathoracic tuberculosis disease. Older children and adolescents more commonly manifest adult-type disease, including pleural tuberculosis and upper lobe consolidations (Marais 2004).

Laboratory confirmation of childhood tuberculosis disease is challenging for two reasons. The first reason is that child tuberculosis most commonly represents a primary disease process, without the formation of cavities (Marais 2006a). The number of acid-fast bacilli in forms of primary tuberculosis such as hilar lymph node or bronchial tuberculosis is generally substantially lower than is found in a pulmonary cavity. Consequently, child tuberculosis is often referred to as being paucibacillary, and it is more difficult to obtain the organisms needed to confirm disease via conventional smear or culture (Dunn 2016). The second reason for the difficulty in confirming childhood tuberculosis disease is that most children younger than six years of age lack the ability to expectorate sputum and are unable to voluntarily produce good-quality specimens. Respiratory specimens are therefore often obtained through sputum induction. Children swallow respiratory secretions, and early-morning gastric aspiration is another well-established approach to specimen collection. The yield of three consecutive morning aspirates is similar to the collection of one induced sputum (Zar 2005). Nasopharyngeal aspiration for respiratory specimens is a less invasive and promising mode of specimen collection (Zar 2012). Stool has also been studied as a child tuberculosis diagnostic specimen; although sensitivity has been lower than with traditional specimens, this specimen has great appeal because collection is non-invasive and requires no training (Nicol 2013). Because laboratory diagnostics for tuberculosis perform poorly in children, algorithms involving signs, symptoms, tuberculosis exposure, laboratory tests, and radiographic findings are commonly used to make a clinical diagnosis of child tuberculosis. However, these algorithms have been shown to perform differently across settings, and their sensitivity and specificity may be site-specific (David 2017).

Index test(s)

The index tests in this review are Xpert MTB/RIF and Xpert Ultra (Cepheid Inc, CA, USA). Xpert MTB/RIF and Xpert Ultra are nucleic acid amplification tests (NAATs) and function as an automated, closed system that performs real-time polymerase chain reaction (PCR). Specimens are processed using Xpert Sample Reagent and incubated for 15 minutes, after which the processed samples are pipetted into the cartridge. The tests can be run by operators (such as laboratory technicians and nurses) with minimal technical expertise. Within two hours, the tests detect both live and dead Mycobacterium tuberculosis complex DNA and simultaneously recognize mutations in the *Mycobacterium tuberculosis* gene encoding the beta subunit of RNA polymerase (*rpoB*) gene, which is the most common site of *M tuberculosis* mutations leading to rifampin resistance. Xpert MTB/RIF and Xpert Ultra require an uninterrupted and stable electrical power supply, temperature control, and yearly calibration of the cartridge modules (WHO 2014b). The WHO has published extensive guidance and practical information on implementing the test (WHO 2014b). There have been five generations of the cartridge: G1, G2, G3, G4, and Xpert Ultra. G1 to G4 cartridges initially improved the detection of tuberculosis and rifampicin resistance. However, in children, Xpert MTB/RIF detects only 11% of all clinical and confirmed cases (Detjen 2015). Xpert Ultra was developed in part to overcome this limitation. There are limited data on the different sensitivity that Xpert Ultra offers as compared to the G4 cartridge; however, existing data suggest it may offer improved sensitivity for tuberculosis detection in hard-to-diagnose populations such as children, HIV-associated tuberculosis, and extrapulmonary tuberculosis (WHO 2017; Dorman 2018). To improve detection of *M tuberculosis*, Xpert Ultra incorporates two different multi-copy amplification targets (IS6110 and IS1081). These revisions resulted in an approximately 1-log improvement in the lower limit of detection compared with Xpert MTB/RIF, including improved

differentiation of certain silent mutations, improved detection of rifampicin resistance in mixed infections, and avoidance of false-positive results for detection of rifampicin resistance in paucibacillary specimens (Chakravorty 2017). As mentioned above, Xpert Ultra also has decreased specificity compared to G4 and may be more likely to identify *M tuberculosis* DNA from prior episodes of tuberculosis disease, particularly in patients classified in the new 'trace' category (Dorman 2018). Trace call corresponds to the lowest bacillary burden for *M tuberculosis* detection, described below (WHO 2017). This Cochrane Review will include studies that used any of the Xpert generations in the diagnosis of tuberculosis (pulmonary tuberculosis, tuberculous meningitis, and lymph node tuberculosis) in children younger than 15 years.

Clinical pathway

Figure 1 is an example of the clinical pathway and the placement of the index tests. A careful clinical history of tuberculosis exposure and symptoms i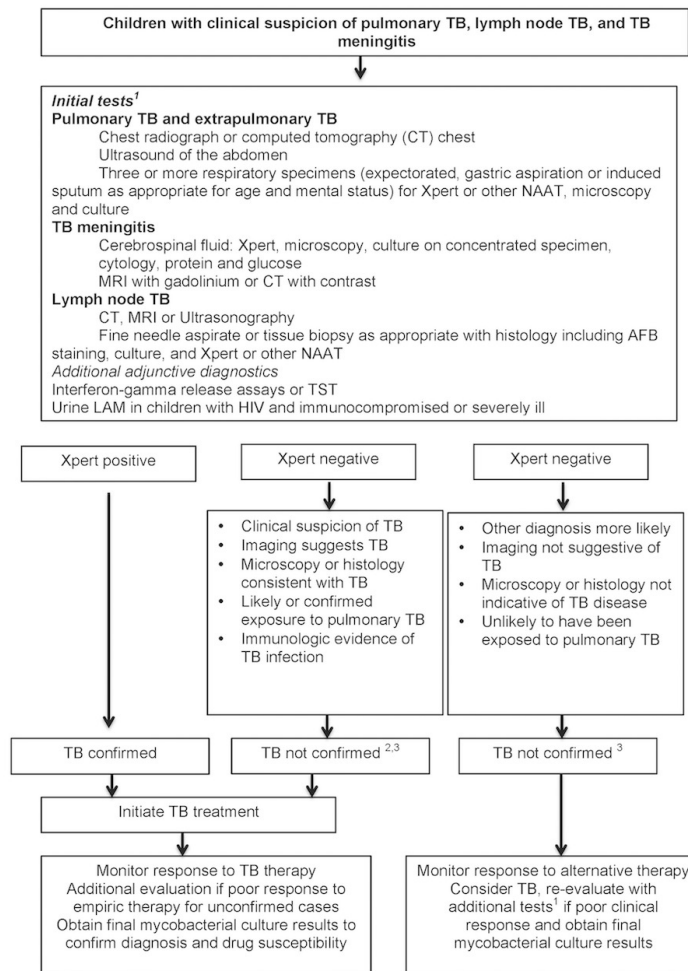s the first step in the diagnostic pathway for childhood tuberculosis. Children with household or other close and persistent exposure to a person with tuberculosis are at increased risk of tuberculosis infection and resultant progression to tuberculosis disease. All children with recent exposure to tuberculosis must be evaluated for clinical symptoms and examination findings consistent with tuberculosis disease. Additional testing depends on the context, but may include chest radiograph and a test of tuberculosis infection. Symptoms of tuberculosis disease are generally persistent for greater than two weeks and are unremitting (Marais 2005). The most common symptoms are cough, fever, decreased appetite, weight loss or failure to thrive, and fatigue or reduced playfulness. Symptoms of extrapulmonary tuberculosis are typically localized, and diagnostic findings are generally obtained from the site of disease (Figure 1). However, no symptom-based diagnostic algorithms have been validated or have been shown to be reliable in multiple contexts. Symptom-based diagnostic algorithms tend to perform poorly in children under three years of age and HIV-positive children, two populations at high risk for disease progression (Marais 2006b).

Figure 1. Abbreviations: AFB: acid-fast bacilli; CT: computed tomography; LAM: mycobacterial lipoarabinomannan antigen; MRI: magnetic resonance imaging; NAAT: nucleic acid amplification test; TB: tuberculosis; TST: tuberculin skin test. The Clinical Pathway. Clinical suspicion of tuberculosis includes persistent cough, fever, weight loss or failure to thrive, lymphadenitis, irritability, lethargy, headache, vomiting or neurological symptoms, history of possible or confirmed exposure to *M tuberculosis*, increased risk for tuberculosis disease due to immunocompromising conditions.

[1] Availability of investigations and tests may be different in high- and low-resource settings and may influence the approach to the diagnosis of child tuberculosis.

[2] Non-microbiological confirmation of *M tuberculosis* does not exclude tuberculosis disease in children, therefore initiation of treatment should be considered empirically if other clinical indications are present.

[3] Mycobacterial culture results are rarely timely to aid the decision to initiate treatment but can confirm or refute clinical decision-making if positive.

Unfortunately, there are no examination features specific to PTB in children. However, the examination findings in extrapulmonary tuberculosis can be quite specific when identified. Clinicians should consider medical comorbidities that increase the risk for tuberculosis disease and modify diagnostic algorithms accordingly. HIV infection not only significantly increases risk of tuberculosis in the paediatric population, but also raises the risk of increased disease severity. HIV-positive children often present with advanced tuberculosis such as disseminated extrapulmonary tuberculosis forms and have high levels of immunosuppression, further complicating diagnosis and management.

Additional diagnostic imaging studies can assist in the diagnosis of nearly all forms of PTB and extrapulmonary tuberculosis. Tests of tuberculosis infection, such as interferon gamma release assays or tuberculin skin tests, can also aid in establishing the diagnosis of tuberculosis in a child but are not necessary to make the diagnosis. Diagnostic recommendations strongly suggest collecting appropriate specimens from the suspected sites of involvement in both PTB and extrapulmonary tuberculosis for microbiological examination. The preferred sample in PTB is sputum, however in young children that cannot expectorate, the sample is commonly obtained via a gastric aspirate or induced sputum. To diagnose extrapulmonary tuberculosis, the collection of samples targets the affected site of disease.

The purpose of Xpert is diagnosis of active tuberculosis (PTB and extrapulmonary tuberculosis) and detection of rifampicin resistance. The results of Xpert can be used as a decision-making tool in the following ways:

- *M tuberculosis* detected/rifampicin resistance not detected: child would start treatment for drug-sensitive tuberculosis;

- *M tuberculosis* detected/rifampicin resistance detected: child would need further resistance testing and would start treatment for drug-resistant tuberculosis according to the country guidelines;

- *M tuberculosis* not detected: a negative Xpert result does not rule out tuberculosis disease, therefore clinicians should still consider initiation of tuberculosis treatment in children with history and clinical features suggestive of tuberculosis disease despite a negative Xpert result. A negative Xpert result may also represent a true negative.

Possible consequences of a false-positive and a false-negative result may include the following:

- false positives (FP): children and their families would likely experience anxiety and morbidity caused by additional testing, unnecessary treatment, and possible adverse effects; possible stigma associated with a tuberculosis or drug-resistant tuberculosis diagnosis and the chance that a false positive may halt further diagnostic evaluation;

- false negatives (FN): would imply an increased risk of

morbidity and mortality and delayed treatment initiation for patients.

Alternative test(s)

Alternative approaches to Xpert for diagnosis of tuberculosis are still used extensively globally. Main tests include the examination of smears for acid-fast bacilli (tuberculosis bacteria) under a microscope (light microscopy, using the classical Ziehl-Neelsen staining technique), fluorescence microscopy, or light-emitting diode (LED)-based fluorescence microscopy. The sensitivity of smear microscopy ranges from 0% to 10% in children (Kunkel 2016). Examination of histology specimens under a microscope following a tissue biopsy targets finding acid-fast bacilli and granulomatous inflammation, frequently with caseous necrosis (necrotizing granulomas), however these options are seldom pursued to diagnose child tuberculosis in low-resource settings due to the invasive nature of the procedures and technical expertise required. Lipoarabinomannan (LAM) antigen is a lipopolysaccharide present in mycobacterial cell wall that can be detected in the urine of people with tuberculosis disease (Shah 2016). This urine test has potential advantages over sputum-based testing due to ease of sample collection. However, due to poor performance (low sensitivity and low specificity), the WHO does not recommend its use in HIV-negative individuals (Nicol 2014). LAM testing is currently recommended in HIV-positive children presenting with tuberculosis symptoms and a CD4 count of less than 100 cells/mm^3 or with severe illness (WHO 2015a).

The quest for novel and more efficient technologies for tuberculosis diagnosis is a cornerstone of the current efforts to reduce the burden of disease worldwide. Over the past decade, there has been unprecedented activity focused on the development of new tools for extrapulmonary tuberculosis diagnosis, largely supported by the engagement of global agencies. As a result there is a strong pipeline of new tools for tuberculosis diagnosis that will complement the use of existing ones and offer improved options (Boyle 2017).

Rationale

Timely and reliable diagnosis of tuberculosis in children remains challenging due to difficulties in collecting sputum samples and the paucibacillary nature of the disease. As a result, undiagnosed cases of disease increase morbidity, mortality, and disease transmission in this key group.

In 2013, informed by a non-Cochrane review (Detjen 2015), the WHO recommended the use of Xpert MTB/RIF in children as front-line test for diagnosis. Detjen 2015 found that evaluated against a reference standard of culture, Xpert MTB/RIF had a sensitivity of 62% (95% credible interval 51% to 73%) and a speci-

ficity of 98% (95% credible interval 97% to 99%). In preparation for a WHO meeting to review recommendations on the use of Xpert, we will perform a Cochrane Review to update the literature, assess the accuracy of both Xpert MTB/RIF and Xpert Ultra, and address previously noted limitations in the prior review, including the following:

- low number of included studies;
- referral bias, with a predominance of studies reporting on hospitalized children;
- heterogeneity in the definition of clinical tuberculosis among the included studies.

We are aware of one other systematic review on the diagnostic accuracy of Xpert MTB/RIF in children (Wang 2015). This review found that Xpert MTB/RIF had a pooled sensitivity of 65% (95% confidence interval 61% to 69%) and pooled specificity of 99% (95% confidence interval 98% to 99%) against a culture reference standard, concluding that Xpert MTB/RIF is an appropriate initial diagnostic. This review had similar limitations to those of Detjen 2015 outlined above.

This review will address these limitations by including additional studies published since 2015 and will include new data on the performance of Xpert Ultra.

OBJECTIVES

Primary objectives

To determine the diagnostic accuracy of Xpert MTB/RIF and Xpert Ultra for a) PTB in children presumed to have tuberculosis; b) lymph node tuberculosis in children presumed to have tuberculosis; c) tuberculous meningitis in children presumed to have tuberculosis; and d) rifampicin resistance in children presumed to have tuberculosis.

For tuberculosis detection, the role of the index tests would be a replacement for standard practice.

For rifampicin resistance detection, the role of the index tests would be an initial replacement test for culture-based drug susceptibility testing.

Secondary objectives

- To compare the accuracy of Xpert MTB/RIF and Xpert Ultra for each of the four target conditions;
- To investigate potential sources of heterogeneity in accuracy estimates. For tuberculosis detection, covariates include age, disease severity, smear-test status, HIV status, clinical setting, specimen type, high tuberculosis burden, and high tuberculosis/ HIV burden. For rifampicin resistance detection, the covariate of interest is multidrug-resistant tuberculosis burden.

METHODS

Criteria for considering studies for this review

Types of studies

We will include cross-sectional studies, cohort studies, and randomized controlled trials from all settings. We will include randomized controlled trials because we may identify studies that evaluate the use of the test on patient health outcomes, but that also report sensitivity and specificity. Although the study design is a randomized trial for the purpose of determining the impact of the test versus a comparator (e.g. usual practice or another test) on health outcomes, the study design is a cross-sectional design for the purpose of determining diagnostic accuracy for the index test in the Cochrane Review protocol. We must be able to extract or derive data on the index test being a true positive, false positive, true negative, or false negative as measured against the reference standards specified below. We will make note of unpublished studies in the 'Ongoing studies' section of the review. We will exclude case-control studies and case reports. We will use abstracts to identify published studies and include those that meet the inclusion criteria.

Participants

We will include studies assessing the index tests for PTB or extrapulmonary tuberculosis in HIV-positive and HIV-negative children aged 0 to 14 years presumed to have tuberculosis. Studies will be eligible for inclusion if they describe the use of Xpert (Xpert MTB/RIF and Xpert Ultra) on routine respiratory specimens such as expectorated or induced sputum, gastric lavage, and nasopharyngeal aspirates. We will include bronchoalveolar lavage. We will also include studies evaluating stool because tuberculosis bacilli are present in swallowed sputum and recoverable from stool samples using Xpert. We will also include studies that assessed several types of specimens.

Index tests

The index tests will be Xpert MTB/RIF and Xpert Ultra.
Index test results are automatically generated, and the user is provided with a printable test result as follows.

- 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 incorporates a semi-quantitative classification for results: trace, very low, low, moderate, and high. 'Trace' corresponds to the lowest bacterial burden for detection of *M tuberculosis* (Chakravorty 2017). Although no rifampicin resistance result will be available for patients with trace results, a trace-positive result is sufficient to initiate anti-tuberculosis therapy in children or HIV-positive patients, according to the WHO report (WHO 2017). Hence, we will consider a trace result to mean *M tuberculosis* DETECTED.

Target conditions

The target conditions are active PTB; two forms of extrapulmonary tuberculosis, lymph node tuberculosis and tuberculous meningitis; and rifampicin resistance.

Reference standards

For PTB, there are two reference standards: (1) a bacteriological reference standard, and (2) a clinical reference standard.

A patient with bacteriologically confirmed PTB is one from whom a biological specimen (such as sputum or gastric aspirate) is positive by smear microscopy or culture. A child with at least one positive smear or culture will be assigned to the 'confirmed tuberculosis' group.

With respect to culture, we will consider studies that use one or more solid media or commercial liquid cultures or both solid and liquid culture to confirm tuberculosis.

A patient with clinical PTB is one from whom a biological specimen is negative by both smear microscopy and culture, and who has been diagnosed with active tuberculosis by a healthcare provider who, following other clinical criteria, has decided to give the patient a full course of tuberculosis treatment. This definition includes patients diagnosed with tuberculosis on the basis of X-ray abnormalities without bacteriological confirmation (WHO 2014a).

For tuberculous meningitis, there two reference standards: (1) bacteriologically confirmed, and (2) clinical diagnosis. A child with bacteriologically confirmed tuberculous meningitis is defined as one with a positive culture or positive acid-fast bacillus smear from a sample of cerebrospinal fluid. Clinical diagnosis may be made based on compatible signs and symptoms, cerebrospinal fluid analysis, cerebral imaging, tuberculosis exposure history, and response to therapy (see Figure 1), leading to initiation of tuberculosis treatment despite not having bacteriological confirmation.

For lymph node tuberculosis, there are two reference standards: (1) bacteriologically confirmed, and (2) clinical diagnosis. A child

with bacteriologically confirmed lymph node tuberculosis is defined as one with a positive culture or positive acid-fast bacillus smear from a sample of a node collected from fine needle biopsy or tissue biopsy or histology consistent with tuberculosis, such as necrotizing granulomas. Clinical diagnosis may be made based on compatible signs and symptoms, tuberculosis exposure history, and response to therapy, leading to initiation of tuberculosis treatment despite not having bacteriological confirmation.

A child will be assigned to the 'not tuberculosis' group if culture and smear (if obtained) are both negative and during the evaluation for tuberculosis an alternative diagnosis was established, their symptoms resolved without tuberculosis treatment, or they did not progress after at least one month. We will not limit the diagnosis of clinical tuberculosis to existing consensus definitions, which would likely be too restrictive (Graham 2012; Graham 2015).

The primary unit of analysis will be the patient, meaning one Xpert test and one reference standard result for a patient. However, we anticipate that some studies may include multiple specimens for a patient. We will perform analyses separately for each unit of analysis and discuss the limitation of using a unit of analysis other than the patient.

The reference standard may be applied to the same child singly or multiple times with a rule, such as at least one positive result, used to determine index test positivity. Ultimately a child will have only one reference standard result against which the result of the index test is cross classified. We will analyse these two ways of applying the reference standard separately (see Statistical analysis and data synthesis) and reflect these differences in the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2), Reference Standard domain.

The reference standard for rifampicin resistance will be phenotypic drug susceptibility testing.

Search methods for identification of studies

We will attempt to identify all relevant studies regardless of language or publication status (published, unpublished, in press, and in progress).

Electronic searches

We will search the following databases using the search terms and strategy described in Appendix 1: Cochrane Infectious Diseases Group Specialized Register; Cochrane Central Register of Controlled Trials (CENTRAL), published in the Cochrane Library; MEDLINE (Ovid); Embase (Ovid); CINAHL (EBSCOHost) (Cumulative Index to Nursing and Allied Health Literature); Science Citation Index-Expanded (Web of Science); and Scopus. We will search the World Health Organization International Clinical Trials Registry Platform (WHO ICTRP) (www.who.int/ictrp/en/) and US National Institutes of Health Ongoing Trials Register

ClinicalTrials.gov (clinicaltrials.gov/) for trials in progress using 'tuberculosis' and 'symptom screening' as search terms.

Searching other resources

We will contact researchers and experts in the field to identify any additional eligible studies. We will also check the references of relevant reviews and studies to identify additional studies.

Data collection and analysis

Selection of studies

Two review authors will independently screen all titles and abstracts in order to identify potentially eligible studies. We will then obtain the full-text articles of potentially eligible studies, and two review authors will independently assess whether they should be included based on predefined inclusion and exclusion criteria. Any disagreements will be resolved by discussion or by consulting a third review author if necessary. We will contact study authors for clarification of methods and other information as needed. We will record and summarize reasons for exclusion of excluded studies in a 'Characteristics of excluded studies' table. We will illustrate the study selection process in a PRISMA diagram.

Data extraction and management

We will design a data extraction form and pilot it on at least two included studies (Appendix 2), and will then finalize the form based on the pilot test. As above, two review authors will independently extract data using this data extraction form, and will discuss inconsistencies to achieve consensus. They will consult a third review author to resolve discrepancies as needed. We will enter abstracted data into an Excel database (Excel 2013) on password-protected computers. We will secure the dataset in a cloud storage workspace (Dropbox), and we will store extracted data for future review updates. A representative list of data extraction fields is listed below.

Study details

- Number of participants after screening for exclusion and inclusion criteria;
- total number of children included in the analysis;
- total number of specimens included with collection methods;
- unit of sample collection: one specimen, multiple specimens, unknown, or unclear;
- did the study include participants with a prior history of tuberculosis;

- if so, % (numerator/denominator) of patients with prior tuberculosis;
- target condition(s)? pulmonary tuberculosis, lymph node tuberculosis, tuberculous meningitis, rifampicin resistance;
- if so, % (numerator/denominator) of patients with prior tuberculosis;
- target condition(s)? pulmonary tuberculosis, lymph node tuberculosis, tuberculous meningitis, rifampicin resistance;
- target condition(s)? pulmonary tuberculosis, lymph node tuberculosis, tuberculous meningitis, rifampicin resistance.

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 of participants;
- percentage and number of HIV-positive or HIV-negative participants, if both were included in the study;
- type of respiratory specimen included: expectorated, induced, nasopharyngeal aspirate, gastric lavage, stool;
- type of non-respiratory specimen included: fine needle aspirate, lymph node biopsy, cerebrospinal fluid, multiple types, other, unknown;
- information obtained from the same specimen;
- number of cultures used to exclude tuberculosis;
- information on smear microscopy: was it used, type, quantification (trace to 4+);
- data on culture performance: number of contaminated cultures with respect to total cultures performed;
- time from specimen collection to diagnosis;
- time from diagnosis to treatment initiation;
- clinical setting: outpatient or inpatient or both;
- description of radiographic findings;
- information on tuberculosis burden in the country.

We will classify 'country' as being high burden or not high burden for tuberculosis, tuberculosis/HIV, or multidrug-resistant tuberculosis according to the WHO post-2015 era classification (WHO 2018a). A country may be classified as high burden for one, two, or all three of the high-burden categories.

Index test

- Xpert version: MTB/RIF or Ultra;
- Xpert platform;
- pretreatment processing procedure for Xpert MTB/RIF or Xpert Ultra;
- specimen treatment: fresh, frozen, mixed;
- number of true positives, false positives, false negatives, and true negatives (see example table in Appendix 3);

- uninterpretable results for tuberculosis detection (invalid, error, or no result);
- indeterminate results for rifampicin resistance detection.

Reference standard

- Reference standard for PTB: details of solid or liquid culture or clinical criteria;
- reference standard for *M tuberculosis*: details of tissue or histopathology or culture of tissue or fluid, and clinical criteria;
- reference standard for rifampicin resistance: details of phenotypic drug susceptibility testing.

Assessment of methodological quality

We will assess the methodological quality of included studies using the QUADAS-2 instrument, which we will adapt for this review (Whiting 2011). The QUADAS-2 tool consists of four domains: (1) patient selection; (2) index test(s); (3) reference standard(s); and (4) flow and timing. We will assess all domains for risk of bias, and the first three domains for concerns regarding applicability. We will first develop guidance on how to appraise each question and interpret this information. One review author will then pilot the tool with two of the included studies. We will finalize the tool based on experience gained from the pilot. Two review authors will independently complete QUADAS-2. Any disagreements will be resolved through discussion or by arbitration with a third review author if necessary. We will present the results of the quality assessment in the text, table, and graphs. The preliminary tool with signalling questions tailored to this review is in Appendix 4.

Statistical analysis and data synthesis

When possible, for each target condition, we will consider one result per index test per child; this ideally corresponds to the first specimen provided. A positive index test is detection of *M tuberculosis* by Xpert.

For tuberculosis detection and rifampicin resistance detection, we will group all analyses by the specific index test, Xpert MTB/RIF or Xpert Ultra. For both PTB and extrapulmonary tuberculosis detection, we will stratify analyses by type of reference standard used, that is bacteriological or clinical. We will also analyse studies that report a single result for the reference standard separately from those that use multiple results for this same reference standard because we suspect that the latter is likely to correctly identify more tuberculosis patients.

For detection of rifampicin resistance, we will include children who:
- were culture-positive;
- had a valid phenotypic drug susceptibility test (DST) result;
- were Xpert tuberculosis-positive; and
- had a valid Xpert rifampicin result.

Sensitivity = Xpert rifampicin resistant/phenotypic DST rifampicin resistant

Specificity = Xpert rifampicin susceptible/phenotypic DST rifampicin susceptible

We will perform descriptive analyses of the included studies and present their key characteristics in 'Characteristics of included studies' tables. 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 2014).

We expect that some studies will consider index test results from more than one specimen. We will use the following two analytic approaches to consider multiple specimen results:
- consider only the first specimen result;
- consider all results and define positive as 'at least one positive result'.

If the sample size supports meaningful analysis, secondary analysis will look at performance of the index test stratified by number of samples tested.

When there are sufficient data, we will perform meta-analyses to estimate an average summary value of sensitivities and specificities using a bivariate model (Chu 2006; Reitsma 2015). We will use the bivariate model because the index tests, Xpert MTB/RIF and Xpert Ultra, each apply a common positivity criterion (Macaskill 2010). If we are unable to fit the bivariate 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). We will perform meta-analyses using the meqrlogit command in Stata version 15 (Stata 15).

If there are sufficient data, we will perform comparative meta-analyses by first including all studies with relevant data, employing indirect comparisons to make use of all available data. When comparative studies that made direct comparisons between Xpert MTB/RIF and Xpert Ultra within the same participants are available, we will perform additional analyses by restricting the analyses to only these comparative studies. We will perform comparative meta-analyses using meta-regression by including test type as a covariate in a bivariate model. 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 bivariate model parameters. We will obtain 95% confidence intervals and P values for the absolute differences using the delta method and Wald tests, respectively.

Approach to uninterpretable index test results

Xpert MTB/RIF and Xpert Ultra report an uninterpretable test result for unexpected results with any of the internal control measures of the assay. The uninterpretable rate for detection of tuberculosis is the number of tests classified as "invalid", "error", or "no result" divided by the total number of Xpert tests performed. The indeterminate rate for detection of rifampicin resistance is the

number of tests classified as "MTB DETECTED; Rif resistance INDETERMINATE" divided by the total number of Xpert-positive results. If we find a low proportion of uninterpretable results for tuberculosis detection or a low proportion of indeterminate results for rifampicin resistance detection, we will exclude these results from the analysis and report the pooled proportion of uninterpretable Xpert MTB/RIF results and Xpert Ultra results for tuberculosis detection and indeterminate Xpert MTB/RIF results and Xpert Ultra results for rifampicin resistance detection. Otherwise, we plan to group these results with either the positive or negative results and perform additional analyses.

Investigations of heterogeneity

We will visually inspect forest and summary receiver operating characteristic (ROC) plots for heterogeneity. If the data allow, we will evaluate sources of heterogeneity using bivariate meta-regression with each source of heterogeneity as a single covariate in a bivariate model. We plan to assess the following as categorical covariates.

Tuberculosis detection

- Smear status (positive or negative);
- HIV status (percentage positive in the study);
- clinical setting (inpatient or outpatient);
- specimen type (collection method);
- age (age group);
- high tuberculosis burden, yes or no;
- high tuberculosis/HIV burden, yes or no.

Rifampicin resistance detection

- High multidrug-resistant tuberculosis burden, yes or no.

Sensitivity analyses

If there are sufficient data, we will perform sensitivity analyses to explore the effect of risk of bias and study characteristics on the accuracy of Xpert MTB/RIF and Xpert Ultra. Specifically, we will limit inclusion in the meta-analyses to the following:

- studies that used consecutive or random selection of participants;
- studies where the reference standard results were interpreted without knowledge of the index test results;
- studies that included only untreated patients;
- studies that only enrolled children aged 0 to 14 years old.

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 (Balshem 2011; Schünemann 2008; Schünemann 2016). 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 will start 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 (AWK, LGF, and KRS) will discuss judgments and apply GRADE in the following way.

Assessment of risk of bias

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

Indirectness

We will us QUADAS-2 for concerns of applicability and look for important differences between the populations studied (for example, in the spectrum of disease), the setting, index test, and outcomes and ask if differences are sufficient to lower certainty in results?

Inconsistency

GRADE recommends downgrading for unexplained inconsistency in sensitivity and specificity estimates. We will carry out pre-specified analyses to investigate potential sources of heterogeneity and will downgrade when we think we can explain inconsistency in the accuracy estimates. Imprecision: we 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. We will also consider whether the number of participants included in the analysis was less than the number generated by a conventional sample size calculation for a single adequately powered study (optimal information size).

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

REFERENCES

Additional references

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.

Boyle 2017

Boyle D. *Tuberculosis Diagnostics Technology and Market Landscape*. 5th Edition. Vernier: World Health Organization Unitaid Secretariat, 2017.

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.

Chiang 2017

Chiang SS, Swanson DS, Starke JR. New diagnostics for childhood tuberculosis. *Infectious Disease Clinics of North America* 2015;**29**(3):477–502.

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.

David 2017

David SG, Lovero KL, Pombo-March MFB, Abreu TG, Ruffino-Netto A, 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.

Detjen 2015

Detjen AK, DiNardo AR, Leyden J, Steingart KR, Menzies D, Schiller I, et al. Xpert MTB/RIF assay for the diagnosis of pulmonary tuberculosis in children: a systematic review and meta-analysis. *Lancet Respiratory Medicine* 2015;**3**(6):451–61.

Dodd 2017

Dodd PJ, Yuen CM, Sismanidis C, Seddon JA, Jenkins HE. The global burden of tuberculosis mortality in children: a mathematical modelling study. *Lancet Global Health* 2017;**5**(9):e898–906.

Dorman 2018

Dorman SE, Schumacher SG, Alland D, Nabeta P, Armstrong DT, King B, et al. Xpert MTB/RIF Ultra for detection of Mycobacterium tuberculosis and rifampicin resistance: a prospective multicentre diagnostic accuracy study. *Lancet Infectious Diseases* 2018;**18**(1):76–84.

Dropbox [Computer program]

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

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 Corporation. Microsoft Excel Software. Version 16.16.4. Seattle: Microsoft Corporation, 2013.

Frigati 2015

Frigati L, Maskew M, Workman L, Munro J, Andronikou S, Nicol MP, et al. Predictors of culture-confirmed pulmonary tuberculosis in children in a high tuberculosis and HIV prevalence area. *Pediatric Infectious Disease Journal* 2015;**34**(9):e206–10.

Furin 2019

Furin J. Advances in the diagnosis, treatment, and prevention of tuberculosis in children. *Expert Review of Respiratory Medicine* 2019;**13**(3):301–11.

Graham 2012

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

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.

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.

Kunkel 2016

Kunkel A, Abel Zur Wiesch P, Nathavitharana RR, Marx FM, Jenkins HE, Cohen T. Smear positivity in paediatric

and adult tuberculosis: systematic review and meta-analysis. *BMC Infectious Diseases* 2016;**16**(282):1–9.

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

Marais 2005

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 Children* 2005;**90**(11):1162–5.

Marais 2006a

Marais BJ, Gie RP, Hesseling AC, Schaaf HS, Enarson DA, Beyers N. Radiographic signs and symptoms in children treated for tuberculosis: possible implications for symptom-based screening in resource-limited settings. *Pediatric Infectious Disease Journal* 2006;**25**(3):237–40.

Marais 2006b

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 2006c

Marais BJ, Gie RP, Schaaf HS, Hesseling AS, Enarson DA, Beyers N. The spectrum of disease in children treated for tuberculosis in a highly endemic area. *International Journal of Tuberculosis and Lung Disease* 2006;**10**:732-8.

Marais 2006d

Marais BJ, Gie RP, Hesseling AC, Schaaf HS, Enarson DA, Beyers N. The bacteriologic yield in children with intrathoracic tuberculosis. *Clinical Infectious Diseases* 2006; **42**:69–71.

Marais 2014

Marais BJ, Schaaf HS. Tuberculosis in children. *Cold Spring Harbor Perspectives in Medicine* 2014;**4**:1–21.

Nicol 2013

Nicol MP, Spiers K, Workman L, Isaacs W, Munro J, Black F, et al. Xpert MTB/RIF testing of stool samples for the diagnosis of pulmonary tuberculosis in children. *Clinical Infectious Diseases* 2013;**57**(3):e18–21.

Nicol 2014

Nicol MP, Allen V, Workman L, Isaacs W, Munro J, Pienaar S, et al. Urine lipoarabinomannan testing for diagnosis of pulmonary tuberculosis in children: a prospective study. *Lancet Global Health* 2014;**2**(5):e278–84.

Reitsma 2015

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**:982-90.

RevMan 2014 [Computer program]

Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

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):1106–10.

Schünemann 2016

Schünemann HJ, Mustafa R, Brozek J, Santesso N, Alonso-Coello P, Guyatt G, et al. GRADE Working Group. GRADE Guidelines: 16. GRADE evidence to decision frameworks for tests in clinical practice and public health. *Journal of Clinical Epidemiology* 2016;**76**:89–98. DOI: 10.1016/j.jclinepi.2016.01.032

Shah 2016

Shah M, Hanrahan C, Wang ZY, Dendukuri N, Lawn SD, Denkinger CM, et al. Lateral flow urine lipoarabinomannan assay for detecting active tuberculosis in HIV-positive adults. *Cochrane Database of Systematic Reviews* 2016, Issue 5. DOI: 10.1002/14651858.CD011420.pub2

Stata 15 [Computer program]

StataCorp. Stata Statistical Software. Version 15. College Station: StataCorp LP, 2017.

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.

Theart 2005

Theart AC, Marais BJ, Gie RP, Hesseling AC, Beyers N. Criteria used for the diagnosis of childhood tuberculosis at primary health care level in a high-burden, urban setting. *International Journal of Tuberculosis and Lung Disease* 2005; **9**(11):1210–4.

Wang 2015

Wang XW, Pappoe F, Huang Y, Cheng XW, Xu DF, Wang H, et al. Xpert MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in children: a meta-analysis. *Clinical Laboratory* 2015;**61**:1775–85.

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. Automated real-time nucleic acid amplification technology for rapid and simultaneous

detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF system for the diagnosis of pulmonary and extrapulmonary TB in adults and children. apps.who.int/iris/handle/10665/112472 (accessed prior to 29 April 2019).

WHO 2013

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 extrapulmonary TB in adults and children. apps.who.int/iris/bitstream/handle/10665/112472/9789241506335_eng.pdf?sequence=1 (accessed prior to 29 April 2019).

WHO 2014a

World Health Organization. Guidance for national tuberculosis programmes on the management of tuberculosis in children. apps.who.int/medicinedocs/documents/s21535en/s21535en.pdf (accessed prior to 29 April 2019).

WHO 2014b

World Health Organization. Xpert MTB/RIF implementation manual. Technical and operational 'how-to' practical considerations. apps.who.int/iris/bitstream/10665/112469/1/9789241506700_eng.pdf 2014 (accessed prior to 29 April 2019).

WHO 2015a

World Health Organization. The use of lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis and screening of active tuberculosis in people living with HIV: policy guidance. www.who.int/tb/areas-of-work/laboratory/policy_statement_lam_web.pdf (accessed prior to 29 April 2019).

WHO 2015b

World Health Organization. Use of high burden country lists for TB by WHO in the post-2015 era. www.who.int/tb/publications/global_report/high_tb_burdencountrylists2016-2020.pdf?ua=1 (accessed prior to 29 April 2019).

WHO 2017

World Health Organization. WHO Meeting Report of a Technical Expert Consultation: Non-inferiority analysis of Xpert MTB/RIF Ultra compared to Xpert MTB/RIF. apps.who.int/iris/bitstream/10665/254792/1/WHO-HTM-TB-2017.04-eng.pdf (accessed prior to 29 April 2019).

WHO 2018a

World Health Organization. *Global tuberculosis report 2018*. Geneva: World Health Organization, 2018.

WHO 2018b

World Health Organization. Rapid Communication: Key changes to treatment of multidrug- and rifampicin-resistant tuberculosis (MDR/RR-TB). August 2018. www.who.int/tb/publications/2018/rapid_communications_MDR/en/ (accessed 5 March 2019).

Zar 2005

Zar HJ, Hanslo D, Apolles P, Swingler G, Hussey G. Induced sputum versus gastric lavage for microbiological confirmation of pulmonary tuberculosis in infants and young children: a prospective study. *Lancet* 2005;**365**(9454):130–4.

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.

* Indicates the major publication for the study

APPENDICES

Appendix 1. Search strategy

1 Mycobacterium tuberculosis/
2 Tuberculosis/ or "Tuberculosis, Multidrug-Resistant"/ or Extensively Drug-Resistant Tuberculosis/
3 (Tuberculosis or MDR-tuberculosis or XDR-tuberculosis or tuberculous).ti. or (Tuberculosis or MDR-tuberculosis or XDR-tuberculosis or tuberculous).ab.
4 1 or 2 or 3
5 Xpert*.ti. or Xpert*.ab.
6 (GeneXpert or cepheid).ti. or (GeneXpert or cepheid).ab.
7 (Xpert* and Ultra).mp.
8 near* patient or near-patient).ti. or (near* patient or near-patient).ab.
9 (pediatric(or paediatric).mp.
10 (child or children or childhood or infant* or newborn or neonate* or toddler*).mp

11 5 or 6 or 7 or 8
12 9 or 10
13 4 and 11 and 12
This is the preliminary search strategy for MEDLINE (Ovid), which will be adapted for other electronic databases. We will report all search strategies in full in the final review version.

Appendix 2. Data extraction form

Diagnostic accuracy of Xpert in the diagnosis of child tuberculosis: data extraction form	
I. ID	
Study ID	First Name/Publication Year
First Author	Name
Corresponding author	Name
Corresponding author email	Email
Was author contacted?	1 - Yes 2 - No If yes, dates(s)
If yes, author response?	
Study data	1 - Published 2 - In-press 3 - Ongoing
Title	
Year (of publication)	YYYY or 9 - Not reported
Year study start date	YYYY or 9 - Not reported
Language	1 - English 2 - Other If other, specify:
II. Study details	
Country where study was conducted	
Country World Bank classification	1 - Low income 2 - Middle income 3 - High income 4 - Low and High income

	5 - Other combination, describe
Country tuberculosis burden (WHO 2015b)	1 - WHO tuberculosis high burden 2 - WHO tuberculosis/HIV high burden 3 - WHO MDR tuberculosis high burden 4 - WHO tuberculosis + MDR tuberculosis high burden 5 - WHO tuberculosis + HIV/tuberculosis high burden 6 - WHO tuberculosis + HIV/tuberculosis + MDR tuberculosis high burden 7 - Not a WHO high-burden country 8 - Both non-high-burden and high-burden countries included 9 - Other
Study design	1 - Randomized controlled trial 2 - Cross-sectional 3 - Cohort 4 - Other, specify 9 - Could not tell If other, describe:
Participant selection	1 - Consecutive 2 - Random 3 - Convenience 7 - Other 9 - Unknown/Not reported
Direction of study data collection	1 - Prospective 2 - Retrospective 9 - Unknown/Not reported
Inclusion criteria	1 - Broad 2 - Rigorous 9 - Unknown/Not reported
Inclusion criteria for presumptive tuberculosis	1 - tuberculosis contact 2 - Cough 3 - Loss of weight 4 - Suggestive chest x-ray 5 - Immunological evidence of tuberculosis infection (TST/IGRA) 6 - Malnutrition 7 - HIV 8 - Other, describe 9 - Unknown/Not reported
Describe inclusion criteria as in study	

Number included after recruitment by inclusion and exclusion criteria	Enter number or 9 - Unknown/Not reported
Total number of children included in systematic review analysis	Enter number or 9 - Unknown/Not reported
Total number of specimens included in analysis with collection method	Enter number or 9 - Unknown/Not reported
Unit of analysis (Xpert)	1 - One specimen per patient 2 - Multiple specimens per patient 3 - Unknown number of specimens per patient 9 - Unknown/Not reported Describe as written in study, if unclear:
Did the study include patients with previous tuberculosis history?	1 - Yes 0 - No 9 - Unknown/Not reported
If so, what is the percentage?	Enter % and specify numerator/denominator
Target condition? Pulmonary tuberculosis?	1 - Yes 0 - No
Target condition? Rifampicin resistance?	1 - Yes 0 - No
Target condition? Lymph node tuberculosis?	1 - Yes 0 - No
Target condition? Tuberculous meningitis?	1 - Yes 0 - No
Comments about study design	
III. Patient characteristics and setting	
Description of study population (age, HIV info, etc.)	1 - All enrolled 2 - All analysed 9 - Unknown/Not reported
Age: median, mean, range by months	Enter number or 9 - Unknown/Not reported
Gender	##/total and % female
HIV status of participants	0 - HIV− 1 - HIV+ 2 - Both HIV+/− 9 - Unknown/Not reported

(Continued)

If HIV-positive participants included, what is the percentage?	% and specify numerator/denominator
Type of respiratory specimen included	1 - All expectorated 2 - All induced 3 - All Bronchoalveolar lavage 4 - All gastric lavage 5 - Nasopharyngeal aspitate 6 - Stool 7 - Multiple types 8 - Other 9 - Unknown/Not reported If 7 or 8, describe types and record numbers:
Type of non-respiratory specimen	1 - Fine needle aspirate 2 - Lymph node biopsy 3 - Cerebrospinal fluid 4 - Multiple types 5 - Other 9 - Unknown/Not reported If 4 or 5, describe types and record numbers:
Were Xpert sample and culture obtained from same specimen?	1 - Yes 0 - No 9 - Unknown/Not reported
Number of cultures used to exclude tuberculosis	Describe
Information on smear microscopy: was it used	1 - Yes 0 - No 9 - Unknown/Not reported
Type of microscopy used	1 - Ziehl-Neelsen 2 - Fluoresence microscopy 3 - Light Emitting Diode-based fluorescence microscopy 4 - Multiple, describe 9 - Unknown/Not reported
Smear type	1 - Direct 2 - Concentrated (processed) 3 - Both direct and concentrated 9 - Unk/NR
Data on culture performance provided?	# of contaminated culture/Total # cultures performed or 9 - Unknown/Not reported
Were patient important outcomes evaluated? (Time to diagnosis, time to treatment, others)	1 - Yes 2 - No 9 - Unknown/Not reported

(Continued)

Time to diagnosis?	Xpert: Culture: 9 - Unknown/Not reported Specify whether time from sample collection to diagnosis in lab or just turnaround time in lab
Time to treatment initiation	Xpert: Culture: 9 - Unknown/Not reported
Clinical setting, describe as written in the paper	1 - Outpatient 2 - Inpatient 3 - Both out- and inpatient 4 - Other, specify 5 - Laboratory based 9 - Unknown/Not reported Describe as in paper:
Laboratory services level	1 - Central (reference) 2 - Intermediate (regional) 3 - Peripheral (microscopy centre, provincial hospital) 4 - Research laboratory 5 - Other, specify
Where were Xpert tests performed? (Tests generally available at different laboratory levels, though tests may overlap) Peripheral: Acid-fast bacilli (Ziehl-Neelsen, Auramine-rhodamine, Auramine-O staining) and Xpert MTB/RIF Intermediate: Peripheral laboratory tests and culture on solid media and line probe assay (LPA) from smear positive sputum Central: Intermediate laboratory tests and culture on liquid media and DST (1st- and 2nd-line anti-tuberculosis drugs) on solid or in liquid media and LPA on positive cultures and rapid speciation tests	1 - Central (reference) 2 - Intermediate (regional) 3 - Peripheral (microscopy centre, provincial hospital) 4 - Other, specify
Was Xpert run outside of a laboratory?	1 - Yes 0 - No
Current treatment: Were patients on treatment (defined as tuberculosis drugs for greater than 7 days) for the current tuberculosis episode? (note: may impact culture results)	1 - Yes 2 - No 9 - Unknown/Not reported
If so, what is the percentage?	% Specify numerator/denominator
IV. Index test	

(Continued)

Xpert version(s) evaluated	1 - Xpert only 2 - Ultra only 3 - Any combination Xpert and Ultra
Xpert platform: Was Omni used? Unless Omni explicitly described, assume standard platform	1 - Yes, only Omni used for Xpert tests 2 - Yes, both Omni and standard platform used for Xpert tests 3 - No
Pretreatment processing procedure for GeneXpert	1 - None 2 - NALC-NaOH 3 - NaOH (Petroff) 4 - Other 9 - Unknown/Not reported
For Xpert specimen, what was the condition of the specimen when tested?	1 - Fresh 2 - Frozen 9 - Unknown/Not reported
Were uninterpretable (invalid error or no result) results reported for Xpert for tuberculosis detection?	1 - Yes 9 - Unknown/Not reported If yes, describe numbers
Were indeterminate results reported for Xpert for rifampicin resistance?	1 - Yes 9 - Unknown/Not reported If yes, describe numbers
V. Reference standard	
For tuberculosis detection, what reference standard(s) was used? Respiratory Samples?	1 - Solid culture (specify 1a) 2 - Liquid culture (specify 2a) 3 - Both solid & liquid culture (specify 1a & 2a) 9 - Unknown/Not reported 1a - Solid culture LJ 7H10 7H11 Other 2a - Liquid culture MGIT 960 Other (specify):
For tuberculosis detection, what reference standard(s) was used? Lymph node?	1 - Solid culture (specify 1a) 2 - Liquid culture (specify 2a) 3 - Both solid & liquid culture (specify 1a & 2a) 9 - Unknown/Not reported

(Continued)

	1a - Solid culture LJ 7H10 7H11 Other 2a - Liquid culture MGIT 960 Other (specify):
For tuberculosis detection, what reference standard(s) was used? Cerebrospinal fluid?	1 - Solid culture (specify 1a) 2 - Liquid culture (specify 2a) 3 - Both solid & liquid culture (specify 1a & 2a) 9 - Unknown/Not reported 1a - Solid culture LJ 7H10 7H11 Other 2a - Liquid culture MGIT 960 Other (specify):
Reference standard pulmonary tuberculosis: clinical	1 - Yes 0 - No Multiple answers, list
If clinical describe as in paper	
For rifampicin resistance detection, what reference standard(s) was used? Respiratory samples?	1 - Solid culture (specify 1a) 2 - Liquid culture (specify 2a) 3 - Both solid & liquid culture (specify 1a & 2a) 4 - *M tuberculosis* DR*plus* 9 - Unknown/NR 1a - Solid culture LJ 7H10 7H11 Other Specify method, e.g. proportion 2a - Liquid culture MGIT 960 Other (specify):

For rifampicin resistance detection, what reference standard(s) was used? Lymph node?	1 - Solid culture (specify 1a) 2 - Liquid culture (specify 2a) 3 - Both solid & liquid culture (specify 1a & 2a) 4 - *M tuberculosis* DR*plus* 9 - Unknown/Not reported 1a - Solid culture LJ 7H10 7H11 Other Specify method, e.g. proportion 2a - Liquid culture MGIT 960 Other (specify):
For rifampicin resistance detection, what reference standard(s) was used? Cerebrospinal fluid?	1 - Solid culture (specify 1a) 2 - Liquid culture (specify 2a) 3 - Both solid & liquid culture (specify 1a & 2a) 4 - *M tuberculosis* DR*plus* 9 - Unknown/Not reported 1a - Solid culture LJ 7H10 7H11 Other Specify method, e.g. proportion 2a - Liquid culture MGIT 960 Other (specify):
If information is available	
Is information on quality assurance of DST available in the study?	1 - Yes 2 - No 9 - Unknown/Not reported If yes, describe potential sources of bias

Abbreviations: DST: Drug susceptibility testing; HIV: human immunodeficiency virus; IGRA: Interferon-gamma release sssays; LJ: Lowenstein Jensen; MDR-TB: Multidrug-resistant tuberculosis; MGIT: Mycobacterial growth indicator tube; TST: Tuberculin skin test

Appendix 3. Example of 2 x 2 result table

Tuberculosis detection, all studies		Confirmed tuberculosis		
Xpert MTB/RIF results		Yes	No	Total
	Positive			
	Negative			
	Total			

Tuberculosis detection, Xpert Ultra		Confirmed tuberculosis		
Xpert Ultra results		Yes	No	Total
	Positive			
	Negative			
	Total			

Extrapulmonary tuberculosis Xpert results

Tuberculous meningitis detection, all studies		Confirmed tuberculosis		
Xpert results		Yes	No	Total
	Positive			
	Negative			
	Total			

Non-traditional respiratory specimens

Tuberculosis detection, all studies		Confirmed tuberculosis		
Xpert results from nasopharyngeal aspirate		Yes	No	Total
	Positive			
	Negative			
	Total			

(Continued)

Clinical tuberculosis				
Tuberculosis detection, all studies		Clinical tuberculosis		
Xpert results		Yes	No	Total
	Positive			
	Negative			
	Total			

Appendix 4. 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?. We will answer 'yes' if the study enrolled a consecutive or random sample of eligible patients; 'no' if the study selected patients by convenience; and '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?

a. Pulmonary tuberculosis - We will score yes for all studies because we do not think there are any inappropriate exclusions for children presumed to have PTB

b. Extrapulmonary tuberculosis (lymph node tuberculosis and tuberculous meningitis) - We will score 'no' if the study excluded specimens based on physical appearance (such as purulence) or a biochemical analysis (e.g. adenosine deaminase (ADA) or cell analysis). We will score 'unclear' if we cannot tell.

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

We are interested in how Xpert performs in patients who were evaluated as they would be in routine practice. Paediatric studies conducted in tertiary centres tend to include a higher number of children with advanced disease, therefore we will answer 'low concern' if patients were evaluated in local hospitals or primary care centres; 'high concern' if patients were evaluated exclusively as inpatients in tertiary care centres; and 'unclear concern' if the clinical setting was not reported or there was insufficient information to make a decision. We will also answer 'unclear concern' if Xpert testing was done at a reference laboratory and the clinical setting was not reported, because it is difficult to tell if a given reference laboratory provides services mainly to very sick patients (inpatients in tertiary care) or to patients with a broad spectrum of disease, including very sick patients and those with less severe disease (primary, secondary, and tertiary care).

Domain 2: Index test

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

Signalling question 1: Were the index test results interpreted without knowledge of the results of the reference standard?. We will answer this question 'yes' for all studies because Xpert MTB/RIF and Xpert Ultra 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.

Signalling question 2: If a threshold was used, was it prespecified? The threshold is prespecified in all versions of Xpert MTB/RIF and Xpert Ultra. We will answer this question 'yes' for all studies.

Applicability: Are there concerns that the index test, its conduct, or its interpretation differ from the review question? Variations in test technology, execution, or interpretation may affect estimates of the diagnostic accuracy of a test. GeneXpert, the test device platform, simplifies molecular testing by fully integrating and automating the three processes (sample preparation, amplification, and detection)

required for real-time polymerase chain reaction (PCR)-based molecular testing. All steps in the Xpert MTB/RIF and Xpert Ultra assays are completely automated and self contained following sample loading. Minimal training is required by operators such as laboratory technicians and nurses to run the index test.

For pulmonary tuberculosis, we will answer 'low concern' if the index test was performed as recommended by the manufacturer. For sputum specimens, we will answer 'unclear concern' if the ratio of the Xpert sample reagent: specimen volume was not 2:1 for a raw specimen or 3:1 for a centrifuged sediment as recommended by the manufacturer or if we cannot tell (WHO 2014a). Central-level laboratories use more highly trained staff than peripheral- and intermediate-level laboratories or health facilities. However, we do not consider this to be a concern about applicability due to the minimal training required to run the index tests.

With respect to extrapulmonary specimens, the WHO has provided detailed information about the processing steps in 'Xpert MTB/ RIF implementation manual. Technical and operational "how-to" practical considerations. Annex 2. Standard Operating Procedure (SOP) for processing extrapulmonary specimens (CSF, lymph nodes and other tissues) for "Xpert MTB/RIF assay"' (WHO 2014b). For extrapulmonary specimens, we will answer 'low concern' if the test was performed according to WHO standard operating procedures. We will score 'high concern' if the test was performed in a way that deviated from these recommendations. We will score 'unclear concern' if we cannot tell.

Domain 3: Reference standard

We have multiple target conditions, each of which each has a different reference standard(s), therefore we have explained how we will assess each signalling question for each reference standard.

Risk of bias: Could the reference standard, its conduct, or its interpretation have introduced bias?

Signalling question 1.a: Is the bacteriological reference standard likely to correctly classify the target condition?

For pulmonary and extrapulmonary tuberculosis, we anticipate that the vast majority of studies will perform culture. Culture is considered the best reference standard for tuberculosis diagnosis. However, particularly in children with paucibacillary disease, it is not 100% accurate. Sensitivity ranges roughly between 20% and 50%, depending on the severity of disease, specimen collection, age of the child, and other factors. The performance of more than one culture increases the diagnostic yield. We will answer 'yes' for studies using multiple specimens and 'unclear' for studies using only one specimen.

Signalling question 1.b: Is the clinical reference standard likely to correctly classify the target condition? A clinical definition of tuberculosis in bacteriologically negative children aims to identify those that were not detected by culture (or histology in the case of lymph node tuberculosis) for the above-mentioned reasons. The accuracy of this reference standard is also not 100%, and its definition is heterogeneous across studies. Irrespective of how clinical tuberculosis was defined in the publications, we will assign children to a group called 'clinical tuberculosis' if they were presumed to have tuberculosis and were started on anti-tuberculosis treatment. We will answer 'unclear' for all studies.

For rifampicin resistance, we will answer 'yes' if a study used phenotypic culture-based drug susceptibility testing as a reference standard. As this is an inclusion criterion for the review, we will only include studies for this target condition that satisfy this reference standard.

Signalling question 2: (tuberculosis) Were the reference standard results interpreted without knowledge of the results of the index test?

For pulmonary tuberculosis we will answer '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. We will answer 'no' if the study stated that the reference standard result was interpreted with knowledge of the Xpert test result. We will answer 'unclear' if we cannot tell.

For extrapulmonary tuberculosis, we will answer '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. We will answer 'no' if the study stated that the reference standard result was interpreted with knowledge of the Xpert test result. We will answer 'unclear' if we cannot tell.

For rifampicin resistance, we will answer 'yes' if the reference test provided an automated result (e.g. MGIT 960 SIRE); blinding was explicitly stated; or it was clear that the reference standard was performed at a separate laboratory or performed by different people. We will answer 'no' if the study stated that the reference standard result was interpreted with knowledge of the Xpert test result. We will answer 'unclear' if we cannot tell.

Applicability: Are there concerns that the target condition as defined by the reference standard does not match the question?

For both pulmonary tuberculosis and extrapulmonary tuberculosis, we will answer 'high concern' if the included studies did not differentiate *Mycobacterium tuberculosis* complex isolated in culture from other mycobacteria using any speciation technique; 'low concern' if speciation was performed using any technique; and 'unclear concern' if we cannot tell.

For rifampicin resistance, we will judge applicability to be of 'low concern' for all studies because the method used (phenotypic culture-based drug susceptibility testing) is appropriate.

Domain 4: Flow and timing

Risk of bias: Could the patient flow have introduced bias?

Signalling question 1: Was there an appropriate interval between the index test and reference standard? We expect to find for most included studies that specimens for Xpert and culture were obtained at the same time, when patients were evaluated for presumed tuberculosis. Even if there were a delay of several days between index test and reference standards, tuberculosis is a chronic disease, and we consider misclassification of disease status to be unlikely, as long as treatment was not initiated in the interim. We will answer 'yes' if the index test and reference standard were performed at the same time or if the time interval was less than or equal to seven days; 'no' if the time interval was greater than seven days; and 'unclear' if we cannot tell.

Signalling question 2: Did all patients receive the same reference standard? We will answer 'yes' if all patients received the same reference standard; 'no' if all patients did not receive the same reference standard; and 'unclear' if we cannot tell.

Signalling question 3: Were all patients included in the analysis? We will determine the answer to this question by comparing the number of patients enrolled with the number of patients included in the 2 x 2 tables. We will answer 'yes' if the numbers matched and 'no' if there were patients enrolled in the study that were not included in the analysis. We will answer 'unclear' if we cannot tell.

Judgements for risk of bias assessments for a given domain:

- If we answer all signalling questions for a domain 'yes', then we will judge risk of bias as 'low'.
- If we answer all or most signalling questions for a domain 'no', then we will judge risk of bias as 'high'.
- If we answer only one signalling question for a domain 'no', we will further discuss the risk of bias judgement.
- If we answer all or most signalling questions for a domain 'unclear', then we will judge risk of bias as 'unclear'.
- If we answer only one signalling question for a domain 'unclear', we will further discuss the risk of bias judgement for the domain.

CONTRIBUTIONS OF AUTHORS

AK and LGF wrote the initial draft of the protocol. AD, AM, and AK designed the draft abstraction forms. AD, AM, AK, and LGF drafted the data analysis sections with input from KRS and YT. KRS drafted the QUADAS-2 section. KRS, YT, AM, and AD contributed methodological advice. All authors (AK, LGF, KRS, ME, YT, RV, AD, AM) provided input for the protocol.

DECLARATIONS OF INTEREST

AK has conducted prior primary research on tuberculosis diagnostics, and has no known conflicts of interest.

LGF has no known conflicts of interest.

YT has no known conflicts of interest.

ME has no known conflicts of interest.

RV has no known conflicts of interest.

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.

AD has conducted prior primary research on tuberculosis diagnostics, and has no known conflicts of interest.

AM has conducted prior primary research on tuberculosis diagnostics, and has no known conflicts of interest.

SOURCES OF SUPPORT