Cochrane Library

Cochrane Database of Systematic Reviews

Xpert® MTB/RIF assay for extrapulmonary tuberculosis and rifampicin resistance (Review)

Kohli M, Schiller I, Dendukuri N, Dheda K, Denkinger CM, Schumacher SG, Steingart KR

Kohli M, Schiller I, Dendukuri N, Dheda K, Denkinger CM, Schumacher SG, Steingart KR.
Xpert® MTB/RIF assay for extrapulmonary tuberculosis and rifampicin resistance.
Cochrane Database of Systematic Reviews 2018, Issue 8. Art. No.: CD012768.
DOI: 10.1002/14651858.CD012768.pub2.

www.cochranelibrary.com

TABLE OF CONTENTS

[Diagnostic Test Accuracy Review]

Xpert® MTB/RIF assay for extrapulmonary tuberculosis and rifampicin resistance

Mikashmi Kohli[1], Ian Schiller[2], Nandini Dendukuri[2], Keertan Dheda[3], Claudia M Denkinger[4], Samuel G Schumacher[4], Karen R Steingart[5]

[1]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada. [2]Division of Clinical Epidemiology, McGill University Health Centre - Research Institute, Montreal, Canada. [3]3 Centre for Lung Infection and Immunity Unit, Department of Medicine and UCT Lung Institute, University of Cape Town, Cape Town, South Africa. [4]FIND, Geneva, Switzerland. [5]Honorary Research Fellow, Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, UK

Contact address: Karen R Steingart, Honorary Research Fellow, Department of Clinical Sciences, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, UK. karen.steingart@gmail.com.

Editorial group: Cochrane Infectious Diseases Group
Publication status and date: Unchanged, published in Issue 8, 2018.

Citation: Kohli M, Schiller I, Dendukuri N, Dheda K, Denkinger CM, Schumacher SG, Steingart KR. Xpert® MTB/RIF assay for extrapulmonary tuberculosis and rifampicin resistance. *Cochrane Database of Systematic Reviews* 2018, Issue 8. Art. No.: CD012768. DOI: 10.1002/14651858.CD012768.pub2.

ABSTRACT

Background

Tuberculosis (TB) is the world's leading infectious cause of death. Extrapulmonary TB accounts for 15% of TB cases, but the proportion is increasing, and over half a million people were newly diagnosed with rifampicin-resistant TB in 2016. Xpert® MTB/RIF (Xpert) is a World Health Organization (WHO)-recommended, rapid, automated, nucleic acid amplification assay that is used widely for simultaneous detection of *Mycobacterium tuberculosis* complex and rifampicin resistance in sputum specimens. This Cochrane Review assessed the accuracy of Xpert in extrapulmonary specimens.

Objectives

To determine the diagnostic accuracy of Xpert a) for extrapulmonary TB by site of disease in people presumed to have extrapulmonary TB; and b) for rifampicin resistance in people presumed to have extrapulmonary TB.

Search methods

We searched the Cochrane Infectious Diseases Group Specialized Register, MEDLINE, Embase, Science Citation Index, Web of Science, Latin American Caribbean Health Sciences Literature (LILACS), Scopus, ClinicalTrials.gov, the WHO International Clinical Trials Registry Platform, the International Standard Randomized Controlled Trial Number (ISRCTN) Registry, and ProQuest up to 7 August 2017 without language restriction.

Selection criteria

We included diagnostic accuracy studies of Xpert in people presumed to have extrapulmonary TB. We included TB meningitis and pleural, lymph node, bone or joint, genitourinary, peritoneal, pericardial, and disseminated TB. We used culture as the reference standard. For pleural TB, we also included a composite reference standard, which defined a positive result as the presence of granulomatous inflammation or a positive culture result. For rifampicin resistance, we used culture-based drug susceptibility testing or MTBDR*plus* as the reference standard.

Data collection and analysis

Two review authors independently extracted data, assessed risk of bias and applicability using the QUADAS-2 tool. We determined pooled predicted sensitivity and specificity for TB, grouped by type of extrapulmonary specimen, and for rifampicin resistance. For TB detection, we used a bivariate random-effects model. Recognizing that use of culture may lead to misclassification of cases of extrapulmonary TB as 'not TB' owing to the paucibacillary nature of the disease, we adjusted accuracy estimates by applying a latent class meta-analysis model. For rifampicin resistance detection, we performed univariate meta-analyses for sensitivity and specificity separately to include studies in which no rifampicin resistance was detected. We used theoretical populations with an assumed prevalence to provide illustrative numbers of patients with false positive and false negative results.

Main results

We included 66 unique studies that evaluated 16,213 specimens for detection of extrapulmonary TB and rifampicin resistance. We identified only one study that evaluated the newest test version, Xpert MTB/RIF Ultra (Ultra), for TB meningitis. Fifty studies (76%) took place in low- or middle-income countries. Risk of bias was low for patient selection, index test, and flow and timing domains and was high or unclear for the reference standard domain (most of these studies decontaminated sterile specimens before culture inoculation). Regarding applicability, in the patient selection domain, we scored high or unclear concern for most studies because either patients were evaluated exclusively as inpatients at tertiary care centres, or we were not sure about the clinical settings.

Pooled Xpert sensitivity (defined by culture) varied across different types of specimens (31% in pleural tissue to 97% in bone or joint fluid); Xpert sensitivity was > 80% in urine and bone or joint fluid and tissue. Pooled Xpert specificity (defined by culture) varied less than sensitivity (82% in bone or joint tissue to 99% in pleural fluid and urine). Xpert specificity was ≥ 98% in cerebrospinal fluid, pleural fluid, urine, and peritoneal fluid.

Xpert testing in cerebrospinal fluid

Xpert pooled sensitivity and specificity (95% credible interval (CrI)) against culture were 71.1% (60.9% to 80.4%) and 98.0% (97.0% to 98.8%), respectively (29 studies, 3774 specimens; moderate-certainty evidence).

For a population of 1000 people where 100 have TB meningitis on culture, 89 would be Xpert-positive: of these, 18 (20%) would not have TB (false-positives); and 911 would be Xpert-negative: of these, 29 (3%) would have TB (false-negatives).

For TB meningitis, ultra sensitivity and specificity against culture (95% confidence interval (CI)) were 90% (55% to 100%) and 90% (83% to 95%), respectively (one study, 129 participants).

Xpert testing in pleural fluid

Xpert pooled sensitivity and specificity (95% CrI) against culture were 50.9% (39.7% to 62.8%) and 99.2% (98.2% to 99.7%), respectively (27 studies, 4006 specimens; low-certainty evidence).

For a population of 1000 people where 150 have pleural TB on culture, 83 would be Xpert-positive: of these, seven (8%) would not have TB (false-positives); and 917 would be Xpert-negative: of these, 74 (8%) would have TB (false-negatives).

Xpert testing in urine

Xpert pooled sensitivity and specificity (95% CrI) against culture were 82.7% (69.6% to 91.1%) and 98.7% (94.8% to 99.7%), respectively (13 studies, 1199 specimens; moderate-certainty evidence).

For a population of 1000 people where 70 have genitourinary TB on culture, 70 would be Xpert-positive: of these, 12 (17%) would not have TB (false-positives); and 930 would be Xpert-negative: of these, 12 (1%) would have TB (false-negatives).

Xpert testing for rifampicin resistance

Xpert pooled sensitivity (20 studies, 148 specimens) and specificity (39 studies, 1088 specimens) were 95.0% (89.7% to 97.9%) and 98.7% (97.8% to 99.4%), respectively (high-certainty evidence).

For a population of 1000 people where 120 have rifampicin-resistant TB, 125 would be positive for rifampicin-resistant TB: of these, 11 (9%) would not have rifampicin resistance (false-positives); and 875 would be negative for rifampicin-resistant TB: of these, 6 (1%) would have rifampicin resistance (false-negatives).

For lymph node TB, the accuracy of culture, the reference standard used, presented a greater concern for bias than in other forms of extrapulmonary TB.

Authors' conclusions

In people presumed to have extrapulmonary TB, Xpert may be helpful in confirming the diagnosis. Xpert sensitivity varies across different extrapulmonary specimens, while for most specimens, specificity is high, the test rarely yielding a positive result for people without TB

(defined by culture). Xpert is accurate for detection of rifampicin resistance. For people with presumed TB meningitis, treatment should be based on clinical judgement, and not withheld solely on an Xpert result, as is common practice when culture results are negative.

2 April 2019

Up to date

All studies incorporated from most recent search

Updated review: all eligible published studies found in the last search (7 Aug, 2017) were included

PLAIN LANGUAGE SUMMARY

Xpert® MTB/RIF test for diagnosing extrapulmonary tuberculosis and rifampicin resistance

Why is improving the diagnosis of extrapulmonary tuberculosis important?

Tuberculosis (TB) is the world's leading infectious cause of death. It mainly affects the lungs (pulmonary TB) but may occur in other body parts than the lungs (extrapulmonary TB). In most people, TB can be cured if the disease is diagnosed and properly treated. One problem involved in treating TB is that the bacteria become resistant to antibiotics. Not recognizing TB early (false-negative result) may result in delayed diagnosis and treatment and increased illness and death. An incorrect TB diagnosis (false-positive result) may result in increased anxiety and unnecessary treatment.

What is the aim of this review?

To find out how accurate Xpert® MTB/RIF (Xpert) is for diagnosing extrapulmonary TB and drug resistance. We included eight forms of extrapulmonary TB: tuberculous meningitis and pleural, lymph node, bone or joint, genitourinary, peritoneal, pericardial, and disseminated TB.

What was studied in this review?

Xpert is a relatively new, automated, rapid test that detects TB and rifampicin resistance at the same time. Rifampicin is an important drug for treating people with TB. Another Cochrane Review showed that Xpert is accurate for diagnosing pulmonary TB. The current review assessed Xpert accuracy for detecting eight forms of extrapulmonary TB, as well as the different specimens that may be collected for diagnosis, for instance, cerebrospinal fluid, pleural fluid, and urine. Xpert results were measured against culture results (benchmark).

What are the main results reported in this review?

We included 66 studies that evaluated 16,213 specimens for extrapulmonary TB and rifampicin resistance. Only one study evaluated the newest test version, Xpert Ultra (Ultra), for tuberculous meningitis.

In urine and bone or joint fluid and tissue, Xpert was sensitive (more than 80%), that is, registered positive in people who actually had TB. In cerebrospinal fluid, pleural fluid, urine, and peritoneal fluid, Xpert was highly specific (98% or more), that is, did not register positive in people who were actually negative.

For a population of 1000 people:

• where 100 have TB meningitis on culture, 89 would be Xpert-positive: of these, 18 (20%) would not have TB; and 911 would be Xpert-negative: of these, 29 (3%) would have TB.

• where 150 have pleural TB on culture, 83 would be Xpert-positive: of these, seven (8%) would not have TB ; and 917 would be Xpert-negative: of these, 74 (8%) would have TB.

• where 70 have genitourinary TB on culture, 70 would be Xpert-positive: of these, 12 (17%) would not have TB; and 930 would be Xpert-negative: of these, 12 (1%) would have TB.

• where 120 have rifampicin-resistant TB, 125 would be positive for rifampicin-resistant TB: of these, 11 (9%) would not have rifampicin resistance; and 875 would be negative for rifampicin-resistant TB: of these, 6 (1%) would have rifampicin resistance.

How confident are we in the review's results?

The diagnosis of extrapulmonary TB was made by assessing patients with culture, generally considered to be the best reference standard. However, it appears that culture did not work well as a reference test for lymph node TB.

Who do the review's results apply to?

People presumed to have extrapulmonary TB. Most studies included only inpatients at tertiary care centres or did not report the clinical setting. Therefore, we could not say how the test would work in primary care.

What are the implications of this review?

Xpert may be helpful in diagnosing extrapulmonary TB. The ability of Xpert to detect TB varies when different specimens are used, while Xpert rarely yields a positive result for people without TB (defined by culture). Xpert is accurate for diagnosing rifampicin resistance. In patients thought to have TB meningitis, which is considered a medical emergency, providers should use clinical judgement and should not rely solely on an Xpert result when deciding to withhold treatment, as is common practice when culture results are negative.

How up-to-date is this review?

The review authors searched for studies published up to 7 August 2017.

SUMMARY OF FINDINGS

Summary of findings 1. Xpert® MTB/RIF in cerebrospinal fluid

Participants: patients presumed to have TB meningitis

Prior testing: patients who received Xpert testing may first have undergone a health examination (history and physical examination) and possibly a chest radiograph

Role: replacement test for usual practice

Settings: primarily tertiary care centres (the index test was often run in reference laboratories)

Index (new) test: Xpert

Studies: cross-sectional studies

Limitations: participants were evaluated exclusively as inpatients at a tertiary care centre, or, if the clinical setting was not reported, Xpert was performed at a reference laboratory rather than at primary care facilities and local hospitals

Pooled sensitivity (95% CrI): 71.1% (60.9 to 80.4); **pooled specificity (95% CrI):** 98.0% (97.0 to 98.8)

Test result	1000 people tested for TB using Xpert® MTB/RIF (95% CrI)			Number of participants (studies)	Certainty of the evidence (GRADE)
	Prevalence of 1%	Prevalence of 5%	Prevalence of 10%		
True-positives (patients with TB meningitis)	7 (6 to 8)	36 (30 to 40)	71 (61 to 80)	433 (29)	⊕⊕⊕⊖ Moderate[a,b]
False-negatives (patients incorrectly classified as not having TB meningitis)	3 (2 to 4)	14 (10 to 20)	29 (20 to 39)		
True-negatives (patients without TB meningitis)	970 (960 to 978)	931 (922 to 939)	882 (873 to 889)	3341 (29)	⊕⊕⊕⊕ High
False-positives (patients incorrectly classified as having TB meningitis)	20 (12 to 30)	19 (11 to 28)	18 (11 to 27)		

Abbreviations: CrI: credible interval; TB: tuberculosis.
The median prevalence in the included studies was 10%. We also included other plausible prevalence estimates for the target condition.
Credible limits were estimated based on those around the point estimates for pooled sensitivity and specificity. The results presented in this table should not be interpreted in isolation from results of the individual included studies contributing to each summary test accuracy measure. These are reported in the main body of the text of the review.
[a]As assessed by QUADAS-2, for the reference standard domain only four studies (14%) had unclear risk of bias because specimens underwent decontamination. We did not downgrade.
[b]The wide CrI around true-positives and false-negatives may lead to different decisions depending on which credible limits are assumed. We downgraded one level.
GRADE certainty of the evidence
High: we are very confident that the true effect lies close to that of the estimate of the effect.

Moderate: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.
The results presented in this table should not be interpreted in isolation from results of the individual included studies contributing to each summary test accuracy measure.

Summary of findings 2. Xpert® MTB/RIF in pleural fluid

Participants: patients presumed to have pleural TB

Prior testing: patients who received Xpert testing may first have undergone a health examination (history and physical examination) and possibly a chest radiograph

Role: replacement test for standard practice, which may include more invasive tests, such as pleural biopsy

Settings: primarily tertiary care centres (the index test was often run in reference laboratories)

Index (new) test: Xpert

Reference standard: solid or liquid culture

Studies: cross-sectional studies

Limitations: in most studies, participants were evaluated at a tertiary care centre, or if the clinical setting was not reported, Xpert was performed at a reference laboratory

Pooled sensitivity (95% CrI): 50.9% (39.7 to 62.8); **pooled specificity (95% CrI):** 99.2% (98.2 to 99.7)

Test result	1000 people tested for TB using Xpert® MTB/RIF (95% CrI)			Number of participants (studies)	Certainty of the evidence (GRADE)
	Prevalence of 10%	Prevalence of 15%	Prevalence of 25%		
True-positives (patients with pleural TB)	25 (20 to 31)	76 (60 to 94)	127 (99 to 157)	606 (27)	⊕⊕⊖⊖ Low[a,b]
False-negatives (patients incorrectly classified as not having pleural TB)	25 (19 to 30)	74 (56 to 90)	123 (93 to 151)		
True-negatives (patients without pleural TB)	942 (933 to 947)	843 (835 to 847)	744 (736 to 748)	3399 (27)	⊕⊕⊕⊕ High
False-positives (patients incorrectly classified as having pleural TB)	8 (3 to 17)	7 (3 to 15)	6 (2 to 14)		

Abbreviations: CrI: credible interval; TB: tuberculosis.
The median prevalence in the included studies was 15%. We also included other plausible prevalence estimates for the target condition.
[a]As assessed by QUADAS-2, for the reference standard domain, ten studies (37%) had unclear risk of bias because specimens underwent decontamination. We did not downgrade.
[b]For individual studies, sensitivity estimates ranged from 10% to 100%. We could not explain heterogeneity by study quality or other factors. We downgraded two levels for inconsistency.

GRADE certainty of the evidence
High: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.
The results presented in this table should not be interpreted in isolation from results of the individual included studies contributing to each summary test accuracy measure.

Summary of findings 3. Xpert® MTB/RIF in urine

Participants: patients presumed to have genitourinary TB

Prior testing: patients who received Xpert testing may first have undergone a health examination (history and physical examination) and possibly a chest radiograph

Role: replacement test for standard practice, which may include more invasive tests, such as biopsy of affected organs

Settings: primarily tertiary care centres (the index test was often run in reference laboratories)

Index (new) test: Xpert

Reference standard: solid or liquid culture

Studies: cross-sectional studies

Limitations: in most studies, participants were evaluated at a tertiary care centre, or if the clinical setting was not reported, Xpert was performed at a reference laboratory

Sensitivity: 82.7% (69.6 to 91.1); **specificity:** 98.7% (94.8 to 99.7)

Test result	1000 people tested for TB using Xpert®MTB/RIF (95% CrI)			Number of participants (studies)	Certainty of the evidence (GRADE)
	Prevalence of 2%	Prevalence of 7%	Prevalence of 15%		
True-positives (patients with genitourinary TB)	17 (14 to 18)	58 (49 to 64)	124 (104 to 137)	73 (13)	⊕⊕⊕⊖ Moderate[a,b]
False-negatives (patients incorrectly classified as not having genitourinary TB)	3 (2 to 6)	12 (6 to 21)	26 (13 to 46)		
True-negatives (patients without genitourinary TB)	967 (929 to 977)	918 (882 to 927)	839 (806 to 847)	1126 (13)	⊕⊕⊕⊖ Moderate[c]
False-positives (patients incorrectly classified as having genitourinary TB)	13 (3 to 51)	12 (3 to 48)	11 (3 to 44)		

Abbreviations: CrI: credible interval; TB: tuberculosis.
The median prevalence in the included studies was 7%. We included what we considered to be plausible prevalence estimates for the target condition.

[a]As assessed by QUADAS-2, for the reference standard domain only four studies (31%) had unclear risk of bias because specimens underwent decontamination.
[b]For individual studies, sensitivity estimates ranged from 0% to 100%. We thought that the small number of culture-positives in studies could explain some, but probably not all, of the variation in sensitivity results. We downgraded one level.
[c]The wide CrI around true-negatives and false-positives may lead to different decisions depending on which credible limits are assumed. We downgraded one level.
GRADE certainty of the evidence
High: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.
The results presented in this table should not be interpreted in isolation from results of the individual included studies contributing to each summary test accuracy measure.

Summary of findings 4. Xpert® MTB/RIF for rifampicin resistance

Participants: patients with TB detected by Xpert® MTB/RIF

Role: replacement test for standard practice, which includes culture-based drug susceptibility testing or MTBDR*plus*

Settings: primarily tertiary care centres (the index test was often run in central (reference laboratories), where drug susceptibility testing for the reference standard could be performed)

Index (new) test: Xpert® MTB/RIF

Reference standard: culture-based drug susceptibility testing using solid or liquid media or MTBDR*plus*

Studies: cross-sectional studies

Pooled sensitivity (95% CrI): 95.0% (89.7 to 97.9); **pooled specificity (95% CrI):** 98.7% (97.8 to 99.4)

Test result	1000 people tested for rifampicin resistance using Xpert®MTB/RIF (95% CrI)		Number of participants (studies)	Certainty of the evidence (GRADE)
	Prevalence of 5%	Prevalence of 12%		
True-positives (patients correctly classified as rifampicin resistant)	48 (45 to 49)	114 (108 to 117)	148 (20)	⊕⊕⊕⊕ High
False-negatives (patients incorrectly classified as rifampicin susceptible)	2 (1 to 5)	6 (3 to 12)		
True-negatives (patients correctly classified as rifampicin susceptible)	938 (929 to 944)	869 (861 to 875)	1088 (39)	⊕⊕⊕⊕ High
False-positives (patients incorrectly classified as rifampicin resistant)	12 (6 to 21)	11 (5 to 19)		

Abbreviations: CrI: credible interval; TB: tuberculosis.
The median prevalence in the included studies was 12%.

GRADE certainty of the evidence
High: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.
The results presented in this table should not be interpreted in isolation from results of the individual included studies contributing to each summary test accuracy measure.

BACKGROUND

Tuberculosis (TB) is caused by infection with *Mycobacterium tuberculosis* (*M. tuberculosis*) bacteria. TB causes tremendous suffering worldwide and has surpassed HIV/AIDS as the world's leading infectious cause of death. The World Health Organization (WHO) estimates that globally in 2016, 1.3 million HIV-negative people and 374,000 HIV-positive people died from TB and 10.4 million people became ill with TB (WHO 2017a). Drug-resistant TB is an enormous threat. In 2016, an estimated 600,000 people were newly diagnosed with rifampicin-resistant TB, 490,000 of whom had multidrug-resistant TB (MDR-TB) (WHO 2017a). MDR-TB is caused by infection with *M. tuberculosis* bacteria that are resistant to at least rifampicin and isoniazid. Rifampicin is the most effective first-line anti-TB drug. When people receive proper treatment, TB is treatable and curable.

TB predominantly affects the lungs (pulmonary TB). Extrapulmonary TB, which refers to TB in parts of the body other than the lungs, is known to affect virtually every part of the body; lymph nodes and the pleura are the most common sites (Sharma 2004). Although active pulmonary TB is transmissible by droplets spread by coughing, extrapulmonary TB is thought to result from hematogenous spread from an initial lung infection and is not infectious. Extrapulmonary TB can occur alone or together with pulmonary TB. Of the 6.3 million new cases of TB notified to WHO in 2016, 15% were cases of extrapulmonary TB (range, 8% in the WHO Western Pacific Region to 24% in the WHO Eastern Mediterranean Region) (WHO 2017a). Among countries in the European Union, extrapulmonary TB was responsible for 19% of all notified cases (range, 6% to 44%) (Sandgren 2013). However, the number of people affected by extrapulmonary TB is likely to be higher, given that, according to WHO, extrapulmonary TB is notified as pulmonary TB when the two forms exist together (WHO 2014b), and diagnosing extrapulmonary TB is challenging, as described below. Additionally, extrapulmonary TB accounts for an increasing proportion of new TB cases in some countries, in part because of host and genetic considerations, and the association of extrapulmonary TB and HIV (Golden 2005; Pai 2016; Perkins 2007; Webster 2014). Based on surveillance and epidemiological data, extrapulmonary TB affects a greater proportion of children than adults (Nelson 2004).

WHO TB treatment guidelines recommend the same drug regimens for extrapulmonary and pulmonary disease with notable mention of other guidelines, which recommend longer treatment for TB meningitis and for bone or joint TB (WHO 2010). An updated guideline, published in 2017, provided recommendations on the use of adjuvant steroids for treatment of TB meningitis (strong recommendation; moderate-certainty evidence), and TB pericarditis (conditional recommendation; very low-certainty evidence) (WHO 2017b). Recent TB treatment guidelines include Index-TB 2016 (India), and those issued by the American Thoracic Society, the Centers for Disease Control and Prevention (CDC), and the Infectious Diseases Society of America (Nahid 2016).

Diagnosis of extrapulmonary TB is challenging for several reasons. Many forms of extrapulmonary TB require invasive diagnostic sampling; gathering adequate specimens can pose risk of harm to the patient and can be costly. Most forms of extrapulmonary TB are paucibacillary (TB disease caused by a small number of bacteria), making diagnosis by the conventional method of smear microscopy less sensitive. This problem particularly affects resource-limited settings, where the more sensitive methods of mycobacterial culture and histological examination are not widely available. Limitations are also associated with culture and histology: culture takes several weeks, requires a highly equipped laboratory, and has reduced sensitivity in paucibacillary disease; histology relies on highly trained operators, and characteristic morphology is shared with other diseases. As a result of these difficulties, diagnosis of extrapulmonary TB is often made on the grounds of clinical suspicion alone, and many people receive the wrong diagnosis, leading to unnecessary TB treatment or poor outcomes from untreated extrapulmonary TB. The need for faster, more reliable diagnostics that are suitable for resource-limited settings is clear and has been defined by the research community (Denkinger 2015). In 2014, the World Health Assembly unanimously approved the End TB Strategy, a 20-year strategy devised to end the global TB epidemic. The END TB strategy calls for early diagnosis of TB and universal drug susceptibility testing (DST) (WHO END TB 2014).

Xpert® MTB/RIF (Xpert) is an automated diagnostic test for the detection of *Mycobacterium tuberculosis* complex (*M. tuberculosis*). It is a DNA-based test that detects the *M. tuberculosis rpoB* gene. Xpert also detects mutations in *rpoB* that may cause rifampicin resistance. Results are available after two hours with minimal hands-on technical time. A Cochrane Review found that Xpert accurately detects*M. tuberculosis* and rifampicin resistance when used on sputum specimens (Steingart 2014). The WHO published updated guidance on use of Xpert in 2013 (WHO 2013). This updated policy statement expanded recommendations for use of Xpert for pulmonary TB in adults and provided additional guidance on use of the test for childhood TB and extrapulmonary TB.

Drawing on a systematic review (Denkinger 2014), and using the GRADE approach, the WHO has issued the following recommendations related to extrapulmonary TB.

- Xpert should be used in preference to conventional microscopy and culture as the initial diagnostic test for cerebrospinal fluid (CSF) specimens from patients presumed to have TB meningitis (strong recommendation given the urgency for rapid diagnosis; very low-certainty evidence).
- Xpert may be used as a replacement test for usual practice (including conventional microscopy, culture, or histopathology) for testing specific non-respiratory specimens (lymph nodes and other tissues) from patients presumed to have extrapulmonary TB (conditional recommendation; very low-certainty evidence).

The use of Xpert has also been incorporated into the International Standards for TB Care 2014 (TB Care I 2014). Clinical practice guidelines on the diagnosis of pulmonary and extrapulmonary TB in adults and children for clinicians in high-resource countries with low TB incidence have recently been published (Lewinsohn 2017).

Currently, the manufacturer, Cepheid Incorporated (Sunnyvale, CA, USA), has made no claim for the use of Xpert in non-sputum specimens (Cepheid 2015); accordingly, Xpert is approved by the US Food and Drug Administration (FDA) for use in raw sputum specimens and concentrated sputum sediment only (FDA 2013).

Target condition being diagnosed

Extrapulmonary TB

The various forms of extrapulmonary TB cause signs and symptoms related to the structures affected. Table 1 describes the forms of extrapulmonary TB included in this Cochrane Review, as well as the different specimens that may be collected for diagnosis.

Rifampicin resistance

Rifampicin inhibits bacterial DNA-dependent RNA polymerase, encoded by the RNA polymerase gene (*rpoB*) (Hartmann 1967). Resistance to this drug has been associated mainly with mutations in a limited region of the *rpoB* gene (Telenti 1993). Rifampicin resistance may occur alone or in association with resistance to isoniazid and other drugs. In settings with a high burden of MDR-TB, the presence of rifampicin resistance alone may serve as a proxy for MDR-TB (WHO 2011).

Index test(s)

Xpert is an automated diagnostic test for the detection of *M. tuberculosis complex* DNA and, when *M. tuberculosis* complex (hereafter expressed to as *M. tuberculosis)* is detected, rifampin-resistance associated mutations of the *rpoB* gene. Test results are available for *M. tuberculosis* and resistance to rifampicin within two hours after the test is begun, with minimal hands-on technical time. Unlike conventional nucleic acid amplification (NAA) tests, Xpert integrates sample processing and PCR amplification and detection into a single self-enclosed test unit, the GeneXpert cartridge (Blakemore 2010). Following sample loading, all steps in the assay are completely automated and self-contained. In addition, the assay's sample reagent, used to liquefy sputum, has potent tuberculocidal (the ability to kill TB bacteria) properties and so largely eliminates biosafety concerns during the test procedure (Banada 2010). Xpert detects both live and dead bacteria (Miotto 2012).

Xpert uses molecular beacon technology to detect rifampicin resistance. Molecular beacons are nucleic acid probes that recognize and report the presence or absence of the normal, rifampicin-susceptible, 'wild-type' sequence of the *rpoB* gene of TB. Beacons of five different colours are used, each covering a separate nucleic acid sequence within the amplified *rpoB* gene.

Xpert provides testing simultaneously for *M. tuberculosis* and rifampicin resistance. Thus, it is really only one test. A rifampicin resistance result is provided whether or not a patient is at risk of resistance. One cannot deselect testing for rifampicin resistance and run only the assay for TB detection. Xpert may be used at all levels of the healthcare system. However, for use of the current device, a stable and uninterrupted electrical supply is required. The WHO has published extensive guidance and practical information on implementing the test (WHO 2014a).

Since Xpert was released, five generations of the cartridge have been developed: G1, G2, G3, G4, and Xpert Ultra (Ultra). Preparation of specimens and the cartridge procedure for Xpert and Ultra are the same (Chakravorty 2017). However, technically, Ultra differs from earlier Xpert generations in several ways. To improve detection of *M. tuberculosis,* Ultra incorporates two different multi-copy amplification targets (IS6110 and IS1081), and to improve detection of rifampicin resistance, Ultra uses melting temperature-based analysis instead of real-time PCR (Chakravorty 2017).

In a multi-country diagnostic accuracy study comparing Ultra and Xpert version G4 in sputum specimens for pulmonary TB (n = 1439), the sensitivity of Ultra was higher than that of Xpert (sensitivity of 63% for Ultra versus 46% for Xpert in people who were smear-negative and culture-positive, 137 participants; sensitivity of 95% for Ultra versus 77% for Xpert in people living with HIV, 115 participants) (Dorman 2018). However, the specificity of Ultra was lower than that of Xpert (specificity of 96% for Ultra versus 98% for Xpert) (Dorman 2018). In additional retrospective studies, Ultra showed improved sensitivity, in particular for TB meningitis and childhood TB. In CSF, Ultra sensitivity was 95% for TB meningitis compared with Xpert sensitivity of 45%. In children, using respiratory specimens, Ultra sensitivity was 71% for TB compared with Xpert sensitivity of 47% (FIND 2017; WHO 2017c). The WHO has recently recommended Ultra as an alternative to Xpert, stating that all recommendations concerning use of Xpert with selected extrapulmonary specimens (CSF, lymph nodes, and tissue specimens) also apply to Ultra (WHO 2017c).

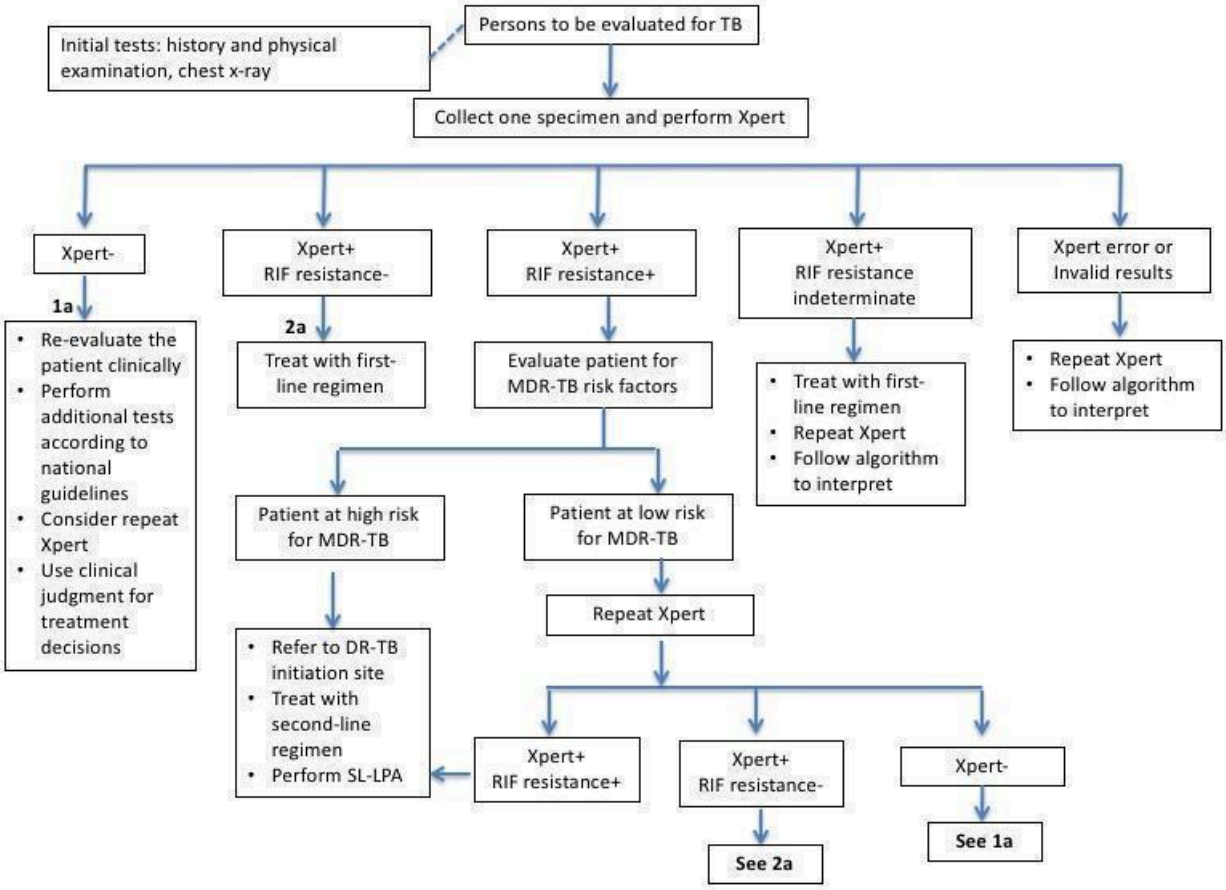
We included in this Cochrane Review studies that used any of the Xpert generations.

Clinical pathway

It is recommended that clinicians who evaluate patients for extrapulmonary TB adhere to Standard 4 of the International Standards for TB Care, which states: "For all patients, including children, presumed to have extrapulmonary TB, appropriate specimens from the presumed sites of involvement should be obtained for microbiological and histological examination. An Xpert test is recommended as the preferred initial microbiological test for presumptive TB meningitis because of the need for a rapid diagnosis" (TB Care I 2014).

Figure 1 shows the clinical pathway and presents the context in which Xpert might be used. The target condition is extrapulmonary TB, of which several forms are known (e.g. pleural TB, TB meningitis).

Figure 1. The clinical pathway describes how patients might present and the point in the pathway at which they would be considered for testing with Xpert. Before a specimen was tested with Xpert, patients presumed of having extrapulmonary TB would have undergone a health examination (history and physical examination) and possibly a chest radiograph. Presentation of extrapulmonary TB varies depending on the body site affected; this condition may imitate other diseases such as cancer and bacterial and fungal infections. Signs and symptoms of extrapulmonary TB are often non-specific and may include fever, night sweats, fatigue, loss of appetite, and weight loss (as seen in pulmonary TB) or specific complaints related to the involved site (e.g. headache for TB meningitis, back pain for TB of the spine). The clinical presentation of extrapulmonary disease may be acute but is more often subacute (falling between acute and chronic) or chronic, meaning that patients may have symptoms for days to months before they seek care. Signs and symptoms for the forms of extrapulmonary TB included in this review are described in Table 1. Standard practice includes obtaining specimens for microscopy, culture, and histological examination. We adapted this algorithm for Xpert from the Global Laboratory Initiative (GLI 2018). Abbreviations: DR-TB: drug-resistant TB; MDR-TB: multidrug-resistant TB; RIF: rifampicin; SL-LPA: line probe assay for second-line drugs; TB: tuberculosis.



Before a specimen is tested with Xpert, patients presumed of having extrapulmonary TB would have undergone a health examination (history and physical examination) and possibly a chest radiograph. The presentation of extrapulmonary TB varies depending on the body site affected, and it may imitate other diseases, such as cancer and bacterial and fungal infections. Signs and symptoms of extrapulmonary TB are often non-specific and may include fever, night sweats, fatigue, loss of appetite, and weight loss (as seen in pulmonary TB) or specific complaints related to the involved site (e.g. headache for TB meningitis, back pain for TB of the spine). The clinical presentation of extrapulmonary disease may be acute but is more often subacute (falling between acute and chronic) or chronic, meaning that patients may have symptoms for days to months before they seek care.

We have described in Table 1 signs and symptoms of the forms of extrapulmonary TB included in this review. The clinician should take a careful history, noting history of TB exposure, prior TB disease, and medical conditions that increase the risk for TB disease (e.g. HIV, diabetes mellitus, low body weight). In comparison with HIV-negative people, HIV-positive people have higher rates of extrapulmonary TB or mycobacteraemia (TB bloodstream infection). HIV-positive patients with signs or symptoms of extrapulmonary TB should have specimens

taken from the suspected site(s) of involvement to increase the likelihood of TB diagnosis. In general, children and adults with extrapulmonary TB present in a similar way. However, infants and young children are at highest risk of developing disseminated TB disease and TB meningitis - the most severe forms of TB. In TB meningitis, diagnosis is often delayed with appalling consequences for patients. For all forms of extrapulmonary TB, patients may be evaluated in primary or secondary care settings. However, if more complex or invasive tests are needed, patients may be referred to a tertiary medical centre (Iseman 2000; Reuter 2009; Sharma 2004). In many countries, district-level and lower-level laboratories offer a range of basic diagnostic tests, including Xpert (GLI 2017).

Xpert is used to diagnose TB and to detect rifampicin resistance. Xpert is performed as a replacement for standard practice, which includes obtaining appropriate specimens from presumed sites of involvement for microbiological (conventional microscopy and culture) and histological examination. An Xpert test is recommended as the preferred initial microbiological test for presumptive TB meningitis because of the need for a rapid diagnosis (TB Care I 2014; WHO 2013). In HIV-positive people with a CD4 cell count of 100 cells/µL or lower, and in HIV-positive people who are seriously ill regardless of CD4 count, the lateral flow urine lipoarabinomannan assay (LF-LAM) (see Alternative test(s)) may be used to facilitate diagnosis of TB (WHO 2015). The WHO further recommends the following: "Individuals presumed of having extrapulmonary TB but who have had a single negative result from Xpert should undergo further diagnostic testing, and those for whom there is a high clinical suspicion for TB (especially children) should be treated even if an Xpert result is negative or if the test is not available" (WHO 2013). The downstream consequences of Xpert testing include the following.

- True-positive (TP): patients would benefit from rapid diagnosis and appropriate treatment.
- True-negative (TN): patients would be spared unnecessary treatment and would benefit from reassurance and pursuit of an alternative diagnosis.
- False-positive (FP): patients would likely experience anxiety and morbidity caused by additional testing, unnecessary treatment, and possible adverse effects; possible stigma associated with a TB or MDR-TB diagnosis; and the chance that a false-positive may halt further diagnostic evaluation.
- False-negative (FN): increased risk of morbidity and mortality and delayed treatment initiation for patients.

Alternative test(s)

For a comprehensive review of new tests not yet in widespread use, we refer the reader to Unitaid 2017.

Smear microscopy (light microscopy (Ziehl-Neelsen), fluorescence microscopy, or light-emitting diode (LED) fluorescence microscopy) is the examination of smears for acid-fast bacilli (TB bacteria) under a microscope. Around 5000 to 10,000 organisms per mL must be present in the specimen for TB bacteria to be visible by microscopy (American Thoracic Society 2000). For extrapulmonary TB, microscopy can be performed in fluid or tissue specimens from sites of disease involvement, for example, in CSF in presumptive TB meningitis or in lymph node tissue in presumptive lymph node TB. For most extrapulmonary sites, because there are usually few organisms, the sensitivity of smear microscopy is generally low. Ranges from studies, some with selected cases, are quoted here: 0% to 10% in pleural fluid; 14% to 39% in pleural tissue; 2% to 30% in CSF; < 5% in peritoneal fluid; and 0% to 42% in pericardial fluid. In contrast, the specificity of smear microscopy tends to be quite high, as can be seen in pulmonary TB (≥ 90%) (Kilpatrick 1986; Lewinsohn 2017).

Mycobacterial culture is a method used to grow bacteria on nutrient-rich media. In comparison with microscopy, a positive culture requires only around 100 organisms per mL and therefore can detect lower numbers of TB bacteria (American Thoracic Society 2000). Additionally, culture is essential for species identification and DST (Van Deun 2004). However, culture takes several weeks and requires a highly equipped laboratory. Culture has reduced sensitivity in paucibacillary disease (reference standards have included culture from a different specimen, such as sputum, smear microscopy, NAA tests, presence of granulomatous inflammation, clinical criteria, imaging studies, and response to anti-TB therapy, done alone or in various combinations): CSF 45% to 70%; pleural fluid 23% to 58%; urine 80% to 90%; peritoneal TB 45% to 69%; pericardial TB 50% to 65% (Lewinsohn 2017); lymph node TB (excisional biopsy) 18% to 93%; and lymph node TB (fine-needle aspirate) 10% to 67% (Fontanilla 2011). Culture is the main reference standard against which the index test was measured in this review.

Histological examination involves examination of tissue specimens under a microscope. Diagnosis of extrapulmonary TB by histological examination is based on finding acid-fast bacilli and granulomatous inflammation, frequently with caseous (cheese-like) necrosis (necrotizing granulomas). The sensitivity of histology has been reported to vary for different forms of extrapulmonary TB (reference standards have included smear microscopy, culture, NAA tests, clinical criteria, and imaging studies, done alone or in various combinations): 59% to 88% for lymph node TB (excisional biopsy) (Fontanilla 2011); 69% to 97% in pleural tissue (closed pleural biopsy); 86% to 94% in urological tissue; 60% to 70% in endometrial curettage; 79% to 100% in peritoneal biopsy; and 73% to 100% in pericardial tissue (Lewinsohn 2017). Sensitivity has also been observed to vary for different diagnostic techniques. Diacon 2003 found thoracoscopy to be more sensitive (sensitivity of 100%) than closed needle biopsy (sensitivity of 66%) for establishing a diagnosis of pleural TB (reference standards have included microscopy smear, culture, or presence of granulomatous inflammation with caseous necrosis). Specificity has been observed to be low because of the presence of granulomas in other diseases, both infectious and non-infectious (Lewinsohn 2017), although the presence of 'necrotizing' granulomatous inflammation increases specificity (Woodard 1982). Histological examination carries the additional concern that invasive procedures that are complex and costly may be required to obtain the necessary specimens (Golden 2005).

Cytopathological examination of fluid specimens (such as pleural and peritoneal fluid) may be performed, first to exclude cancer, and then to obtain material for additional analyses, such as measurement of levels of adenosine deaminase and free interferon-gamma (IFN-γ) and cell counts (Lewinsohn 2017; Wright 2009a). Advantages of these tests include that they are rapid and simple and can be performed in most clinical laboratories (Dinnes 2007). In pleural, pericardial, and peritoneal fluid, a predominance of lymphocytes, especially in the absence of mesothelial cells,

is highly suggestive of TB (Wright 2009a). However, in HIV-positive people, this pattern may not be observed (Wright 2009a). Adenosine deaminase, an enzyme involved in purine metabolism, has been extensively studied for its potential role in the diagnosis of pleural TB, peritoneal TB, and TB meningitis (Lewinsohn 2017). IFN-γ is released after it is sensitized by T cells in response to specific *M. tuberculosis* antigens. A recent review of the evidence using GRADE provides the following recommendations.

- "...cell counts and chemistries be performed on amenable fluid specimens (including include pleural, cerebrospinal, ascitic, and joint fluid) collected from sites of suspected extrapulmonary TB (conditional recommendation, very low-quality evidence).
- ...adenosine deaminase levels be measured, rather than not measured, on fluid collected from patients with suspected pleural TB, TB meningitis, peritoneal TB, or pericardial TB (conditional recommendation, low-quality evidence).
- ...free IFN-γ levels be measured, rather than not measured, on fluid collected from patients with suspected pleural TB or peritoneal TB (conditional recommendation, low-quality evidence)" (Lewinsohn 2017).

NAA test is a molecular technique that can detect small quantities of genetic material (DNA or RNA) from micro-organisms, such as *M. tuberculosis.* The key advantage of NAA tests is that they are rapid diagnostic tests, potentially providing results in a few hours. This is a particularly important feature of the test in life-threatening forms of extrapulmonary TB, such as TB meningitis. A variety of molecular amplification methods are available, of which PCR is the most common. NAA tests are available as commercial kits and in-house tests (based on a protocol developed in a laboratory) and are used routinely in high-income countries for TB detection. In-house PCR is widely used in low-income countries because these tests are less expensive than commercial kits. An older editorial summarizing three systematic reviews (140 studies) of commercial and in-house NAA tests (other than Xpert) for different forms of extrapulmonary TB found relatively low sensitivity and underscored concerns about the cost and feasibility of this technology in resource-limited areas (Pai 2008). Similarly, another systematic review found that NAA tests have relatively low sensitivity for extrapulmonary TB but high specificity (e.g. for TB meningitis, for pleural TB), indicating that these tests cannot be used reliably to rule out TB (Dinnes 2007). A recent evidence synthesis reported sensitivities of 72% to 88% in lymph node tissue, 28% to 81% in pleural fluid, 90% in pleural tissue, and 31% to 56% in CSF. Specificity ranged from 90% to 100% (Lewinsohn 2017).

GenoType MTBDR*plus* (Hain Lifescience, Nehren, Germany) is a commercial NAA test that belongs to a category of molecular tests called 'line probe assay'. MTBDR*plus* detects the presence of mutations associated with drug resistance to isoniazid and rifampicin (Nathavitharana 2017). The WHO recommends that MTBDR*plus* should be used for cultured isolates of *M. tuberculosis* from both pulmonary and extrapulmonary sites (WHO 2016b).

LF-LAM (Alere Determine™ TB LAM Ag, Alere Inc, Waltham, USA) is a commercially available point-of-care test for active TB (pulmonary and extrapulmonary TB). The test detects lipoarabinomannan (LAM), a component of the bacterial cell wall, which is present in some people with active TB. LF-LAM is performed by placing urine on one end of a test strip, with results appearing as a line (i.e. a band) on the strip if TB is present. The test is simple, requires no special equipment, and shows results in 25 minutes (Shah 2016b). Of note, the presence of LAM in the urine of HIV-positive adults undergoing treatment for TB has been found to be associated with increased risk of mortality (Gupta-Wright 2016). In randomized trials, use of LF-LAM in HIV-positive inpatients has been shown to reduce mortality (Gupta-Wright 2018; Peter 2016). Based in part on evidence from a Cochrane Review (Shah 2016b), the WHO recommends that LF-LAM should be used to assist in the diagnosis of TB in adult inpatients, specifically, "people living with HIV who have signs or symptoms of TB and a CD4 cell count less than or equal to 100 cells/μL, and people living with HIV who are 'seriously ill' regardless of CD4 count or if the CD4 count is unknown. This recommendation also applies to HIV-positive children with signs and symptoms of TB (pulmonary and/or extrapulmonary) based on the generalisation of data from adults while acknowledging very limited data and concern regarding low specificity of the LF-LAM assay in children" (WHO 2015). The WHO does not recommend LF-LAM for TB screening or diagnosis of active TB disease in most population groups (WHO 2015).

Rationale

Existing diagnostic tests for extrapulmonary TB are not sensitive enough or are invasive and costly. This Cochrane Review estimated sensitivity and specificity of Xpert for detection of extrapulmonary TB and rifampicin resistance. We are aware of six systematic reviews previously published on this topic: Chang 2012; Denkinger 2014; Li Y 2017; Maynard-Smith 2014; Penz 2015; Sehgal 2016 (Table 2). These reviews found different pooled accuracy estimates for different forms of extrapulmonary TB and noted several limitations, including the following: small number of samples for a given specimen type, incomplete information on HIV status, concerns about accuracy of the reference standards used, limited data for assessing the accuracy of Xpert for detection of rifampicin resistance, and considerable differences in the preparation of specimens for testing. Concerning the latter, the WHO has provided standard operating procedures for preparation of non-respiratory specimens for use with Xpert (WHO 2014a). This Cochrane Review updates the literature and provides an opportunity to address some of the noted limitations.

OBJECTIVES

To determine the diagnostic accuracy of Xpert a) for extrapulmonary TB by site of disease in people presumed to have extrapulmonary TB; and b) for rifampicin resistance in people presumed to have extrapulmonary TB.

Secondary objectives

- To investigate the effects of potential sources of heterogeneity on test accuracy across the included studies.

For extrapulmonary TB, covariates of interest were microscopy smear status, HIV status, anti-TB treatment, past history of TB, reference standard used to verify pleural TB, and prevalence of extrapulmonary TB (culture confirmed) in included studies. For CSF, we considered the presence of a concentration step and specimen volume. For tissue specimens, we considered whether the WHO standard operating procedure was followed.

In addition, for TB meningitis, pleural TB, and lymph node TB, we adjusted accuracy estimates by applying a latent class meta-

analysis model to account for the imperfect nature of culture as the reference standard.

For detection of rifampicin resistance, the covariate of interest in included studies was the prevalence of rifampicin resistance.

METHODS

Criteria for considering studies for this review

Types of studies

We included randomized controlled trials, cross-sectional studies, and observational cohort studies. We included primary studies that compared results of the index test with results of the reference standard and reported data from which we could extract TP, FP, FN, and TN. We excluded case-control studies and case reports. We used abstracts to identify published studies and included these when they met the inclusion criteria.

Participants

We included participants of all ages from all settings and countries who were thought to have extrapulmonary TB. We included non-respiratory specimens (such as lymph node aspirate or tissue, pleural fluid, and CSF), except as noted. We excluded sputum and other respiratory specimens, such as fluid obtained from bronchial alveolar lavage and tracheal aspiration. As we anticipated finding many studies, we set a bar to exclude smaller studies to reduce unnecessary work. Therefore, we required studies to provide data for at least five specimens for a form of extrapulmonary TB included in the review. We excluded studies that evaluated Xpert by aspiration of gastric fluid, as this specimen is used most often to investigate pulmonary TB in children. We also excluded stool specimens because TB bacteria may be swallowed and passed into stool as a marker of pulmonary TB. We excluded studies evaluating the use of Xpert to diagnose relapse of previously treated extrapulmonary TB, so as to avoid the selection bias that may arise by limiting to a group that is already at elevated risk of extrapulmonary TB. We attempted to identify studies that included patients who were not taking anti-TB drugs or had taken anti-TB drugs for less than seven days. For those studies that included some patients on TB drugs, we addressed this concern in a sensitivity analysis.

Index tests

The index tests were the Xpert assay and the Ultra assay. 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).

Indeterminate results for detection of extrapulmonary TB refer to 'invalid', 'error', or 'no result'. Indeterminate results for detection of rifampicin resistance refer to 'MTB detected; rifampicin resistance indeterminate'.

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). For extrapulmonary specimens, based on retrospective studies that enrolled selected participants, the WHO recommends that "trace calls should be considered to be true-positive results for use in clinical decisions and patient follow-up" (WHO 2017c). We summarized the findings for Xpert and Ultra separately.

Target conditions

The target condition was extrapulmonary TB. We included eight common forms and considered subcategories of the target condition as separate diagnostic classifications (CDC 2015; Sandgren 2013; Sharma 2004).

- TB meningitis.
- Pleural TB.
- Lymph node TB.
- Genitourinary TB.
- Bone or joint TB.
- Peritoneal TB.
- Pericardial TB.
- Disseminated TB.

Table 1 lists the forms of extrapulmonary TB and specimens used for diagnosis in the review. We excluded less common forms, such as cutaneous TB, ocular TB, female genital TB, and TB of the breast, ear, and paranasal sinuses (Sharma 2004).

Reference standards

Detection of all forms of extrapulmonary TB

The primary reference standard was solid or liquid mycobacterial culture.

- 'TB' was defined as a positive *M. tuberculosis* culture.
- 'Not TB' was defined as a negative *M. tuberculosis* culture.

For pleural TB, we also included a composite reference standard that defined a positive result as the presence of granulomatous inflammation *or* a positive culture. We found evidence to support including histopathological examination in the composite reference standard for pleural TB. Around 60% of patients undergoing pleural biopsy will show granulomatous inflammation (American Thoracic Society 2000). In a prospective cohort study of patients with clinical and radiological findings consistent with pleural TB, Conde 2003 found that histological examination of tissue obtained from pleural biopsy had a higher diagnostic yield (78%; 66/84) than that of culture (62%; 52/84). For other forms of TB, we decided against use of a composite reference standard owing to the differing definitions of the composite reference standards, difficulty involved in interpreting them, concern for bias (Schiller 2016), and difficulty and impracticality in obtaining biopsy specimens in some forms of extrapulmonary TB (e.g. pericardial TB).

Culture is considered the best reference standard for TB, and we calculated sensitivity and specificity by measuring the results of Xpert against those of culture. Both culture sensitivity and specificity are expected to be better than those of Xpert, and culture specificity is expected to be perfect. However, culture may lead to misclassification of some cases of extrapulmonary TB as 'not TB' owing to the paucibacillary nature of the disease. This means that culture may have low sensitivity for extrapulmonary TB overall and further that culture sensitivity may differ for different forms of extrapulmonary TB. This misclassification by culture may lead to biased estimates (overestimation or underestimation) of the diagnostic accuracy of Xpert. The extent of bias will depend on the frequency of errors by culture and the degree of correlation in errors by culture and Xpert because both culture and Xpert are likely to pick up cases with a higher bacterial load, and both are likely to miss cases with a lower bacterial load. Ignoring this dependence could lead to an overestimation of the sensitivity of Xpert.

- Effect of low sensitivity of culture on Xpert specificity: the low sensitivity of culture means that index test TPs may be misclassified as FPs when culture is used as the reference standard. Therefore, when Xpert is evaluated against culture, the number of FPs (classified as positive by the index test and negative by the reference test) may be increased and Xpert specificity may be underestimated.
- Effect of low sensitivity of culture on Xpert sensitivity: the low sensitivity of culture means that index test FNs may be misclassified as TNs when culture is used as the reference standard. Therefore, when Xpert is evaluated against culture, the number of FNs (classified as positive by the index test and negative by the reference test) may be decreased and Xpert sensitivity may be overestimated.

In an attempt to improve the estimation of diagnostic accuracy, we applied a latent class meta-analysis model to the three most commonly studied forms of extrapulmonary TB. We discuss this approach further in the Statistical analysis and data synthesis section.

Detection of rifampicin resistance

The reference standard was culture-based DST using solid or liquid media or MTBDR*plus* as recommended by the WHO (WHO 2012; WHO 2016b).

Search methods for identification of studies

We attempted to identify all relevant studies regardless of language or publication status (published, unpublished, in press, or ongoing). We monitored abstracts to see if these studies were published during the time we performed the review. We included only published studies in the review.

Electronic searches

We searched the following databases up to 7 August 2017 using the search terms and strategy described in Appendix 1: Cochrane Infectious Diseases Group Specialized Register; MEDLINE (OVID, from 1966); Embase (OVID, from 1974); Science Citation Index - Expanded (from 1900), Conference Proceedings Citation Index - Science (CPCI-S, from 1990), and BIOSIS Previews (from 1926), all three from the Web of Science; Scopus (Elsevier, from 1970); and Latin American Caribbean Health Sciences Literature (LILACS) (BIREME, from 1982). We also searched ClinicalTrials.gov, the WHO International Clinical Trials Registry (ICTRP) Platform (www.who.int/trialsearch), and the International Standard Randomized Controlled Trials Number (ISRCTN) registry (www.isrctn.com/) for trials in progress, and ProQuest Dissertations & Theses A&I (1990 to 7 August 2017) for dissertations.

Searching other resources

We reviewed reference lists of included articles and any relevant review articles identified through the above methods. We contacted the test manufacturer (Cepheid Inc.) to identify unpublished studies. We also contacted researchers at FIND, members of the Stop TB Partnership's New Diagnostics Working Group, and other experts in the field of TB diagnostics for information on ongoing and unpublished studies.

Data collection and analysis

Selection of studies

We used Covidence to manage the selection of studies (Covidence 2017). Two review authors independently scrutinized titles and abstracts identified by electronic literature searching to identify potentially eligible studies. We selected any citation identified by either review author as potentially eligible for full-text review. The same review authors independently assessed full-text papers for study eligibility using predefined inclusion and exclusion criteria and resolved any discrepancies by discussion. We recorded all studies excluded after full-text assessment and their reasons for exclusion in the Characteristics of excluded studies table. We illustrated the study selection process in a PRISMA diagram.

Data extraction and management

Two review authors piloted a data extraction form with five studies and, based on the pilot, finalized the form (Appendix 2). Next, two review authors worked independently to extract data on the following characteristics.

- Author; publication year; country; setting (outpatient, inpatient, or both outpatient and inpatient); study design; manner of participant selection; number of participants enrolled; number of participants for whom results are available.
- Characteristics of participants: gender; age; HIV status; history of TB; receipt of anti-TB treatment.
- Index test.
- Target condition and subcategories.
- Reference standard.
- Quality Assessment of Studies of Diagnostic Accuracy - Revised (QUADAS-2) items.
- Details of specimen: type (such as CSF, pleural fluid, lymph node aspirate or tissue); condition (fresh or frozen); smear-positive or smear-negative.
- Specimen preparation; homogenization step (for tissue specimens); concentration step and specimen volume (for CSF); adherence to WHO standard operating procedures.
- Number of TP, FP, FN, and TN (i.e. true-positives, false-positives, false-negatives, and true-negatives, with respect to culture); number of indeterminate results for detection of extrapulmonary TB; number of indeterminate results for detection of rifampicin resistance.
- Number of missing or unavailable test results.

We classified country income status as either low- and middle-income or high-income, according to the World Bank List of Economies (World Bank 2017).

We extracted TP, FP, FN, and TN values for the following specimens: CSF, pleural fluid and tissue, lymph node aspirate and tissue (the latter specimen acquired by surgical biopsy), bone or joint fluid and tissue, urine, peritoneal fluid and tissue, pericardial fluid and tissue, and blood. We extracted these values for each of the specimen types separately. For example, we used one 2 × 2 table for lymph node aspirate, and another 2 × 2 table for lymph node tissue. In situations in which a participant contributed more than one specimen but of different types, we extracted data for all specimens. When a study included data for both raw specimens and concentrated sediment involving the same participants, we preferentially extracted data for raw specimens, except in the case of CSF, for which we extracted data for concentrated sediment as recommended by the WHO (WHO 2014a). We extracted accuracy data according to the defined reference standard, which was an inclusion criterion for the Review (see Reference standards). We did not encounter any situations in which a subset of participants in a study received the reference standard but others did not. Hence, there was no need to make corrections for verification bias in the statistical analysis (Begg 1983).

In most studies, the number of specimens was the same as the number of participants. However, in some studies, the number of specimens exceeded the number of participants or study authors reported only the number of specimens. Hence the unit of analysis in this review should be considered "specimen". We added post hoc a sensitivity analysis limiting inclusion to studies that included one specimen per participant.

We contacted authors of primary studies for missing data or clarifications. We entered all data into Microsoft Excel 2014.

As recommended for reporting of systematic reviews of diagnostic test accuracy, we extracted information on manufacturers' involvement and funding (McGrath 2017). This information included donation of the index test; financial support for non-test-related study costs; and design, analysis, or production of the manuscript.

Assessment of methodological quality

We used the QUADAS-2 tool, tailored to this review, to assess the quality of the included studies (Appendix 3) (Whiting 2011). QUADAS-2 consists of four domains: patient selection, index test, reference standard, and flow and timing. We assessed all domains for the potential for risk of bias and the first three domains for concerns regarding applicability. Two review authors independently completed QUADAS-2 and resolved disagreements through discussion. We present the results of this quality assessment in Review text, tables, and graphs.

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

Statistical analysis and data synthesis

We performed descriptive analyses of the characteristics of included studies using Stata 12 (Stata 2011), and we presented key study characteristics in the Characteristics of included studies table. We used data reported in the TP, FP, FN, and TN format to calculate sensitivity and specificity estimates and 95% confidence intervals (CIs) for individual studies and presented individual study results graphically by plotting the estimates of sensitivity and specificity (and their 95% CIs) in forest plots and receiver operating characteristic (ROC) space using Review Manager 5 (RevMan 5) (RevMan 2014).

When data were sufficient, we performed meta-analyses to estimate pooled sensitivity and specificity and corresponding 95% credible (CrI, defined below) and prediction intervals using an adaptation of the bivariate random-effects approach of Reitsma and colleagues (Reitsma 2005), which uses the exact binomial likelihood for the observed proportions (Chu 2006). The bivariate random-effects approach allowed us to calculate the pooled estimates of sensitivity and specificity while dealing with potential sources of variation caused by (1) imprecision of sensitivity and specificity estimates within individual studies; (2) correlation between sensitivity and specificity across studies; and (3) variation in sensitivity and specificity between studies. The model has a hierarchical structure, with the logit sensitivity in individual studies assumed to come from a common probability distribution whose mean is the pooled logit sensitivity, and whose standard deviation is the between-study standard deviation, and likewise for the specificity. This structure allows for borrowing strength across studies. In the absence of sufficient studies, we simply presented descriptive statistics.

We performed separate analyses grouped by type of extrapulmonary specimen (e.g. CSF, pleural fluid, peritoneal fluid) rather than determine summary accuracy estimates for all forms of extrapulmonary TB combined, because we considered the former approach to be most clinically meaningful. We performed additional analyses for three forms of extrapulmonary TB: lymph node and pleural TB - these being two of the most common forms - and TB meningitis - although less common, this form has high mortality. For analysis of Xpert accuracy for rifampicin resistance detection, we included patients who (1) were culture-positive; (2) had a valid phenotypic DST (or MTBDR*plus)* result; (3) were Xpert TB-positive; and (4) had a valid Xpert Rif result.

- Sensitivity = Xpert Rif resistant/DST Rif resistant.
- Specificity = Xpert Rif susceptible/DST Rif susceptible.

For detection of rifampicin resistance, when a study included multiple types of specimens, we based our determination of Xpert sensitivity and specificity on all available data in the study, including data for specimens that we did not include in the primary analyses for detection of extrapulmonary TB. For example, if a study provided data for several specimen types combined (e.g. all tissue specimens) and we could not disaggregate the data for a specific specimen type, we included all data (for all tissue specimens) in the analysis for rifampicin resistance detection. We did this because we did not expect the accuracy of Xpert for rifampicin resistance to vary by specimen type. In addition, for detection of rifampicin resistance, we performed univariate meta-analyses (using all available data) to determine sensitivity and specificity estimates separately. We did this because in many

studies, all participants were rifampicin susceptible (rifampicin resistance-negatives), thus contributing data for specificity but not for sensitivity. We also performed a sensitivity analysis using the bivariate random-effects model for the subset of studies that provided data for both sensitivity and specificity.

Culture-negative specimens found to be Xpert-positive for rifampicin resistance have rarely been described in the literature (Boyles 2014; Kelly 2014). When reported in the included studies, we extracted and included this information in the Findings and Discussion sections of the review.

We estimated all models using a Bayesian approach with low-information prior distributions using OpenBUGS software (Version 3.2.3) (Lunn 2009), along with R (Version 3.3.2) (R Core Team 2016). Under the Bayesian approach, all unknown parameters must be provided a prior distribution that defines the range of possible values of the parameter and the weight of each of those values, based on information external to the data. To allow observed data to dominate the final results, we chose to use low-information prior distributions. We defined prior distributions on the log-odds scale over the pooled sensitivity and specificity parameters, their corresponding between-study standard deviations, and the correlation between the sensitivities and specificities across studies. For the pooled log odds of the sensitivity or the pooled log odds of the specificity, we used a normal prior distribution with mean 0 and a wide variance of 4 (or a precision of 0.25). This corresponds to a roughly uniform distribution over the pooled sensitivity and pooled specificity on the probability scale. For the between-study precision, we used a gamma distribution with a shape parameter of 2 and a rate parameter of 0.5. This corresponds to a 95% prior credible interval (CrI) for the between-study standard deviation in the log odds of sensitivity or the log odds of specificity ranging from roughly 0.29 to 1.44, corresponding to moderate to high values of between-study heterogeneity. Covariance terms followed a uniform prior distribution whose upper and lower limits were determined by the sensitivity of the two tests. The OpenBUGS model used appears in Appendix 4. It is known that meta-analysis models can be sensitive to the choice of prior distributions over between-study standard deviation parameters. Therefore, we carried out sensitivity analyses and considered alternative prior distributions that are less informative, allowing a wider range of possible values. To study the sensitivity of all results to the choice of prior distributions given above, we considered alternative prior distributions that were less informative, allowing a wider range of possible values. We increased the variance of the normal distributions over the pooled log odds of sensitivity or specificity to 100. We used a uniform prior distribution ranging from 0 to 3 over the between-study standard deviation on the log odds scale (see programme in Appendix 4). We noted no appreciable change in pooled accuracy parameters but found that the posterior CrIs and prediction intervals were slightly wider, as expected.

We combined information from the prior distribution with the likelihood of the observed data, in accordance with Bayes' theorem, using the OpenBUGS programme, which provides a sample from the posterior distribution of each unknown parameter. We were particularly interested in the pooled sensitivity and specificity of Xpert and between-study variance in the sensitivity and specificity of Xpert on the log-odds scale. Using a sample from the posterior distribution, we calculated various descriptive statistics of interest. We estimated the median pooled sensitivity and specificity and their 95% CrI. The median or the 50% quantile is the value below which 50% of the posterior sample lies. We report the median because the posterior distributions of some parameters may be skewed and the median would be considered a better point estimate of the unknown parameter than the mean in such cases. The 95% CrI is the Bayesian equivalent of the classical (frequentist) 95% CI (we will indicate 95% CI for individual study estimates and 95% CrI for pooled study estimates as appropriate). The 95% CrI may be interpreted as an interval that has a 95% probability of capturing the true value of the unknown parameter, given observed data and prior information. We prepared summary receiver operating characteristic (SROC) curves for each meta-analysis model using the methods described in Harbord 2007.

We also determined the predicted sensitivity and specificity of Xpert and their 95% CrIs. Predicted values represent our best guess for sensitivity and specificity in a future study and will be close to the pooled estimates. However, their CrIs may be different. If there is no heterogeneity at all between studies, the CrI around the predicted estimate will be the same as the CrI around the pooled estimate. On the other hand, if considerable heterogeneity is observed between studies, the CrI around the predicted estimate will be much wider than the CI around the pooled estimate.

In addition, in a secondary analysis for three forms of extrapulmonary TB - TB meningitis (CSF), pleural TB (pleural fluid), and lymph node TB (lymph node aspirate) - we adjusted accuracy estimates by applying a latent class meta-analysis model to account for the imperfect nature of culture as the reference standard (Chu 2009; Dendukuri 2012).

Latent class analysis is a statistical modelling technique that allows estimation of test accuracy in the absence of an adequate reference standard to define the presence or absence of disease (Van Smeden 2014). The latent class meta-analysis model expanded the traditional meta-analysis model in two ways: (1) we added parameters for the sensitivity and specificity of culture; and (2) we added covariance terms to adjust for the dependence between Xpert and culture among disease-positive and disease-negative participants in each study. We used hierarchical prior distributions over the logit sensitivity and logit specificity of culture. In other words, we assumed that the logit sensitivities in the individual studies come from a common probability distribution whose mean is the pooled mean logit sensitivity of culture and whose standard deviation is the between-study standard deviation. Likewise for the specificities. We used the same low-information prior distributions over the pooled logit mean and between-study standard deviation parameters as we had for the corresponding parameters for the Xpert test. We used uniform prior distributions for covariance terms over their ranges, which are determined by the sensitivities and the specificities of the two tests in each study (see Appendix 4 for the OpenBUGS model). We found that we did not need to augment observed data with prior information from other sources for most models. However, in a post hoc analysis of lymph node aspirate in which we suspected a systematic bias in the performance of culture, we used informative prior distributions over the specificity of culture (ranging from 99% to 100%) and the specificity of Xpert (ranging from 98% to 100%) (see Appendix 4). We added the SROC plots of the three latent class meta-analyses to the SROC plots resulting from the models in which culture was treated as a perfect test, so they could be compared.

Based on recent work evaluating Xpert for childhood TB (Schumacher 2016), we anticipated that latent class meta-analyses would lead to a decrease in the estimated pooled sensitivity of Xpert and an increase in the estimated pooled specificity of Xpert compared with the primary analyses. In other words, this method should help to correct the biases in Xpert sensitivity and specificity resulting from treating culture as a perfect reference standard, which we detailed earlier in the section on the reference standard.

Approach to indeterminate index test results

Xpert reports an indeterminate test result for unexpected results with any of the internal control measures of the assay. The indeterminate rate for detection of extrapulmonary TB was 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 was the number of tests classified as "MTB detected; Rif resistance INDETERMINATE" divided by the total number of Xpert-positive results. As we found very few indeterminate results reported, we excluded these results from the quantitative analysis. We used a Bayesian hierarchical model for a single proportion to estimate the pooled proportion of uninterpretable Xpert results.

Investigations of heterogeneity

Initially, we investigated heterogeneity through visual examination of forest plots of sensitivities and specificities and through visual examination of the ROC space of the raw data. We assessed heterogeneity through meta-regression modelling. We included the prevalence of extrapulmonary TB (confirmed by culture) as a covariate because changes in disease prevalence have often been found to be associated with other important changes, such as changes in the disease spectrum, which may affect diagnostic accuracy estimates (Leeflang 2013). We planned to include the following categorical covariates in the model, one at a time.

- Smear status.
- HIV status.
- Prior history of TB.
- For TB meningitis, concentration step used for preparing specimen (yes or no).
- CSF specimen volume used for Xpert testing.
- For pleural TB, culture reference standard versus composite reference standard.
- Prevalence of extrapulmonary TB, defined as the percentage of TB confirmed by culture in the study.
- Prevalence of rifampicin resistance, defined as the percentage of rifampicin resistance confirmed by the reference standard in the study.

However, we had insufficient data to investigate smear status, prior history of TB, and whether WHO standard procedures for preparing tissue specimens were followed.

For analyses involving the prevalence of extrapulmonary TB and rifampicin resistance, we compared the sensitivity or specificity between groups of interest by calculating the difference between groups together with a 95% CrI. We also calculated the probability that the difference was greater than zero.

Sensitivity analyses

For Xpert testing in CSF, pleural fluid, and lymph node aspirate, we performed sensitivity analyses to explore the contributions of risk of bias and patient characteristics on Xpert accuracy by limiting inclusion in the meta-analysis to the following.

- Studies that used consecutive or random selection of participants.
- Studies in which the reference standard results were interpreted without knowledge of the index test results.
- Studies that included only untreated patients.
- Studies that included only one specimen per patient.
- For lymph node aspirate, studies that involved only adults.

Other analyses

Non-tuberculous mycobacteria (NTM), such as *M. avium* complex and *M. intracellulare*, constitute a multi-species group of human pathogens that are ubiquitous in water and soil. NTM can cause severe diseases that share clinical signs with TB but are treated differently. People infected with HIV with severe immunosuppression are particularly vulnerable to infections caused by NTM (Gopinath 2010). Although previous studies have shown that Xpert does not cross-react with other mycobacterial species (Blakemore 2010; Helb 2010), we thought it important to summarize data for NTM separately by determining the percentage of false-positive Xpert results in specimens that grew NTMs.

Assessment of reporting bias

We did not perform a formal assessment of publication bias using methods such as funnel plots or regression tests because such techniques have not been helpful for diagnostic test accuracy studies (Macaskill 2010).

Assessment of certainty of the evidence

Two review authors assessed the certainty of the evidence (also called quality of the evidence) using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach (Balshem 2011; GRADE 2013; Schünemann 2008), along with GRADEpro Guideline Development Tool (GDT) software (GRADEpro GDT 2015). In the context of a systematic review, ratings of the certainty of the evidence reflect the extent of our confidence that the estimates of effect (including test accuracy and associations) are correct. As recommended, we rated the certainty of the evidence as high (not downgraded), moderate (downgraded by one level), low (downgraded by two levels), or very low (downgraded by more than two levels) for five domains: risk of bias, indirectness, inconsistency, imprecision, and publication bias.

For each outcome, we considered the certainty of the evidence to begin as high when high-quality observational studies (cross-sectional or cohort studies) enrolled participants with diagnostic uncertainty. If we had a reason for downgrading, we used our judgement to classify the reason as serious (downgraded by one level) or very serious (downgraded by two levels). We summarized this information in the 'Summary of findings' tables (Schünemann 2011). As recommended, we determined the overall certainty of the evidence by using the lowest grade for any of the outcomes deemed critical (sensitivity and specificity) (Brozek 2009).

We applied GRADE in the following ways.

- Risk of bias: we used QUADAS-2 to assess risk of bias.
- Indirectness: we used QUADAS-2 for concerns of applicability and looked for important differences between the populations studied (e.g. patient characteristics, study setting) and the review questions.
- Inconsistency: GRADE recommends downgrading for unexplained inconsistency in sensitivity and specificity estimates. We carried out prespecified analyses to investigate potential sources of heterogeneity and did not downgrade when we believed we could explain inconsistency in the accuracy estimates.
- Imprecision: we considered a precise estimate to be one that would allow a clinically meaningful decision. We considered the width of the CrI and asked ourselves, "Would we make a different decision if the lower or upper boundary of the CrI represented the truth?" In addition, we worked out projected ranges for TP, FN, TN, and FP for a given prevalence of TB and made judgements on imprecision from these calculations.
- Publication bias: we rated publication bias as undetected (not serious) because of the comprehensiveness of the literature search and following extensive outreach to TB researchers to identify studies.

RESULTS

Results of the search

We identified 66 unique studies that met the inclusion criteria (Ablanedo-Terrazas 2014; Al-Ateah 2012; Arockiaraj 2017; Bahr 2015; Bahr 2017; Bera 2015; Bholla 2016; Biadglegne 2014; Blaich 2014; Causse 2011; Che 2017; Christopher 2013; Coetzee 2014; Dhasmana 2014; Dhooria 2016; Diallo 2016; Du 2015; Feasey 2013; Friedrich 2011; Ghariani 2015; Gu 2015; Gursoy 2016; Hanif 2011; Held 2014; Held 2016; Hillemann 2011; Ioannidis 2011; Iram 2015; Jing 2017; Kim 2015a; Li 2017; Ligthelm 2011; Lusiba 2014; Malbruny 2011; Massi 2017; Mazzola 2016; Meldau 2014; Nataraj 2016; Nhu 2014; Ozkutuk 2014; Pandey 2017; Pandie 2014; Patel 2013; Penata 2016; Pink 2016; Pohl 2016; Rufai 2015; Rufai 2017a; Rufai 2017b; Saeed 2017a; Safianowska 2012; Scott 2014; Sharma 2014; Sharma 2016; Solomons 2016; Suzana 2016; Tadesse 2015; Teo 2011; Tortoli 2012; Trajman 2014; Ullah 2017; Vadwai 2011; Van Rie 2013; Wang 2016a; Zeka 2011; Zmak 2013). Only one study evaluated Ultra; this study compared Ultra and Xpert for TB meningitis (Bahr 2017). All studies but four (one written in French - Diallo 2016, one in Spanish - Penata 2016, and two in Turkish - Gursoy 2016; Ozkutuk 2014), were written in English. Figure 2 shows the flow of studies in the review. We recorded the excluded studies and the reasons for their exclusion in the Characteristics of excluded studies table.

Figure 2. Study flow diagram. *See Table 3.

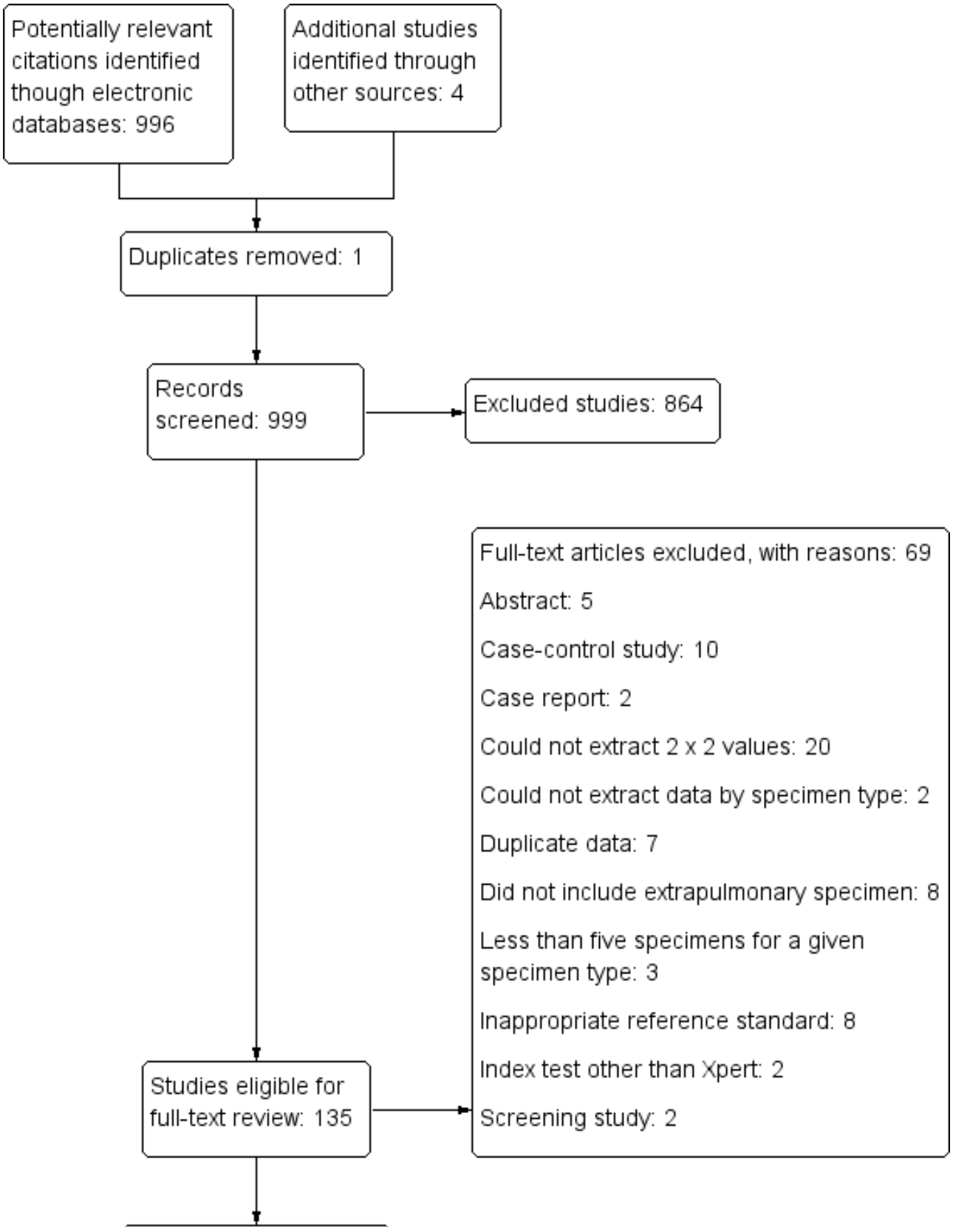
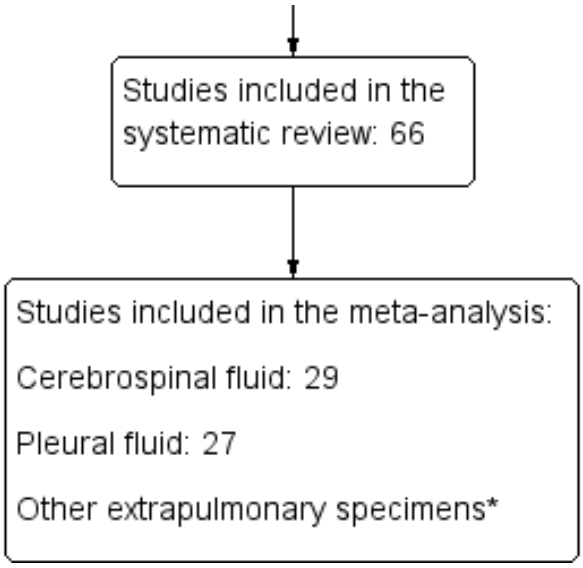


Figure 2. (Continued)



Methodological quality of included studies

Figure 3 and Figure 4 show risk of bias and applicability concerns for each of the 66 included studies. In the patient selection domain, we thought that 51 studies (77%) had low risk of bias, and six studies (9%) had high risk of bias for the following reasons: four studies selected participants by convenience (Bholla 2016; Ioannidis 2011; Malbruny 2011; Pandey 2017), and two studies had inappropriate exclusions (Saeed 2017a; Ullah 2017). We thought that nine studies (14%) had unclear risk of bias for the following reasons: the manner of patient selection was unclear - eight studies (Diallo 2016; Gu 2015; Li 2017; Massi 2017; Rufai 2015; Rufai 2017a; Rufai 2017b; Zmak 2013), and it was unclear whether the study avoided inappropriate exclusions - one study (Bera 2015). Regarding applicability (patient characteristics and setting), we thought that three studies (4%) had low concern because participants were evaluated in local hospitals or primary health settings (Bholla 2016; Pandie 2014; Trajman 2014); nine studies (14%) had high concern because participants were evaluated exclusively as inpatients at a tertiary care centre (Bahr 2015; Bahr 2017; Causse 2011; Che 2017; Du 2015; Feasey 2013; Gu 2015; Held 2014; Held 2016); and 54 studies had unclear concern because we could not tell the clinical setting.

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

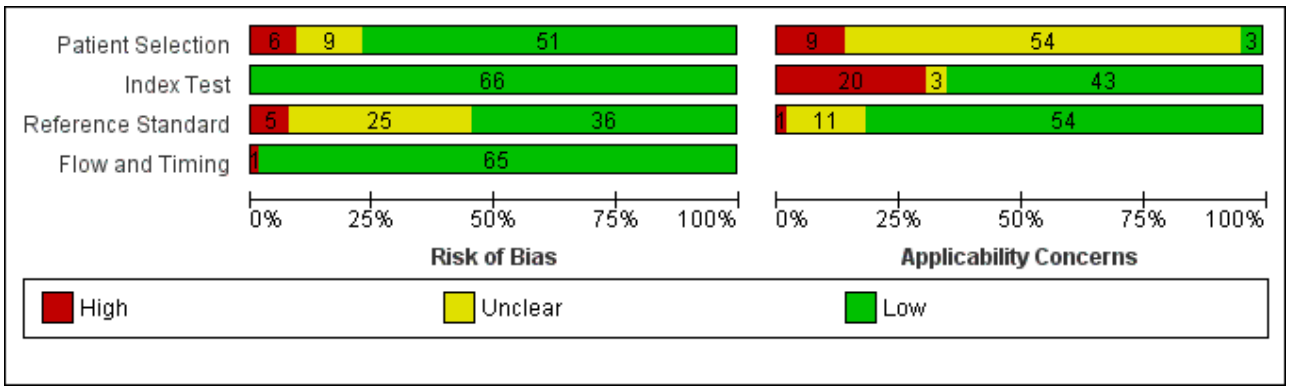


Figure 4. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.

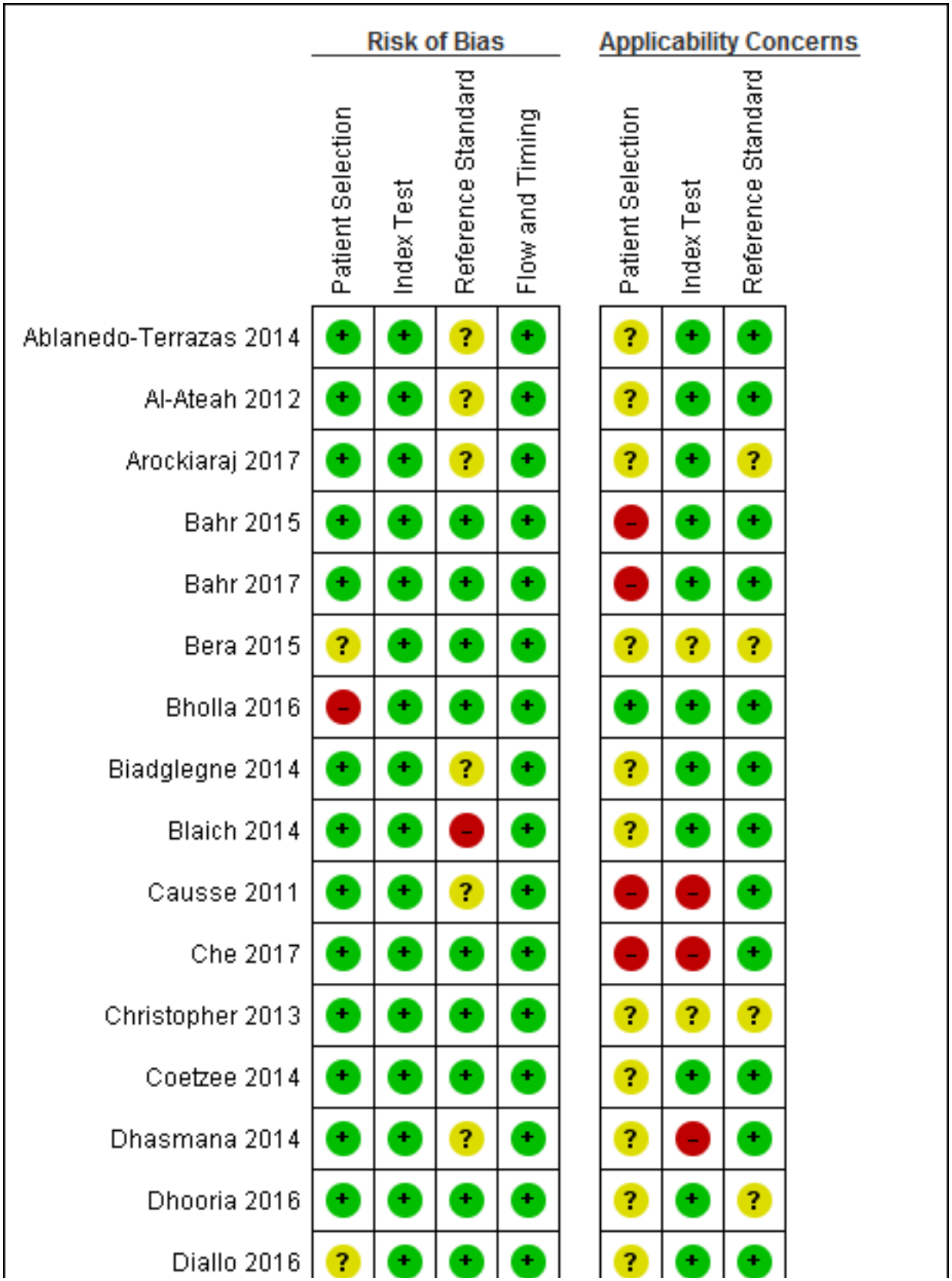


Figure 4. (Continued)

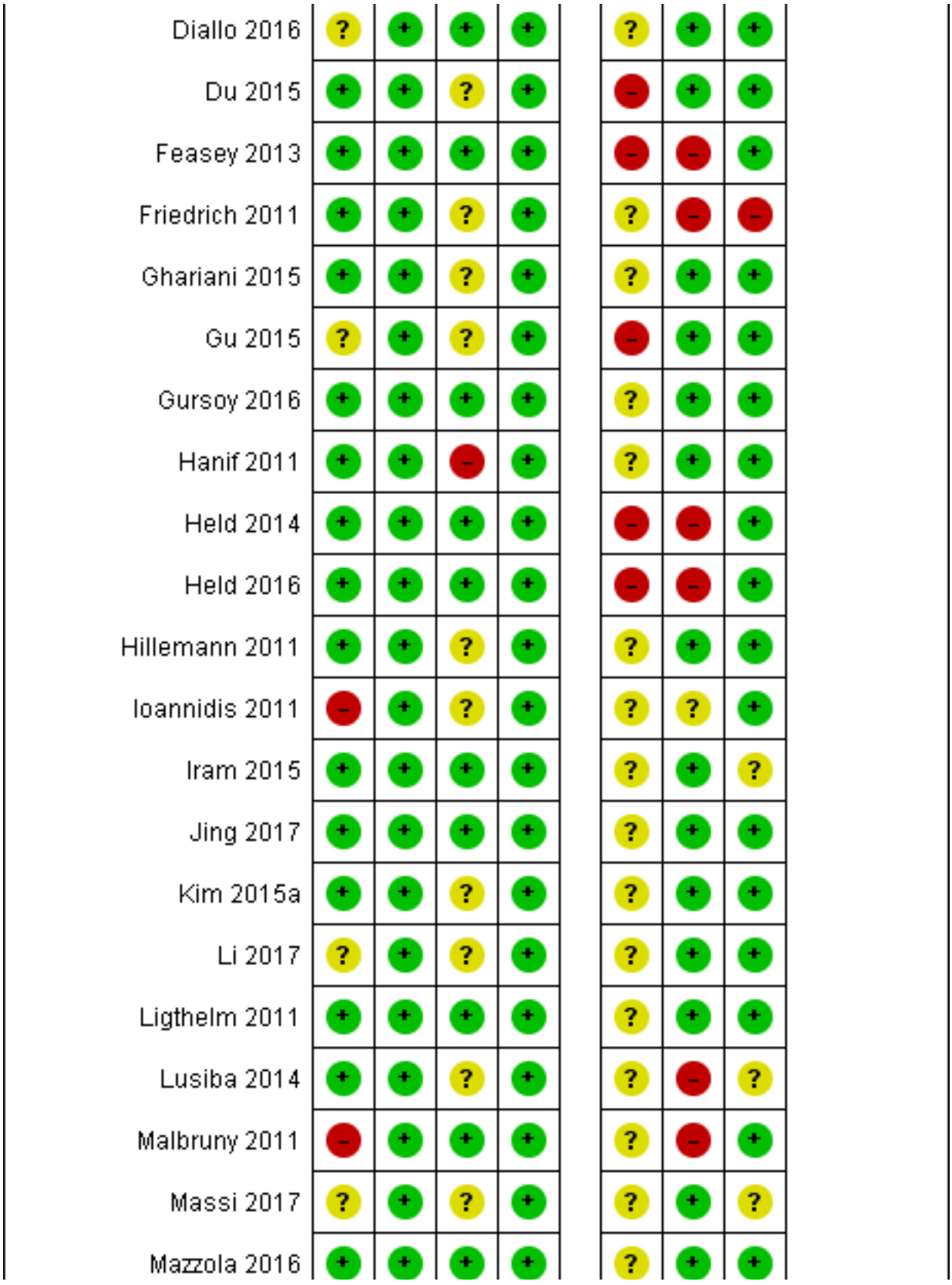


Figure 4. (Continued)

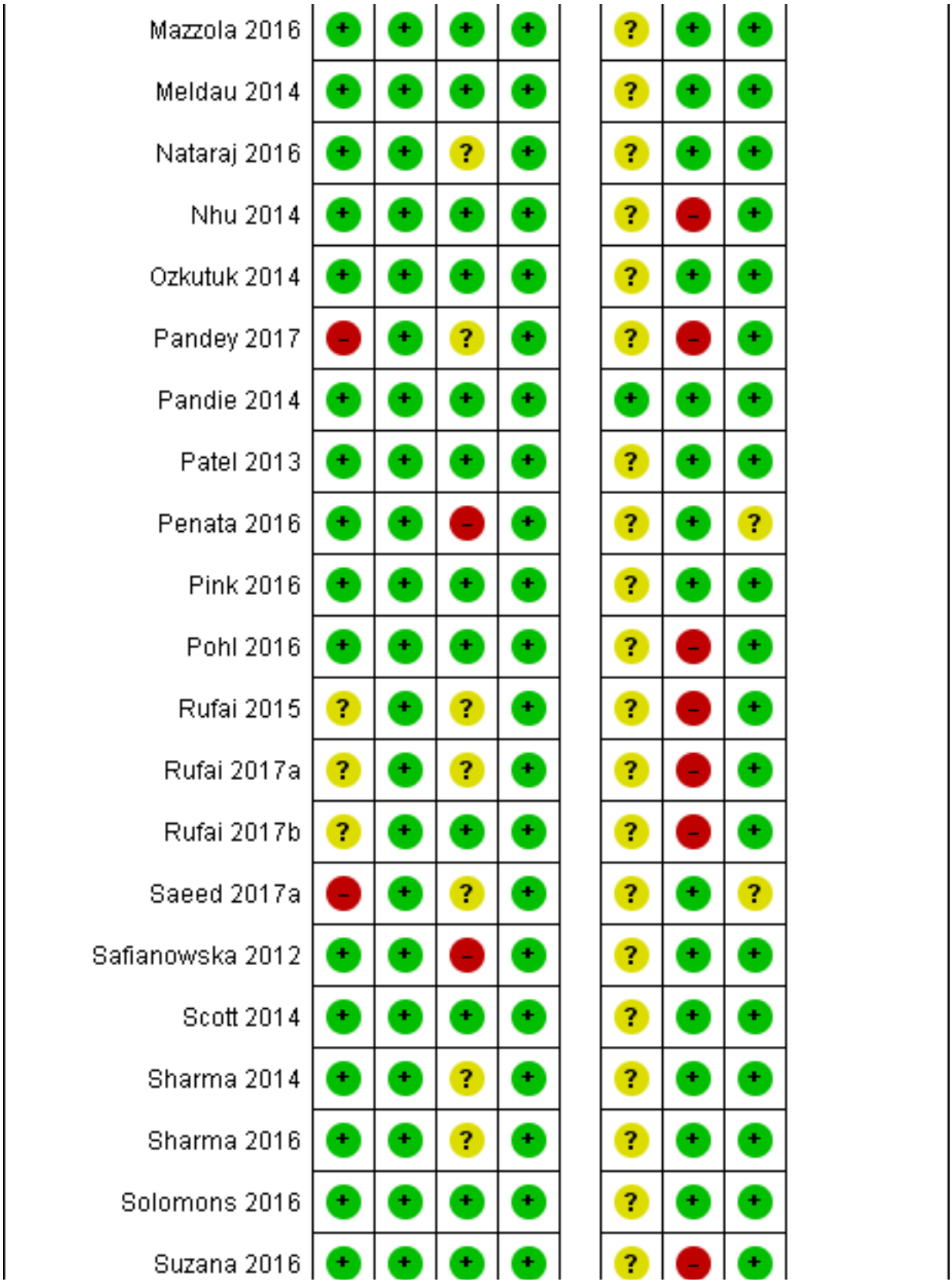
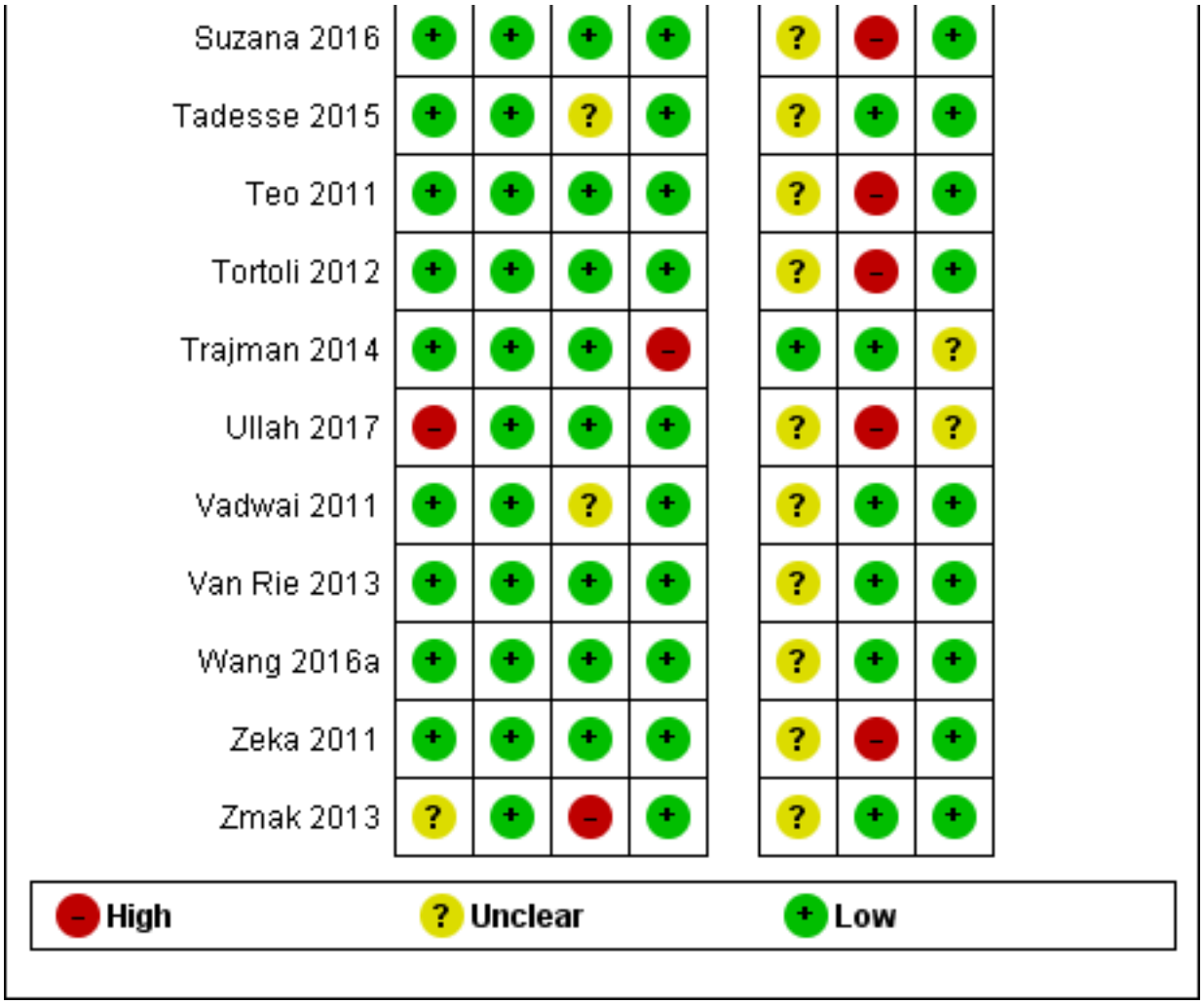


Figure 4. (Continued)



In the index test domain, we thought that all studies had low risk of bias because Xpert test results are automatically generated, the user is provided with printable test results, and the test threshold is prespecified. Regarding applicability, we thought that 42 studies (64%) had low concern because at least 75% of the specimen types in these studies were processed according to WHO recommendations, and 21 studies (32%) had high concern because less than 50% of the specimen types in these studies were processed according to WHO recommendations (Arockiaraj 2017; Causse 2011; Che 2017; Dhasmana 2014; Feasey 2013; Friedrich 2011; Held 2014; Held 2016; Lusiba 2014; Malbruny 2011; Nhu 2014; Pandey 2017; Pohl 2016; Rufai 2015; Rufai 2017a; Rufai 2017b; Suzana 2016; Teo 2011; Tortoli 2012; Ullah 2017; Zeka 2011). Three studies (5%) had unclear concern because the manner of specimen processing was not reported (Bera 2015; Ioannidis 2011), or only 50% of the specimen types were processed according to WHO recommendations (Christopher 2013).

In the reference standard domain, 36 studies (55%) had low risk of bias because results of the reference standard were interpreted without knowledge of results of the index test and only non-sterile specimens were decontaminated (Bahr 2015; Bahr 2017; Bera 2015; Bholla 2016; Che 2017; Christopher 2013; Coetzee 2014; Dhooria 2016; Diallo 2016; Feasey 2013; Gursoy 2016; Held 2014; Held 2016; Iram 2015; Jing 2017; Ligthelm 2011; Malbruny 2011; Mazzola 2016; Meldau 2014; Nhu 2014; Ozkutuk 2014; Pandie 2014; Patel 2013; Pink 2016; Pohl 2016; Rufai 2017b; Scott 2014; Solomons 2016; Suzana 2016; Teo 2011; Tortoli 2012; Trajman 2014; Ullah 2017; Van Rie 2013; Wang 2016a; Zeka 2011). Five studies (8%) had high risk of bias because results of the reference standard were interpreted with knowledge of results of the index test (Blaich 2014; Hanif 2011; Penata 2016; Safianowska 2012; Zmak 2013). Twenty-five studies (38%) had unclear risk of bias for the following reasons: two studies did not report whether there was blinding of the reference standard (Lusiba 2014; Saeed 2017a); 21 studies decontaminated specimens generally considered to be sterile (Al-Ateah 2012; Biadglegne 2014; Causse 2011; Dhasmana 2014; Du 2015; Friedrich 2011; Ghariani 2015; Gu 2015; Hillemann 2011; Ioannidis 2011; Kim 2015a; Li 2017; Massi 2017; Nataraj 2016; Pandey 2017; Rufai 2015; Rufai 2017a; Safianowska 2012; Sharma 2014; Tadesse 2015; Vadwai 2011); and two studies did not report blinding and decontaminated specimens generally considered to be sterile (Ablanedo-Terrazas 2014; Arockiaraj 2017).

Breaking this down by type of specimen, we found that before culture inoculation, four studies reported decontaminating CSF specimens (Kim 2015a; Li 2017; Nataraj 2016; Vadwai 2011); 10 studies reported decontaminating pleural fluid specimens (Al-Ateah 2012; Du 2015; Friedrich 2011; Ioannidis 2011; Kim 2015a; Li 2017; Nataraj 2016; Rufai 2015; Safianowska 2012; Vadwai 2011); and nine studies reported decontaminating lymph node aspirates (Al-Ateah 2012; Biadglegne 2014; Blaich 2014; Dhasmana 2014; Ghariani 2015; Nataraj 2016; Pandey 2017; Sharma 2014; Tadesse 2015). (Some studies are mentioned more than once because they evaluated more than one type of specimen.) We think decontamination of sterile specimens may have led to a decrease in viable TB bacteria and consequently false-negative cultures.

Regarding applicability of the reference standard, we thought that 54 studies (82%) had low concern because these studies performed a test to identify *M. tuberculosis* species (speciation). However, we thought that one study (2%) had high concern because this study did not do speciation (Friedrich 2011), and 11 studies (17%) had unclear concern because we could not tell whether the study performed speciation (Arockiaraj 2017; Bera 2015; Christopher 2013; Dhooria 2016; Iram 2015; Lusiba 2014; Massi 2017; Penata 2016; Saeed 2017a; Trajman 2014; Ullah 2017).

In the flow and timing domain, we considered almost all studies to have low risk of bias, noting that all participants were included in the analysis except in one study, which included less than 50% of eligible participants in the analysis (Trajman 2014).

We noted manufacturer involvement in five studies (8%), and this included the following.

- Donation of the index test (four studies; Hillemann 2011; Ioannidis 2011; Nhu 2014; Tortoli 2012).
- Involvement in manuscript design, analysis, or production (one study; Vadwai 2011).

We are also aware that studies located in low- and middle-income countries may have received index test cartridges at a reduced price. However, most studies did not report this information.

Findings

We included 66 unique studies that evaluated 16,213 specimens for detection of extrapulmonary TB and rifampicin resistance. Thirty-three studies (50%) included only one specimen type: TB meningitis (CSF) nine studies; pleural TB (fluid) six studies; lymph node TB (aspirate) eight studies; bone or joint TB five studies (fluid one study, tissue four studies); genitourinary TB (urine) zero studies; peritoneal TB (fluid) one study; pericardial TB two studies (fluid one study, tissue one study); and disseminated TB (blood) two studies. The remaining studies included different types of specimens in varying percentages. Fifty studies (76%) were conducted in low- or middle-income countries. Thirty studies (45%) included children in their study population; however, only five studies were conducted exclusively in children (Bholla 2016; Coetzee 2014; Held 2016; Pohl 2016; Solomons 2016). Forty-one studies (62%) reported the HIV status of participants. Of these, five studies exclusively or largely included HIV-positive participants (Ablanedo-Terrazas 2014; Bahr 2015; Bahr 2017; Feasey 2013; Van Rie 2013). In the remaining studies, the percentages of included HIV-positive patients ranged from 1% to 87%.

Fifty-eight studies (88%) evaluated fresh specimens, six studies (9%) evaluated only archived frozen samples (Patel 2013; Tadesse 2015; Tortoli 2012; Trajman 2014; Wang 2016a; Zeka 2011), and one study (2%) evaluated both fresh and frozen specimens (Malbruny 2011). Bahr 2017 compared Xpert in fresh specimens versus Ultra in frozen specimens. For the reference standard, seven studies (11%) used only solid culture, 29 studies (44%) used only liquid culture, and 30 studies (45%) used both solid and liquid cultures. Most studies performed Xpert and culture on the same specimen type, except two studies in which Xpert was performed on blood and culture was performed on sputum (Feasey 2013; Pohl 2016). Most studies did not report the precise number of cultures used to confirm a diagnosis of TB; however, it is likely that many studies used a single culture. We presented key characteristics of the included studies in the Characteristics of included studies table.

I. Detection of extrapulmonary TB

Table 3 presents pooled (summary) and predicted sensitivity and specificity results with respect to culture for all forms of extrapulmonary TB and specimen types included in the review.

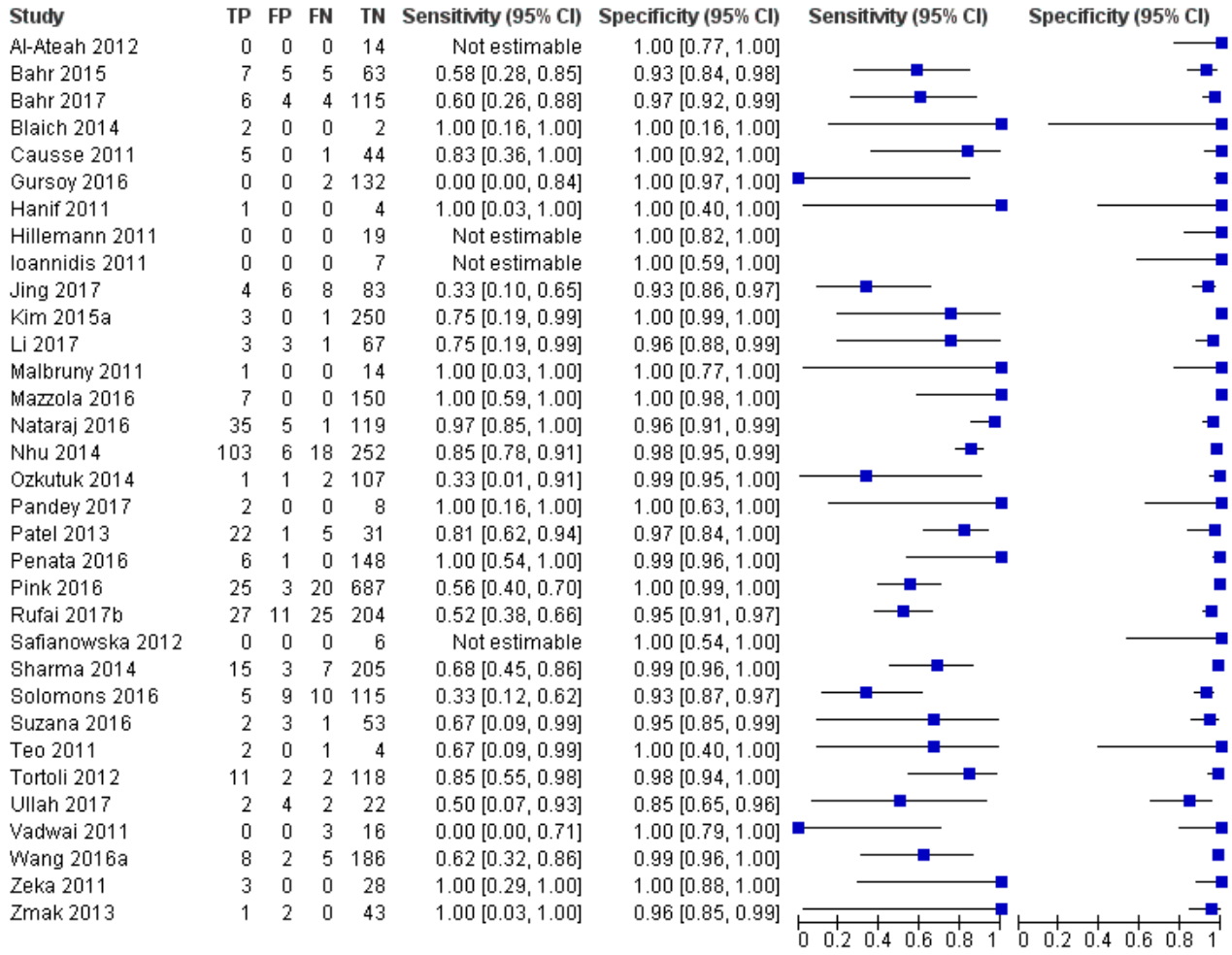
Xpert pooled sensitivity varied greatly by type of specimen, ranging from 50.9% (95% CrI 39.7 to 62.8) in pleural fluid to 97.2% (95% CrI 89.5 to 99.6) in bone or joint fluid. Pooled specificity ranged from 85.3% (58.7 to 96.4) in bone or joint tissue to 99.2% (98.2 to 99.7) in pleural fluid. In urine, pooled sensitivity and specificity were 82.7% (69.6 to 91.1) and 98.7% (94.8 to 99.7), respectively (13 studies, 1199 specimens).

A. Xpert testing in cerebrospinal fluid for TB meningitis

1. Primary analysis, Xpert

A total of 33 studies evaluated CSF specimens (Al-Ateah 2012; Bahr 2015; Bahr 2017; Blaich 2014; Causse 2011; Gursoy 2016; Hanif 2011; Hillemann 2011; Ioannidis 2011; Jing 2017; Kim 2015a; Li 2017; Malbruny 2011; Mazzola 2016; Nataraj 2016; Nhu 2014; Ozkutuk 2014; Pandey 2017; Patel 2013; Penata 2016; Pink 2016; Rufai 2017b; Safianowska 2012; Sharma 2014; Solomons 2016; Suzana 2016; Teo 2011; Tortoli 2012; Ullah 2017; Vadwai 2011; Wang 2016a; Zeka 2011; Zmak 2013). The median sample size (interquartile range (IQR)) was 74 (19 to 155) specimens. In individual studies, Xpert sensitivity ranged from 33% to 100% and specificity ranged from 93% to 100% (Figure 5). Pooled sensitivity and specificity (95% CrI) were 71.1% (60.9 to 80.4) and 98.0% (97.0 to 98.8), respectively (29 studies, 3774 specimens) (Table 3; Appendix 5).

Figure 5. Forest plots of Xpert® MTB/RIF sensitivity and specificity in cerebrospinal fluid. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive.



1.a. Primary analysis, Ultra

In a study on the treatment of HIV-associated cryptococcal meningitis in Uganda, Bahr 2017 compared the accuracy of Ultra and Xpert in 129 CSF specimens. Measured against culture as the reference standard, sensitivity was considerably higher with Ultra at 90% (95% CI 55 to 100) than with Xpert at 60% (95% CI 26 to 88). However, specificity was lower with Ultra at 90% (95% CI 83 to 95) versus Xpert at 97% (95% CI 92 to 99).

2. Investigations of heterogeneity

a. Xpert testing in HIV-positive and HIV-negative participants

We identified three studies that included mainly HIV-positive people (Bahr 2015; Bahr 2017; Patel 2013) and three studies that included mainly HIV-negative people (Hanif 2011; Jing 2017; Wang 2016a). In studies involving HIV-positive people, sensitivity ranged from 58% to 81% compared with 33% to 100% in studies involving HIV-negative people. In all studies, specificity was ≥ 93%.

b. Specimen concentration

We found that concentrating CSF improved both sensitivity and specificity. Pooled sensitivity in concentrated specimens was 74.8% (95% CrI 63.1 to 84.4) (15 studies, 2758 specimens) versus 66.2% (95% CrI 48.5 to 81.4) (12 studies, 905 specimens) in unconcentrated specimens. Pooled specificity in concentrated specimens was 98.3% (95% CrI 97.1 to 99.1) versus 97.7% (95% CrI 95.4 to 99.0) in unconcentrated specimens (Appendix 6).

c. Cerebrospinal fluid collection volumes

Five studies reported the volume of CSF collected for Xpert testing. Starting from the largest collection volume, at 7 mL, Nhu 2014 found the highest sensitivity of85%; at 6 mL, Bahr 2015 found sensitivity of 58%; at 6 mL, Bahr 2017 found sensitivity of 60%; at 3 mL, Patel 2013 found sensitivity of 81%; and at 2 mL, Rufai 2017b found the lowest sensitivity of 52%. Specificities in the five studies were ≥ 93% (Figure 5).

d. TB prevalence

See Table 4. The median prevalence of TB meningitis (as measured by culture positivity) in these studies was 10%. We found higher

Xpert sensitivity in settings with higher TB prevalence than in those with lower TB prevalence, with pooled sensitivity of 72.0% (95% CrI 59.7 to 82.8) versus 68.2% (95% CrI 50.9 to 82.4). We found lower specificity in settings with higher TB prevalence than in those with lower TB prevalence, with pooled specificity of 96.8% (95% CrI 95.0 to 98.2) versus 98.9% (95% CrI 97.9 to 99.4). In the case of specificity, accuracy in the two groups was significantly different (probability of specificity higher in low TB prevalence group = 0.008).

3. Sensitivity analysis

See Table 5. In comparison with all studies, studies that evaluated only one specimen per participant had lower pooled sensitivity at 63.5% (47.6 to 76.3) and lower pooled specificity at 96.1% (94.2 to 97.4). The other sensitivity analyses made little difference in any of these findings.

4. Indeterminate Xpert results

Fourteen studies (42%) reported the number of indeterminate Xpert results. Nine of these studies reported zero indeterminate results (Al-Ateah 2012; Bahr 2015; Blaich 2014; Causse 2011; Hanif 2011; Ioannidis 2011; Sharma 2014; Teo 2011; Zeka 2011). For CSF, of 2096 tests performed, the pooled percentage of indeterminate Xpert results was 0.9% (95% CrI 0.3 to 1.9).

5. Latent class meta-analysis

Based on the latent class meta-analysis model, Xpert pooled sensitivity and specificity (95% CrI) were 63.2% (53.8 to 73.6) and 99.6% (98.5 to 99.9), respectively (29 studies, 3774 specimens) (Table 6). Xpert pooled sensitivity was lower and pooled specificity higher than when culture was treated as having perfect accuracy. This analysis also provided accuracy estimates of culture. The pooled sensitivity of culture at 68.6% (59.0 to 78.0) was estimated to be lower than 100%, although it remained greater than that of Xpert. The pooled specificity of culture was estimated to be 99.3% (98.1 to 99.8) (Table 6). Appendix 5 shows the summary receiver operating characteristic (SROC) curves from the meta-analysis treating culture as a perfect reference standard and from the latent class meta-analysis. The latent class meta-analysis resulted in low heterogeneity in the specificity of Xpert across studies, as would be expected of an automated, commercial test. This was the result of adjustments for the imperfect and heterogeneous accuracy of culture across studies.

B. Xpert testing in pleural fluid for pleural TB

1. Primary analysis, culture reference standard

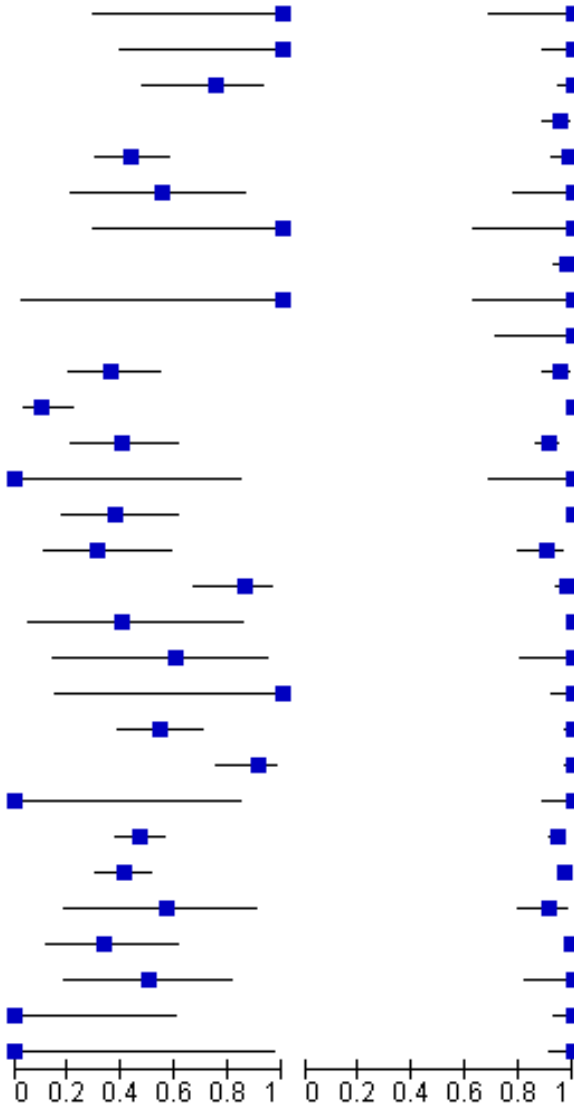
Thirty studies evaluated pleural fluid with respect to a culture reference standard (Al-Ateah 2012; Causse 2011; Che 2017; Christopher 2013; Du 2015; Friedrich 2011; Hanif 2011; Hillemann 2011; Ioannidis 2011; Iram 2015; Jing 2017; Kim 2015a; Li 2017; Malbruny 2011; Mazzola 2016; Meldau 2014; Nataraj 2016; Ozkutuk 2014; Pandey 2017; Penata 2016; Rufai 2015; Saeed 2017a; Safianowska 2012; Scott 2014; Sharma 2014; Suzana 2016; Tortoli 2012; Vadwai 2011; Zeka 2011; Zmak 2013). The median sample size (IQR) was 77 (30 to 166) specimens. In individual studies, Xpert sensitivity ranged from 0% to 100% and specificity ranged from 90% to 100% (Figure 6). Pooled sensitivity and specificity (95% CrI) against culture were 50.9% (39.7 to 62.8) and 99.2% (98.2 to 99.7), respectively (27 studies, 4006 specimens) (Table 3; Appendix 7).

Figure 6. Forest plots of Xpert® MTB/RIF sensitivity and specificity in pleural fluid with respect to a culture reference standard (upper plots) and a composite reference standard (lower plots). The squares represent the

sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive.

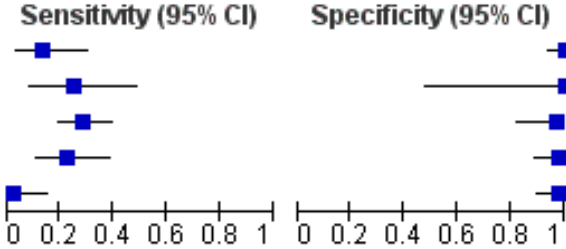
Pleural fluid, culture

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Al-Ateah 2012	3	0	0	10	1.00 [0.29, 1.00]	1.00 [0.69, 1.00]
Causse 2011	4	0	0	30	1.00 [0.40, 1.00]	1.00 [0.88, 1.00]
Che 2017	12	0	4	62	0.75 [0.48, 0.93]	1.00 [0.94, 1.00]
Christopher 2013	0	4	0	83	Not estimable	0.95 [0.89, 0.99]
Du 2015	24	1	31	70	0.44 [0.30, 0.58]	0.99 [0.92, 1.00]
Friedrich 2011	5	0	4	15	0.56 [0.21, 0.86]	1.00 [0.78, 1.00]
Hanif 2011	3	0	0	8	1.00 [0.29, 1.00]	1.00 [0.63, 1.00]
Hillemann 2011	0	2	0	103	Not estimable	0.98 [0.93, 1.00]
Ioannidis 2011	1	0	0	8	1.00 [0.03, 1.00]	1.00 [0.63, 1.00]
Iram 2015	0	0	0	11	Not estimable	1.00 [0.72, 1.00]
Jing 2017	12	4	21	87	0.36 [0.20, 0.55]	0.96 [0.89, 0.99]
Kim 2015a	5	0	44	339	0.10 [0.03, 0.22]	1.00 [0.99, 1.00]
Li 2017	10	18	15	178	0.40 [0.21, 0.61]	0.91 [0.86, 0.94]
Malbruny 2011	0	0	2	10	0.00 [0.00, 0.84]	1.00 [0.69, 1.00]
Mazzola 2016	8	0	13	693	0.38 [0.18, 0.62]	1.00 [0.99, 1.00]
Meldau 2014	5	6	11	54	0.31 [0.11, 0.59]	0.90 [0.79, 0.96]
Nataraj 2016	24	3	4	136	0.86 [0.67, 0.96]	0.98 [0.94, 1.00]
Ozkutuk 2014	2	0	3	227	0.40 [0.05, 0.85]	1.00 [0.98, 1.00]
Pandey 2017	3	0	2	17	0.60 [0.15, 0.95]	1.00 [0.80, 1.00]
Penata 2016	2	0	0	46	1.00 [0.16, 1.00]	1.00 [0.92, 1.00]
Rufai 2015	23	0	19	119	0.55 [0.39, 0.70]	1.00 [0.97, 1.00]
Saeed 2017a	30	0	3	125	0.91 [0.76, 0.98]	1.00 [0.97, 1.00]
Safianowska 2012	0	0	2	30	0.00 [0.00, 0.84]	1.00 [0.88, 1.00]
Scott 2014	59	21	66	336	0.47 [0.38, 0.56]	0.94 [0.91, 0.96]
Sharma 2014	37	8	54	265	0.41 [0.30, 0.51]	0.97 [0.94, 0.99]
Suzana 2016	4	4	3	42	0.57 [0.18, 0.90]	0.91 [0.79, 0.98]
Tortoli 2012	5	3	10	312	0.33 [0.12, 0.62]	0.99 [0.97, 1.00]
Vadwai 2011	5	0	5	19	0.50 [0.19, 0.81]	1.00 [0.82, 1.00]
Zeka 2011	0	0	4	52	0.00 [0.00, 0.60]	1.00 [0.93, 1.00]
Zmak 2013	0	0	1	41	0.00 [0.00, 0.97]	1.00 [0.91, 1.00]



Pleural fluid, composite reference standard

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Christopher 2013	4	0	26	61	0.13 [0.04, 0.31]	1.00 [0.94, 1.00]
Friedrich 2011	5	0	15	5	0.25 [0.09, 0.49]	1.00 [0.48, 1.00]
Lusiba 2014	25	1	62	28	0.29 [0.20, 0.39]	0.97 [0.82, 1.00]
Meldau 2014	9	1	31	47	0.23 [0.11, 0.38]	0.98 [0.89, 1.00]
Trajman 2014	1	1	32	51	0.03 [0.00, 0.16]	0.98 [0.90, 1.00]



2. Investigations of heterogeneity

a. Composite reference standard

Five studies evaluated pleural fluid with respect to the composite reference standard (Christopher 2013; Friedrich 2011; Lusiba 2014; Meldau 2014; Trajman 2014) (Figure 6). With a composite reference standard, we found lower pooled sensitivity at 18.4% (9.9 to 30.7) compared with a culture reference standard at 50.9% (39.7 to 62.8). We found similar specificity with a composite reference standard at 98.2% (94.8 to 99.5) versus with a culture reference standard at 99.2% (98.2 to 99.7) (Table 3).

b. TB prevalence

See Table 4. The median prevalence of pleural TB (as measured by culture positivity) in these studies was 15%. We found higher sensitivity in settings with higher TB prevalence than in those with lower TB prevalence, with pooled sensitivity of 58.0% (95% CrI 45.0 to 70.2) versus 38.0% (23.9 to 55.5) (probability of higher sensitivity in settings with higher TB prevalence = 0.97). We found similar specificity in settings with higher and lower TB prevalence at 99.0% (95% CrI 97.5 to 99.8) versus 99.3% (98.1 to 99.8).

3. Sensitivity analysis

See Table 5. Overall, the sensitivity analyses made little difference in any of the findings.

4. Indeterminate Xpert results

Thirteen studies (43%) reported the number of indeterminate Xpert results. Eight of these studies reported zero indeterminate results (Al-Ateah 2012; Causse 2011; Christopher 2013; Friedrich 2011; Hanif 2011; Ioannidis 2011; Sharma 2014; Zeka 2011). For pleural fluid, of 1416 tests performed, the pooled percentage of indeterminate Xpert results was 1.2% (95% CrI 0.4 to 2.6).

5. Latent class meta-analysis

Based on the latent class meta-analysis model, Xpert pooled sensitivity and specificity (95% CrI) were 56.4% (44.7 to 68.9) and 99.7% (98.1 to 100.0), respectively (27 studies, 4006 specimens) (Table 6). The pooled sensitivity of Xpert was slightly higher and its pooled specificity was comparable to what was obtained when culture was treated as having perfect accuracy. The pooled sensitivity and specificity of culture were estimated to be 81.8% (69.5 to 91.2) and 98.1% (95.9 to 99.5). The decrease in the estimated specificity of culture under the latent class meta-analysis model resulted in an increase in the estimated sensitivity of Xpert. The apparent between-study heterogeneity in the specificity of Xpert based on the primary meta-analysis was reduced after adjustments for the imperfect and heterogeneous accuracy of culture across studies (Appendix 7).

B.1. Xpert testing in pleural tissue for pleural TB

1. Primary analysis, culture reference standard

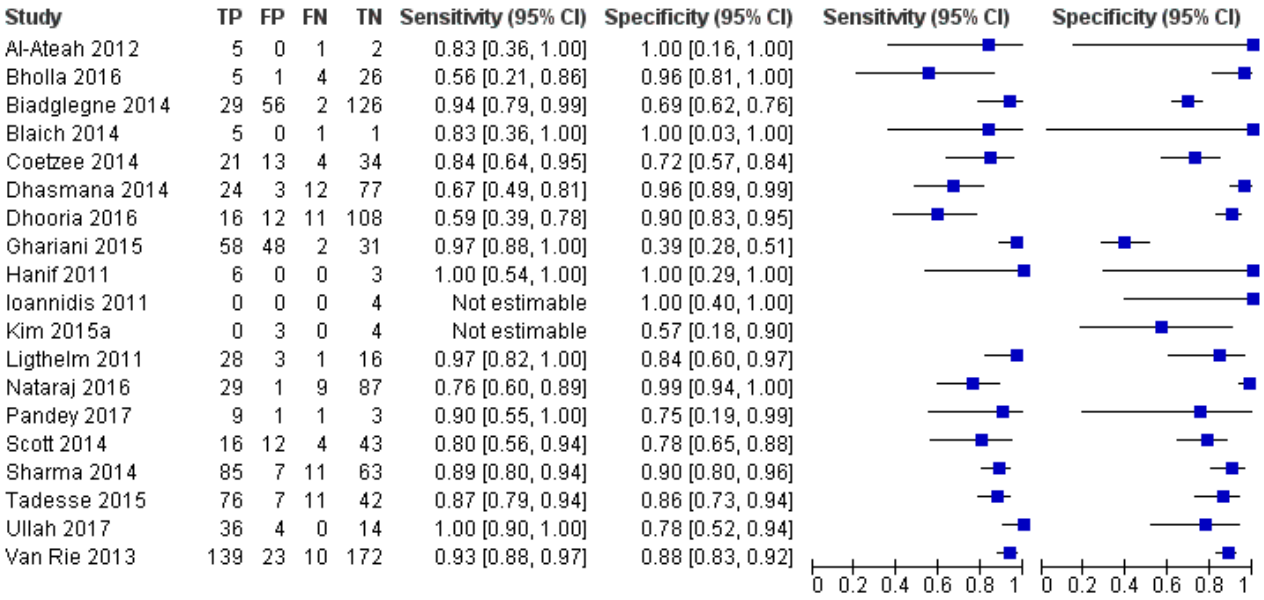
Four studies evaluated pleural tissue with respect to a culture reference standard (Christopher 2013; Du 2015; Ozkutuk 2014; Suzana 2016). The median sample size (IQR) was 41 (21 to 73) specimens. In individual studies, Xpert sensitivity ranged from 0% to 85% and specificity ranged from 97% to 100%. Pooled sensitivity and specificity (95% CrI) against culture were 30.5% (3.5 to 77.8) and 97.4% (92.1 to 99.3), respectively (three studies, 207 specimens) (Table 3).

C. Xpert testing in lymph node aspirate for lymph node TB

1. Primary analysis

Nineteen studies evaluated Xpert in lymph node aspirates (Al-Ateah 2012; Bholla 2016; Biadglegne 2014; Blaich 2014; Coetzee 2014; Dhasmana 2014; Dhooria 2016; Ghariani 2015; Hanif 2011; Ioannidis 2011; Kim 2015a; Ligthelm 2011; Nataraj 2016; Pandey 2017; Scott 2014; Sharma 2014; Tadesse 2015; Ullah 2017; Van Rie 2013). The median sample size (IQR) was 72 (12 to 138) specimens. In individual studies, Xpert sensitivity ranged from 56% to 100% and specificity from 39% to 100% (Figure 7). Xpert specificity in lymph node aspirates was considerably more heterogeneous than in CSF and pleural fluid (Figure 7). The variability in Xpert specificity in lymph node aspirates was unexpected and was suspected to be the result of a systematic, unexplained bias in some studies. Pooled sensitivity and specificity (95% CrI) against culture were 87.6% (81.7 to 92.0) and 86.0% (78.4 to 91.5), respectively (17 studies, 1710 specimens) (Table 3; Appendix 8). We discuss potential reasons for low pooled Xpert specificity in the Discussion section.

Figure 7. Forest plot of Xpert® MTB/RIF sensitivity and specificity in lymph node aspirates with respect to a culture reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive.



2. Investigations of heterogeneity

a. TB prevalence

See Table 4. The median prevalence of lymph node TB (as measured by culture positivity) in the included studies was 43%. We found higher sensitivity in settings with higher TB prevalence than in those with lower TB prevalence, with pooled sensitivity of 92.6% (95% CrI 88.1 to 95.7) versus 78.5% (95% CrI 69.2 to 86.4) (probability of higher sensitivity in the higher TB prevalence group = 0.999).

3. Sensitivity analysis

See Table 5. In comparison with all studies, studies that evaluated only adults had lower pooled sensitivity at 83.1% (69.2 to 91.5) and higher pooled specificity at 91.2% (85.2 to 95.0). In comparison with all studies, studies that evaluated only participants not receiving TB treatment had lower pooled sensitivity at 83.2% (69.2 to 90.3) and higher pooled specificity at 88.8% (80.9 to 93.8). The other sensitivity analyses made little difference in any of the findings.

4. Indeterminate Xpert results

Twelve studies (62%) reported the number of indeterminate Xpert results. Eight of these studies reported zero indeterminate results (Al-Ateah 2012; Bholla 2016; Blaich 2014; Hanif 2011; Ioannidis 2011; Ligthelm 2011; Scott 2014; Sharma 2014). For lymph node aspirate, in the 1134 tests performed, the pooled percentage of indeterminate Xpert results was 1.0% (95% CrI 0.4 to 2.0).

5. Latent class meta-analysis

Based on the latent class meta-analysis model using non-informative priors, Xpert pooled sensitivity and specificity (95% CrI) were 92.2% (82.9 to 98.1) and 89.2% (78.9 to 98.2). Unlike in the meta-analyses of Xpert in CSF and pleural fluid, adjustment for the imperfect and heterogeneous nature of culture across studies did not bring down the heterogeneity in Xpert specificity. The pooled sensitivity of culture at 88.5% (75.2 to 98.1) was estimated to be lower than 100%, although it remained greater than that of Xpert. The pooled specificity of culture was estimated to be 91.6% (84.6 to 97.1) (Table 6). As explained in the Discussion section, we believe this unusually low estimate of culture specificity was possibly the result of a systematic bias. However, when informative prior distributions were used over Xpert and culture specificity, the pooled sensitivity of both Xpert and culture was close to 80% (Table 6; Appendix 8).

C.1. Xpert testing in lymph node tissue for lymph node TB

1. Primary analysis

Ten studies evaluated lymph node tissue with respect to a culture reference standard (Blaich 2014; Causse 2011; Ghariani 2015; Kim 2015a; Ozkutuk 2014; Pandey 2017; Penata 2016; Sharma 2014; Suzana 2016; Zeka 2011). The median sample size (IQR) was 43 (15 to 82) specimens. In individual studies, Xpert sensitivity ranged from 50% to 100% and specificity ranged from 0% to 100%. Pooled sensitivity and specificity (95% CrI) against culture were 84.4% (74.7 to 91.0) and 78.9% (52.6 to 91.5), respectively (10 studies, 484 specimens) (Table 3).

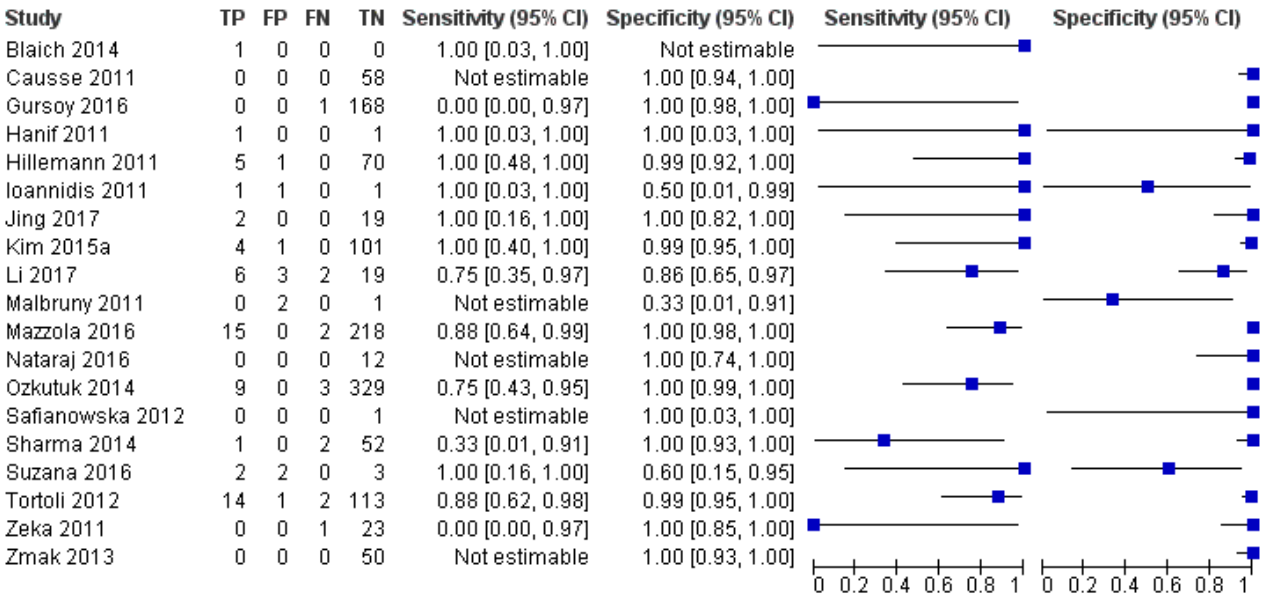
D. Xpert testing in urine for genitourinary TB

1. Primary analysis, Xpert

Nineteen studies evaluated urine (Blaich 2014; Causse 2011; Gursoy 2016; Hanif 2011; Hillemann 2011; Ioannidis 2011; Jing 2017; Kim 2015a; Li 2017; Malbruny 2011; Mazzola 2016; Nataraj 2016; Ozkutuk 2014; Safianowska 2012; Sharma 2014; Suzana 2016; Tortoli 2012; Zeka 2011; Zmak 2013). The median sample size (IQR) was 30 (five to 91) specimens. In individual studies, Xpert sensitivity ranged from 33% to 100% and specificity ranged from 33% to 100% (Figure 8). Pooled sensitivity and specificity (95% CrI) were 82.7% (69.6 to 91.1)

and 98.7% (94.8 to 99.7), respectively (13 studies, 1199 specimens) (Table 3; Appendix 9).

Figure 8. Forest plots of Xpert® MTB/RIF sensitivity and specificity in urine with respect to a culture reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive.



2. Investigations of heterogeneity

a. Specimen concentration

Five of the total 19 studies (26%) concentrated urine specimens. In one study, sensitivity and specificity (95% CI) were 88% (62 to 98) and 99% (95 to 100) (Tortoli 2012). Of the remaining four studies, three studies had zero TB culture-positives (Malbruny 2011; Nataraj 2016; Safianowska 2012), and one study had only one TB culture-positive (Zeka 2011).

b. TB prevalence

See Table 4. The median prevalence of genitourinary TB (as measured by culture positivity) in these studies was 7%. We found higher sensitivity in settings with higher TB prevalence than in those with lower TB prevalence, with pooled sensitivity of 87.9% (95% CrI 75.1 to 95.1) versus 69.6% (95% CrI 45.3 to 87.1). We found lower specificity in settings with higher TB prevalence than in those with lower TB prevalence at 98.1% (95% CrI 93.5 to 99.6) versus 99.3% (95% CrI 96.3 to 99.8). In the case of sensitivity (probability = 0.963) and specificity (probability = 0.137), accuracy in the two groups was not significantly different.

E. Xpert testing for bone or joint TB

1. Primary analysis, Xpert in bone or joint fluid

Twelve studies evaluated bone or joint fluid (Al-Ateah 2012; Blaich 2014; Gu 2015; Ioannidis 2011; Kim 2015a; Li 2017; Malbruny 2011; Nataraj 2016; Ozkutuk 2014; Penata 2016; Safianowska 2012; Suzana 2016). The median sample size (IQR) was five (two to 14) specimens. The median prevalence of TB in these studies was 50%. In individual studies, Xpert sensitivity ranged from 96% to 100% and specificity ranged from 53% to 100% (Appendix 10). Pooled sensitivity and specificity (95% CrI) were 97.2% (89.5 to 99.6) and 90.2% (55.6 to 98.5), respectively (five studies, 385 specimens) (Table 3).

2. Primary analysis, Xpert in bone or joint tissue

Seven studies evaluated bone or joint tissue (Arockiaraj 2017; Held 2014; Held 2016; Malbruny 2011; Massi 2017; Ozkutuk 2014; Penata 2016). The median sample size (IQR) was 70 (13 to 90) specimens. The median prevalence of TB in these studies was 20%. In individual studies, Xpert sensitivity ranged from 50% to 100% and specificity ranged from 17% to 100% (Appendix 10). Pooled sensitivity and specificity (95% CrI) were 94.6% (84.6 to 98.5) and 85.3% (58.7 to 96.4), respectively (six studies, 280 specimens) (Table 3).

F. Xpert testing for peritoneal TB

1. Primary analysis, Xpert in peritoneal fluid

Twenty studies evaluated peritoneal fluid (Al-Ateah 2012; Causse 2011; Iram 2015; Jing 2017; Kim 2015a; Li 2017; Malbruny 2011; Mazzola 2016; Ozkutuk 2014; Penata 2016; Rufai 2017a; Safianowska 2012; Scott 2014; Sharma 2014; Suzana 2016; Tortoli 2012; Ullah 2017; Vadwai 2011; Zeka 2011; Zmak 2013). The median sample size (IQR) was 18 (nine to 59) specimens. The median prevalence of TB in these studies was 16%. In individual studies, Xpert sensitivity ranged from 33% to 100% and specificity ranged from 90% to 100% (Appendix 11). Pooled sensitivity and specificity (95% CrI) were 59.2% (45.2 to 73.5) and 97.9% (96.2 to 99.1), respectively (16 studies, 712 specimens) (Table 3).

2. Primary analysis, Xpert in peritoneal tissue

One study evaluated peritoneal tissue (Bera 2015). Xpert sensitivity and specificity (95% CI) were 50% (7 to 93) and 92% (73 to 99) (Appendix 11).

G. Xpert testing in fluid for pericardial TB

1. Primary analysis, Xpert

Eighteen studies evaluated pericardial fluid (Al-Ateah 2012; Blaich 2014; Causse 2011; Ioannidis 2011; Kim 2015a; Mazzola 2016; Ozkutuk 2014; Pandie 2014; Penata 2016; Saeed 2017a; Safianowska 2012; Sharma 2014; Suzana 2016; Tortoli 2012; Ullah 2017; Vadwai 2011; Zeka 2011; Zmak 2013). The median sample size (IQR) was 13 (three to 19) specimens. The median prevalence of TB in these studies was 20%. In individual studies, Xpert sensitivity ranged from 25% to 100% and specificity ranged from 69% to 100% (Appendix 12). Pooled sensitivity and specificity (95% Crl) were 65.7% (46.3 to 81.4) and 96.0% (85.8 to 99.3), respectively (seven studies, 324 specimens) (Table 3).

H. Xpert testing in blood for disseminated TB

1. Primary analysis, Xpert

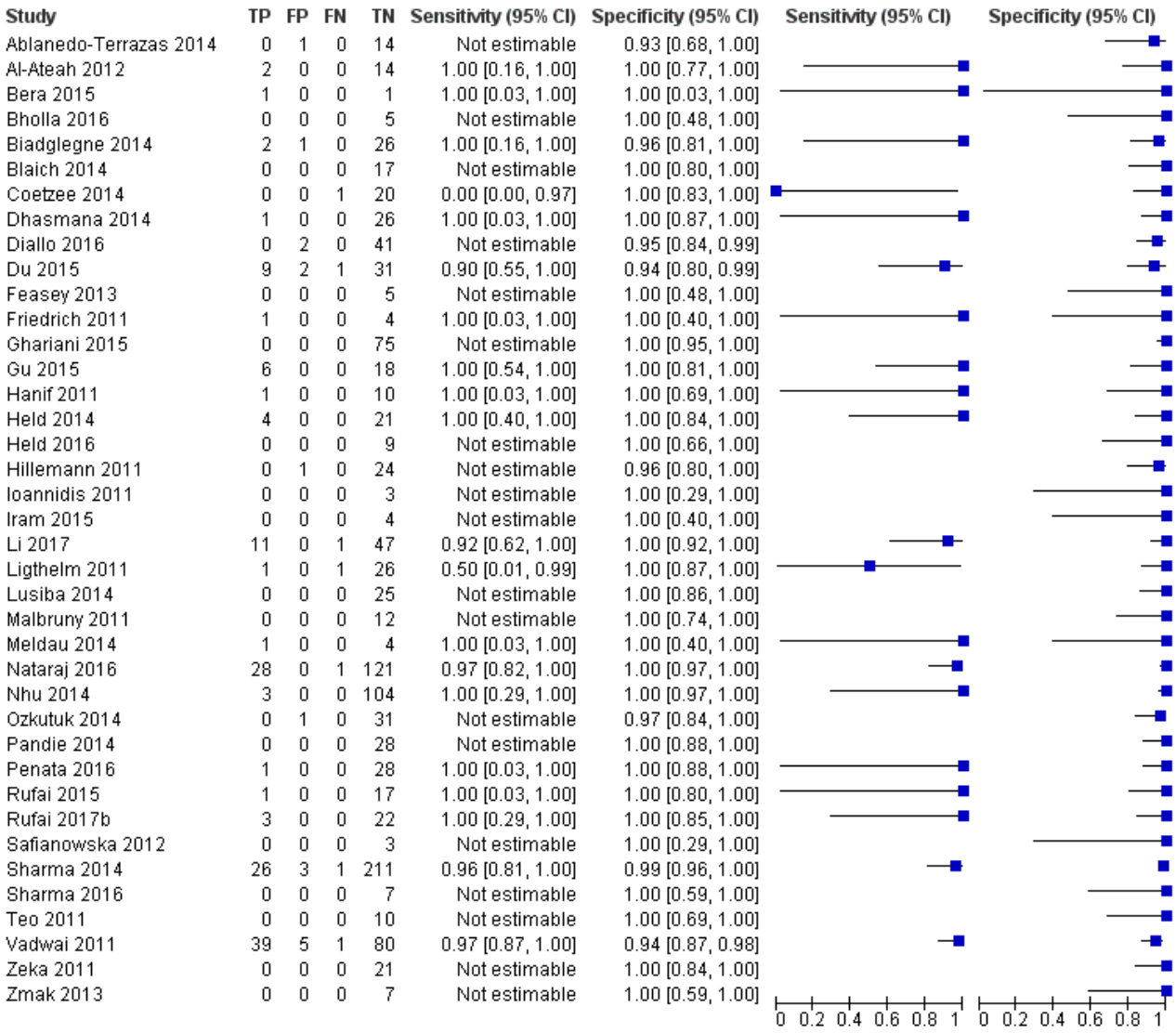
Three studies evaluated blood (Feasey 2013; Pohl 2016; Zmak 2013); however only two of these studies reported TB culture-positives. In Feasey 2013, Xpert sensitivity and specificity (95%CI) were 56% (21 to 86) and 94% (85 to 98). In Pohl 2016, Xpert sensitivity and specificity were 7% (0 to 34) and 98% (94 to 99) (Appendix 13).

II. Detection of rifampicin resistance

A. Primary analysis

Thirty-nine studies contributed data for rifampicin resistance. In individual studies, sensitivity estimates varied from 50% to 100%; specificity varied less than sensitivity (93% to 100%), (Figure 9). Three studies accounted for most of the rifampicin-resistant specimens (65%; 96/148) (Nataraj 2016; Sharma 2014; Vadwai 2011). By univariate analysis, pooled sensitivity and specificity (95% Crl) were 95.0% (89.7 to 97.9) and 98.7% (97.8 to 99.4) (Table 3). We also performed a sensitivity analysis using the bivariate random-effects model for the subset of studies that provided data for both sensitivity and specificity and found nearly identical results; the pooled sensitivity and specificity were 95.0% (89.9 to 97.9) and 98.8% (97.7 to 99.6), respectively (20 studies) (Al-Ateah 2012; Bera 2015; Biadglegne 2014; Coetzee 2014; Dhasmana 2014; Du 2015; Friedrich 2011; Gu 2015; Hanif 2011; Held 2014; Li 2017; Ligthelm 2011; Meldau 2014; Nataraj 2016; Nhu 2014; Penata 2016; Rufai 2015; Rufai 2017b; Sharma 2014; Vadwai 2011).

Figure 9. Forest plot of Xpert® MTB/RIF sensitivity and specificity for rifampicin resistance. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive.



Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ablanedo-Terrazas 2014	0	1	0	14	Not estimable	0.93 [0.68, 1.00]
Al-Ateah 2012	2	0	0	14	1.00 [0.16, 1.00]	1.00 [0.77, 1.00]
Bera 2015	1	0	0	1	1.00 [0.03, 1.00]	1.00 [0.03, 1.00]
Bholla 2016	0	0	0	5	Not estimable	1.00 [0.48, 1.00]
Biadglegne 2014	2	1	0	26	1.00 [0.16, 1.00]	0.96 [0.81, 1.00]
Blaich 2014	0	0	0	17	Not estimable	1.00 [0.80, 1.00]
Coetzee 2014	0	0	1	20	0.00 [0.00, 0.97]	1.00 [0.83, 1.00]
Dhasmana 2014	1	0	0	26	1.00 [0.03, 1.00]	1.00 [0.87, 1.00]
Diallo 2016	0	2	0	41	Not estimable	0.95 [0.84, 0.99]
Du 2015	9	2	1	31	0.90 [0.55, 1.00]	0.94 [0.80, 0.99]
Feasey 2013	0	0	0	5	Not estimable	1.00 [0.48, 1.00]
Friedrich 2011	1	0	0	4	1.00 [0.03, 1.00]	1.00 [0.40, 1.00]
Ghariani 2015	0	0	0	75	Not estimable	1.00 [0.95, 1.00]
Gu 2015	6	0	0	18	1.00 [0.54, 1.00]	1.00 [0.81, 1.00]
Hanif 2011	1	0	0	10	1.00 [0.03, 1.00]	1.00 [0.69, 1.00]
Held 2014	4	0	0	21	1.00 [0.40, 1.00]	1.00 [0.84, 1.00]
Held 2016	0	0	0	9	Not estimable	1.00 [0.66, 1.00]
Hillemann 2011	0	1	0	24	Not estimable	0.96 [0.80, 1.00]
Ioannidis 2011	0	0	0	3	Not estimable	1.00 [0.29, 1.00]
Iram 2015	0	0	0	4	Not estimable	1.00 [0.40, 1.00]
Li 2017	11	0	1	47	0.92 [0.62, 1.00]	1.00 [0.92, 1.00]
Ligthelm 2011	1	0	1	26	0.50 [0.01, 0.99]	1.00 [0.87, 1.00]
Lusiba 2014	0	0	0	25	Not estimable	1.00 [0.86, 1.00]
Malbruny 2011	0	0	0	12	Not estimable	1.00 [0.74, 1.00]
Meldau 2014	1	0	0	4	1.00 [0.03, 1.00]	1.00 [0.40, 1.00]
Nataraj 2016	28	0	1	121	0.97 [0.82, 1.00]	1.00 [0.97, 1.00]
Nhu 2014	3	0	0	104	1.00 [0.29, 1.00]	1.00 [0.97, 1.00]
Ozkutuk 2014	0	1	0	31	Not estimable	0.97 [0.84, 1.00]
Pandie 2014	0	0	0	28	Not estimable	1.00 [0.88, 1.00]
Penata 2016	1	0	0	28	1.00 [0.03, 1.00]	1.00 [0.88, 1.00]
Rufai 2015	1	0	0	17	1.00 [0.03, 1.00]	1.00 [0.80, 1.00]
Rufai 2017b	3	0	0	22	1.00 [0.29, 1.00]	1.00 [0.85, 1.00]
Safianowska 2012	0	0	0	3	Not estimable	1.00 [0.29, 1.00]
Sharma 2014	26	3	1	211	0.96 [0.81, 1.00]	0.99 [0.96, 1.00]
Sharma 2016	0	0	0	7	Not estimable	1.00 [0.59, 1.00]
Teo 2011	0	0	0	10	Not estimable	1.00 [0.69, 1.00]
Vadwai 2011	39	5	1	80	0.97 [0.87, 1.00]	0.94 [0.87, 0.98]
Zeka 2011	0	0	0	21	Not estimable	1.00 [0.84, 1.00]
Zmak 2013	0	0	0	7	Not estimable	1.00 [0.59, 1.00]

B. Investigations of heterogeneity

1. TB prevalence

See Table 4. The median prevalence of rifampicin resistance in these studies was 12%. We found higher sensitivity in settings with higher prevalence of rifampicin resistance than in those with lower prevalence, with pooled sensitivity of 96.2% (95% CrI 91.1 to 98.7) versus 92.0% (95% CrI 80.0 to 97.4). We found similar specificity in settings with higher and lower prevalence at 98.7% (95% CrI 96.8 to 99.6) versus 99.1% (95% CrI 97.7 to 99.7). In the case of sensitivity (probability = 0.878) and specificity (probability = 0.310), accuracy in the two groups was not significantly different.

C. Indeterminate Xpert results for rifampicin resistance

Eighteen studies reported the number of indeterminate Xpert results, of which six studies reported zero indeterminate results - Al-Ateah 2012 (0/17); Blaich 2014 (0/15); Held 2016 (0/17); Li 2017 (0/76); Ligthelm 2011(0/31); Teo 2011 (0/13). For rifampicin resistance testing, of the 1003 tests performed, the pooled percentage of indeterminate Xpert results was 2.6% (95% CrI 1.4 to 4.3).

D. Special topics: culture-negative specimens found to be Xpert-positive for rifampicin resistance

Culture-negative Xpert rifampicin-resistance results were infrequently reported. Three studies each reported one culture-negative, Xpert rifampicin-resistant result (Biadglegne 2014; Held 2014; Nhu 2014), and one study reported six cases (Scott 2014).

Other analyses

Non-tuberculous mycobacteria

Ten studies involving 6975 specimens provided data on a variety of NTM that grew from the specimens tested to look for evidence

of cross-reactivity: five NTM in Ablanedo-Terrazas 2014; 17 NTM in Hillemann 2011; nine NTM in Li 2017; one NTM in Malbruny 2011; 49 NTM in Mazzola 2016; three NTM in Pandey 2017; one NTM in Pink 2016; eight NTM in Sharma 2014; one NTM in Tadesse 2015; and 47 NTM inn Tortoli 2012. Among these 10 studies comprising 141 NTM, Xpert was negative in all specimens.

DISCUSSION

Summary of main results

This systematic review summarizes the current literature and includes 66 unique studies on the accuracy of Xpert for extrapulmonary tuberculosis (TB) and rifampicin resistance. Seventy-six per cent of these studies were conducted in low- and middle-income countries. Major findings from our review include the following.

- Xpert sensitivity for TB in extrapulmonary specimens varied across different types of specimens (from 31% in pleural tissue to 97% in bone or joint fluid) (Table 3).
- Xpert specificity varied less than sensitivity and in cerebrospinal fluid, pleural fluid, urine, and peritoneal fluid was ≥ 98%, with all results measured against culture as the reference standard (Table 3).
- In cerebrospinal fluid, Xpert sensitivity and specificity were 71% and 98% against culture (Summary of findings 1).
- In pleural fluid, Xpert sensitivity and specificity were 51% and 99% against culture (Summary of findings 2).
- In urine, Xpert sensitivity and specificity were 83% and 99% against culture (Summary of findings 3).
- For rifampicin resistance, Xpert sensitivity and specificity were 95% and 99% (Summary of findings 4).
- The percentage of indeterminate Xpert results was 2% for TB detection.
- The percentage of indeterminate Xpert results was 3% for rifampicin resistance detection.

For most forms of extrapulmonary TB investigated, pooled sensitivity was higher in settings with higher TB prevalence and specificity was similar or lower in settings with lower TB prevalence (Table 4).

Xpert testing in cerebrospinal fluid

(Summary of findings 1)

Results of these studies indicate that in theory, for a population of 1000 people where 100 have TB meningitis on culture, 89 would be Xpert-positive: of these, 18 (20%) would not have TB (false-positives); and 911 would be Xpert-negative: of these, 29 (3%) would have TB (false-negatives).

Rapid diagnosis of TB meningitis is critical so that lifesaving treatment can be started promptly. Around 50% of those affected die or experience disabling consequences (Thwaites 2013). In this review, we found Xpert to have a pooled sensitivity of 71% and a pooled specificity of 98% for TB meningitis. In a meta-regression analysis, we found improved Xpert accuracy in studies that concentrated the cerebrospinal fluid (CSF): pooled sensitivity concentrated 75% versus unconcentrated 66%, and identical pooled specificity of 98% in both concentrated and unconcentrated specimens. The Tuberculous Meningitis International Research Consortium has recommended increasing the volume of CSF collected for diagnosis followed by centrifugation as a way of improving Xpert sensitivity (Bahr 2016); however, we did not have sufficient data to investigate CSF collection volume. Increased Xpert sensitivity in HIV-positive people compared with HIV-negative people has been reported, with the increased bacterial burden in TB and HIV co-infection proposed as the reason (Patel 2013). We had limited data to investigate this as we identified only three studies in HIV-positive people, with Xpert sensitivities of 58% (Bahr 2015), 60% (Bahr 2017), and 81% (Patel 2013). In a sensitivity analysis in which we limited the studies to those using one specimen per participant, accuracy estimates decreased (sensitivity 64% and specificity 96%).

Xpert Ultra testing in CSF

Ultra was designed to improve TB detection, in particular in people with paucibacillary disease. The limit of detection is lower with Ultra (16 bacterial colony-forming units (cfu) per mL) than with Xpert (131 cfu per mL) (Chakravorty 2017). We identified one study that evaluated Ultra for TB meningitis in HIV-positive patients. This study found considerably higher sensitivity with Ultra (90%) compared with Xpert (60%) based on a culture reference standard (Bahr 2017). Notwithstanding Ultra's high sensitivity, given the disastrous consequences of missing a diagnosis of TB meningitis, providers should use clinical judgement and should not rely solely on an Ultra result when deciding to withhold treatment.

Bahr 2017 found the specificity of Ultra (90%) for TB meningitis to be considerably lower than that of Xpert (97%). We considered several reasons in trying to explain this finding. One reason that has been proposed is the lingering presence of dead TB bacteria (or bacterial components) from previous TB (WHO 2017c). In a study of pulmonary TB, Ultra had lower specificity than Xpert, and of interest, the difference was more pronounced in previously treated patients (Chakravorty 2017). However, Bahr and colleagues considered that this reason may not apply to Ultra for TB meningitis because it is unlikely that TB bacilli in CSF are derived from prior TB (either TB bacteria are no longer present or the patient has died). A second reason for the lower specificity with Ultra is linked to 'trace-calls' (Chakravorty 2017). For extrapulmonary specimens, the World Health Organization (WHO) recommends that 'trace calls' should be considered to be true-positive results for use in clinical decisions and patient follow-up" (WHO 2017c).

Xpert testing in pleural fluid

(Summary of findings 2)

Results of these studies indicate that in theory, for a population of 1000 people where 150 have pleural TB on culture, 83 would be Xpert-positive: of these, seven (8%) would not have TB (false-positives); and 917 would be Xpert-negative: of these, 74 (8%) would have TB (false-negatives).

We found Xpert to have low sensitivity (51%) in pleural fluid when measured against a culture reference standard and even lower sensitivity (18%) when measured against a composite reference standard. By design, we expected to find higher pooled sensitivity with the culture reference standard than with the composite reference standard. One reason for the low sensitivity of Xpert could be the paucibacillary nature of pleural TB. Other possible reasons are contamination of blood or the presence of certain polymerase chain reaction (PCR) inhibitors in the pleural fluid (Pai

2004; Woods 2001). However, in a study by Theron and colleagues, extrapulmonary specimens showed less evidence of PCR inhibition than pulmonary specimens, with the bacterial load more important for a positive Xpert result (Theron 2014).

Xpert specificity in pleural fluid was 99%. However, given that false-negative results were common (low sensitivity), a negative Xpert result may not be relied on to exclude TB. The WHO recommends that pleural biopsy tissue is the preferred specimen type for diagnosing pleural TB using Xpert (WHO 2013). However, we had insufficient data to determine summary accuracy of Xpert in pleural tissue (three studies, 207 specimens).

Xpert testing in lymph node aspirates

In 76% of the included studies (13 of 17 studies contributing both sensitivity and specificity data), Xpert achieved a sensitivity of 80% or higher, suggesting that Xpert could improve the diagnosis of lymph node TB. It is important to point out that although tissue biopsy provides material for histological examination, which may be of substantial diagnostic value, a fluid specimen may be collected more easily. In addition, fine-needle aspiration of lymph nodes is well suited for use in resource-limited settings because the procedure is simple, easy to learn, minimally invasive, and inexpensive (Wright 2009b). Thus clinicians may want to consider fine-needle aspiration of lymph nodes before surgical biopsy.

In our review, using a standard bivariate meta-analysis model, Xpert specificity (defined by culture) in lymph node aspirate was 86%, whereas with a latent class meta-analysis model with informative priors, Xpert specificity increased to 99%. In previous meta-analyses, Xpert specificity for lymph node TB (aspirate and tissue) against culture as a reference standard was 94% (Denkinger 2014), 93% (Maynard-Smith 2014), and 92% (Penz 2015). See Table 2. Using a composite reference standard (defined by the primary study authors), Denkinger and colleagues found increased Xpert specificity of 99% for lymph node TB (five studies, 728 specimens) (Denkinger 2014). Thus, it appears that accuracy results depend in part on the choice of reference standard. In our review, we used culture as the reference standard and adjusted accuracy estimates with a latent class meta-analysis model rather than using a composite reference standard owing to differing definitions of the composite reference standards, difficulty in interpreting them, and concern for bias (Schiller 2016) (see section Strengths and weaknesses of the review).

We considered several reasons why Xpert specificity would be lower for lymph node TB than for other forms of extrapulmonary TB. Lymph node aspirates may be of lesser quality when collected from children (Coetzee 2014), and we included participants of all ages in the review. In a post hoc sensitivity analysis limiting inclusion to studies that involved only adults, specificity increased from 86% to 91% (Table 5). Although not always reported, studies may have included patients receiving TB treatment. In a sensitivity analysis limiting inclusion to studies that involved participants not receiving TB treatment, specificity increased from 86% to 89% (Table 5). Theron and colleagues found Xpert-positive, culture-negative results to be more common in people with a history of TB (Theron 2016); however, we had insufficient data to evaluate this factor. We considered the type of culture used in the included studies because liquid culture is more sensitive than solid culture (American Thoracic Society 2000). Most studies did use liquid culture or a combination of solid and liquid culture; only two of the 17 studies (12%) exclusively used solid culture. Culture results may also be negative owing to inefficient specimen collection or errors in sampling, differing bacterial load, and contamination (Wright 2009b). Negative culture results in lymph node TB have previously been reported (Fontanilla 2011).

Another reason for negative culture results is that there may have been a decrease in live TB bacteria during processing with N-acetyl-L-cysteine-sodium hydroxide, which is routinely used to homogenize, decontaminate, and liquefy non-sterile specimens, such as sputum, for TB culture (American Thoracic Society 2000). Harsh decontamination practices have been noted to contribute to false-negative culture results, especially in paucibacillary specimens (FIND 2017). Standards specify, "specimens collected from normally sterile sites may be placed directly into the culture medium" (American Thoracic Society 2000). CSF, pleural fluid, and lymph node aspirates are usually considered to be sterile specimens. It is our understanding that some laboratories do decontaminate sterile site specimens as a precaution against non-sterile collection procedures. In this review, 47% of the studies reported decontaminating lymph node aspirates before culture inoculation. We did not have sufficient data to further investigate laboratory practices.

In sum, several factors probably contributed to low Xpert specificity in lymph node aspirate. The "true" specificity of Xpert in lymph node aspirate is likely to be higher, similar to that found in CSF, pleural fluid, and other specimens (Table 3). For all of the aforementioned reasons, we recommend caution in interpreting the results of Xpert accuracy for lymph node TB.

Xpert testing in urine

(Summary of findings 3)

Results of these studies indicate that in theory, for a population of 1000 people where 70 have genitourinary TB on culture, 70 would be Xpert-positive: of these, 12 (17%) would not have TB (false-positives); and 930 would be Xpert-negative: of these, 12 (1%) would have TB (false-negatives).

Xpert was sensitive and specific for genitourinary TB. Urine is an attractive specimen for TB diagnosis because of its availability, accessibility (it is easily collected from adults and children), few processing requirements, and low risk of infection risk to healthcare workers during specimen collection (Peter 2010). It has been proposed that concentrated urine increases the sensitivity of Xpert (Peter 2012). However, we had insufficient data to investigate this proposition.

Xpert specificity in patients with a prior history of TB or TB treatment

For detection of extrapulmonary TB, we intended to determine Xpert specificity in patients with a prior history of TB. However, this information was infrequently reported: lymph node, five studies (24%); pleural fluid, four studies (13%); and CSF, seven studies (21%).

Xpert testing for rifampicin resistance

(Summary of findings 4)

For detection of rifampicin resistance, we found a sensitivity of 95% and a specificity of 99%, similar to the estimates in the review for

pulmonary TB: sensitivity (95%) and specificity (98%) (Steingart 2014). These findings suggest that use of Xpert could assist in rapid diagnosis of rifampicin-resistant TB and early initiation of treatment for multidrug-resistant TB (MDR-TB).

Results of these studies indicate that in theory, for a population of 1000 people where 120 have rifampicin-resistant TB, 125 would be positive for rifampicin-resistant TB: of these, 11 (9%) would not have rifampicin resistance (false-positives); and 875 would be negative for rifampicin-resistant TB: of these, six (1%) would have rifampicin resistance (false-negatives).

Culture-negative specimens found to be Xpert-positive for rifampicin resistance have been described in the literature for pulmonary TB (Boyles 2014; Kelly 2014). In the included studies, we looked for information on this topic but found only a few cases.

Of note, concerns have been raised about rapid drug susceptibility testing (DST) methods, in particular automated mycobacteria growth indicator tube (MGIT) 960 for TB drug resistance using the recommended critical concentrations. As a priority, the WHO is planning to re-evaluate the critical concentrations for rifampicin (WHO 2018).

Strengths and weaknesses of the review

Completeness of evidence

This is a reasonably complete data set. We included any non-English studies that we found from which we could obtain accuracy data. However, we acknowledge that we may have missed some studies despite the comprehensive search and our outreach to investigators. We included eight common forms of extrapulmonary TB in the review. However, for some of these forms, such as disseminated TB, data were insufficient to allow us to determine summary accuracy estimates. We did not include less common forms, such as cutaneous TB, ocular TB, female genital TB, and TB of the breast.

Accuracy of the reference standards used

In a systematic review of diagnostic test accuracy studies, the reference standard is the best available test to determine the presence or absence of the target condition. In this review, we used culture as the reference standard for all forms of extrapulmonary TB. Although culture is the best available reference standard, it is not a perfect reference standard for extrapulmonary TB owing to the paucibacillary nature of the disease. Therefore, we applied a latent class model to correct the biases in Xpert sensitivity and specificity resulting from treating culture as a perfect reference standard. We added parameters for the sensitivity and specificity of culture and terms for conditional dependence to adjust for the dependence between Xpert and culture among disease-positive and disease-negative patients. In this way, we were able to improve estimation of both the pooled sensitivity and specificity of Xpert, as well as between-study variability.

In terms of accuracy of the reference standard for lymph node aspirate in particular, several factors may have contributed to false-negative culture results, including inefficient specimen collection and overly harsh decontamination. For this particular analysis, we were able to take advantage of the Bayesian estimation approach to incorporate prior information on Xpert and culture specificity. This allowed us to make the best use of data from the included studies and our knowledge of the performance of Xpert.

Establishing a diagnosis of extrapulmonary TB would ideally include pursuing the diagnosis of pulmonary TB as well because patients with TB may have both pulmonary and extrapulmonary TB and the lung may be the only site where the presence of TB may be established. For example, for lymph node TB in children, specimens would include lymph node aspirate or tissue, sputum, gastric washings, and possibly stool. It is necessary to pursue every avenue of diagnosis because of the paucibacillary nature of extrapulmonary TB and the varying sensitivity of culture among different specimen types. As another example, because of the difficulties involved in diagnosing HIV-associated TB, it is recommended that multiple cultures from sputum and other types of specimens be evaluated in HIV-positive people (Shah 2016b). Given these limitations in the reference standard, we recommend that future studies consider utilizing liquid culture because liquid culture is more sensitive than solid culture and that researchers obtain multiple specimens for culture to confirm the diagnosis of extrapulmonary TB.

In terms of detection of rifampicin resistance, most studies included in this review used culture-based DST (either Löwenstein-Jensen (LJ) or mycobacteria growth indicator tube (MGIT) 960) as the reference standard. Of note, concerns have been raised about rapid DST methods, in particular automated MGIT 960, for TB drug resistance using the recommended critical concentrations. As a priority, the WHO is planning to re-evaluate the critical concentrations for rifampicin (WHO 2018).

Quality and quality of reporting of the included studies

Risk of bias was low for the patient selection, index test, and flow and timing domains and was high or unclear for the reference standard domain (most of these studies performed specimen decontamination before culture inoculation). A limitation was that several studies included more than one specimen per participant, which artificially inflated the sample size of the study and may have led to overestimation or underestimation of the accuracy estimates. In general, studies were fairly well reported, although we corresponded with almost all primary study authors to ask for additional data and missing information. In several studies, accuracy data by site of extrapulmonary disease were not reported, and in a minority of studies, blinding was not reported. We strongly encourage the authors of future studies to follow the recommendations provided in the updated Standards for Reporting Diagnostic Accuracy (STARD) statement to improve the quality of reporting (Bossuyt 2015).

Interpretability of subgroup analyses

We investigated potential sources of heterogeneity in the different extrapulmonary specimens. Generally, we found increased sensitivity in settings with higher TB prevalence (culture-confirmed TB cases in the study) and similar or slightly lower specificity. In pleural fluid, with use of a composite reference standard, as expected, Xpert sensitivity was lower in comparison with culture (composite 18% vs culture 51%). Specificity was similar (composite 98% versus culture 99%).

Comparison with other systematic reviews

We are aware of six systematic reviews previously published on this topic that estimated summary accuracy with respect to a culture reference standard, as we did in our review (Table 2). Chang 2012 (seven studies) and Li Y 2017 (26 studies) determined the diagnostic accuracy of Xpert for multiple forms of extrapulmonary TB combined, and Denkinger 2014 (18 studies), Maynard-Smith 2014 (27 studies), Penz 2015 (37 studies), and Sehgal 2016 (24 studies) determined Xpert accuracy for specific forms of extrapulmonary TB. In these reviews, sensitivities ranged from 69% to 85% for CSF (our review: 71%). In pleural fluid, sensitivities ranged from 34% to 51% (our review: 51%). Specificities ranged from 97% to 100% (our review: CSF 98%, pleural fluid 99%).

Wen 2017 (12 studies) determined Xpert accuracy for bone or joint TB measured against culture, histology, or a composite reference standard and found pooled sensitivity and specificity of 81% and 83%, respectively (our review, against a culture reference standard: sensitivity 97%, specificity 92%).

Compared with previous systematic reviews, our review extended the date of the search for potential studies for inclusion. Our strict inclusion criteria - for example, including only studies that used culture as the reference standard and excluding case-control studies - meant that some of the studies included in other reviews were excluded from our review.

Applicability of findings to the review question

For the patient selection domain, most studies had high or unclear risk because either patients were evaluated exclusively as inpatients in tertiary care or we were not sure about the clinical settings. Therefore, we cannot be sure of the applicability of our findings to primary care. Studies that take place in referral settings may include patients whose condition is more difficult to diagnose than are seen at lower levels of the health system. However, we recognize that classifying studies with respect to primary, secondary, or tertiary care may not adequately account for differences in disease spectrum (Leeflang 2013). For the index and reference test domains, most studies had low concern for applicability.

AUTHORS' CONCLUSIONS

Implications for practice

In people presumed to have extrapulmonary TB, Xpert may be helpful in confirming the diagnosis. Xpert sensitivity varies across different extrapulmonary specimens, while for most specimens, specificity is high, the test rarely yielding a positive result for people without TB (defined by culture). Xpert is accurate for detection of rifampicin resistance. For people thought to have TB meningitis, treatment should be based on clinical judgement, and not withheld solely on an Xpert result, as is common practice when culture results are negative.

Implications for research

Future studies should perform comparisons of different tests, including Xpert Ultra, as this approach will reveal which tests (or strategies) yield superior diagnostic accuracy. For these studies, the preferred study design is one in which all participants receive all available diagnostic tests or are randomly assigned to receive one or another of the tests. Studies should include children and HIV-positive people. Future research should acknowledge the concern associated with culture as a reference standard in paucibacillary specimens and should consider ways to address this limitation.

Rapid point-of-care diagnostic tests for extrapulmonary TB are critically needed. Research groups should focus on developing diagnostic tests and strategies that use readily available clinical specimens such as urine, rather than specimens that require invasive procedures for collection.

ACKNOWLEDGEMENTS

The Academic Editors on this review were Dr Michael Eisenhut (Cochrane Infectious Diseases Group; CIDG) and Dr Mia Schmidt-Hansen (Cochrane Diagnostic Test Accuracy).

We are grateful to Vittoria Lutje, the CIDG Information Specialist, for help with the search strategy. The CIDG editorial base is funded by UK aid from the UK government for the benefit of low- and middle-income countries (project number 300342-104). The views expressed do not necessarily reflect the UK government's official policies.

We thank Hannah Ryan, from the Tropical and Infectious Diseases Unit, Royal Liverpool University Hospital, for her help with the protocol. We thank Chunli Lu, from the Centre for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing and Marcela Perlwitz, Ivy Tech Community College of Indiana, West Lafayette, for help with translation. We thank all authors of the included studies for providing answers to our questions along with additional data. We thank the referees for their helpful comments.

REFERENCES

References to studies included in this review

Ablanedo-Terrazas 2014 *{published data only}*

Ablanedo-Terrazas Y, Alvarado-de la Barrera C, Hernandez-Juan R, Ruiz-Cruz M, Reyes-Teran G. Xpert MTB/RIF for diagnosis of tuberculous cervical lymphadenitis in HIV-infected patients. *Laryngoscope* 2014;**124**(6):1382-5.

Al-Ateah 2012 *{published data only}*

Al-Ateah SM, Al-Dowaidi MM, El-Khizzi NA. Evaluation of direct detection of Mycobacterium tuberculosis complex in respiratory and non-respiratory clinical specimens using the Cepheid Gene Xpert(R) system. *Saudi Medical Journal* 2012;**33**(10):1100-5.

Arockiaraj 2017 *{published data only}*

Arockiaraj J, Michael JS, Amritanand R, David KS, Krishnan V. The role of Xpert MTB/RIF assay in the diagnosis of tubercular spondylodiscitis. *European Spine Journal* 2017;**26**(12):3162-9.

Bahr 2015 *{published data only}*

Bahr NC, Tugume L, Rajasingham R, Kiggundu R, Williams DA, Morawski B, et al. Improved diagnostic sensitivity for tuberculous meningitis with Xpert((R)) MTB/RIF of centrifuged CSF. *International Journal of Tuberculosis and Lung Disease* 2015;**19**(10):1209-15.

Bahr 2017 *{published data only}*

Bahr NC, Nuwagira E, Evans EE, Cresswell FV, Bystrom PV, Byamukama A, et al. Diagnostic accuracy of Xpert MTB/Rif Ultra for TB meningitis in HIV-infected adults: a prospective cohort study. *Lancet Infectious Diseases* 2017;**18**(1):68-75.

Bera 2015 *{published data only}*

Bera C, Michael JS, Burad D, Shirly SB, Gibikote S, Ramakrishna B, et al. Tissue Xpert^{TM} MTB/Rif assay is of limited use in diagnosing peritoneal tuberculosis in patients with exudative ascites. *Indian Journal of Gastroenterology* 2015;**34**(5):395-8.

Bholla 2016 *{published data only}*

Bholla M, Kapalata N, Masika E, Chande H, Jugheli L, Sasamalo M, et al. Evaluation of Xpert(R) MTB/RIF and Ustar EasyNAT TB IAD for diagnosis of tuberculous lymphadenitis of children in Tanzania: a prospective descriptive study. *BMC Infectious Diseases* 2016;**16**:246. [DOI: 10.1186/s12879-016-1578-z]

Biadglegne 2014 *{published data only}*

Biadglegne F, Mulu A, Rodloff A C, Sack U. Diagnostic performance of the Xpert MTB/RIF assay for tuberculous lymphadenitis on fine needle aspirates from Ethiopia. *Tuberculosis (Edinburgh, Scotland)* 2014;**94**(5):502-5.

Blaich 2014 *{published data only}*

Blaich A, Frei R. Performance of the Xpert MTB/RIF assay on nonrespiratory specimens and accuracy of this assay for detection of rifampin resistance in a low-prevalence setting. *Journal of Clinical Microbiology* 2014;**52**(2):706.

Causse 2011 *{published data only}*

Causse M, Ruiz P, Gutierrez-Aroca JB, Casal M. Comparison of two molecular methods for rapid diagnosis of extrapulmonary tuberculosis. *Journal of Clinical Microbiology* 2011;**49**(8):3065-7.

Che 2017 *{published data only}*

Che N, Yang X, Liu Z, Li K, Chen X. Rapid detection of cell-free Mycobacterium tuberculosis DNA in tuberculous pleural effusion. *Journal of Clinical Microbiology* 2017;**55**(5):1526-32.

Christopher 2013 *{published data only}*

Christopher DJ, Schumacher SG, Michael JS, Luo R, Balamugesh T, Duraikannan P, et al. Performance of Xpert MTB/RIF on pleural tissue for the diagnosis of pleural tuberculosis. *European Respiratory Journal* 2013;**42**(5):1427-9.

Coetzee 2014 *{published data only}*

Coetzee L, Nicol MP, Jacobson R, Schubert PT, van Helden PD, Warren RM, et al. Rapid diagnosis of pediatric mycobacterial lymphadenitis using fine needle aspiration biopsy. *Pediatric Infectious Disease Journal* 2014;**33**(9):893-6.

Dhasmana 2014 *{published data only}*

Dhasmana DJ, Ross C, Bradley CJ, Connell DW, George PM, Singanayagam A, et al. Performance of Xpert MTB/RIF in the diagnosis of tuberculous mediastinal lymphadenopathy by endobronchial ultrasound. *Annals of the American Thoracic Society* 2014;**11**(3):392-6.

Dhooria 2016 *{published data only}*

Dhooria S, Gupta N, Bal A, Sehgal IS, Aggarwal A, Sethi S, et al. Role of Xpert MTB/RIF in differentiating tuberculosis from sarcoidosis in patients with mediastinal lymphadenopathy undergoing EBUS-TBNA: a study of 147 patients. *Sarcoidosis Vasculitis and Diffuse Lung Disease* 2016;**33**:258-266.

Diallo 2016 *{published data only}*

Diallo AB, Kollo AI, Camara M, Lo S, Ossoga GW, Mbow M, et al. Performance of GeneXpert MTB / RIF in the diagnosis of extrapulmonary tuberculosis in Dakar: 2010-2015 [Performance du GeneXpert MTB/RIF® dans le diagnostic de la tuberculose extra-pulmonaire à Dakar: 2010-2015]. *Pan African Medical Journal* 2016;**25**:129. [DOI: 10.11604/pamj.2016.25.129.10065]

Du 2015 *{published data only}*

Du J, Huang Z, Luo Q, Xiong G, Xu X, Li W, et al. Rapid diagnosis of pleural tuberculosis by Xpert MTB/RIF assay using pleural biopsy and pleural fluid specimens. *Journal of Research In Medical Sciences* 2015;**20**(1):26-31.

Feasey 2013 *{published data only}*

Feasey NA, Banada PP, Howson W, Sloan DJ, Mdolo A, Boehme C, et al. Evaluation of Xpert MTB/RIF for detection of tuberculosis from blood samples of HIV-infected adults confirms Mycobacterium tuberculosis bacteremia as an indicator of poor prognosis. *Journal of Clinical Microbiology* 2013;**51**(7):2311-6.

Friedrich 2011 *{published data only}*

Friedrich SO, von Groote-Bidlingmaier F, Diacon AH. Xpert MTB/RIF assay for diagnosis of pleural tuberculosis. *Journal of Clinical Microbiology* 2011;**49**(12):4341-2.

Ghariani 2015 *{published data only}*

Ghariani A, Jaouadi T, Smaoui S, Mehiri E, Marouane C, Kammoun S, et al. Diagnosis of lymph node tuberculosis using the GeneXpert MTB/RIF in Tunisia. *International Journal of Mycobacteriology* 2015;**4**(4):270-5.

Gu 2015 *{published data only}*

Gu Y, Wang G, Dong W, Li Y, Ma Y, Shang Y, et al. Xpert MTB/RIF and GenoType MTBDRplus assays for the rapid diagnosis of bone and joint tuberculosis. *International Journal of Infectious Diseases* 2015;**36**:27-30.

Gursoy 2016 *{published data only}*

Gursoy NC, Yakupogullari Y, Tekerekoglu MS, Otlu B. Evaluation of the diagnostic performance of Xpert MTB/RIF test for the detection of Mycobacterium tuberculosis and rifampin resistance in clinical samples [Klinik Örneklerden Mycobacterium tuberculosis Saptanması ve Rifampin Direnci Tespitinde Xpert MTB/RIF Testinin Tanısal Performansının Değerlendirilmesi]. *Mikrobiyoloji Bulteni* 2016;**50**(2):196-204.

Hanif 2011 *{published data only}*

Hanif SN, Eldeen HS, Ahmad S, Mokaddas E. GeneXpert(R) MTB/RIF for rapid detection of Mycobacterium tuberculosis in pulmonary and extra-pulmonary samples. *International Journal of Tuberculosis and Lung Disease* 2011;**15**(9):1274-5.

Held 2014 *{published data only}*

Held M, Laubscher M, Zar HJ, Dunn RN. GeneXpert polymerase chain reaction for spinal tuberculosis: an accurate and rapid diagnostic test. *Bone & Joint Journal* 2014;**96-b**(10):1366-9.

Held 2016 *{published data only}*

Held M, Laubscher M, Mears S, Dix-Peek S, Workman L, Zar H, et al. Diagnostic accuracy of the Xpert MTB/RIF assay for extrapulmonary tuberculosis in children with musculoskeletal infections. *Pediatric Infectious Disease Journal* 2016;**35**(11):1165-8.

Hillemann 2011 *{published data only}*

Hillemann D, Rusch-Gerdes S, Boehme C, Richter E. Rapid molecular detection of extrapulmonary tuberculosis by the automated GeneXpert MTB/RIF system. *Journal of Clinical Microbiology* 2011;**49**(4):1202-5.

Ioannidis 2011 *{published data only}*

Ioannidis P, Papaventsis D, Karabela S, Nikolaou S, Panagi M, Raftopoulou E, et al. Cepheid GeneXpert MTB/RIF assay for Mycobacterium tuberculosis detection and rifampin resistance identification in patients with substantial clinical indications of tuberculosis and smear-negative microscopy results. *Journal of Clinical Microbiology* 2011;**49**(8):3068-70.

Iram 2015 *{published data only}*

Iram S, Zeenat A, Hussain S, Wasim Yusuf N, Aslam M. Rapid diagnosis of tuberculosis using Xpert MTB/RIF assay - report from a developing country. *Pakistan Journal of Medical Sciences* 2015;**31**(1):105-10.

Jing 2017 *{published data only}*

Jing H, Lu ZM, Deng YF, Gao DC, Li L, Graviss EA, et al. Evaluation of Xpert MTB/RIF in detection of pulmonary and extrapulmonary tuberculosis cases in China. *International Journal of Clinical and Experimental Pathology* 2017;**10**(4):4847-51.

Kim 2015a *{published data only}*

Kim YW, Kwak N, Seong MW, Kim EC, Yoo CG, Kim YW, et al. Accuracy of the Xpert(R) MTB/RIF assay for the diagnosis of extra-pulmonary tuberculosis in South Korea. *International Journal of Tuberculosis and Lung Disease* 2015;**19**(1):81-6.

Li 2017 *{published data only}*

Li Y, Pang Y, Zhang T, Xian X, Wang X, Yang J, et al. Rapid diagnosis of extrapulmonary tuberculosis with Xpert Mycobacterium tuberculosis/rifampicin assay. *Journal of Medical Microbiology* 2017;**66**(7):910-4.

Ligthelm 2011 *{published data only}*

Ligthelm LJ, Nicol MP, Hoek KG, Jacobson R, van Helden PD, Marais BJ, et al. Xpert MTB/RIF for rapid diagnosis of tuberculous lymphadenitis from fine-needle-aspiration biopsy specimens. *Journal of Clinical Microbiology* 2011;**49**(11):3967-70.

Lusiba 2014 *{published data only}*

Lusiba JK, Nakiyingi L, Kirenga BJ, Kiragga A, Lukande R, Nsereko M, et al. Evaluation of Cepheid's Xpert MTB/Rif test on pleural fluid in the diagnosis of pleural tuberculosis in a high prevalence HIV/TB setting. *PLOS One* 2014;**9**(7):e102702. [DOI: 10.1371/journal.pone.0102702]

Malbruny 2011 *{published data only}*

Malbruny B, Le Marrec G, Courageux K, Leclercq R, Cattoir V. Rapid and efficient detection of Mycobacterium tuberculosis in respiratory and non-respiratory samples. *International Journal of Tuberculosis and Lung Disease* 2011;**15**(4):553-5.

Massi 2017 *{published data only}*

Massi MN, Biatko KT, Handayani I, Pratama MY, Septriani S, Nurdin GM, et al. Evaluation of rapid GeneXpert MTB/RIF method using DNA tissue specimens of vertebral bones in patients with suspected spondylitis TB. *Journal of Orthopaedics* 2017;**14**(1):189-91.

Mazzola 2016 *{published data only}*

Mazzola E, Arosio M, Nava A, Fanti D, Gesu G, Farina C. Performance of real-time PCR Xpert (R)MTB/RIF in diagnosing extrapulmonary tuberculosis. *Infezioni in Medicina* 2016;**24**(4):304-9.

Meldau 2014 *{published data only}*

Meldau R, Peter J, Theron G, Calligaro G, Allwood B, Symons G, et al. Comparison of same day diagnostic tools including Gene Xpert and unstimulated IFN-gamma for the evaluation of pleural tuberculosis: a prospective cohort study. *BMC Pulmonary Medicine* 2014;**14**:58. [DOI: 10.1186/1471-2466-14-58]

Nataraj 2016 *{published data only}*

Nataraj G, Kanade S, Mehta P. Xpert((R)) MTB/RIF for improved case detection of extra-pulmonary TB in a tertiary care setting in urban India. *International Journal of Tuberculosis and Lung Disease* 2016;**20**(7):890-4.

Nhu 2014 *{published data only}*

Nhu NT, Heemskerk D, Thu do DA, Chau TT, Mai NT, Nghia HD, et al. Evaluation of GeneXpert MTB/RIF for diagnosis of tuberculous meningitis. *Journal of Clinical Microbiology* 2014;**52**(1):226-33.

Ozkutuk 2014 *{published data only}*

Ozkutuk N, Surucuoglu S. Evaluation of the Xpert MTB/RIF assay for the diagnosis of pulmonary and extrapulmonary tuberculosis in an intermediate-prevalence setting [Orta Prevalanslı Bölgede Akciğer ve Akciğer DışıTüberküloz Tanısında Xpert MTB/RIF Testinin Değerlendirilmesi]. *Mikrobiyoloji Bulteni* 2014;**48**(2):223-32.

Pandey 2017 *{published data only}*

Pandey S, Congdon J, McInnes B, Pop A, Coulter C. Evaluation of the GeneXpert MTB/RIF assay on extrapulmonary and respiratory samples other than sputum: a low burden country experience. *Pathology* 2017;**49**(1):70-4.

Pandie 2014 *{published data only}*

Pandie S, Peter JG, Kerbelker ZS, Meldau R, Theron G, Govender U, et al. Diagnostic accuracy of quantitative PCR (Xpert MTB/RIF) for tuberculous pericarditis compared to adenosine deaminase and unstimulated interferon-gamma in a high burden setting: a prospective study. *BMC Medicine* 2014;**12**:101. [DOI: 10.1186/1741-7015-12-101]

Patel 2013 *{published data only}*

Patel VB, Theron G, Lenders L, Matinyena B, Connolly C, Singh R, et al. Diagnostic accuracy of quantitative PCR (Xpert MTB/RIF) for tuberculous meningitis in a high burden setting: a prospective study. *PLOS Medicine* 2013;**10**(10):e1001536. [DOI: 10.1371/journal.pmed.1001536]

Penata 2016 *{published data only}*

Peñata A, Salazar R, Castano T, Bustamante J, Ospina S. Molecular diagnosis of extrapulmonary tuberculosis and sensitivity to rifampicin with an automated real-time method [Diagnóstico molecular de tuberculosis extrapulmonary sensibilidad a rifampicina con un método automatizado en tiempo real]. *Biomedica* 2016;**36 Suppl 1**:78-89.

Pink 2016 *{published data only}*

Pink F, Brown TJ, Kranzer K, Drobniewski F. Evaluation of Xpert MTB/RIF for detection of Mycobacterium tuberculosis in cerebrospinal fluid. *Journal of Clinical Microbiology* 2016;**54**(3):809-11.

Pohl 2016 *{published data only}*

Pohl C, Rutaihwa LK, Haraka F, Nsubuga M, Aloi F, Ntinginya NE, et al. Limited value of whole blood Xpert (R) MTB/RIF for diagnosing tuberculosis in children. *Journal of Infection* 2016;**73**(4):326-35.

Rufai 2015 *{published data only}*

Rufai SB, Singh A, Kumar P, Singh J, Singh S. Performance of Xpert MTB/RIF Assay in diagnosis of pleural tuberculosis by use of pleural fluid samples. *Journal of Clinical Microbiology* 2015;**53**(11):3636-8.

Rufai 2017a *{published data only}*

Rufai SB, Singh S, Singh A, Kumar P, Singh J, Vishal A. Performance of Xpert MTB/RIF on ascitic fluid samples for detection of abdominal tuberculosis. *Journal of Laboratory Physicians* 2017;**9**(1):47-52.

Rufai 2017b *{published data only}*

Rufai SB, Singh A, Singh J, Kumar P, Sankar MM, Singh S. Diagnostic usefulness of Xpert MTB/RIF assay for detection of tuberculous meningitis using cerebrospinal fluid. *Journal of Infection* 2017;**75**(2):125-31.

Saeed 2017a *{published data only}*

Saeed M, Ahmad M, Iram S, Riaz S, Akhtar M, Aslam M. GeneXpert technology. A breakthrough for the diagnosis of tuberculous pericarditis and pleuritis in less than 2 hours. *Saudi Medical Journal* 2017;**38**(7):699-705.

Safianowska 2012 *{published data only}*

Safianowska A, Walkiewicz R, Nejman-Gryz P, Grubek-Jaworska H. Two selected commercially based nucleic acid amplification tests for the diagnosis of tuberculosis. *Pneumonologia Alergologia Polska* 2012;**80**(1):6-12.

Scott 2014 *{published data only}*

Scott LE, Beylis N, Nicol M, Nkuna G, Molapo S, Berrie L, et al. Diagnostic accuracy of Xpert MTB/RIF for extrapulmonary tuberculosis specimens: establishing a laboratory testing algorithm for South Africa. *Journal of Clinical Microbiology* 2014;**52**(6):1818-23.

Sharma 2014 *{published data only}*

Sharma SK, Kohli M, Chaubey J, Yadav RN, Sharma R, Singh BK, et al. Evaluation of Xpert MTB/RIF assay performance in diagnosing extrapulmonary tuberculosis among adults in a tertiary care centre in India. *European Respiratory Journal* 2014;**44**(4):1090-3.

Sharma 2016 *{published data only}*

Sharma JB, Kriplani A, Dharmendra S, Chaubey J, Kumar S, Sharma SK. Role of Gene Xpert in diagnosis of female genital tuberculosis: a preliminary report. *European Journal of Obstetrics & Gynecology and Reproductive Biology* 2016;**207**:237-8.

Solomons 2016 *{published data only}*

Solomons RS, Visser DH, Marais BJ, Schoeman JF, van Furth AM. Diagnostic accuracy of a uniform research case definition for TBM in children: a prospective study. *International Journal of Tuberculosis and Lung Disease* 2016;**20**(7):903-8.

Suzana 2016 *{published data only}*

Suzana S, Ninan MM, Gowri M, Venkatesh K, Rupali P, Michael JS. Xpert MTB/Rif for the diagnosis of extrapulmonary tuberculosis - an experience from a tertiary care centre

in South India. *Tropical Medicine & International Health* 2016;**21**(3):385-92.

Tadesse 2015 *{published data only}*

Tadesse M, Abebe G, Abdissa K, Aragaw D, Abdella K, Bekele A, et al. GeneXpert MTB/RIF assay for the diagnosis of tuberculous lymphadenitis on concentrated fine needle aspirates in high tuberculosis burden settings. *PLOS One* 2015;**10**(9):e0137471. [DOI: 10.1371/journal.pone.0137471]

Teo 2011 *{published data only}*

Teo J, Jureen R, Chiang D, Chan D, Lin R. Comparison of two nucleic acid amplification assays, the Xpert MTB/RIF assay and the amplified mycobacterium tuberculosis direct assay, for detection of Mycobacterium tuberculosis in respiratory and nonrespiratory specimens. *Journal of Clinical Microbiology* 2011;**49**(10):3659-62.

Tortoli 2012 *{published data only}*

Tortoli E, Russo C, Piersimoni C, Mazzola E, Dal Monte P, Pascarella M, et al. Clinical validation of Xpert MTB/RIF for the diagnosis of extrapulmonary tuberculosis. *European Respiratory Journal* 2012;**40**(2):442-7.

Trajman 2014 *{published data only}*

Trajman A, da Silva Santos Kleiz de Oliveira EF, Bastos ML, Belo Neto E, Silva EM, da Silva Lourenco MC, et al. Accuracy of polymerase chain reaction for the diagnosis of pleural tuberculosis. *Respiratory Medicine* 2014;**108**(6):918-23.

Ullah 2017 *{published data only}*

Ullah I, Javaid A, Masud H, Ali M, Basit A, Ahmad W, et al. Rapid detection of Mycobacterium tuberculosis and rifampicin resistance in extrapulmonary tuberculosis and sputum smear-negative pulmonary suspects using Xpert MTB/RIF. *Journal of Medical Microbiology* 2017;**66**(4):412-8.

Vadwai 2011 *{published data only}*

Vadwai V, Boehme C, Nabeta P, Shetty A, Alland D, Rodrigues C. Xpert MTB/RIF: a new pillar in diagnosis of extrapulmonary tuberculosis?. *Journal of Clinical Microbiology* 2011;**49**(7):2540-5.

Van Rie 2013 *{published data only}*

Van Rie A, Page-Shipp L, Mellet K, Scott L, Mkhwnazi M, Jong E, et al. Diagnostic accuracy and effectiveness of the Xpert MTB/RIF assay for the diagnosis of HIV-associated lymph node tuberculosis. *European Journal of Clinical Microbiology & Infectious Diseases* 2013;**32**(11):1409-15.

Wang 2016a *{published data only}*

Wang T, Feng GD, Pang Y, Yang YN, Dai W, Zhang L, et al. Sub-optimal specificity of modified Ziehl-Neelsen staining for quick identification of tuberculous meningitis. *Frontiers in Microbiology* 2016;**7**:2096-103.

Zeka 2011 *{published data only}*

Zeka AN, Tasbakan S, Cavusoglu C. Evaluation of the GeneXpert MTB/RIF assay for rapid diagnosis of tuberculosis and detection of rifampin resistance in pulmonary and extrapulmonary specimens. *Journal of Clinical Microbiology* 2011;**49**(12):4138-41.

Zmak 2013 *{published data only}*

Zmak L, Jankovic M, Jankovic VK. Evaluation of Xpert MTB/RIF assay for rapid molecular diagnosis of tuberculosis in a two-year period in Croatia. *International Journal of Mycobacteriology* 2013;**2**(3):179-82.

References to studies excluded from this review

Alvarez Uria 2012 *{published data only}*

Alvarez-Uria G, Azcona JM, Midde M, Naik PK, Reddy S, Reddy R. Rapid diagnosis of pulmonary and extrapulmonary tuberculosis in HIV-infected patients. Comparison of LED fluorescent microscopy and the GeneXpert MTB/RIF assay in a district hospital in India. *Tuberculosis Research and Treatment* 2012;**2012**:932862. [DOI: 10.1155/2012/932862]

Andrey 2015 *{published data only}*

Andrey DO, Hinrikson H, Renzi G, Hibbs J, Adler D, Schrenzel J. Xpert((R)) MTB/RIF assay sensitivity with different methods of CSF processing for the diagnosis of TB meningitis. *International Journal of Tuberculosis and Lung Disease* 2015;**19**(12):1555-6.

Armand 2011 *{published data only}*

Armand S, Vanhuls P, Delcroix G, Courcol R, Lemaitre N. Comparison of the Xpert MTB/RIF test with an IS6110-TaqMan real-time PCR assay for direct detection of Mycobacterium tuberculosis in respiratory and nonrespiratory specimens. *Journal of Clinical Microbiology* 2011;**49**(5):1772-6.

Arockiaraj 2015 *{published data only}*

Arockiaraj J, Amritanand R, Krishnan V, Sundararaj G, Michael J. GeneXpert polymerase chain reaction (PCR) Test: role in the diagnosis of tubercular spondylodiscitis. *Spine Journal* 2015;**15**(10 Suppl 1):200S-201S.

Bablishvili 2015 *{published data only}*

Bablishvili N, Tukvadze N, Avaliani Z, Blumberg HM, Kempker RR. A comparison of the Xpert MTB/RIF and GenoTypeMTBDRplus assays in Georgia. *International Journal of Tuberculosis and Lung Disease* 2015;**19**(6):676-8.

Bajrami 2016 *{published data only}*

Bajrami R, Mulliqi G, Kurti A, Lila G, Raka L. Comparison of GeneXpert MTB/RIF and conventional methods for the diagnosis of tuberculosis in Kosovo. *Journal of Infection in Developing Countries* 2016;**10**(4):418-22.

Balcha 2014 *{published data only}*

Balcha TT, Sturegard E, Winqvist N, Skogmar S, Reepalu A, Jemal ZH, et al. Intensified tuberculosis case-finding in HIV-positive adults managed at Ethiopian health centers: diagnostic yield of Xpert MTB/RIF compared with smear microscopy and liquid culture. *PLOS One* 2014;**9**(1):e85478. [DOI: 10.1371/journal.pone.0085478]

Bemba 2017 *{published data only}*

Bemba ELP, Moukassa D, Ouedraogo AR, Okombi FHO, Bopaka RG, Koumeka PP, et al. Performance of GeneXpert MTB / RIF in the diagnosis of pleural tuberculosis in Brazzaville: a preliminary report [Performance du GeneXpert MTB/RIF dans

le Diagnostic de laTuberculose Pleurale à Brazzaville: Étude Préliminaire]. *Health Sciences and Diseases* 2017;**18**(3):21-7.

Bhatia 2016 *{published data only}*

Bhatia R, Dayal R, Jindal S, Agarwal D, Goyal A. GeneXpert for diagnosis of tubercular meningitis. *Indian Journal of Pediatrics* 2016;**83**(11):1353-5.

Biadglegne 2013 *{published data only}*

Biadglegne F, Tesfaye W, Sack U, Rodloff AC. Tuberculous lymphadenitis in Northern Ethiopia: in a public health and microbiological perspectives. *PLOS One* 2013;**8**(12):e81918. [DOI: 10.1371/journal.pone.0081918]

Bilgin 2016 *{published data only}*

Bilgin K, Yanik K, Karadag A, Odabasi H, Tas H, Gunaydin M. Comparison of a real-time polymerase chain reaction-based system and Erlich-Ziehl-Neelsen method with culture in the identification of Mycobacterium tuberculosis. *Turkish Journal of Medical Sciences* 2016;**46**(1):203-6.

Bunsow 2014 *{published data only}*

Bunsow E, Ruiz-Serrano MJ, Lopez Roa P, Kestler M, Viedma DG, Bouza E. Evaluation of GeneXpert MTB/RIF for the detection of Mycobacterium tuberculosis and resistance to rifampin in clinical specimens. *Journal of Infection* 2014;**68**(4):338-43.

Celik 2015 *{published data only}*

Celik C, Gozel MG, Bakici MZ, Berk S, Ozsahin SL, Gulturk E. Applicability of Xpert MTB/RIF assay for routine diagnosis of tuberculosis: a four-year single-center experience. *Turkish Journal of Medical Sciences* 2015;**45**(6):1329-34.

Chen 2016 *{published data only}*

Chen ZF, Lao HL, Li XH, Wang J, Chen Q, Wang ZX, et al. Experimental study of GeneXpert((R)) system in the diagnosis of extra-pulmonary tuberculosis [Chinese]. *Chinese Journal of Tuberculosis and Respiratory Diseases* 2016;**39**(7):529-33.

Coleman 2015 *{published data only}*

Coleman M, Finney LJ, Komrower D, Chitani A, Bates J, Chipungu GA, et al. Markers to differentiate between Kaposi's sarcoma and tuberculous pleural effusions in HIV-positive patients. *International Journal of Tuberculosis and Lung Disease* 2015;**19**(2):144-50.

Deggim 2013 *{published data only}*

Deggim V, Somoskovi A, Voit A, Bottger EC, Bloemberg GV. Integrating the Xpert MTB/RIF assay into a diagnostic workflow for rapid detection of Mycobacterium tuberculosis in a low-prevalence area. *Journal of Clinical Microbiology* 2013;**51**(7):2396-9.

Dharan 2016 *{published data only}*

Dharan NJ, Blakemore R, Sloutsky A, Kaur D, Alexander RC, Ghajar M, et al. Performance of the G4 Xpert MTB/RIF assay for the detection of Mycobacterium tuberculosis and rifampin resistance: a retrospective case-control study of analytical and clinical samples from high- and low-tuberculosis prevalence settings. *BMC Infectious Diseases* 2016;**16**:764. [DOI: 10.1186/s12879-016-2039-4]

Diop 2016 *{published data only}*

Diop SA, Massaly A, Ka D, Manga NM, Fortes-Deguenonvo L, Ndour CT, et al. Use of GeneXpert test for the diagnosis of tuberculosis in the department of infectious diseases of FANN University Hospital [Utilisation du test GeneXpert pour le diagnostic de la tuberculose au service desmaladies infectieuses du CHNU de Fann]. *Pan African Medical Journal* 2016;**23**:244. [DOI: 10.11604/pamj.2016.23.244.7442]

Edwards 2016 *{published data only}*

Edwards S, Glynn P, David MD, Kamesh L. Diagnosing tuberculous peritonitis early in patients on peritoneal dialysis: use of Xpert MTB/RIF assay. *Peritoneal Dialysis International* 2016;**36**(4):461.

Erdem 2014 *{published data only}*

Erdem H, Ozturk-Engin D, Elaldi N, Gulsun S, Sengoz G, Crisan A, et al. The microbiological diagnosis of tuberculous meningitis: results of Haydarpasa-1 study. *Clinical Microbiology and Infection* 2014;**20**(10):O600-8. [DOI: 10.1111/1469-0691.12478]

Fanosie 2016 *{published data only}*

Fanosie A, Gelaw B, Tessema B, Tesfay W, Admasu A, Yitayew G. Mycobacterium tuberculosis complex and HIV co-infection among extrapulmonary tuberculosis suspected cases at the University of Gondar Hospital, Northwestern Ethiopia. *PLOS One* 2016;**11**(3):e0150646. [DOI: 10.1371/journal.pone.0150646]

Gascoyne-Binzi 2012 *{published data only}*

Gascoyne-Binzi DM, Robinson GA, English A, Collyns TA. Comparison of on-demand PCR testing for Mycobacterium tuberculosis with culture and batched PCR. *Clinical Microbiology and Infection* 2012;**18**:543. [DOI: https://doi.org/10.1111/j.1469-0691.2012.03802.x]

Habeenzu 2017 *{published data only}*

Habeenzu C, Nakajima C, Solo E, Bwalya P, Kajino K, Miller M, et al. Evaluation of in-house loop-mediated isothermal amplification for tuberculosis diagnosis compared with Xpert MTB/RIF. *Journal of Infection in Developing Countries* 2017;**11**(6):440-4.

Ioannidis 2010 *{published data only}*

Ioannidis P, Papaventsis D, Nikolaou S, Karabela S, Konstantinidou E, Marinou I, et al. Tuberculosis resistance detection rate to the two main anti-TB drugs, isoniazid and rifampicin, using molecular techniques: experience of the Hellenic National Reference Center for Mycobacteria. [Greek]. *Acta Microbiologica Hellenica* 2010;**55**(2):175-82.

Jain 2017 *{published data only}*

Jain A, Singh PK, Singh U, Kumar V. Initial screening of extra-pulmonary tuberculosis using the Xpert MTB/RIF assay improves case detection rates. *International Journal of Tuberculosis and Lung Disease* 2017;**21**(4):478-80.

Kilfoil 2015 *{published data only}*

Kilfoil KM, Mayne E, Scott L, Stevens W. A high burden human immunodeficiency virus and tuberculosis resource limited setting, gains from Including Xpert MTB/RIF in the diagnostic algorithm of fluid specimens submitted for exclusion of

lymphoma by immunophenotypic analysis. *PLOS One* 2015;**10**(8):e0134404. [DOI: 10.1371/journal.pone.0134404]

Kim 2014 *{published data only}*

Kim CH, Woo H, Hyun IG, Kim C, Choi JH, Jang SH, et al. A comparison between the efficiency of the Xpert MTB/RIF assay and nested PCR in identifying Mycobacterium tuberculosis during routine clinical practice. *Journal of Thoracic Disease* 2014;**6**(6):625-31.

Kim 2015b *{published data only}*

Kim MJ, Nam YS, Cho SY, Park TS, Lee HJ. Comparison of the Xpert MTB/RIF Assay and real-time PCR for the detection of Mycobacterium tuberculosis. *Annals of Clinical and Laboratory Science* 2015;**45**(3):327-32.

Kim 2015c *{published data only}*

Kim CH, Hyun IG, Hwang YI, Kim DG, Lee CY, Lee MG, et al. Identification of Mycobacterium tuberculosis and rifampin resistance in clinical specimens using the Xpert MTB/RIF assay. *Annals of Clinical and Laboratory Science* 2015;**45**(1):32-8.

Kumar 2017 *{published data only}*

Kumar S, Bopanna S, Kedia S, Mouli P, Dhingra R, Padhan R, et al. Evaluation of Xpert MTB/RIF assay performance in the diagnosis of abdominal tuberculosis. *Intestinal Research* 2017;**15**(2):187-94.

Kurbaniyazova 2017 *{published data only}*

Kurbaniyazova G, Joncevska M, Kalon S, Kalmambetova G, Mohr T, Toktogonova A, et al. Results of Xpert MTB/RIF implementation in Kyrgyzstan. *International Journal of Tuberculosis and Lung Disease* 2017;**21**(3):333-7.

Kwak 2015 *{published data only}*

Kwak N, Seong MW, Kim EC, Yoo CG, Kim YW, Han SK, et al. Accuracy of the Xpert MTB/RIF assay for the diagnosis of extrapulmonary tuberculosis in South Korea. *International Journal of Tuberculosis and Lung Disease* 2015;**19**(1):81-6.

Lawn 2012 *{published data only}*

Lawn SD, Kerkhoff AD, Vogt M, Wood R. High diagnostic yield of tuberculosis from screening urine samples from HIV-infected patients with advanced immunodeficiency using the Xpert MTB/RIF assay. *Journal of Acquired Immune Deficiency Syndromes* 2012;**60**(3):289-94.

Lawn 2013 *{published data only}*

Lawn SD, Kerkhoff AD, Vogt M, Wood R. HIV-associated tuberculosis: relationship between disease severity and the sensitivity of new sputum-based and urine-based diagnostic assays. *BMC Medicine* 2013;**11**:231. [DOI: 10.1186/1741-7015-11-231]

Lawn 2015 *{published data only}*

Lawn SD, Kerkhoff AD, Burton R, Schutz C, van Wyk G, Vogt M, et al. Rapid microbiological screening for tuberculosis in HIV-positive patients on the first day of acute hospital admission by systematic testing of urine samples using Xpert MTB/RIF: a prospective cohort in South Africa. *BMC Medicine* 2015;**13**:192. [DOI: 10.1186/s12916-015-0432-2]

Lawn 2017 *{published data only}*

Lawn SD, Kerkhoff AD, Burton R, Schutz C, Boulle A, Vogt M, et al. Diagnostic accuracy, incremental yield and prognostic value of Determine TB-LAM for routine diagnostic testing for tuberculosis in HIV-infected patients requiring acute hospital admission in South Africa: a prospective cohort. *BMC Medicine* 2017;**15**(1):67. [DOI: 10.1186/s12916-017-0822-8]

Lee 2017 *{published data only}*

Lee J, Choi SM, Lee CH, Lee SM, Yim JJ, Yoo CG, et al. The additional role of Xpert MTB/RIF in the diagnosis of intrathoracic tuberculous lymphadenitis. *Journal of Infection and Chemotherapy* 2017;**23**(6):381-4.

Liu 2015 *{published data only}*

Liu X, Huang Z, Du J. Rapid diagnosis of pleural tuberculosis by Xpert MTB/RIF assay. *Zhonghua Jie He He Hu Xi Za Zhi [Chinese Journal of Tuberculosis and Respiratory Diseases]* 2015;**38**(10):741-5.

Lombardi 2017 *{published data only}*

Lombardi G, Di Gregori V, Girometti N, Tadolini M, Bisognin F, Dal Monte P. Diagnosis of smear-negative tuberculosis is greatly improved by Xpert MTB/RIF. *PLOS One* 2017;**12**(4):e0176186. [DOI: 10.1371/journal.pone.0176186]

Marouane 2014 *{published data only}*

Marouane C, Smaoui S, Kammoun S, Slim L, Messadi-Akrout F. Evaluation of GeneXpert MTB/RIF for the detection of Mycobacterium tuberculosis and resistance to rifampin in extra-pulmonary specimens. *International Journal of Mycobacteriology* 2014;**4 Suppl 1**:101.

Marouane 2016 *{published data only}*

Marouane C, Smaoui S, Kammoun S, Slim L, Messadi-Akrout F. Evaluation of molecular detection of extrapulmonary tuberculosis and resistance to rifampicin with GeneXpert(R) MTB/RIF. *Medecine et Maladies Infectieuses* 2016;**46**(1):20-4.

Miller 2011 *{published data only}*

Miller MB, Popowitch EB, Backlund MG, Ager EP. Performance of Xpert MTB/RIF RUO assay and IS6110 real-time PCR for Mycobacterium tuberculosis detection in clinical samples. *Journal of Clinical Microbiology* 2011;**49**(10):3458-62.

Mishra 2017 *{published data only}*

Mishra DR, Bhatta N, Lamsal M, Bhattarai N, Maskey R. Study of diagnostic utility of Xpert MTB/RIF test on pleural fluid in the evaluation of patients presenting with pleural tuberculosis in high prevalence setting. *American Journal of Respiratory and Critical Care Medicine* 2017;**195**:A2088. [https://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2017.195.1_MeetingAbstracts.A2088]

Moure 2011 *{published data only}*

Moure R, Munoz L, Torres M, Santin M, Martin R, Alcaide F. Rapid detection of Mycobacterium tuberculosis complex and rifampin resistance in smear-negative clinical samples by use of an integrated real-time PCR method. *Journal of Clinical Microbiology* 2011;**49**(3):1137-9.

Moure 2012 *{published data only}*

Moure R, Martin R, Alcaide F. Effectiveness of an integrated real-time PCR method for detection of the Mycobacterium tuberculosis complex in smear-negative extrapulmonary samples in an area of low tuberculosis prevalence. *Journal of Clinical Microbiology* 2012;**50**(2):513-5.

Nhu 2013 *{published data only}*

Nhu NTQ, Ha DTM, Anh ND, Thu DDA, Duong TN, Quang ND, et al. Evaluation of Xpert MTB/RIF and MODS assay for the diagnosis of pediatric tuberculosis. *BMC Infectious Diseases* 2013;**13**:31. [DOI: 10.1186/1471-2334-13-31]

Patel 2014 *{published data only}*

Patel VB, Connolly C, Singh R, Lenders L, Matinyenya B, Theron G, et al. Comparison of amplicor and GeneXpert MTB/RIF tests for diagnosis of tuberculous meningitis. *Journal of Clinical Microbiology* 2014;**52**(10):3777-80.

Peter 2012 *{published data only}*

Peter JG, Theron G, Muchinga TE, Govender U, Dheda K. The diagnostic accuracy of urine-based Xpert MTB/RIF in HIV-infected hospitalized patients who are smear-negative or sputum scarce. *PLOS One* 2012;**7**(7):e39966. [DOI: 10.1371/journal.pone.0039966]

Porcel 2013 *{published data only}*

Porcel JM, Palma R, Valdes L, Bielsa S, San-Jose E, Esquerda A. Xpert(R) MTB/RIF in pleural fluid for the diagnosis of tuberculosis. *International Journal of Tuberculosis and Lung Disease* 2013;**17**(9):1217-9.

Rachow 2012 *{published data only}*

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

Raizada 2015 *{published data only}*

Raizada N, Sachdeva KS, Swaminathan S, Kulsange S, Khaparde SD, Nair SA, et al. Piloting upfront Xpert MTB/RIF testing on various specimens under programmatic conditions for diagnosis of TB & DR-TB in paediatric population. *PLOS One* 2015;**10**(10):e0140375. [DOI: 10.1371/journal.pone.0140375]

Ramamurthy 2016 *{published data only}*

Ramamurthy K, Bhat S, Shenoy S, Rangnekar A. Xpert Mycobacterium tuberculosis/rifampicin assay: a boon in tuberculosis diagnostics. *Asian Journal of Pharmaceutical and Clinical Research* 2016;**9**(5):225-7.

Razack 2014 *{published data only}*

Razack R, Louw M, Wright CA. Diagnostic yield of fine needle aspiration biopsy in HIV-infected adults with suspected mycobacterial lymphadenitis. *South African Medical Journal* 2014;**104**(1):27-8.

Saeed 2017b *{published data only}*

Saeed M, Iram S, Hussain S, Ahmed A, Akbar M, Aslam M. GeneXpert: a new tool for the rapid detection of rifampicin resistance in mycobacterium tuberculosis. *Journal of the Pakistan Medical Association* 2017;**67**(2):270-4.

Salvador 2015 *{published data only}*

Salvador F, Los-Arcos I, Sanchez-Montalva A, Tortola T, Curran A, Villar A, et al. Epidemiology and diagnosis of tuberculous lymphadenitis in a tuberculosis low-burden country. *Medicine* 2015;**94**(4):e509. [DOI: 10.1097/MD.0000000000000509]

Sanjuan Jimenez 2015 *{published data only}*

Sanjuan-Jimenez R, Toro-Peinado I, Bermudez P, Colmenero JD, Morata P. Comparative study of a real-time PCR assay targeting senX3-regX3 versus other molecular strategies commonly used in the diagnosis of tuberculosis. *PLOS One* 2015;**10**(11):e0143025. [DOI: 10.1371/journal.pone.0143025]

Shah 2016a *{published data only}*

Shah I, Gupta Y, Chhina AS, Shenoi A, Kumar RK, Patel A. Xpert MTB/RIF for diagnosis of tuberculosis and drug resistance in Indian children. *Indian Pediatrics* 2016;**53**(9):837-9.

Singanayagam 2014 *{published data only}*

Singanayagam A, Donaldson H, Kon OM. GeneXpert™ MTB/RIF in low prevalence settings: a UK laboratory perspective. *Journal of Infection* 2014;**69**(2):199-200.

Singh 2016 *{published data only}*

Singh UB, Pandey P, Mehta G, Bhatnagar AK, Mohan A, Goyal V, et al. Genotypic, phenotypic and clinical validation of GeneXpert in extra-pulmonary and pulmonary tuberculosis in India. *PLOS One* 2016;**11**(2):e0149258. [DOI: 10.1371/journal.pone.0149258]

Smith 2014 *{published data only}*

Smith P, van Esch A, Wallace M, Wood R, Bekker LG. GeneXpert TB 8: a point-of-care diagnostic pilot. *South African Medical Journal* 2014;**104**(8):524.

Solomons 2015 *{published data only}*

Solomons RS, Visser DH, Friedrich SO, Diacon AH, Hoek KG, Marais BJ, et al. Improved diagnosis of childhood tuberculous meningitis using more than one nucleic acid amplification test. *International Journal of Tuberculosis and Lung Disease* 2015;**19**(1):74-80.

Theron 2014 *{published data only}*

Theron G, Peter J, Calligaro G, Meldau R, Hanrahan C, Khalfey H, et al. Determinants of PCR performance (Xpert MTB/RIF), including bacterial load and inhibition, for TB diagnosis using specimens from different body compartments. *Scientific Reports* 2014;**4**:5658. [DOI: 10.1038/srep05658]

Toure 2017 *{published data only}*

Toure NO, Wayzani M, Thiam K, Cisse MF, Mbaye FB. Contribution of the Xpert MTB/RIF to the etiological diagnosis of tuberculous pleurisy [Apport de l'Xpert MTB/RIF dans le diagnostic étiologique des pleurésies tuberculeuses]. *Revue des Maladies Respiratoires* 2017;**34**(7):758-64. [DOI: 10.1016/j.rmr.2017.01.003]

Vallejo 2015 *{published data only}*

Vallejo VP, Searle MA, Rodriguez DJC, Farga CV. Xpert® MTB/RIF assay for tuberculosis diagnosis [Ensayo Xpert MTB/RIF en el diagnóstico de tuberculosis]. *Revista Chilena de Enfermedades Respiratorias* 2015;**31**(2):127-31.

Verghese 2016 *{published data only}*

Verghese VP, Thomas L, Michael JS, Rose W, Jeyaseelan V. Accuracy of the Xpert MTB/RIF assay compared to the "gold standard" AFB culture in the diagnosis of tuberculosis in children in India. *International Journal of Infectious Diseases* 2016;**45**:343.

Wang 2016 *{published data only}*

Wang T, Feng GD, Pang Y, Liu JY, Zhou Y, Yang YN, et al. High rate of drug resistance among tuberculous meningitis cases in Shaanxi province, China. *Scientific Reports* 2016;**6**:25251. [DOI: 10.1038/srep25251]

Wei 2016 *{published data only}*

Wei G, Mu J, Wang G, Huo F, Dong L, Li Y, et al. The reliability analysis of Xpert-positive result for smear-negative and culture-negative specimen collected from bone and joint tuberculosis suspects. *Journal of Thoracic Disease* 2016;**8**(6):1205-9.

Yuan 2016 *{published data only}*

Yuan M, Lyu Y, Chen ST, Cai C, Li Y, Zhang ZG, et al. Evaluation of Xpert MTB/RIF for the diagnosis of extrapulmonary tuberculosis in China. *Biomedical and Environmental Sciences* 2016;**29**(8):599-602.

Zhang 2016 *{published data only}*

Zhang AM, Li F, Liu XH, Xia L, Lu SH. Application of Gene Xpert Mycobacterium tuberculosis DNA and resistance to rifampicin assay in the rapid detection of tuberculosis in children. *Zhonghua Er Ke Za Zhi [Chinese Journal of Pediatrics]* 2016;**54**(5):370-4.

Additional references

American Thoracic Society 2000

American Thoracic Society, the Centers for Disease Control and Prevention, Infectious Disease Society of America. Diagnostic Standards and Classification of Tuberculosis in Adults and Children. This official statement of the American Thoracic Society and the Centers for Disease Control and Prevention was adopted by the ATS Board of Directors, July 1999. This statement was endorsed by the Council of the Infectious Disease Society of America, September 1999. *American Journal Respiratory and Critical Care Medicine* 2000;**161**(4 Pt 1):1376-95.

Bahr 2016

Bahr NC, Marais S, Caws M, van Crevel R, Wilkinson RJ, Tyagi JS, et al. Tuberculous Meningitis International Research Consortium. GeneXpert MTB/Rif to diagnose tuberculous meningitis: perhaps the first test but not the last. *Clinical Infectious Diseases* 2016;**62**(9):1133-5.

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.

Banada 2010

Banada PP, Sivasubramani SK, Blakemore R, Boehme C, Perkins MD, Fennelly K, et al. Containment of bioaerosol infection risk by the Xpert MTB/RIF assay and its applicability to point-of-care settings. *Journal of Clinical Microbiology* 2010;**48**(10):3551-7.

Begg 1983

Begg CB, Greenes RA. Assessment of diagnostic tests when disease verification is subject to selection bias. *Biometrics* 1983;**39**(1):207-15.

Blakemore 2010

Blakemore R, Story E, Helb D, Kop J, Banada P, Owens MR, et al. Evaluation of the analytical performance of the Xpert MTB/RIF assay. *Journal of Clinical Microbiology* 2010;**48**(7):2495-501.

Bossuyt 2015

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015;**351**:h5527. [DOI: 10.1136/bmj.h5527]

Boyles 2014

Boyles TH, Hughes J, Cox V, Burton R, Meintjes G, Mendelson M. False-positive Xpert® MTB/RIF assays in previously treated patients: need for caution in interpreting results. *International Journal Tuberculosis and Lung Disease* 2014;**18**(7):876-8.

Brozek 2009

Brozek JL, Akl EA, Jaeschke R, Lang DM, Bossuyt P, Glasziou P, et al. GRADE Working Group. Grading quality of evidence and strength of recommendations in clinical practice guidelines: Part 2 of 3. The GRADE approach to grading quality of evidence about diagnostic tests and strategies. *Allergy* 2009;**64**(8):1109-16.

CDC 2015

Centers for Disease Control and Prevention. Reported Tuberculosis in the United States, 2015. www.cdc.gov/tb/statistics/reports/2015/default.htm (accessed 8 November 2017).

Cepheid 2015

Cepheid. Xpert® MTB/RIF. Two-hour detection of MTB and rifampin resistance mutations. www.cepheid.com/us/cepheid-solutions/clinical-ivd-tests/critical-infectious-diseases/xpert-mtb-rif (accessed 21 September 2016).

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):e00812-17. [DOI: 10.1128/mBio.00812-17]

Chang 2012

Chang K, Lu W, Wang J, Zhang K, Jia S, Li F, et al. Rapid and effective diagnosis of tuberculosis and rifampicin resistance with Xpert MTB/RIF assay: a meta-analysis. *Journal of Infection* 2012;**64**(6):580-8.

Chow 2002

Chow KM, Chow VC, Hung LC, Wong SM, Szeto CC. Tuberculous peritonitis-associated mortality is high among patients waiting for the results of mycobacterial cultures of ascitic fluid samples. *Clinical Infectious Diseases* 2002;**35**(4):409-13.

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.

Chu 2009

Chu H, Chen S, Louis TA. Random effects models in a meta-analysis of the accuracy of two diagnostic tests without a gold standard. *Journal of the American Statistical Association* 2009;**104**(486):512–23.

Conde 2003

Conde MB, Loivos AC, Rezende VM, Soares SL, Mello FC, Reingold AL, et al. Yield of sputum induction in the diagnosis of pleural tuberculosis. *American Journal of Respiratory and Critical Care Medicine* 2003;**167**(5):723-5.

Covidence 2017 [Computer program]

Veritas Health Innovation. Covidence systematic review software. Available at www.covidence.org. Melbourne: Veritas Health Innovation, 2017.

Dendukuri 2012

Dendukuri N, Schiller I, Joseph L, Pai M. Bayesian meta-analysis of the accuracy of a test for tuberculous pleuritis in the absence of a gold standard reference. *Biometrics* 2012;**68**(4):1285–93.

Denkinger 2014

Denkinger CM, Schumacher SG, Boehme CC, Dendukuri N, Pai M, Steingart KR. Xpert MTB/RIF assay for the diagnosis of extrapulmonary tuberculosis: a systematic review and meta-analysis. *European Respiratory Journal* 2014;**44**(2):435-46.

Denkinger 2015

Denkinger CM, Kik SV, Cirillo DM, Casenghi M, Shinnick T, Weyer K, et al. Defining the needs for next generation assays for tuberculosis. *Journal of Infectious Diseases* 2015;**211 Suppl 2**:S29-38.

Diacon 2003

Diacon AH, Van de Wal BW, Wyser C, Smedema JP, Bezuidenhout J, Bolliger CT, et al. Diagnostic tools in tuberculous pleurisy: a direct comparative study. *European Respiratory Journal* 2003;**22**(4):589-91.

Dinnes 2007

Dinnes J, Deeks J, Kunst H, Gibson A, Cummins E, Waugh N, et al. A systematic review of rapid diagnostic tests for the detection of tuberculosis infection. *Health Technology Assessment* 2007;**11**(3):1-196.

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.

FDA 2013

United States Food, Drug Administration. Decision summary K131706. www.accessdata.fda.gov/cdrh_docs/reviews/k131706.pdf (accessed 29 October 2016).

FIND 2017

FIND. Report for WHO: a multicentre non-inferiority diagnostic accuracy study of the Ultra assay compared to the Xpert MTB/RIF assay. 2017. Version 1.8. www.finddx.org/publication/ultra-report/ (accessed 17 January 2018):50-1.

Fontanilla 2011

Fontanilla JM, Barnes A, von Reyn CF. Current diagnosis and management of peripheral tuberculous lymphadenitis. *Clinical Infectious Diseases* 2011;**53**(6):555-62.

GLI 2017

Global Laboratory Initiative. GLI practical guide to TB laboratory strengthening, 2017. www.stoptb.org/wg/gli/gat.asp (accessed 10 October 2017).

GLI 2018

Global Laboratory Initiative. GLI model TB diagnostic algorithms, 2018. www.stoptb.org/wg/gli/gat.asp (accessed 11 August 2018).

Golden 2005

Golden MP, Vikram HR. Extrapulmonary tuberculosis: an overview. *American Family Physician* 2005;**72**(9):1761-8.

Gopinath 2010

Gopinath K, Singh S. Non-tuberculous mycobacteria in TB-endemic countries: are we neglecting the danger?. *PLOS Neglected Tropical Diseases* 2010;**4**(4):e615. [DOI: 10.1371/journal.pntd.0000615]

GRADE 2013

GRADE Working Group, Schünemann H, Brozek J, Guyatt G, Oxman A, editors. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. http://gdt.guidelinedevelopment.org/app/handbook/handbook.html (accessed 18 July 2017).

GRADEpro GDT 2015 [Computer program]

McMaster University (developed by Evidence Prime). GRADEpro GDT. Version accessed 29 October 2016. Hamilton (ON): McMaster University (developed by Evidence Prime), 2015.

Gupta 2015

Gupta RK, Lucas SB, Fielding KL, Lawn SD. Prevalence of tuberculosis in post-mortem studies of HIV-infected adults and

children in resource-limited settings: a systematic review and meta-analysis. *AIDS* 2015;**29**(15):1987-2002.

Gupta-Wright 2016

Gupta-Wright A, Peters JA, Flach C, Lawn SD. Detection of lipoarabinomannan (LAM) in urine is an independent predictor of mortality risk in patients receiving treatment for HIV-associated tuberculosis in sub-Saharan Africa: a systematic review and meta-analysis. *BMC Medicine* 2016;**141**(1):53.

Gupta-Wright 2018

Gupta-Wright A, Corbett EL, van Oosterhout JJ, Wilson DK, Grint D, Alufandika-Moyo M, et al. Urine-based screening for tuberculosis: a randomized trial in HIV-positive inpatients (the STAMP trial). Conference on retroviruses and opportunistic infections (CROI). www.croiconference.org/sessions/urine-based-screening-tuberculosis-randomized-trial-hiv-positive-inpatients. Boston, 2018:38LB.

Harbord 2007

Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007;**8**(2):239-51.

Hartmann 1967

Hartmann G, Honikel KO, Knüsel F, Nüesch J. The specific inhibition of the DNA-directed RNA synthesis by rifamycin. *Biochimica et Biophysica Acta* 1967;**145**(3):843–4.

Helb 2010

Helb D, Jones M, Story E, Boehme C, Wallace E, Ho K, et al. Rapid detection of Mycobacterium tuberculosis and rifampin resistance by use of on-demand, near-patient technology. *Journal of Clinical Microbiology* 2010;**48**(1):229-37.

Index-TB 2016

Central TB Division, Ministry of Health and Family Welfare, Government of India. Index-TB guidelines: guidelines on extra-pulmonary TB for India. http://icmr.nic.in/guidelines/TB/Index-TB%20Guidelines%20-%20green%20colour%202594164.pdf (accessed 6 October 2016).

Iseman 2000

Iseman MD. Extrapulmonary tuberculosis in adults. In: Iseman MD editor(s). A Clinician's Guide to Tuberculosis. Philadelphia: Lippincott Williams and Wilkins, 2000:145–97.

Kelly 2014

Kelly JD, Grace Lin SY, Barry PM, Keh C, Higashi J, Metcalfe JZ. Xpert MTB/RIF false detection of rifampin-resistant tuberculosis from prior infection. *American Journal of Respiratory and Critical Care Medicine* 2014;**190**(11):1316-8.

Kilpatrick 1986

Kilpatrick ME, Girgis NI, Yassin MW, Abu el Ella AA. Tuberculous meningitis-clinical and laboratory review of 100 patients. *Journal of Hygiene (London)* 1986;**96**(2):231–8.

Leeflang 2013

Leeflang MM, Rutjes AW, Reitsma JB, Hooft L, Bossuyt PM. Variation of a test's sensitivity and specificity with disease prevalence. *Canadian Medical Association Journal* 2013;**185**(11):E537-44. [DOI: 10.1503/cmaj.121286]

Lewinsohn 2017

Lewinsohn DM, Leonard MK, LoBue PA, Cohn DL, Daley CL, Desmond E, et al. Official American Thoracic Society/Infectious Diseases Society of America/Centers for Disease Control and Prevention Clinical Practice Guidelines: Diagnosis of Tuberculosis in Adults and Children. *Clinical Infectious Diseases* 2017;**64**(2):e1–33. [DOI: 10.1093/cid/ciw694]

Li Y 2017

Li S, Liu B, Peng M, Chen M, Yin W, Tang H, et al. Diagnostic accuracy of Xpert MTB/RIF for tuberculosis detection in different regions with different endemic burden: a systematic review and meta-analysis. *PLOS One* 2017;**12**(7):e0180725. [DOI: 10.1371/journal.pone.0180725]

Lunn 2009

Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: evolution, critique, and future directions. *Statistics in Medicine* 2009;**28**(25):3049-67.

Macaskill 2010

Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and presenting results. In: Deeks JJ, Bossuyt PM, C Gatsonis (editors). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 2010. http://srdta.cochrane.org/ (accessed 29 October 2016).

Maynard-Smith 2014

Maynard-Smith L, Larke N, Peters JA, Lawn SD. Diagnostic accuracy of the Xpert MTB/RIF assay for extrapulmonary and pulmonary tuberculosis when testing non-respiratory samples: a systematic review. *BMC Infectious Diseases* 2014;**14**:709. [DOI: 10.1186/s12879-014-0709-7]

McGrath 2017

McGrath TA, Alabousi M, Skidmore B, Korevaar DA, Bossuyt PMM, Moher D, et al. Recommendations for reporting of systematic reviews and meta-analyses of diagnostic test accuracy: a systematic review. *Systematic Reviews* 2017;**6**(1):194. [DOI: 10.1186/s13643-017-0590-8]

Miotto 2012

Miotto P, Bigoni S, Migliori GB, Matteelli A, Cirillo DM. Early tuberculosis treatment monitoring by Xpert(R) MTB/RIF. *European Respiratory Journal* 2012;**39**(5):1269-71.

Nahid 2016

Nahid P, Dorman SE, Alipanah N, Barry PM, Brozek JL, Cattamanchi A, et al. Official American Thoracic Society/Centers for Disease Control and Prevention/Infectious Diseases Society of America clinical practice guidelines: treatment of drug-susceptible tuberculosis. *Clinical Infectious Diseases* 2016;**63**(7):e147-95. [DOI: 10.1093/cid/ciw376]

Nathavitharana 2017

Nathavitharana RR, Cudahy PG, Schumacher SG, Steingart KR, Pai M, Denkinger CM. Accuracy of line probe assays for the

diagnosis of pulmonary and multidrug-resistant tuberculosis: a systematic review and meta-analysis. *European Respiratory Journal* 2017;**49**(1):pii: 1601075.

Nelson 2004

Nelson LJ, Wells CD. Global epidemiology of childhood tuberculosis. *International Journal of Tuberculosis and Lung Disease* 2004;**8**(5):636-47.

Pai 2004

Pai M, Flores LL, Hubbard A, Riley LW, Colford Jr JM. Nucleic acid amplification tests in the diagnosis of tuberculous pleuritis: a systematic review and meta-analysis. *BMC Infectious Diseases* 2004;**4**:6. [DOI: 10.1186/1471-2334-4-6]

Pai 2008

Pai M, Ling DI. Rapid diagnosis of extrapulmonary tuberculosis using nucleic acid amplification tests: what is the evidence?. *Future Microbiology* 2008;**3**(1):1-4.

Pai 2016

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

Penz 2015

Penz E, Boffa J, Roberts DJ, Fisher D, Cooper R, Ronksley PE, et al. Diagnostic accuracy of the Xpert® MTB/RIF assay for extra-pulmonary tuberculosis: a meta-analysis. *International Journal of Tuberculosis and Lung Disease* 2015;**19**(3):278-84.

Perkins 2007

Perkins MD, Cunningham J. Facing the crisis: improving the diagnosis of tuberculosis in the HIV era. *Journal of Infectious Diseases* 2007;**196 Suppl 1**:S15–27.

Peter 2010

Peter J, Green C, Hoelscher M, Mwaba P, Zumla A, Dheda K. Urine for the diagnosis of tuberculosis: current approaches, clinical applicability, and new developments. *Current Opinion in Pulmonary Medicine* 2010;**16**(3):262-70.

Peter 2016

Peter JG, Zijenah LS, Chanda D, Clowes P, Lesosky M, Gina P, et al. Effect on mortality of point-of-care, urine-based lipoarabinomannan testing to guide tuberculosis treatment initiation in HIV-positive hospital inpatients: a pragmatic, parallel-group, multicountry, open-label, randomised controlled trial. *Lancet* 2016;**387**(10024):1187-97.

R Core Team 2016 [Computer program]

R Core Team. R: A language and environment for statistical computing. www.R-project.org. Vienna: R Foundation for Statistical Computing, 2016.

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 2009

Reuter H, Wood R, Schaaf HS, Donald PR. Overview of extrapulmonary tuberculosis in adults and children. In: Schaaf HS, Zumla A editor(s). Tuberculosis: A Comprehensive Clinical Reference. 1st Edition. Amsterdam: Elsevier Science Publishers, 2009:377-90.

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.

Sandgren 2013

Sandgren A, Hollo V, van der Werf MJ. Extrapulmonary tuberculosis in the European Union and European economic area, 2002 to 2011. *Euro Surveillance* 2013;**18**(12):pii: 20431.

Schiller 2016

Schiller I, van Smeden M, Hadgu A, Libman M, Reitsma JB, Dendukuri N. Bias due to composite reference standards in diagnostic accuracy studies. *Statistics in Medicine* 2016;**35**(9):1454-70.

Schumacher 2016

Schumacher S, van Smeden M, Dendukuri N, Joseph L, Nicol M, Pai M, et al. Diagnostic test accuracy in childhood pulmonary tuberculosis: a Bayesian latent class analysis. *American Journal of Epidemiology* 2016;**184**(9):690-700.

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 2011

Schünemann HJ, Oxman AD, Gunn EV, Higgins JPT, Deeks JL, Glasziou P, et al. Chapter 11. Presenting results and "Summary of findings" tables. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Sehgal 2016

Sehgal IS, Dhooria S, Aggarwal AN, Behera D, Agarwal R. Diagnostic performance of Xpert MTB/RIF in tuberculous pleural effusion: systematic review and meta-analysis. *Journal of Clinical Microbiology* 2016;**54**(4):1133-6.

Shah 2016b

Shah M, Hanrahan C, Wang ZY, Dendukuri N, Lawn SD, Denkinger CM, Steingart KR. 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]

Sharma 2004

Sharma SK, Mohan A. Extrapulmonary tuberculosis. *Indian Journal of Medical Research* 2004;**120**(4):316-53.

Shu 2011

Shu CC, Wang JT, Wang JY, Lee LN, Yu CJ. In-hospital outcome of patients with culture-confirmed tuberculous pleurisy: clinical impact of pulmonary involvement. *BMC Infectious Diseases* 2011;**11**:46. [DOI: 10.1186/1471-2334-11-46]

Stata 2011 [Computer program]

StataCorp. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP, 2011.

Steingart 2014

Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database of Systematic Reviews* 2014, Issue 1. [DOI: 10.1002/14651858.CD009593.pub3]

TB Care I 2014

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

Telenti 1993

Telenti A, Imboden P, Marchesi F, Lowrie D, Cole S, Colston MJ, et al. Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. *Lancet* 1993;**341**(8846):647-50.

Theron 2016

Theron G, Venter R, Calligaro G, Smith L, Limberis J, Meldau R, et al. Xpert MTB/RIF results in patients with previous tuberculosis: can we distinguish true from false positive results?. *Clinical Infectious Diseases* 2016;**62**:995–1001.

Thwaites 2013

Thwaites GE, van Toorn R, Schoeman J. Tuberculous meningitis: more questions, still too few answers. *Lancet Neurology* 2013;**12**(10):999-1010.

Unitaid 2017

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

Van Deun 2004

Van Deun. What is the role of mycobacterial culture in diagnosis and case definition?. Frieden T, ed. Toman's Tuberculosis: Case Detection, Treatment, and Monitoring - Questions and Answers. 2nd Edition. Geneva: World Health Organization, 2004:7–13.

Van Smeden 2014

van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KG, de Groot JA. Latent class models in diagnostic studies when there is no reference standard - a systematic review. *Amercian Journal of Epidemiology* 2014;**179**(4):423-31.

Webster 2014

Webster AS, Shandera WX. The extrapulmonary dissemination of tuberculosis: a meta-analysis. *International Journal of Mycobacteriology* 2014;**3**:9-16.

Wen 2017

Wen H, Li P, Ma H, Lv G. Diagnostic accuracy of Xpert MTB/RIF assay for musculoskeletal tuberculosis: a meta-analysis. *Infection and Drug Resistance* 2017;**10**:299-305.

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 2010

World Health Organization. Guidelines for treatment of tuberculosis. www.who.int/tb/publications/2010/9789241547833/en/. Fourth edition. Geneva: World Health Organization, 2010 (accessed 11 May 2017).

WHO 2011

World Health Organization. Rapid implementation of the Xpert MTB/RIF diagnostic test. Technical and operational 'How-to'. Practical considerations. www.who.int/tb/publications/tb-amplificationtechnology-implementation/en/ 2011 (accessed 10 October 2016).

WHO 2012

World Health Organization. Updated interim critical concentrations for first-line and second-line DST (as of May 2012). www.stoptb.org/wg/gli/assets/documents/Updated%20critical%20concentration%20table_1st%20and%202nd%20line%20drugs.pdf. Geneva: World Health Organization, (accessed 5 October 2016).

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 system for the diagnosis of pulmonary and extrapulmonary TB in adults and children. http://apps.who.int/iris/handle/10665/112472. Geneva: World Health Organization, 2013 (accessed 23 September 2016).

WHO 2014a

World Health Organization. Xpert MTB/RIF implementation manual. Technical and operational 'how-to' practical considerations. http://apps.who.int/iris/bitstream/10665/112469/1/9789241506700_eng.pdf. Geneva: World Health Organization, 2014 (accessed 23 September 2016).

WHO 2014b

World Health Organization. Definitions and reporting framework for tuberculosis. 2013 revision, updated December 2014. www.who.int/tb/publications/definitions/en/. Geneva: World Health Organization, 2014 (accessed 7 October 2016).

WHO 2015

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 update. www.who.int/tb/publications/use-of-lf-lam-tb-hiv/en/. Geneva: World Health Organization, 2015 (accessed 21 March 2017).

WHO 2016b

World Health Organization. The use of molecular line probe assays for the detection of resistance to isoniazid and rifampicin: policy update. www.who.int/tb/publications/molecular-test-resistance/en/. Geneva: World Health Organization, 2016 (accessed 17 July 2017).

WHO 2017a

World Health Organization. Global tuberculosis report. Global Tuberculosis Report 2017. Geneva: World Health Organization, 2017.

WHO 2017b

World Health Organization. Guidelines for the treatment of drug-susceptible tuberculosis and patient care, 2017 update. www.who.int/tb/publications/2017/dstb_guidance_2017/en/. Geneva: World Health Organization, 2017 (accessed 11 May 2017).

WHO 2017c

World Health Organization. WHO meeting report of a technical expert consultation: non-inferiority analysis of Xpert MTF/RIF Ultra compared to Xpert MTB/RIF. www.who.int/tb/publications/2017/XpertUltra/en/. Geneva: World Health Organization, 2017 (accessed 29 March 2017).

WHO 2018

World Health Organization. Technical report on critical concentrations for drug susceptibility testing of medicines used in the treatment of drug-resistant tuberculosis (WHO/CDS/TB/2018.5). www.who.int/tb/publications/2018/WHO_technical_report_concentrations_TB_drug_susceptibility/en/. Geneva: World Health Organization, (accessed 23 April 2018).

WHO END TB 2014

World Health Organization. The END TB strategy. www.who.int/tb/strategy/end-tb/en/. Geneva: World Health Organization, 2014 (accessed 19 September 2016).

Woodard 1982

Woodard BH, Rosenberg SI, Farnham R, Adams DO. Incidence and nature of primary granulomatous inflammation in surgically removed material.. *American Journal of Surgical Pathology* 1982;**6**(2):119-29.

Woods 2001

Woods GL. Molecular techniques in mycobacterial detection. Archives of Pathology & Laboratory Medicine. Vol. **125**, Geneva: World Health Organization, 2001:122-6.

World Bank 2017

World Bank. World Bank list of economies (June 2017). databank.worldbank.org/data/download/site-content/CLASS.xls (accessed 18 July 2017).

Wright 2009a

Wright CA, Bezuidenhout J. Histopathology and cytopathology. In: Schaaf HS, Zumla A editor(s). Tuberculosis: A Comprehensive Clinical Reference. 1st Edition. Amsterdam: Elsevier Science Publishers, 2009:205-15.

Wright 2009b

Wright CA, Hesseling AC, Bamford C, Burgess SM, Warren R, Marais BJ. Fine-needle aspiration biopsy: a first-line diagnostic procedure in paediatric tuberculosis suspects with peripheral lymphadenopathy?. *International Journal of Tuberculosis and Lung Disease* 2009;**13**(11):1373-9.

References to other published versions of this review

Kohli 2017

Kohli M, Schiller I, Dendukuri N, Ryan H, Dheda K, Denkinger CM, Schumacher SG, Steingart KR. Xpert® MTB/RIF assay for extrapulmonary tuberculosis and rifampicin resistance. *Cochrane Database of Systematic Reviews* 2017, Issue 8. [DOI: 10.1002/14651858.CD012768]

CHARACTERISTICS OF STUDIES

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

Ablanedo-Terrazas 2014

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: HIV-positive patients with palpable cervical lymph nodes Age: median 29 years [interquartile range (IQR) 24 to 36] Sex, female: 12% Children: no HIV infection: 100% Clinical setting: tertiary care centre (inpatient and outpatient)

Ablanedo-Terrazas 2014 *(Continued)*

	Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 15 Laboratory level: central Country: Mexico World Bank Income Classification: middle income TB incidence rate: 22 per 100,000 Per cent MDR-TB among new TB cases: 2.6%; among retreatment cases: 11% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO standard operating procedure (SOP) or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: lymph node (LN) TB Reference standard for TB detection: Löwenstein–Jensen (LJ) and Mycobacterium growth indicator tube (MGIT) Reference standard for rifampicin resistance: not reported Speciation: yes Decontamination: yes, N-acetyl-L-cysteine-sodium hydroxide (NALC-NaOH)
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			

Ablanedo-Terrazas 2014 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Al-Ateah 2012

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients suspected of having extrapulmonary TB Age: median 35 years Sex, female: 45% Children: 3% HIV infection: 0% Clinical setting: tertiary care centre (laboratory-based evaluation) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 67 Laboratory level: central

Al-Ateah 2012 *(Continued)*

	Country: Saudi Arabia World Bank Income Classification: high income TB incidence rate: 10 per 100,000 Per cent MDR-TB among new TB cases: 2.6%; among retreatment cases: 20% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: lymph node TB, pleural TB Reference standards for TB detection: LJ and MGIT Reference standard for rifampicin resistance: MGIT-drug susceptibility testing (DST) Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	Site of extrapulmonary disease was not reported for 16 tissue specimens and 10 abscesses

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			

Al-Ateah 2012 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Arockiaraj 2017

Study characteristics	
Patient sampling	Cross-sectional, retrospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: people with back pain for longer than 3 months and radiological features suggestive of spondylodiscitis (refers to infection of the intervertebral disc and neighbouring vertebral bodies) Age: mean 42 years, range 5 to 82 years Sex, female: 40% Children: not reported HIV infection: not reported Clinical setting: tertiary care centre Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 338 Laboratory level: central Country: India World Bank Income Classification: middle income TB incidence rate: 211 per 100,000 Per cent MDR-TB among new TB cases: 2.8%; among retreatment cases: 12% (source: WHO Global TB report, 2017)

Arockiaraj 2017 *(Continued)*

Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: bone and joint TB Reference standard for TB detection: LJ and MGIT Reference standard for rifampicin resistance: not reported Speciation: not reported Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Unclear		

Arockiaraj 2017 *(Continued)*

		Unclear	**Unclear**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Bahr 2015

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: HIV-infected patients presenting with symptoms of meningitis being evaluated for cryptococcal meningitis. All persons who were CSF cryptococcal antigen-negative had a TB workup Age: median 40 years (IQR 30 to 45) Sex, female: 34% Children: no HIV infection: 98% Clinical setting: tertiary care centre (Inpatient) Past history of TB: 22% Participants on anti-TB treatment: yes, 11% Number of specimens evaluated: 80 Laboratory level: central Country: Uganda World Bank Income Classification: low income TB incidence rate: 201 per 100,000 Per cent MDR-TB among new TB cases: 1.6%; among retreatment cases: 12% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: TB meningitis Reference standard for TB detection: LJ and MGIT

Bahr 2015 *(Continued)*

	Reference standard for rifampicin resistance: MGIT-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	Reference standards were culture and a TB meningitis uniform case definition

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Bahr 2015 *(Continued)*

Were all patients included in the analysis?	Yes
	Low

Bahr 2017

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: HIV-infected patients presenting with symptoms of meningitis being evaluated for cryptococcal meningitis. All persons who were CSF cryptococcal antigen-negative had a TB workup Age: TB meningitis: median 32 years (IQR 30 to 34); other meningitis: 34 years (IQR 29 to 43) Sex, female: 45% Children: no HIV infection: 100% Clinical setting: tertiary care centre (inpatient) Past history of TB: 6% Participants on anti-TB treatment: yes, 2% Number of specimens evaluated: 129 Laboratory level: central Country: Uganda World Bank Income Classification: low income TB incidence rate: 201 per 100,000 Per cent MDR-TB among new TB cases: 1.6%; among retreatment cases: 12% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF and Xpert® MTB/RIF Ultra WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: TB meningitis Reference standard for TB detection: MGIT Reference standard for rifampicin resistance: MGIT-DST Speciation: yes Decontamination: no
Flow and timing	

Bahr 2017 *(Continued)*

Comparative	
Notes	This study evaluated Xpert® MTB/RIF and Xpert® MTB/RIF Ultra Reference standards were culture and a TB meningitis uniform case definition

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Bera 2015

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with exudative ascites (lymphocytic ascites and ascitic fluid protein content > 2.5 g/dL) Age: mean 43 years (standard deviation (SD) 15 years) Sex, female: 29% Children: no HIV infection: not reported Clinical setting: tertiary care centre (outpatient) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 28 Laboratory level: central Country: India World Bank Income Classification: middle income TB incidence rate: 211 per 100,000 Per cent MDR-TB among new TB cases: 2.8%; among retreatment cases: 12% (source: WHO Global TB report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: not reported Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: peritoneal TB Reference standard for TB detection: LJ and MGIT Reference standard for rifampicin resistance: LJ and MGIT-DST Speciation: not reported Decontamination: no
Flow and timing	
Comparative	
Notes	"The study included only smear-negative specimens, however, the study excluded specimens that were negative for malignant cells on prior testing (i.e. cytology)"
Methodological quality	

Bera 2015 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		Unclear	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Unclear**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Unclear**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Bholla 2016

Study characteristics	
Patient sampling	Cross-sectional, prospective, manner of participant selection by convenience

Bholla 2016 *(Continued)*

Patient characteristics and setting	Presenting signs and symptoms: 1 or more palpable lymph nodes of 1 cm or larger persisting for longer than 4 weeks in spite of oral antibiotic therapy and a strong clinical suspicion or microbiological confirmation of mycobacterial infection Age: 6 weeks to 16 years Sex, female: 39% Children: 100% HIV infection: 20% Clinical setting: local hospital (outpatient) Past history of TB: 3% Patients on anti-TB treatment: yes, 11% Number of specimens evaluated: 36 Laboratory level: central Country: Tanzania World Bank Income Classification: low income TB incidence rate: 287 per 100,000 Per cent MDR-TB among new TB cases: 1.3%; among retreatment cases: 6.2% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: lymph node TB Reference standard for TB detection: MGIT Reference standard for rifampicin resistance: MGIT-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	Exclusions: children who had received TB treatment in the preceding 12 months Culture contamination rate was high

Methodological quality

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

Bholla 2016 *(Continued)*

Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		High	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Biadglegne 2014

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with enlarged lymph nodes not responding to a 2-week course of antibiotics and clinically suspected for TB lymphadenitis Age: ≤ 14 years: 15%; > 14 years: 85% Sex, female: 57%

Biadglegne 2014 *(Continued)*

	Children: 15% HIV infection: not reported Clinical setting: tertiary care centres (multi-centre study) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 213 Laboratory level: intermediate Country: Ethiopia World Bank Income Classification: low income TB incidence rate: 177 per 100,000 Per cent MDR-TB among new TB cases: 2.7%; among retreatment cases: 14% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: lymph node TB Reference standard for TB detection: LJ and Gottsascker and BacT/ALERT 3D Reference standard for rifampicin resistance: MTBDR*plus* and BacT-DST Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	Total number of patients: 231; included: 213 (excluded: contaminated = 11; invalid/error = 7)

Methodological quality

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

Biadglegne 2014 *(Continued)*

DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Blaich 2014

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with suspicion of extrapulmonary TB Age: median 34 (IQR 30 to 52) Sex, female: 46% Children: no HIV infection: yes, 8% Clinical setting: university hospital (inpatient and outpatient) Past history of TB: yes, 11% Patients on anti-TB treatment: no

Blaich 2014 *(Continued)*

	Number of specimens evaluated: 20 Laboratory level: central Country: Switzerland World Bank Income Classification: high income TB incidence rate: 7.8 per 100,000 Per cent MDR-TB among new TB cases: 3.2%; among retreatment cases: 26% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes for lymph node aspirate, bone and joint fluid, urine, peritoneal fluid, and lymph node tissue; no for CSF Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, TB meningitis, lymph node TB, pericardial TB, genitourinary TB, bone and joint TB Reference standard for TB detection: LJ and MGIT Reference standard for rifampicin resistance: MGIT-DST Speciation: yes Decontamination: yes, NALC-NaOH for all specimens except pleural fluid and CSF
Flow and timing	
Comparative	
Notes	Study included 1 bone marrow specimen that consisted of both aspirate and tissue

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Blaich 2014 *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		High	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Causse 2011

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: not reported Age: median 45 years, range 5 to 83 years Sex, female: 31% Children: yes, 15% HIV infection: not reported Clinical setting: tertiary care centre (inpatient) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 261 Laboratory level: central Country: Spain

Causse 2011 *(Continued)*

	World Bank Income Classification: high income TB incidence rate: 10 per 100,000 Per cent MDR-TB among new TB cases: 4.2%; among retreatment cases: 18% (source: WHO Global TB report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, lymph node TB, TB meningitis, peritoneal TB, pericardial TB, genitourinary TB Reference standard for TB detection: LJ and MGIT Reference standard for rifampicin resistance: not reported Speciation: yes Decontamination: yes, NALC-NaOH for all specimens except pleural fluid and CSF
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			

Causse 2011 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Che 2017

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with evidence of pleural effusion demonstrated by X-ray, suspected to have tuberculosis pleurisy Age: median 44 years, range 18 to 83 years Sex, female: 31% Children: no HIV infection: 1% Clinical setting: tertiary care centre (inpatient) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 78 Laboratory level: central Country: China World Bank Income Classification: middle income TB incidence rate: 64 per 100,000 Per cent MDR-TB among new TB cases: 7.1%; among retreatment cases: 24% (source: WHO Global TB Report, 2017)

Che 2017 *(Continued)*

Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB Reference standard for TB detection: MGIT Reference standard for rifampicin resistance: not reported Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Unclear		

Che 2017 *(Continued)*

		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Christopher 2013

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: clinical symptoms and radiographic evidence of a pleural effusion Age: median 46 years (IQR 33 to 57) Sex, female: 20% Children: no HIV infection: not reported Clinical setting: tertiary care centre (Inpatient and outpatient) Past history of TB: yes, 18% Patients on anti-TB treatment: not reported Number of specimens evaluated against culture: 142 Number of specimens evaluated against composite reference standard: 146 Laboratory level: central Country: India World Bank Income Classification: middle income TB incidence rate: 211 per 100,000 Per cent MDR-TB among new TB cases: 2.8%; among retreatment cases: 12% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes for pleural tissue, no for pleural fluid Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB

Christopher 2013 *(Continued)*

	Reference standard for TB detection: LJ and MGIT Reference standard for rifampicin resistance: not reported Speciation: not reported Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Unclear**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Unclear		
		Low	**Unclear**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Christopher 2013 *(Continued)*

Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Coetzee 2014

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: children with persistent superficial lymphadenopathy and clinical suspicion of mycobacterial infection Age: < 1 year 33%, 1 to 4 years 42%, 5 to 9 years 18%, ≥ 10 years 7% Sex, female: 40% Children: 100% HIV infection: 8% Clinical setting: tertiary care centre (inpatient and outpatient) Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 72 Laboratory level: central Country: South Africa World Bank Income Classification: middle income TB incidence rate: 781 per 100,000 Per cent MDR-TB among new TB cases: 3.4%; among retreatment cases: 7.1% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: lymph node TB Reference standard for TB detection: MGIT and Middlebrook 7H9 Reference standard for rifampicin resistance: MTBDR*plus* Speciation: yes Decontamination: no
Flow and timing	

Coetzee 2014 *(Continued)*

Comparative
Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Dhasmana 2014

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: all participants undergoing endobronchial ultrasound (EBUS) for mediastinal lymphadenopathy Age: median 46 years, range 14 to 85 years Sex, female: 37% Children: no HIV infection: 7% Clinical setting: tertiary care centre (inpatient and outpatient) Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 116 Laboratory level: central Country: United Kingdom World Bank Income Classification: high income TB incidence rate: 9.9 per 100,000 Per cent MDR-TB among new TB cases: 1.4%; among retreatment cases: 3.4% (source: WHO Global TB report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: lymph node TB Reference standard for TB detection: MGIT Reference standard for rifampicin resistance: MGIT-DST Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	

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

Dhasmana 2014 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Dhooria 2016

Study characteristics	
Patient sampling	Cross-sectional, retrospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with enlarged mediastinal or hilar lymph nodes (≥ 1 cm in short axis) on computed tomography of the chest who underwent EBUS-guided transbronchial needle aspiration Age: median 40 years, range 30 to 53 years Sex, female: 43%

Dhooria 2016 *(Continued)*

	Children: no HIV infection: 0% Clinical setting: tertiary care centre (outpatient) Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 147 Laboratory level: central Country: India World Bank Income Classification: middle income TB incidence rate: 211 per 100,000 Per cent MDR-TB among new TB cases: 2.8%; among retreatment cases: 12% (source: WHO Global TB report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: lymph node TB Reference standard for TB detection: MGIT Reference standard for rifampicin resistance: not reported Speciation: not reported Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			

Dhooria 2016 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Unclear		
		Low	**Unclear**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Diallo 2016

Study characteristics	
Patient sampling	Cross-sectional, retrospective, and manner of participant selection not reported
Patient characteristics and setting	Presenting signs and symptoms: patients with clinical suspicion of EPTB Age: < 18 years 30%, ≥ 18 years 70% Sex, female: 45% Children: 30% HIV infection: not reported Clinical setting: university hospital (laboratory-based evaluation) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 43

Diallo 2016 *(Continued)*

	Laboratory level: central Country: Senegal World Bank Income Classification: low income TB incidence rate: 140 per 100,000 Per cent MDR-TB among new TB cases: 0.9%; among retreatment cases: 19% (source: WHO Global TB report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, pericardial TB, genitourinary TB Reference standard for TB detection: MGIT Reference standard for rifampicin resistance: MGIT-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Unclear	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			

Diallo 2016 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Du 2015

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients found to be smear-negative on prior testing with radiographic evidence of pleural effusion and those subsequently undergoing thoracocentesis and pleural biopsy Age: mean 39 years, SD 13 Sex, female: 44% Children: 0% HIV infection: 4% Clinical setting: 4 tertiary care centres (inpatient) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 126 Laboratory level: central Country: China World Bank Income Classification: middle income TB incidence rate: 64 per 100,000 Per cent MDR-TB among new TB cases: 7.1%; among retreatment cases: 24% (source: WHO Global TB report, 2017)

Du 2015 *(Continued)*

Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB Reference standard for TB detection: LJ and MGIT Reference standard for rifampicin resistance: MGIT-DST Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	Study included specimens found to be smear-negative on prior testing. In the present study, 4 specimens were smear-positive specimens for pleural fluid and 15 were smear-positive for pleural tissue The reference standard for both pleural fluid and pleural tissue was pleural biopsy culture

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		

Du 2015 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Feasey 2013

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: HIV-infected patients with clinical suspicion of TB Age: mean 37 years, SD 11 years Sex, female: 33% Children: no HIV infection: 100% Clinical setting: tertiary care centre (inpatient) Past history of TB: no Patients on anti-TB treatment: no Number of specimens evaluated: 74 Laboratory level: central Country: Malawi World Bank Income Classification: low income TB incidence rate: 159 per 100,000 Per cent MDR-TB among new TB cases: 0.75%; among retreatment cases: 6.4% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF

Feasey 2013 *(Continued)*

	WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: disseminated TB (blood) Reference standard for TB detection: Bactec Myco/F Lytic culture Reference standard for rifampicin resistance: not reported Speciation: yes Decontamination: yes, NALC-NaOH for sputum specimens
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Unclear		
		Low	**Low**

Feasey 2013 *(Continued)*

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
		Low

Friedrich 2011

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with undiagnosed pleural effusion and high clinical suspicion of pleural TB Age: not reported Sex, female: 36% Children: 0% HIV infection: 28% Clinical setting: tertiary care centre (inpatient and outpatient) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated against culture: 24 Number of specimens evaluated against composite reference standard: 25 Laboratory level: central Country: South Africa World Bank Income Classification: middle income TB incidence rate: 781 per 100,000 Per cent MDR-TB among new TB cases: 3.4%; among retreatment cases: 7.1% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB Reference standard for TB detection: MGIT

Friedrich 2011 *(Continued)*

	Reference standard for rifampicin resistance: not reported Speciation: no Decontamination: yes, NALC-NaOH		
Flow and timing			
Comparative			
Notes			

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Unclear		
		Unclear	**High**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Friedrich 2011 *(Continued)*

Were all patients included in the analysis?	Yes
	Low

Ghariani 2015

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with clinical suspicion of TB Age: mean 32 years, range 3 to 79 years Sex, female: 68% Children: 13% HIV infection: no Clinical setting: tertiary care centre (inpatient and outpatient) Past history of TB: 18% Patients on anti-TB treatment: yes, 3% Number of specimens evaluated: 174 Laboratory level: central Country: Tunisia World Bank Income Classification: middle income TB incidence rate: 38 per 100,000 Per cent MDR-TB among new TB cases: 0.93%; among retreatment cases: 4.2% (source: WHO global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: lymph node TB Reference standard for TB detection: LJ and MGIT Reference standard for rifampicin resistance: MGIT-DST Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	

Ghariani 2015 *(Continued)*

Notes

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Gu 2015

Study characteristics

Gu 2015 *(Continued)*

Patient sampling	Cross-sectional, prospective; manner of participant selection not reported
Patient characteristics and setting	Presenting signs and symptoms: patients with suspicion of bone and joint TB Age: median 42 years for TB patients, range 18 to 82 years Sex, female: 54% Children: no HIV infection: not reported Clinical setting: tertiary care centre (inpatient) Past history of TB: not reported Patients on anti-TB treatment: yes, 100% Number of specimens evaluated: 60 Laboratory level: central Country: China World Bank Income Classification: middle income TB incidence rate: 64 per 100,000 Per cent MDR-TB among new TB cases: 7.1%; among retreatment cases: 24% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: bone and joint TB Reference standard for TB detection: MGIT Reference standard for rifampicin resistance: MGIT-DST Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	

Methodological quality

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

Gu 2015 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Unclear	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Gursoy 2016

Study characteristics	
Patient sampling	Cross-sectional, retrospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: not reported Age: not reported Sex, female: not reported Children: not reported

Gursoy 2016 *(Continued)*

	HIV infection: not reported Clinical setting: tertiary care centre (inpatient and outpatient) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 303 Laboratory level: central Country: Turkey World Bank Income Classification: middle income TB incidence rate: 18 per 100,000 Per cent MDR-TB among new TB cases: 2.9%; among retreatment cases: 16% (source: WHO Global TB report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: TB meningitis, genitourinary TB Reference standard for TB detection: LJ and VersaTrek Reference standard for rifampicin resistance: VersaTrek Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			

Gursoy 2016 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Hanif 2011

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with suspicion of TB due to symptoms such as fever, cough, and/or weight loss, or because they were not responding to initial therapy for other diseases Age: range 20 to 57 years Sex, female: 39% Children: no HIV infection: no Clinical setting: national reference laboratory Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 29

Hanif 2011 *(Continued)*

	Laboratory level: central Country: Kuwait World Bank Income Classification: middle income TB incidence rate: 24 per 100,000 Per cent MDR-TB among new TB cases: 2.7%; among retreatment cases: 0% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol: yes for lymph node aspirate, pleural fluid, and urine; no for CSF Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: TB meningitis, lymph node TB, pleural TB, genitourinary TB Reference standard for TB detection: LJ and MGIT Reference standard for rifampicin resistance: LJ-DST and MGIT-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**

Hanif 2011 *(Continued)*

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		High	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Held 2014

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: history of chronic pain for longer than 3 months and presence of constitutional symptoms: low-grade fever, night sweats, loss of appetite, weight loss; loss of anterior vertebral height Age: median 40 years, IQR 27 to 60 years Sex, female: 55% Children: no HIV infection: 32% Clinical setting: tertiary care centre (inpatient) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 71 Laboratory level: central Country: South Africa World Bank Income Classification: middle income TB incidence rate: 781 per 100,000

Held 2014 *(Continued)*

	Per cent MDR-TB among new TB cases: 3.4%; among retreatment cases: 7.1% (source: WHO Global TB report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: bone and joint TB Reference standard for TB detection: MGIT Reference standard for rifampicin resistance: MGIT-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		

Held 2014 *(Continued)*

For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Held 2016

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients under 13 years of age who presented with suspected musculoskeletal TB were included. Symptoms and signs suspicious for musculoskeletal TB included joint or back pain of insidious onset associated with elevated inflammatory markers, TB contact, constitutional symptoms, chronic cough, and HIV. Suspicious radiological signs were a chest radiograph suggestive of TB, or a radiograph of an affected joint showing erosions and osteopenia involving both sides of the joint Age: median 6 years, IQR 2 to 9 years Sex, female: 41% Children: 100% HIV infection: 10% Clinical setting: tertiary care centre (inpatient) Past history of TB: no Patients on anti-TB treatment: no Number of specimens evaluated: 109 Laboratory level: central Country: South Africa World Bank Income Classification: middle income TB incidence rate: 781 per 100,000 Per cent MDR-TB among new TB cases: 3.4%; among retreatment cases: 7.1% (source: WHO Global TB report, 2017)
Index tests	Xpert® MTB/RIF

Held 2016 *(Continued)*

	WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: bone and joint TB Reference standard for TB detection: MGIT Reference standard for rifampicin resistance: MGIT-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**

Held 2016 *(Continued)*

DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Hillemann 2011

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with suspected *M tuberculosis* or non-tuberculous mycobacterial infection on the basis of clinical criteria Age: not reported Sex, female: not reported Children: 5% HIV infection: not reported Clinical setting: national reference laboratory Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 200 Laboratory level: central Country: Germany World Bank Income Classification: high income TB incidence rate: 8.1 per 100,000 Per cent MDR-TB among new TB cases: 2.2%; among retreatment cases: 23% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: yes, donation of index test
Target condition and reference standard(s)	Target condition: pleural TB, TB meningitis, genitourinary TB Reference standard for TB detection: LJ and MGIT Reference standard for rifampicin resistance: MGIT-DST

Hillemann 2011 *(Continued)*

	Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Hillemann 2011 *(Continued)*

	Low

Ioannidis 2011

Study characteristics	
Patient sampling	Cross-sectional, prospective, manner of participant selection by convenience
Patient characteristics and setting	Presenting signs and symptoms: patients with high clinical suspicion of TB Age: not reported Sex, female: not reported Children: not reported HIV infection: not reported Clinical setting: national reference laboratory Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 26 Laboratory level: central Country: Greece World Bank Income Classification: high income TB incidence rate: 4.4 per 100,000 Per cent MDR-TB among new TB cases: 1.5%; among retreatment cases: 9.1% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: not reported Manufacturer's involvement: yes, donation of index test
Target condition and reference standard(s)	Target condition: pleural TB, lymph node TB, TB meningitis, pericardial TB, bone and joint TB, genitourinary TB Reference standard for TB detection: LJ and MGIT Reference standard for rifampicin resistance: LJ-DST, MGIT-DST, MTBDR-*plus* Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	

Ioannidis 2011 *(Continued)*

Notes	Specimens were primarily smear-negative		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		High	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Unclear**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Iram 2015

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with clinical presentation, radiological findings, and histopathological evidence of extrapulmonary TB Age: mean 37 years, range 10 to 80 years Sex, female: 41% Children: 3% HIV infection: 2% Clinical setting: teaching hospital Past history of TB: 53% Patients on anti-TB treatment: yes, 3% Number of specimens evaluated: 18 Laboratory level: intermediate Country: Pakistan World Bank Income Classification: middle income TB incidence rate: 268 per 100,000 Per cent MDR-TB among new TB cases: 4.2%; among retreatment cases: 16% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, peritoneal TB Reference standard for TB detection: LJ Reference standard for rifampicin resistance: LJ-DST Speciation: not reported Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**

Iram 2015 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Unclear**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Jing 2017

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with suspicion of EPTB Age: not reported Sex, female: not reported

Jing 2017 *(Continued)*

	Children: not reported HIV infection: not reported Clinical setting: tertiary care centre Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 277 Laboratory level: central Country: China World Bank Income Classification: middle income TB incidence rate: 64 per 100,000 Per cent MDR-TB among new TB cases: 7.1%; among retreatment cases: 24% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, TB meningitis, genitourinary TB, peritoneal TB Reference standard for TB detection: MGIT Reference standard for rifampicin resistance: MGIT-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			

Jing 2017 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Kim 2015a

Study characteristics	
Patient sampling	Cross-sectional, retrospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: not reported Age: median 59 years (IQR 44 to 71 years) Sex, female: 47% Children: 7% HIV infection: 1% Clinical setting: tertiary care centre Past history of TB: 9% Patients on anti-TB treatment: no Number of specimens evaluated: 1209 Laboratory level: central

Kim 2015a *(Continued)*

	Country: Korea World Bank Income Classification: high income TB incidence rate: 77 per 100,000 Per cent MDR-TB among new TB cases: 3.4%; among retreatment cases: 11% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: lymph node TB, pleural TB, TB meningitis, peritoneal TB, pericardial TB, bone and joint TB, genitourinary TB Reference standard for TB detection: MGIT Reference standard for rifampicin resistance: LJ-DST Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		

Kim 2015a *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Li 2017

Study characteristics	
Patient sampling	Cross-sectional, prospective; manner of participant selection not reported
Patient characteristics and setting	Presenting signs and symptoms: patients with suspected EPTB Age: mean 48 years, SD 10 years Sex, female: 39% Children: no HIV infection: not reported Clinical setting: tertiary care centre Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 414 Laboratory level: central Country: China World Bank Income Classification: middle income TB incidence rate: 64 per 100,000 Per cent MDR-TB among new TB cases: 7.1%; among retreatment cases: 24% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF

Li 2017 *(Continued)*

	WHO SOP or manufacturer's protocol followed: yes for pleural fluid, bone and joint TB fluid, urine, and peritoneal fluid; no for CSF Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, TB meningitis, peritoneal TB, bone and joint TB, genitourinary TB Reference standard for TB detection: LJ Reference standard for rifampicin resistance: LJ-DST Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Unclear	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		

Li 2017 *(Continued)*

		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Ligthelm 2011

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with suspicion of lymph node TB Age: < 5 years 4%; 5 to 20 years 13%; > 20 years 83% Sex, female: 58% Children: 4% HIV infection: 19% Clinical setting: university hospital (inpatient and outpatient) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 48 Laboratory level: central Country: South Africa World Bank Income Classification: middle income TB incidence rate: 781 per 100,000 Per cent MDR-TB among new TB cases: 3.4%; among retreatment cases: 7.1% (source: WHO Global TB Report, 2017
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: lymph node TB Reference standard for TB detection: MGIT Reference standard for rifampicin resistance: MTBDR*plus*

Ligthelm 2011 *(Continued)*

	Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	"It is unlikely that our patient cohort had exacerbated disease compared to patients presenting at primary health care clinics, as these patients are routinely referred from the primary health care clinic to the referral centre for FNAB (fine needle aspiration biopsy)"

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Ligthelm 2011 *(Continued)*

Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
		Low

Lusiba 2014

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with suspected pleural TB based on clinical signs and symptoms and radiological evidence of a pleural effusion that was considered large enough for a pleural biopsy Age: mean 34 years, SD 13 years Sex, female: 43% Children: no HIV infection: 45% Clinical setting: tertiary care centre (inpatient and outpatient) Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 116 Laboratory level: central Country: Uganda World Bank Income Classification: low income TB incidence rate: 201 per 100,000 Per cent MDR-TB among new TB cases: 1.6%; among retreatment cases: 12% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB Reference standard for TB detection: LJ and MGIT Reference standard for rifampicin resistance: MGIT-DST Speciation: not reported Decontamination: no
Flow and timing	

Lusiba 2014 *(Continued)*

Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Unclear**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Malbruny 2011

Study characteristics	
Patient sampling	Cross-sectional, prospective, manner of participant selection by convenience
Patient characteristics and setting	Presenting signs and symptoms: patients with clinical suspicion of TB Age: median 52 years Sex, female: 40% Children: 7% HIV infection: not reported Clinical setting: university hospital Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 67 Laboratory level: central Country: France World Bank Income Classification: high income TB incidence rate: 7.7 per 100,000 Per cent MDR-TB among new TB cases: 1%; among retreatment cases: 10% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, TB meningitis, bone and joint TB, peritoneal TB, genitourinary TB Reference standard for TB detection: MGIT and Coletsos slants Reference standard for rifampicin resistance: MGIT-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**

Malbruny 2011 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		High	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Massi 2017

Study characteristics	
Patient sampling	Cross-sectional, prospective; manner of participant selection not reported
Patient characteristics and setting	Presenting signs and symptoms: not reported Age: not reported Sex, female: not reported

Massi 2017 *(Continued)*

	Children: not reported HIV infection: not reported Clinical setting: university hospital Past history of TB: not reported Patients not on anti-TB treatment: not reported Number of specimens evaluated: 70 Laboratory level: central Country: Indonesia World Bank Income Classification: middle income TB incidence rate: 391 per 100,000 Per cent MDR-TB among new TB cases: 2.8%; among retreatment cases: 16% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: bone and joint TB Reference standard for TB detection: MGIT Reference standard for rifampicin resistance: MGIT-DST Speciation: not reported Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	Study observed low specificity, which was probably due to inclusion of many participants on anti-TB treatment, considered standard procedure in this setting before surgery for spondylitis TB

Methodological quality

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

Massi 2017 *(Continued)*

DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Unclear**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Mazzola 2016

Study characteristics	
Patient sampling	Cross-sectional, retrospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: not reported Age: not reported Sex, female: 40% Children: not reported HIV infection: not reported Clinical setting: reference laboratories Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 1201

Mazzola 2016 *(Continued)*

	Laboratory level: central Country: Italy World Bank Income Classification: high income TB incidence rate: 6.1 per 100,000 Per cent MDR-TB among new TB cases: 2.8%; among retreatment cases: 13% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, TB meningitis, peritoneal TB, pericardial TB, genitourinary TB Reference standard for TB detection: LJ and MGIT Reference standard for rifampicin resistance: LJ-DST and MGIT-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			

Mazzola 2016 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Unclear		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Meldau 2014

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients presumed to have pleural TB with any symptoms, including cough, fever, night sweats, loss of weight, haemoptysis, and chest pain, along with features consistent with a pleural effusion on chest X-ray Age: definitive TB: median 39 years (IQR 29 to 55 years); non-TB: median 61 years (IQR 54 to 69 years) Sex, female: 40% Children: no HIV infection: 15% Clinical setting: tertiary care hospital Past history of TB: 13% Patients on anti-TB treatment: no Number of specimens evaluated against culture: 76 Number of specimens evaluated against a composite reference standard: 88 Laboratory level: central Country: South Africa World Bank Income Classification: middle income

Meldau 2014 *(Continued)*

	TB incidence rate: 781 per 100,000 Per cent MDR-TB among new TB cases: 3.4%; among retreatment cases: 7.1% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB Reference standard for TB detection: MGIT Reference standard for rifampicin resistance: MGIT-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		

Meldau 2014 *(Continued)*

For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Nataraj 2016

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with clinical suspicion of extrapulmonary TB Age: < 14 years 13%; 15 to 45 years 52%; > 45 years 34%; range 2 months to 78 years Sex, female: 44% Children: 13% HIV infection: not reported Clinical setting: tertiary care centre (inpatient and outpatient) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 494 Laboratory level: intermediate Country: India World Bank Income Classification: middle income TB incidence rate: 211 per 100,000 Per cent MDR-TB among new TB cases: 2.8%; among retreatment cases: 12% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no

Nataraj 2016 *(Continued)*

Target condition and reference standard(s)	Target condition: pleural TB, lymph node TB, TB meningitis, bone and joint TB, genitourinary TB Reference standard TB detection: LJ Reference standard rifampicin resistance detection: LJ-DST Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	Patients on treatment may have been included, although the number was not reported: "Of the two specimens that were smear-positive and smear-negative on both culture and Xpert, one was pleural fluid from a patient who had been receiving Category II anti-tuberculosis treatment for 2 months and the other was pus aspirated from an axillary lymph node"

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		

Nataraj 2016 *(Continued)*

		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Nhu 2014

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients suspected of having TB meningitis with at least 5 days of meningitis symptoms, nuchal rigidity, and CSF abnormalities Age: > 18 years Sex, female: not reported Children: no HIV infection: 21% Clinical setting: university hospital Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 379 Laboratory level: central Country: Vietnam World Bank Income Classification: middle income TB incidence rate: 133 per 100,000 Per cent MDR-TB among new TB cases: 4.1%; among retreatment cases: 26% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: yes, donation of index test
Target condition and reference standard(s)	Target condition: TB meningitis Reference standard TB detection: MGIT

Nhu 2014 *(Continued)*

	Reference standard rifampicin resistance detection: MGIT-DST and MTBDR*plus* Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	Analysis by uniform case definition also included

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Nhu 2014 *(Continued)*

Were all patients included in the analysis?	Yes
	Low

Ozkutuk 2014

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: not reported Age: median 54 years, range 1 to 99 years Sex, female: 47% Children: 3% HIV infection: not reported Clinical setting: tertiary care centre (inpatient and outpatient) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 1022 Laboratory level: central Country: Turkey World Bank Income Classification: middle TB incidence rate: 18 per 100,000 Per cent MDR-TB among new TB cases: 2.9%; among retreatment cases: 16% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, lymph node TB, TB meningitis, genitourinary TB, bone and joint TB, pericardial TB, peritoneal TB Reference standard TB detection: LJ and MGIT Reference standard rifampicin resistance detection: MGIT-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	

Ozkutuk 2014 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Pandey 2017

Study characteristics

Pandey 2017 *(Continued)*

Patient sampling	Cross-sectional, prospective, manner of participant selection by convenience
Patient characteristics and setting	Presenting signs and symptoms: not reported Age: not reported Sex, female: not reported Children: not reported HIV infection: not reported Clinical setting: reference laboratory Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 57 Laboratory level: central Country: Australia World Bank Income Classification: high income TB incidence rate: 6.1 per 100,000 Per cent MDR-TB among new TB cases: 3.6%; among retreatment cases: 24% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol: no for lymph node aspirate, pleural fluid, and CSF; yes for lymph node tissue Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: lymph node TB, pleural TB, TB meningitis Reference standard TB detection: LJ and MGIT Reference standard rifampicin resistance detection: MGIT-DST Speciation: yes Decontamination: yes for lymph node aspirate
Flow and timing	
Comparative	
Notes	

Methodological quality

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

Pandey 2017 *(Continued)*

Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		High	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Pandie 2014

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with presence of a large pericardial effusion amenable to safe pericardiocentesis (> 10 mm echo-free space around the heart in diastole) Age: median 34 years (IQR 29 to 42) Sex, female: 38%

Pandie 2014 *(Continued)*

	Children: no HIV infection: 74% Clinical setting: 4 district hospitals and 1 tertiary centre (inpatient) Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 134 Laboratory level: central Country: South Africa World Bank Income Classification: middle income TB incidence rate: 781 per 100,000 Per cent MDR-TB among new TB cases: 3.4%; among retreatment cases: 7.1% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pericardial TB Reference standard TB detection: MGIT Reference standard rifampicin resistance detection: MTBDR*plus* Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Low**
DOMAIN 2: Index Test All tests			

Pandie 2014 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Patel 2013

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with clinical suspicion of meningitis Age: mean 33 years (SD 9) Sex, female: 61% Children: 2% HIV infection: 87% Clinical setting: tertiary care centre (inpatient and outpatient) Past history of TB: 31% Patients on anti-TB treatment: no Number of specimens evaluated: 59 Laboratory level: central

Patel 2013 *(Continued)*

	Country: South Africa World Bank Income Classification: middle income TB incidence rate: 781 per 100,000 Per cent MDR-TB among new TB cases: 3.4%; among retreatment cases: 7.1% (source: WHO Global TB Report, 2017)		
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no		
Target condition and reference standard(s)	Target condition: TB meningitis Reference standard TB detection: MGIT Reference standard rifampicin resistance detection: MGIT-DST Speciation: yes Decontamination: no		
Flow and timing			
Comparative			
Notes	Study used frozen specimens		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Patel 2013 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Penata 2016

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with clinical suspicion of extrapulmonary tuberculosis Age: mean 42 years (SD 19), range 1 to 91 years Sex, female: 39% Children: 7% HIV infection: 40% Clinical setting: university hospital (inpatient) Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 236 Laboratory level: intermediate Country: Colombia World Bank Income Classification: middle income TB incidence rate: 32 per 100,000 Per cent MDR-TB among new TB cases: 2.4%; among retreatment cases: 14% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes

Penata 2016 *(Continued)*

	Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: lymph node TB, pleural TB, TB meningitis, peritoneal TB, pericardial TB, bone and joint TB Reference standard TB detection: Ogawa medium Reference standard rifampicin resistance detection: Ogawa-DST Speciation: not reported Decontamination: unclear
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	No		
		High	**Unclear**
DOMAIN 4: Flow and Timing			

Penata 2016 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Pink 2016

Study characteristics	
Patient sampling	Cross-sectional, retrospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: not reported Age: median 46 years; range 0 to 93 years Sex, female: 41% Children: not reported HIV infection: not reported Clinical setting: national reference laboratory Past history of TB: mot reported Patients on anti-TB treatment: no Number of specimens evaluated: 735 Laboratory level: central Country: United Kingdom World Bank Income Classification: high income TB incidence rate: 9.9 per 100,000 Per cent MDR-TB among new TB cases: 1.4%; among retreatment cases: 3.4% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: TB meningitis Reference standard TB detection: MGIT and Kirchner media Reference standard rifampicin resistance detection: not reported Speciation: yes Decontamination: no
Flow and timing	

Pink 2016 *(Continued)*

Comparative
Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Unclear		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Pohl 2016

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with presumptive TB meeting 1 or more of the following criteria: persistent, non-remitting cough longer than 14 days not responding to course of antibiotics; repeated episodes of fever within the previous 14 days not responding to course of antibiotics and after malaria has been excluded; weight loss or failure to thrive within the previous 3 months; signs and symptoms suggestive of extrapulmonary tuberculosis: non-painful enlarged lymph nodes; gibbus (form of structural kyphosis), especially of recent onset; lethargy; convulsions; meningism (symptoms and signs of meningitis, but without actual inflammation of the meninges); pleural effusion; pericardial effusion; distended abdomen with ascites; non-painful enlarged joint; signs of tuberculin hypersensitivity Age: median 5 years (IQR 3 to 10 years) Sex, female: 56% Children: 100% HIV infection: 36% Clinical setting: multi-centre tertiary care centres (inpatient and outpatient) Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 192 Laboratory level: central Country: Tanzania and Uganda World Bank Income Classification: low income TB incidence rate: 287 per 100,000 (Tanzania); 201 per 100,000 (Uganda) Per cent MDR-TB among new TB cases: 1.3% (Tanzania), 1.6% (Uganda); among retreatment cases: 6.2% (Tanzania); 12% (Uganda) (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: disseminated TB Reference standard TB detection: LJ and MGIT Reference standard rifampicin resistance detection: not reported Speciation: yes Decontamination: yes, NALC-NaOH for sputum specimens
Flow and timing	
Comparative	
Notes	This study performed Xpert on blood and culture on sputum specimens

Pohl 2016 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Unclear		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Rufai 2015

Study characteristics	
Patient sampling	Cross-sectional, prospective, manner of participant selection not reported
Patient characteristics and setting	Presenting signs and symptoms: patients with high suspicion of pleural TB. Enrolment was based on standard clinical and radiological criteria, including a persistent cough of 2 weeks or longer, unexplained fever for 2 weeks or longer, unexplained weight loss with or without night sweats, chest pain, and radiological evidence of pleural effusion Age: males: mean 42 years (SD 19 years); females: mean 39 years (SD 19 years) Sex, female: 28% Children: 6% HIV infection: no Clinical setting: tertiary care centre Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 161 Laboratory level: central Country: India World Bank Income Classification: middle income TB incidence rate: 211 per 100,000 Per cent MDR-TB among new TB cases: 2.8%; among retreatment cases: 12% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB Reference standard TB detection: MGIT Reference standard rifampicin resistance detection: MGIT-DST Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	
Methodological quality	

Rufai 2015 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Unclear	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Rufai 2017a

Study characteristics	
Patient sampling	Cross-sectional, prospective, manner of participant selection not reported

Rufai 2017a *(Continued)*

Patient characteristics and setting	Presenting signs and symptoms: patients with clinical or radiological suspicion of abdominal TB Age: males: mean 41 years (SD 19 years); females: mean 46 years (SD 20 years) Sex, female: 36% Children: no HIV infection: no Clinical setting: tertiary care centre Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 67 Laboratory level: central Country: India World Bank Income Classification: middle income TB incidence rate: 211 per 100,000 Per cent MDR-TB among new TB cases: 2.8%; among retreatment cases: 12% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: peritoneal TB Reference standard TB detection: MGIT Reference standard rifampicin resistance detection: MGIT-DST Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	

Methodological quality

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

Rufai 2017a *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
		Unclear	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Rufai 2017b

Study characteristics	
Patient sampling	Cross-sectional, prospective, manner of participant selection not reported
Patient characteristics and setting	Presenting signs and symptoms: fatigue, malaise, low-grade fever, confusion, nausea and vomiting, lethargy, irritability, and unconsciousness Age: males: mean 38 years (SD 10 years); females: mean 34 years (SD 22 years) Sex, female: 41% Children: 6% HIV infection: not reported

Rufai 2017b *(Continued)*

	Clinical setting: tertiary care centre Past history of TB: not reported Patients on anti-TB treatment: yes, 4% Number of specimens evaluated: 267 Laboratory level: central Country: India World Bank Income Classification: middle income TB incidence rate: 211 per 100,000 Per cent MDR-TB among new TB cases: 2.8%; among retreatment cases: 12% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: TB meningitis Reference standard TB detection: MGIT Reference standard rifampicin resistance detection: MGIT-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Unclear	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Rufai 2017b *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Saeed 2017a

Study characteristics	
Patient sampling	Cross-sectional, prospective; manner of participant selection not reported
Patient characteristics and setting	Presenting signs and symptoms: patients with strong suspicion of TB on the basis of (a) clinical presentation, (b) relative laboratory investigation, (c) echocardiography, and (d) radiological finding Age: not reported Sex, female: not reported Children: not reported HIV infection: not reported Clinical setting: tertiary care centre Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 286 Laboratory level: intermediate Country: Pakistan

Saeed 2017a *(Continued)*

	World Bank Income Classification: middle income TB incidence rate: 268 per 100,000 Per cent MDR-TB among new TB cases: 4.2%; among retreatment cases: 16% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, pericardial TB Reference standard TB detection: LJ Reference standard rifampicin resistance detection: LJ-DST Speciation: not reported Decontamination: no
Flow and timing	
Comparative	
Notes	Study authors report, "In this study, strict patient selection criteria in which *strong suspicion* of TB patients were included on the basis of clinical and radiological evidence could have [been] attributed to high sensitivity"

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		High	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Saeed 2017a *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Unclear		
		Unclear	**Unclear**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Safianowska 2012

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: not reported Age: not reported Sex, female: 46% Children: no HIV infection: no Clinical setting: university hospital Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 51 Laboratory level: intermediate Country: Poland World Bank Income Classification: high income TB incidence rate: 18 per 100,000 Per cent MDR-TB among new TB cases: 0.83%; among retreatment cases: 4.4% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF

Safianowska 2012 *(Continued)*

	WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, lymph node TB, TB meningitis, peritoneal TB, pericardial TB, genitourinary TB, bone and joint TB Reference standard TB detection: LJ Reference standard rifampicin resistance detection: LJ-DST Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	No		
		High	**Low**

Safianowska 2012 *(Continued)*

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
		Low

Scott 2014

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: not reported Age: median 39 years, range < 1 year to 96 years Sex, female: 45% Children: 4% HIV infection: not reported Clinical setting: reference laboratory Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 696 Laboratory level: central Country: South Africa World Bank Income Classification: middle income TB incidence rate: 781 per 100,000 Per cent MDR-TB among new TB cases: 3.4%; among retreatment cases: 7.1% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes for lymph node aspirate, pleural fluid, and peritoneal fluid; no for CSF Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, lymph node TB, TB meningitis, peritoneal TB Reference standard TB detection: MGIT Reference standard rifampicin resistance detection: MGIT-DST and MTBDR*plus*

Scott 2014 *(Continued)*

	Speciation: yes Decontamination: no		
Flow and timing			
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Scott 2014 *(Continued)*

	Low

Sharma 2014

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with clinical suspicion of EPTB Age: mean 35 years (SD 15 years) Sex, female: 50% Children: no HIV infection: not reported Clinical setting: tertiary care centre Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 1139 Laboratory level: central Country: India World Bank Income Classification: middle income TB incidence rate: 211 per 100,000 Per cent MDR-TB among new TB cases: 2.8%; among retreatment cases: 12% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol: yes for body fluids and LN tissue; no for CSF Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, lymph node TB, TB meningitis, peritoneal TB, pericardial TB, genitourinary TB Reference standard TB detection: LJ and MGIT Reference standard rifampicin resistance detection: LJ-DST Speciation: yes Decontamination: yes, NALC-NaOH (for all specimens except CSF, pleural fluid, and urine)
Flow and timing	
Comparative	

Sharma 2014 *(Continued)*

Notes

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Sharma 2016

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: women being evaluated for infertility and suspected to have TB Age: mean 29 years, range 19 to 41 years Sex, female: 100% Children: no HIV infection: not reported Clinical setting: tertiary care centre Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 240 Laboratory level: central Country: India World Bank Income Classification: middle income TB incidence rate: 211 per 100,000 Per cent MDR-TB among new TB cases: 2.8%; among retreatment cases: 12% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: genitourinary TB Reference standard TB detection: LJ and MGIT Reference standard rifampicin resistance detection: MGIT-DST Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	

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

Sharma 2016 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Solomons 2016

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with clinically suspected meningitis including fever, irritability, lethargy, bulging fontanelle, nuchal rigidity, fever with or without headache, or photophobia, confirmed by CSF analysis Age: TBM median 31 months (IQR 21 to 54 months); bacterial meningitis median 29 months (IQR 20 to 81 months); viral meningitis median 62 months (IQR 22 to 92 months)

Solomons 2016 *(Continued)*

	Sex, female: TBM 48%; bacterial meningitis 50%; viral meningitis 24% Children: 100% HIV infection: 11% Clinical setting: university hospital Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 139 Laboratory level: central Country: South Africa World Bank Income Classification: middle income TB incidence rate: 781 per 100,000 Per cent MDR-TB among new TB cases: 3.4%; among retreatment cases: 7.1% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: TB meningitis Reference standard TB detection: MGIT Reference standard rifampicin resistance detection: MGIT-DST and MTBDR*plus* Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	This study was performed at a single hospital, which may limit generalization of study findings to other settings; however, Tygerberg Children's Hospital serves a population that shares a similar disease burden and health challenges experienced in other TB-endemic areas.

Methodological quality

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

Solomons 2016 *(Continued)*

		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Suzana 2016

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients with signs and symptoms suggestive of extrapulmonary TB Age: median 34 years Sex, female: 39% Children: 0.06% HIV infection: 7% Clinical setting: tertiary care centre

Suzana 2016 *(Continued)*

	Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 215 Laboratory level: central Country: India World Bank Income Classification: middle income TB incidence rate: 211 per 100,000 Per cent MDR-TB among new TB cases: 2.8%; among retreatment cases: 12% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes for lymph node tissue and pleural tissue; no for pleural fluid, bone and joint fluid, urine, peritoneal fluid, pericardial fluid, and CSF Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, lymph node TB, TB meningitis, peritoneal TB, pericardial TB, genitourinary TB, bone and joint TB Reference standard TB detection: LJ and MGIT Reference standard rifampicin resistance detection: LJ-DST and MGIT-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Suzana 2016 *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Tadesse 2015

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: people with presumptive lymph node TB Age: ≤ 15 years 15%; > 15 years 85% Sex, female: 53% Children: 15% HIV infection: not reported Clinical setting: university hospital (outpatient) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 136 Laboratory level: central Country: Ethiopia

Tadesse 2015 *(Continued)*

	World Bank Income Classification: low income TB incidence rate: 177 per 100,000 Per cent MDR-TB among new TB cases: 2.7%; among retreatment cases: 14% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: lymph node TB Reference standard TB detection: LJ Reference standard rifampicin resistance detection: not reported Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	
Notes	Study used frozen specimens

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		

Tadesse 2015 *(Continued)*

For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Unclear		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Teo 2011

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: not reported Age: not reported Sex, female: not reported Children: not reported HIV infection: not reported Clinical setting: university hospital (laboratory-based evaluation) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 7 Laboratory level: central Country: Singapore World Bank Income Classification: high income TB incidence rate: 51 per 100,000 Per cent MDR-TB among new TB cases: 1.4%; among retreatment cases: 2.3% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: TB meningitis

Teo 2011 *(Continued)*

	Reference standard TB detection: LJ and MGIT
	Reference standard rifampicin resistance detection: LJ-DST
	Speciation: yes
	Decontamination: no
Flow and timing	
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Teo 2011 *(Continued)*

Were all patients included in the analysis?	Yes
	Low

Tortoli 2012

Study characteristics	
Patient sampling	Cross-sectional, retrospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: not reported Age: not reported Sex, female: not reported Children: 34% HIV infection: 10% Clinical setting: 8 Italian laboratories Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 668 Laboratory level: central Country: Italy World Bank Income Classification: high income TB incidence rate: 6.1 per 100,000 Per cent MDR-TB among new TB cases: 2.8%; among retreatment cases: 13% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes for CSF; no for pleural fluid, urine, peritoneal fluid, and pericardial fluid Manufacturer's involvement: yes, donation of the index test
Target condition and reference standard(s)	Target condition: pleural TB, TB meningitis, peritoneal TB, pericardial TB, genitourinary TB Reference standard TB detection: LJ and MGIT Reference standard rifampicin resistance detection: MGIT-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	

Tortoli 2012 *(Continued)*

Notes	Study used frozen specimens		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Trajman 2014

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patientsParticipants with a pleural effusion needing thoracentesis Age: median 50 years (IQR 40 to 57) Sex, female: 20% Children: no HIV infection: 5% Clinical setting: secondary health facility (inpatient) Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 85 Laboratory level: central Country: Brazil World Bank Income Classification: middle income TB incidence rate: 42 per 100,000 Per cent MDR-TB among new TB cases: 1.5%; among retreatment cases: 8% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB Reference standard TB detection: MGIT Reference standard rifampicin resistance detection: MGIT-DST Speciation: not reported Decontamination: no
Flow and timing	
Comparative	
Notes	Patients were excluded if they had bleeding disorders contraindicating thoracentesis, if the fluid volume was insufficient for storage, or if a final diagnosis could not be ascertained. One of the main limitations of the study was the high number of presumptive (non-confirmed) cases. The number of exclusions was also high - out of 203 eligible patients, 110 were excluded: 21 did not have a final diagnosis and 89 did not have sufficient fluid to store. "Cultures of pleural tissue, which could significantly improve accuracy of diagnosis, were not performed" Study used frozen specimens

Trajman 2014 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Unclear**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Unclear		
Were all patients included in the analysis?	No		
		High	

Ullah 2017

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients meeting the following criteria: previously TB-treated cases with both positive and negative smears; failure of Cat-I and Cat-II TB drugs; all smear-positive cases that remained positive by the end of the second month of TB treatment; TB/HIV co-infection cases; seriously ill patients; contacts of MDR-TB patients Age: mean 34 years (SD 19 years), range 3 to 80 years Sex, female: 51% Children: 14% HIV infection: not reported Clinical setting: tertiary care centre Past history of TB: 60% Patients on anti-TB treatment: yes, percentage not reported Number of specimens evaluated: 168 Laboratory level: central Country: Pakistan World Bank Income Classification: middle income TB incidence rate: 268 per 100,000 Per cent MDR-TB among new TB cases: 4.2%; among retreatment cases: 16% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: lymph node TB, TB meningitis, peritoneal TB, pericardial TB Reference standard TB detection: Middlebrook 7H10 Reference standard rifampicin resistance detection: Middlebrook 7H10 Speciation: not reported Decontamination: no
Flow and timing	
Comparative	
Notes	Study included a highly selective population that met specified criteria: previously TB-treated cases with both positive and negative smears; failure of Cat-I and Cat-II TB drugs; all smear-positive cases that remained positive by the end of the second month of TB treatment; TB/HIV co-infection cases; seriously ill patients; contacts of MDR-TB patients

Ullah 2017 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
		High	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Unclear**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Vadwai 2011

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: suspected extrapulmonary TB based on symptoms: brain: irritability, restlessness, neck stiffness, headache persistent for 2 to 3 weeks, vomiting, seizures, changes in mental condition or behaviour; intestinal tract, abdomen: abdominal pain, diarrhoea; lymph nodes: enlargement of lymph nodes, mass formation in the neck; cardiorespiratory: shortness of breath, hypertension, chest pain, dyspnoea; endometrium: pelvic pain, pelvic mass, irregular periods, infertility; skin (cutaneous): visible presence of ulcers or lesions, tender nodules Age: median 37 years Sex, female: 15% Children: 3% HIV infection: 3% Clinical setting: tertiary care centre Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 60 Laboratory level: central Country: India World Bank Income Classification: middle income TB incidence rate: 211 per 100,000 Per cent MDR-TB among new TB cases: 2.8%; among retreatment cases: 12% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes for pleural fluid, peritoneal fluid, pericardial fluid; no for CSF Manufacturer's involvement: yes, in design, analysis, or manuscript production (David Alland is among a group of co-investigators who invented molecular beacons and receive income from licensees, including to Cepheid, for *M tuberculosis* detection)
Target condition and reference standard(s)	Target condition: pleural TB, TB meningitis, peritoneal TB, pericardial TB Reference standard TB detection: LJ and MGIT Reference standard rifampicin resistance detection: MGIT-DST Speciation: yes Decontamination: yes, NALC-NaOH
Flow and timing	
Comparative	

Vadwai 2011 *(Continued)*

Notes	"Patients were enrolled only if they could provide detailed clinical history and radiological and histology/cytology reports, along with an adequate amount of specimen material"

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Vadwai 2011 *(Continued)*

Were all patients included in the analysis?	Yes
	Low

Van Rie 2013

Study characteristics	
Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: HIV-infected patients with suspicion of LNTB Age: mean 36 years, range 18 to 73 years Sex, female: 49% Children: no HIV infection: 100% Clinical setting: tertiary care centre (inpatient and outpatient) Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 344 Laboratory level: central Country: South Africa World Bank Income Classification: middle income TB incidence rate: 781 per 100,000 Per cent MDR-TB among new TB cases: 3.4%; among retreatment cases: 7.1% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: lymph node TB Reference standard TB detection: MGIT Reference standard rifampicin resistance detection: MGIT-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	

Van Rie 2013 *(Continued)*

Notes

Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Wang 2016a

Study characteristics

Wang 2016a *(Continued)*

Patient sampling	Cross-sectional, prospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: patients presenting with symptoms of meningitis (fever, headache, seizure, vomiting, nuchal rigidity, or abnormal CSF parameters) Age: mean 31 years, range 1 to 80 years Sex, female: 38% Children: not reported HIV infection: no Clinical setting: 11 tertiary care centres Past history of TB: 2% Patients on anti-TB treatment: not reported Number of specimens evaluated: 316 Laboratory level: central Country: China World Bank Income Classification: middle income TB incidence rate: 64 per 100,000 Per cent MDR-TB among new TB cases: 7.1%; among retreatment cases: 24% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: TB meningitis Reference standard TB detection: MGIT Reference standard rifampicin resistance detection: not reported Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	Includes analysis by uniform case definition also. Study used frozen specimens

Methodological quality

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

Wang 2016a *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	Unclear		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Zeka 2011

Study characteristics	
Patient sampling	Cross-sectional, retrospective, and consecutive
Patient characteristics and setting	Presenting signs and symptoms: clinical findings of possible TB Age: median 48 years Sex, female: 42% Children: 13%

Zeka 2011 *(Continued)*

	HIV infection: 1% Clinical setting: tertiary care centre Past history of TB: not reported Patients on anti-TB treatment: no Number of specimens evaluated: 149 Laboratory level: central Country: Turkey World Bank Income Classification: middle income TB incidence rate: 18 per 100,000 Per cent MDR-TB among new TB cases: 2.9%; among retreatment cases: 16% (source: WHO Global TB Report, 2017)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: no Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, TB meningitis, genitourinary TB, peritoneal TB, pericardial TB Reference standard TB detection: LJ and BacT liquid medium Reference standard rifampicin resistance detection: 7H10 agar media Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	Study used frozen specimens

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			

Zeka 2011 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**High**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	No		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Zmak 2013

Study characteristics	
Patient sampling	Cross-sectional, prospective; manner of participant selection not reported
Patient characteristics and setting	Presenting signs and symptoms: patients suspected of EPTB Age: not reported Sex, female: not reported Children: not reported HIV infection: not reported Clinical setting: reference laboratory Past history of TB: not reported Patients on anti-TB treatment: not reported Number of specimens evaluated: 176 Laboratory level: central

Zmak 2013 *(Continued)*

	Country: Croatia World Bank Income Classification: high income TB incidence rate: 12 per 100,000 Per cent MDR-TB among new TB cases: 0%; among retreatment cases: 0% (source: WHO Global TB Report, 2016)
Index tests	Xpert® MTB/RIF WHO SOP or manufacturer's protocol followed: yes for pleural fluid, urine, peritoneal fluid, pericardial fluid, and blood; no for CSF Manufacturer's involvement: no
Target condition and reference standard(s)	Target condition: pleural TB, TB meningitis, peritoneal TB, pericardial TB, genitourinary TB, disseminated TB Reference standard TB detection: LJ, Stonebrink, and MGIT Reference standard rifampicin resistance detection: LJ-DST Speciation: yes Decontamination: no
Flow and timing	
Comparative	
Notes	"Although the NRL performs a third-level laboratory service for the whole country, it is actually also involved in first and second-level laboratory work for several counties"

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
		Unclear	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			

Zmak 2013 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
For rifampicin resistance testing, were the reference standard results interpreted without knowledge of the results of the index test?	No		
		High	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

CSF: cerebrospinal fluid; DST: drug susceptibility testing; EBUS: endobronchial ultrasound; EPTB: extrapulmonary tuberculosis: IQR: interquartile ratio; LJ: Löwenstein-Jensen; LN: lymph node; MDR-TB: multi-drug-resistant tuberculosis; MGIT: mycobacteria growth indicator tube; NALC-NaOH: N-acetyl-L-cysteine-sodium hydroxide; SD: standard deviation; SOP: standard operating procedure; TB: tuberculosis; TBM: tuberculous meningitis; WHO: World Health Organization.

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

Study	Reason for exclusion
Alvarez Uria 2012	Inappropriate reference standard
Andrey 2015	Case report
Armand 2011	Case-control study
Arockiaraj 2015	Abstract; we included the published study, Arockiaraj 2017, in the review
Bablishvili 2015	Did not contain specimen for extrapulmonary TB
Bajrami 2016	Could not extract 2 × 2 values
Balcha 2014	Did not contain specimen for extrapulmonary TB
Bemba 2017	Inappropriate reference standard
Bhatia 2016	Could not extract 2 × 2 values
Biadglegne 2013	Could not extract 2 × 2 values
Bilgin 2016	Could not extract 2 × 2 values

Study	Reason for exclusion
Bunsow 2014	Could not extract 2 × 2 values
Celik 2015	Could not extract 2 × 2 values
Chen 2016	Could not extract 2 × 2 values
Coleman 2015	Case-control study
Deggim 2013	Fewer than 5 specimens for a given type of specimen (only 1 pleural fluid specimen)
Dharan 2016	Did not contain specimen for extrapulmonary TB
Diop 2016	Inappropriate reference standard
Edwards 2016	Case report
Erdem 2014	Index test other than Xpert MTB/RIF
Fanosie 2016	Did not contain specimen for extrapulmonary TB
Gascoyne-Binzi 2012	Abstract; we could not extract data by form of extrapulmonary TB
Habeenzu 2017	Did not contain specimen for extrapulmonary TB
Ioannidis 2010	Duplicate data
Jain 2017	Inappropriate reference standard
Kilfoil 2015	Could not extract 2 × 2 values
Kim 2014	Could not extract 2 × 2 values; unclear if culture-positive; pleural fluid (3), CSF (2); peritoneal fluid (1)
Kim 2015b	Case-control study
Kim 2015c	Could not extract 2 × 2 values
Kumar 2017	Case-control study
Kurbaniyazova 2017	Did not contain specimen for extrapulmonary TB
Kwak 2015	Duplicate data
Lawn 2012	Screening study
Lawn 2013	Could not extract 2 × 2 values
Lawn 2015	Screening study
Lawn 2017	Could not extract 2 × 2 values
Lee 2017	Duplicate data
Liu 2015	Duplicate data

Study	Reason for exclusion
Lombardi 2017	Could not extract data by site of extrapulmonary TB
Marouane 2014	Abstract; we excluded the publication, Marouane 2016, because we could not extract 2 × 2 values
Marouane 2016	Could not extract 2 × 2 values
Miller 2011	Fewer than 5 specimens for a given type of specimen; lymph node biopsy (3 specimens, of which 1 was culture-positive) and endometrial biopsy (1 specimen that was culture-positive)
Mishra 2017	Abstract; we did not identify a published study
Moure 2011	Fewer than 5 specimens for a given type of specimen: CSF (3 specimens, all culture-negative); pleural fluid (4 specimens, 2 culture-positive); lymph node aspirate (1 specimen, culture-negative); urine (2 specimens, both culture-positive); peritoneal fluid (2, both culture-negative)
Moure 2012	Case-control study
Nhu 2013	Inappropriate reference standard
Patel 2014	Duplicate data
Peter 2012	Case-control study
Porcel 2013	Case-control study
Rachow 2012	Did not contain specimen for extrapulmonary TB
Raizada 2015	Inappropriate reference standard
Ramamurthy 2016	Could not extract data by site of extrapulmonary TB
Razack 2014	Index test other than Xpert MTB/RIF
Saeed 2017b	Could not extract 2 × 2 values
Salvador 2015	Case-control study
Sanjuan Jimenez 2015	Case-control study
Shah 2016a	Case-control study
Singanayagam 2014	Could not extract 2 × 2 values
Singh 2016	Could not extract 2 × 2 values
Smith 2014	Did not contain specimen for extrapulmonary TB
Solomons 2015	Duplicate data
Theron 2014	Duplicate data
Toure 2017	Could not extract 2 × 2 values
Vallejo 2015	Could not extract 2 × 2 values

Study	Reason for exclusion
Verghese 2016	Abstract; we did not identify a published study
Wang 2016	Could not extract 2 × 2 values
Wei 2016	Inappropriate reference standard
Yuan 2016	Inappropriate reference standard
Zhang 2016	Could not extract 2 × 2 values

TB: tuberculosis.

DATA

Presented below are all the data for all of the tests entered into the review.

Table Tests. Data tables by test

Test	No. of studies	No. of participants
1 Cerebrospinal fluid	33	3820
2 Cerebrospinal fluid, Ultra	1	129
3 Pleural fluid, culture	30	4209
4 Pleural fluid, composite reference standard	5	405
5 Pleural tissue, culture	4	214
6 Pleural tissue, composite reference standard	1	55
7 Lymph node aspirate	19	1721
8 Lymph node tissue	10	484
9 Urine	19	1324
10 Bone or joint fluid	12	407
11 Bone or joint tissue	7	618
12 Peritoneal fluid	20	751
13 Peritoneal tissue	1	28
14 Pericardial fluid	18	435
15 Blood	3	277
16 Rifampicin resistance testing	39	1336

Test 1. Cerebrospinal fluid.

Test 2. Cerebrospinal fluid, Ultra.

Test 3. Pleural fluid, culture.

Test 4. Pleural fluid, composite reference standard.

Test 5. Pleural tissue, culture.

Test 6. Pleural tissue, composite reference standard.

Test 7. Lymph node aspirate.

Test 8. Lymph node tissue.

Test 9. Urine.

Test 10. Bone or joint fluid.

Test 11. Bone or joint tissue.

Test 12. Peritoneal fluid.

Test 13. Peritoneal tissue.

Test 14. Pericardial fluid.

Test 15. Blood.

Test 16. Rifampicin resistance testing.

ADDITIONAL TABLES

Table 1. Forms of extrapulmonary TB

Form of extrapulmonary TB	Characteristics	Diagnostic specimens and means of collection
TB meningitis, also called tuberculous meningitis	TB infection of the meninges affects people of all ages but is most common among children and people with untreated HIV infection. In adults, TB meningitis presents with gradual onset of headache, neck stiffness, malaise, and fever, and, if untreated, can progress to altered sensorium, focal neurological deficits, coma, and death. Young children may present with poor weight gain, low-grade fever, and listlessness. Infants may present with fever, cough (related to the primary pulmonary infection that occurs before TB meningitis develops), change of consciousness at presentation, bulging anterior fontanel, and seizures (Thwaites 2013). TB meningitis is sometimes associated with a concurrent cerebral tuberculoma, or, more rarely, a tuberculous abscess	Cerebrospinal fluid, acquired by lumbar puncture with or without radiological guidance; biopsy of tuberculoma, acquired surgically
Pleural TB, also called TB pleurisy	TB infection of the pleura presents with gradual onset of pleuritic chest pain, shortness of breath, fever, night sweats, and weight loss. Chest X-ray may demonstrate unilateral or occasionally bilateral pleural effusion. The severity of symptoms is highly variable, with many patients experiencing spontaneous resolution of symptoms, while others may develop severe pleural effusions requiring drainage. Pleuro-pulmonary TB, in which parenchymal lung involvement is visible on a chest X-ray, is associated with higher mortality than isolated pleural infection, which appears to be rarely fatal (Shu 2011)	Pleural fluid; pleural biopsy, which may be performed via thoracoscopy or percutaneously with Abram's needle, with or without ultrasound guidance
Lymph node TB, also called TB lymphadenitis	TB infection of lymph nodes may affect 1 node or a group of nodes, or multiple groups within a chain. Lymph node TB is relatively more common among children than adults. The most common presentation is of a single, firm, non-tender enlarged node in the neck, although any lymph node group can be affected. This may be accompanied by fever, weight loss, and night sweats, particularly in people with HIV. Patients with TB in deep lymph nodes, such as the mediastinal or mesenteric lymph nodes, may present with fever, night sweats, and weight loss, or, more rarely, with symptoms related to compression of adjacent structures. Over time lymph nodes become fluctuant and may discharge via a sinus to the skin or an adjacent viscus. It should be noted that lymphadenopathy may also be seen in other forms of TB as part of the immune response, but this is not usually caused by direct infection of the lymph nodes	Fine-needle aspiration of fluid from affected lymph node, with or without radiological guidance; surgical biopsy of superficial lymph nodes; endoscopic biopsy of deep lymph nodes with ultrasound guidance
Bone or joint TB	TB infection of bones or joints or both causes chronic pain, deformity, and disability, and TB of the cervical spine can be life-threatening. The usual presenting symptom is pain. Fever and weight loss, with or without signs of spinal cord compression, may be present. Patients with advanced disease may have severe pain, spinal deformity, paraspinal muscle wasting, and neurological deficit. Children may have failure to thrive and difficulty walking	Aspiration of joint fluid or periarticular abscesses; percutaneous computed tomography-guided biopsy of lesions is preferred, but some

Table 1. Forms of extrapulmonary TB *(Continued)*

		patients may require open biopsy
Genitourinary TB	TB infection of the genitourinary tract includes renal TB and TB of the reproductive system. Renal TB presents with flank pain, haematuria, and dysuria. Female genital TB presents with infertility (and may be otherwise asymptomatic), pelvic pain, and vaginal bleeding. Testicular TB presents with a scrotal mass and infertility	Urine; biopsy of affected organs, acquired under radiological guidance or surgically
Pericardial TB, also called TB pericarditis	TB infection of the pericardium presents with fever, malaise, night sweats, and weight loss. Chest pain and shortness of breath are also commonly experienced symptoms. Pericardial TB may be associated with pericardial effusion, which can be severe and lead to life-threatening tamponade. Some patients go on to develop pericardial constriction, which can lead to heart failure and death and may require surgical intervention even after mycobacterial cure	Pericardial fluid acquired by pericardiocentesis; pericardial biopsy, acquired under radiological guidance or surgically
Peritoneal TB	TB infection of the peritoneum usually presents with pain and abdominal swelling, which may be accompanied by fever, weight loss, and anorexia	Ascitic fluid acquired by paracentesis; peritoneal biopsy (Chow 2002)
Disseminated TB, also called miliary TB. It has been proposed that the designation 'miliary TB' be restricted to disseminated TB with miliary shadows on chest radiograph (Reuter 2009)	Disseminated TB refers to TB that involves 2 or more distinctly separate sties. Manifestations may be varied, ranging from acute fulminant disease to non-specific symptoms of fever, weight loss, and weakness. HIV-positive people are more likely to have disseminated TB than HIV-negative people. In a systematic review of the prevalence of TB in postmortem evaluations of HIV-positive people, among adults disseminated TB was found in 88% of TB cases and was considered the cause of death in 91% of TB cases (Gupta 2015)	Blood; specimens acquired from affected extrapulmonary sites

Abbreviations: HIV: human immunodeficiency virus; TB: tuberculosis.
We adapted the table from Index-TB 2016.

Table 2. Systematic reviews of Xpert® MTB/RIF for extrapulmonary TB *(Continued)*

Systematic review	Search period	Number of studies (total number of extrapulmonary specimens)	Forms of extrapulmonary TB or types of specimens	Accuracy against culture reference standard		
				TB meningitis	Pleural TB (pleural fluid)	Lymph node TB
Chang 2012[a]	Up to 1 October 2011	7 (1058)	Multiple forms combined	Not reported	Not reported	Not reported
Denkinger 2014[b]	Up to 15 October 2013	18 (4461)	Lymph node, pleural fluid, CSF	Sensitivity 81%; specificity 98%	Sensitivity 46%; specificity 99%	Sensitivity 83%; specificity 94%

Table 2. Systematic reviews of Xpert® MTB/RIF for extrapulmonary TB *(Continued)*

Maynard-Smith 2014	Up to 6 November 2013	27 (6026)	Lymph node, pleural fluid, CSF, other forms	Median sensitivity 85% (IQR 75% to 100%); median specificity 100% (IQR 98% to 100%)	Sensitivity 34%; specificity 98%	Sensitivity 96%; specificity 93%
Penz 2015	Up to 15 August 2014	36 (9523)	Lymph node, pleural fluid, CSF, other forms	Sensitivity 69%; specificity 97%	Sensitivity 37%; specificity 98%	Sensitivity 87%; specificity 92%
Sehgal 2016	Up to 31 August 2015	24 (2486)	Pleural fluid	Not applicable	Sensitivity 51%; specificity 99%	Not applicable
Li Y 2017[c]	Up to 20 June 2015	26 (not reported)	Multiple forms combined	Not reported	Not reported	Not reported

Abbreviations: CI: confidence interval; CSF: cerebrospinal fluid: IQR: interquartile range; TB: tuberculosis.
[a]For all forms of extrapulmonary TB combined, Chang 2012 reported pooled sensitivity and specificity of 80.4% (95% confidence interval (CI) 75.0 to 85.1) and 86.1% (95% CI 83.5 to 88.4), respectively.
[b]Using a composite reference standard, Denkinger 2014 found the following pooled sensitivity and specificity estimates: lymph node TB (aspirate or tissue) 81.2% (95% CI 72.4 to 87.7) and 99.1% (95% CI 94.5 to 99.9); pleural TB 21.4% (95% CI 8.8 to 33.9) and 100% (95% CI 99.4 to 100); and meningeal TB 62.8% (95% CI 47.7 to 75.8) and 98.8% (95% CI 95.7 to 100), respectively.
[c]For both pulmonary and extrapulmonary TB, review authors included 106 studies involving 52,410 samples. For all forms of extrapulmonary TB combined, Li Y 2017 reported pooled sensitivity and specificity of 80% (95% CI 69 to 88) and 97% (95% CI 94 to 98), respectively.

Table 3. Accuracy of Xpert® MTB/RIF for detection of extrapulmonary TB and rifampicin resistance

Form of extrapulmonary TB, type of specimen	Number of studies (specimens)	Number of specimens with culture-confirmed TB (%)	Pooled sensitivity (95% credible interval)	Pooled specificity (95% credible interval)	Predicted sensitivity (95% credible interval)	Predicted specificity (95% credible interval)
TB meningitis, cerebrospinal fluid	29 (3774)	433 (11.5)	71.1% (60.9 to 80.4)	98.0% (97.0 to 98.8)	71.1% (27.8 to 94.8)	98.0% (88.1 to 99.7)
Pleural TB, fluid[a]	27 (4006)	607 (15.2)	50.9% (39.7 to 62.8)	99.2% (98.2 to 99.7)	50.9% (12.3 to 88.8)	99.2% (81.6 to 100)
Pleural TB, tissue	3 (207)	71 (34.3)	30.5% (3.5 to 77.8)	97.4% (92.1 to 99.3)	30.9% (0.2 to 98.2)	97.4% (87.3 to 99.6)
Lymph node, aspirate	17 (1710)	671 (39.2)	87.6% (81.7 to 92.0)	86.0% (78.4 to 91.5)	87.7% (58.1 to 97.4)	86.0% (46.5 to 97.9)
Lymph node, tissue	10 (484)	147 (30.4)	84.4% (74.7 to 91.0)	78.9% (52.6 to 91.5)	78.9% (52.6 to 91.5)	78.9% (9.1 to 99.2)
Genitourinary TB, urine	13 (1199)	73 (6.1)	82.7% (69.6 to 91.1)	98.7% (94.8 to 99.7)	82.7% (54.3 to 95.1)	98.8% (45.2 to 100)

Table 3. Accuracy of Xpert® MTB/RIF for detection of extrapulmonary TB and rifampicin resistance *(Continued)*

Bone or joint TB, fluid	5 (385)	58 (15.1)	97.2% (89.5 to 99.6)	90.2% (55.6 to 98.5)	97.3% (83.9 to 99.7)	90.5% (6.1 to 99.9)
Bone or joint TB, tissue	7 (618)	179 (29.0)	91.8% (82.5 to 96.8)	82.0% (56.6 to 94.9)	91.8% (70.1 to 98.4)	82.0% (10.4 to 99.5)
Peritoneal TB, fluid	16 (712)	115 (16.2)	59.2% (45.2 to 73.5)	97.9% (96.2 to 99.1)	59.1% (23.3 to 88.8)	97.9% (93.4 to 99.6)
Pericardial TB, fluid	7 (324)	76 (23.5)	65.7% (46.3 to 81.4)	96.0% (85.8 to 99.3)	65.7% (30.7 to 89.3)	95.9% (41.8 to 99.9)
Disseminated TB, blood	2 (266)	23 (8.6)	-	-	-	-
Rifampicin resistance	[b]	[b]	95.0% (89.7 to 97.9)	98.7% (97.8 to 99.4)	95.0% (82.3 to 98.8)	98.7% (95.3 to 99.7)

Abbreviations: CrI: credible interval; TB: tuberculosis.
Studies included in the table are limited to those that report data for both sensitivity and specificity; thus the number of studies (specimens) may differ slightly from those reported in the main text of the review. For TB detection, the reference standard was culture. For rifampicin resistance detection, the reference standard was culture-based drug susceptibility testing or MTBDR*plus*. Pooled sensitivity and pooled specificity are posterior median estimates.
[a]For pleural fluid measured against the composite reference standard, pooled sensitivity and specificity (95% CrI) were 18.4% (9.9 to 30.7) and 98.2% (94.8 to 99.5).
[b]Univariate analyses: pooled sensitivity included 20 studies (148 specimens); pooled specificity included 39 studies (1088 specimens). Bivariate analysis: pooled sensitivity and specificity (95% CrI) were 95.0% (89.9 to 97.9) and 98.8% (97.7 to 99.6) (20 studies, 990 specimens). We did not perform a meta-analysis for blood owing to sparse data.

Table 4. Impact of TB prevalence on sensitivity and specificity

Analysis (number of studies, specimens)	Pooled sensitivity (95% credible interval)	Pooled specificity (95% credible interval)
Cerebrospinal fluid		
Among studies with prevalence ≥ 10% (17, 1704)	72.0% (59.7 to 82.8)	96.8% (95.0 to 98.2)
Among studies with prevalence < 10% (12, 2070)	68.2% (50.9 to 82.4)	98.9% (97.9 to 99.4)
Difference (≥ 10% group minus < 10% group)	3.8% (-13.8 to 23.5)	-2.0% (-3.8 to -0.4)
Probability (difference > 0)	0.658	0.008
Pleural fluid		
Among studies with prevalence ≥ 15% (15, 1847)	58.0% (45.0 to 70.2)	99.0% (97.5 to 99.8)
Among studies with prevalence < 15% (12, 2159)	38.0% (23.9 to 55.5)	99.3% (98.1 to 99.8)
Difference (≥ 15% group minus < 15% group)	19.8% (-0.9 to 37.9)	-0.3% (-1.8 to 0.9)
Probability (difference > 0)	0.970	0.296
Lymph node aspirate		

Table 4. Impact of TB prevalence on sensitivity and specificity *(Continued)*

Among studies with prevalence ≥ 43% (10, 925)	92.6% (88.1 to 95.7)	84.0% (72.0 to 92.1)
Among studies with prevalence < 43% (7, 785)	78.5% (69.2 to 86.4)	89.3% (80.6 to 94.5)
Difference (≥ 43% group minus < 43% group)	14.0% (5.3 to 23.6)	-5.1% (-17.7 to 6.0)
Probability (difference > 0)	0.999	0.248
Urine		
Among studies with prevalence ≥ 7% (8, 504)	87.9% (75.1 to 95.1)	98.1% (93.5 to 99.6)
Among studies with prevalence < 7% (5, 695)	69.6% (45.3 to 87.1)	99.3 % (96.3 to 99.8)
Difference (≥ 7% group minus < 7% group)	18.0% (-1.5 to 41.5)	-1.1% (-5.0 to 1.4)
Probability (difference > 0)	0.963	0.137
Rifampicin resistance		
Among studies with prevalence ≥ 12% (10, 536)	96.2% (91.1 to 98.7)	98.7% (96.8 to 99.6)
Among studies with prevalence < 12% (11, 479)	92.0% (80.0 to 97.4)	99.1% (97.7 to 99.7)
Difference (≥ 12% group minus < 12% group)	4.0% (-2.6 to 15.9)	-0.3% (-2.2 to 1.1)
Probability (difference > 0)	0.878	0.310

Abbreviations: TB: tuberculosis.
Prevalence refers to the percentage of culture-confirmed TB specimens or confirmed rifampicin-resistant specimens in the study. We used median prevalence in the studies.

Table 5. Sensitivity analyses

Type of specimen	Number of studies (specimens)	Pooled sensitivity (95% credible interval)	Pooled specificity (95% credible interval)	Predicted sensitivity (95% credible interval)	Predicted specificity (95% credible interval)
Cerebrospinal fluid					
All participants	29 (3774)	71.1% (60.9 to 80.4)	98.0% (97.0 to 98.8)	71.1% (27.8 to 94.8)	98.0% (88.1 to 99.7)
Consecutive participant selection	25 (3408)	71.2% (59.8 to 80.9)	98.2% (97.0 to 99.0)	71.1% (24.9 to 95.1)	98.2% (87.6 to 99.8)
Reference standard blinding	27 (3723)	70.5% (59.8 to 79.8)	98.0% (96.9 to 98.8)	70.4% (26.7 to 94.2)	98.0% (87.6 to 99.7)
Participants not on anti-TB treatment	12 (2257)	72.8% (60.5 to 83.4)	98.6% (97.4 to 99.3)	72.7% (36.0 to 93.2)	98.6% (92.2 to 99.8)
Single specimen per patient	15 (1835)	63.5% (47.6 to 76.3)	96.1% (94.2 to 97.4)	63.7% (17.9 to 93.1)	96.1% (87.6 to 98.9)

Table 5. Sensitivity analyses *(Continued)*

Pleural fluid					
All participants	27 (4006)	50.9% (39.7 to 62.8)	99.2% (98.2 to 99.7)	50.9% (12.3 to 88.8)	99.2% (81.6 to 100.0)
Consecutive participant selection	20 (3381)	48.2% (36.6 to 61.5)	98.8% (97.7 to 99.6)	48.2% (12.9 to 86.2)	98.9% (81.7 to 100.0)
Reference standard blinding	19 (3301)	48.8% (37.9 to 60.8)	98.5% (96.8 to 99.5)	48.8% (15.0 to 84.0)	98.5% (75.4 to 100.0)
Participants not on anti-TB treatment	9 (1822)	43.1% (25.0 to 64.1)	97.9% (94.3 to 99.4)	43.2% (6.3 to 90.0)	97.9% (65.9 to 99.9)
Single specimen per patient	13 (1160)	51.0% (39.2 to 63.0)	97.4% (95.2 to 98.9)	50.9% (20.0 to 81.3)	97.4% (84.9 to 99.7)
Lymph node aspirate					
All participants	17 (1710)	87.6% (81.7 to 92.0)	86.0% (78.4 to 91.5)	87.7% (58.1 to 97.4)	86.0% (46.5 to 97.9)
Consecutive participant selection	15 (1660)	88.4% (82.7 to 92.8)	85.1% (76.9 to 91.2)	88.4% (61.0 to 97.6)	85.1% (44.6 to 97.7)
Reference standard blinding	15 (1694)	87.5% (81.0 to 92.2)	85.6% (77.6 to 91.5)	87.5% (56.4 to 97.5)	85.5% (44.3 to 97.9)
Participants not on anti-TB treatment	6 (852)	82.3% (69.2 to 90.3)	88.8% (80.9 to 93.8)	82.3% (46.1 to 96.1)	88.8% (65.5 to 97.2)
Single specimen per patient	11 (1183)	90.5% (84.7 to 94.4)	84.4% (72.9 to 92.3)	90.5% (68.3 to 97.7)	84.3% (36.9 to 98.3)
Adults only	6 (789)	83.1% (69.2 to 91.5)	91.2% (85.2 to 95.0)	83.0% (44.5 to 96.9)	91.2% (75.7 to 97.3)

Abbreviations: TB: tuberculosis.
Pooled sensitivity and pooled specificity are posterior median estimates.

Table 6. Latent class meta-analysis

Form of extrapulmonary TB, type of specimen	Number of studies (specimens)	Culture-confirmed TB (%)	Pooled sensitivity (95% credible interval)	Pooled specificity (95% credible interval)	Predicted sensitivity (95% credible interval)	Predicted specificity (95% credible interval)
Accuracy estimates of Xpert® MTB/RIF						
TB meningitis, cerebrospinal fluid	29 (3774)	433 (11.5)	63.2% (53.8 to 73.6)	99.6% (98.5 to 99.9)	63.1% (39.9 to 83.0)	99.6% (98.3 to 99.9)
Pleural TB, fluid	27 (4006)	607 (15.2)	56.4% (44.7 to 68.9)	99.7% (98.1 to 100.0)	56.5% (25.6 to 83.5)	99.7% (99.0 to 99.9)
Lymph node TB, aspirate[a]	17 (1710)	671 (39.2)	92.2% (82.9 to 98.1)	89.2% (78.9 to 98.2)	92.3% (72.6 to 98.8)	90.1% (57.9 to 98.6)

Table 6. Latent class meta-analysis *(Continued)*

Lymph node TB, aspirate[b]	17 (1710)	671 (39.2)	81.5% (73.4 to 88.3)	99.0% (98.1 to 99.5)	81.4% (54.4 to 94.9)	99.0% (78.4 to 100)
Accuracy estimates of culture						
TB meningitis, cerebrospinal fluid	29 (3774)	433 (11.5)	68.6% (59.0 to 78.0)	99.3% (98.1 to 99.8)	68.5% (44.9 to 86.5)	99.3% (97.7 to 99.8)
Pleural TB, fluid	27 (4006)	607 (15.2)	81.8% (69.5 to 91.2)	98.1% (95.9 to 99.5)	81.5% (43.7 to 97.1)	98.1% (95.0 to 99.5)
Lymph node TB, aspirate[a]	17 (1710)	671 (39.2)	88.5% (75.2 to 98.1)	91.6% (84.6 to 97.1)	89.6% (51.5 to 98.8)	91.7% (81.3 to 97.5)
Lymph node TB, aspirate[b]	17 (1710)	671 (39.2)	78.8% (68.9 to 89.8)	99.6% (99.4 to 99.8)	79.3% (45.5 to 94.8)	99.6% (98.5 to 97.9)
Accuracy estimates of Xpert® MTB/RIF against culture as a reference standard[c]						
TB meningitis, cerebrospinal fluid	29 (3774)	433 (11.5)	71.1% (60.9 to 80.4)	98.0% (97.0 to 98.8)	71.1% (27.8 to 94.8)	98.0% (88.1 to 99.7)
Pleural TB, fluid	27 (4006)	607 (15.2)	50.9% (39.7 to 62.8)	99.2% (98.2 to 99.7)	50.9% (12.3 to 88.8)	99.2% (81.6 to 100)
Lymph node TB, aspirate	17 (1710)	671 (39.2)	87.6% (81.7 to 92.0)	86.0% (78.4 to 91.5)	87.7% (58.1 to 97.4)	86.0% (46.5 to 97.9)

Abbreviations: TB: tuberculosis.
We generally used non-informative priors in the latent class model.
[a]The model used non-informative priors.
[b]The model used informative priors.
[c]Accuracy estimates were determined via a bivariate random-effects approach for comparison.

APPENDICES

Appendix 1. Detailed search strategies

MEDLINE (OVID)

1 Mycobacterium tuberculosis/

2 Tuberculosis/ or "Tuberculosis, Multidrug-Resistant"/ or Extensively Drug-Resistant Tuberculosis/

3 (Tuberculosis or MDR-TB or XDR-TB or "Multidrug Resistant Tuberculosis" or "Extensively Drug Resistant Tuberculosis" or tuberculous).ti. ab .

4 (extrapulmonary or extra-pulmonary or EPTB).ti. ab .

5 (lymphadenitis or disseminated or miliary or pleur* or skeletal or spine or mening* or intracranial or intra-ocular or ocular or abdominal or splenic or genitourinary or pericardial).ti. ab .

6 "Tuberculosis, Central Nervous System"/ or "Tuberculosis, Urogenital"/ or "Tuberculosis, Splenic"/ or "Tuberculosis, Spinal"/ or "Tuberculosis, Renal"/ or "Tuberculosis, Pleural"/ or "Tuberculosis, Osteoarticular"/ or "Tuberculosis, Oral"/ or "Tuberculosis, Ocular"/ or "Tuberculosis, Meningeal"/ or "Tuberculosis, Lymph Node"/ or "Tuberculosis, Laryngeal"/ or "Tuberculosis, Hepatic"/ or "Tuberculosis,

Gastrointestinal"/ or "Tuberculosis, Female Genital"/ or "Tuberculosis, Endocrine"/ or "Tuberculosis, Cutaneous"/ or "Tuberculosis, Cardiovascular"/ or Tuberculosis, Miliary/ or Tuberculosis, Male Genital/

7 1 or 2 or 3

8 4 or 5

9 7 and 8

10 9 or 6

11 Xpert*.ti. ab .

12 (GeneXpert or cepheid).ti.ab .

13 (near* patient or near-patient).ti.ab

14 11 or 12 or 13

15 10 and 14

Embase (OVID)

1 Tuberculosis, Multidrug-Resistant/ or Extensively Drug-Resistant Tuberculosis/ or Tuberculosis/ or tuberculosis.mp. or Mycobacterium tuberculosis/

2 (MDR-TB or XDR-TB).mp.

3 1 or 2

4 (extrapulmonary or extra-pulmonary or EPTB).ti. or (extrapulmonary or extra-pulmonary or EPTB).ab.

5 (lymphadenitis or disseminated or miliary or pleur* or skeletal or spine or mening* or intracranial or intra-ocular or ocular or abdominal or genitourinary or pericardial).ti. or (lymphadenitis or disseminated or miliary or pleur* or skeletal or spine or mening* or intracranial or intra-ocular or ocular or abdominal or genitourinary or pericardial).ab.

6 tuberculous.ti. or tuberculous.ab.

7 3 or 6

8 Tuberculosis, Central Nervous System/ or Tuberculosis, Hepatic/ or Tuberculosis, Male Genital/ or Tuberculosis, Spinal/ or Tuberculosis, Cutaneous/ or Tuberculosis, Urogenital/ or Tuberculosis, Osteoarticular/ or Tuberculosis, Endocrine/ or Tuberculosis, Renal/ or Tuberculosis, Splenic/ or Tuberculosis, Ocular/ or Tuberculosis, Laryngeal/ or Tuberculosis, Gastrointestinal/ or Tuberculosis/ or Tuberculosis, Meningeal/ or Tuberculosis, Oral/ or Tuberculosis, Pleural/ or Tuberculosis, Lymph Node/ or Tuberculosis, Female Genital/ or Tuberculosis, Miliary/ or Tuberculosis, Cardiovascular/

9 4 or 5 or 8

10 7 and 9

11 xpert*TB.mp.

12 Xpert* MTB RIF.ti. or Xpert* MTB RIF.ab.

13 (GeneXpert or cepheid).ti. or (GeneXpert or cepheid).ab.

14 (near* patient or near-patient).ti. or (near* patient or near-patient).ab.

15 12 or 13 or 14

16 10 and 15

Indexes=SCI-EXPANDED, CPCI-S, Biosis previews

TOPIC

(tuberculosis or tuberculous) *AND* **TOPIC:** (extrapulmonary or extra-pulmonary or EPTB or lymphadenitis or disseminated or miliary or pleur* or skeletal or spine or mening* or intracranial or intra-ocular or ocular or abdominal or genitourinary or pericardial) *AND* **TOPIC:** (Xpert* or Genexpert or cepheid)

LILACS

tuberculosis or tuberculous [Words] and Xpert$ or Genexpert or cepheid [Words]

SCOPUS

(TITLE-ABS-KEY (tuberculosis OR tuberculous) AND TITLE-ABS-KEY (extrapulmonary OR extra-pulmonary OR eptb OR lymphadenitis OR disseminated OR miliary OR pleur* OR skeletal OR spine OR mening* OR intracranial OR intra-ocular OR ocular OR abdominal OR genitourinary OR pericardial) AND TITLE-ABS-KEY (xpert* OR genexpert OR cepheid))

Cochrane Infectious Diseases Group Specialist Register; ClinicalTrials.gov, WHO ICTRP, ISRCTN registry, ProQuest Dissertations & Theses A&I

tuberculosis and Xpert$; tuberculosis and Genexpert; tuberculosis and Cepheid.

Appendix 2. Data extraction form

Data extractor	MK KRS
First study author	
Corresponding study author and email	
Title of paper	
Journal	
Language if other than English	
Year	

I. Study details

Type of study: randomized controlled trial; cross-sectional cohort (with follow-up); case-control (exclude); unclear/not reported

Study data collection: prospective; retrospective; unclear/not reported

Participant selection: convenience; consecutive; random; other; unclear/not reported

Country:

Country income status: low; middle; high

II. Presenting signs and symptoms, setting

Presenting signs and symptoms?

Clinical setting: inpatient; outpatient; both; unclear/not reported

Level of laboratory running Xpert? peripheral; intermediate; central (reference)

Comments, describe exclusions

(Tests at laboratory levels)

Peripheral: AFB (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 (first- and second-line anti-TB drugs) on solid or in liquid media and LPA on positive cultures and rapid speciation tests

III. Other demographics

HIV patients included? yes; no; unclear/not reported; if yes ## and percentage? (*denominator is number tested, when possible*)

Age? Median age in years (IQR); mean (SD); range; unclear/not reported

Children (< 15 years old) included: yes; no; unclear/not reported; if yes, percentage?

Percentage female included? Unclear/not reported

Past history of TB? yes; no; unclear/not reported; if yes, percentage?

Only patients who received TB treatment for ≤ 7 days were included? yes; no; unclear/not reported; if no, percentage on treatment included?

IV. Reference standard

A. Reference standard for TB detection

Solid culture (specify): LJ 7H10 7H11; other

Liquid culture (specify): MGIT Bactec 460; other

Solid and liquid culture (indicate which kind above)

Were reference standard results interpreted without knowledge of index test results? yes; no; unclear/not reported

B. Composite reference standard for pleural TB

Solid culture (specify): LJ 7H10 7H11; other

Liquid culture (specify): MGIT Bactec 460; other

Solid and liquid culture (indicate which kind above)

Histopathology (specify): granulomas; caseating granulomas

Were reference standard results interpreted without knowledge of index test results? yes; no; unclear/not reported

Did all patients receive the same reference standard? yes; no; unclear/not reported; if no, describe

C. Reference standard for rifampicin resistance

LJ DST MGIT DST MTBDR*plus*

Were reference standard results interpreted without knowledge of index test results? yes; no; unclear/not reported

V. Sites with more than five specimens (check all that apply)

A. Lymph node TB fluid; tissue; both fluid and tissue

B. Pleural TB fluid; tissue; both fluid and tissue

C. TB meningitis CSF

D. Bone or joint TB fluid; tissue; both fluid and tissue

E. Genitourinary TB urine; other, specify

F. Peritoneal TB fluid; tissue; both fluid and tissue

G. Pericardial TB fluid; tissue; both fluid and tissue

H. Disseminated TB blood

I. Other, specify

VI. Specimen processing for Xpert

Condition of specimens: fresh frozen

If frozen for > 7 days, indicate WHO not followed

For a given site, how many specimens were collected per patient? one; multiple; unclear/not reported

A. Lymph node tissue, other tissue

Was the WHO standard operating procedure (SOP) followed for each specimen type?

1a. Lymph node tissue WHO followed: yes; no; unclear

1b. Lymph node tissue homogenization step for tissue specimens: yes; no; unclear/not reported

2a. Other tissue, specify WHO followed: yes; no; unclear

2b. Other tissue homogenization step for tissue specimens: yes; no; unclear/not reported

(For tissue, if WHO SOP not followed, briefly describe specimen processing in comments.)

WHO SOPs for specimen processing; lymph node and other tissue; sterile specimen

1. *Cut the tissue specimen into small pieces in a sterile mortar.*
2. *Add approximately 2 mL of sterile phosphate buffered saline (PBS).*
3. *Grind solution of tissue and PBS until homogeneous suspension has been obtained.*
4. *Place approximately 0.7 mL of the homogenized tissue in a sterile, conical screw-capped tube.*
5. *Double volume of specimen with Xpert® Sample Reagent (1.4 mL Sample Reagent to 0.7 mL of homogenized tissue).*
6. *Shake tube vigorously 10 to 20 times or vortex for at least 10 seconds.*
7. *Incubate specimen for 10 minutes at room temperature, and again shake specimen 10 to 20 times or vortex for at least 10 seconds.*
8. *Incubate specimen at room temperature for an additional 5 minutes.*
9. *Transfer 2mL to Xpert® MTB/RIF cartridge.*
10. *Load into GeneXpert and per manufacturer's instructions.*

(Note: For specimens **not collected in a sterile manner,** WHO SOP suggests an NaOH decontamination/concentration protocol similar to that used for sputum.)

B. CSF

3a. CSF WHO followed: yes; no; unclear

3b. CSF concentration step: yes; no; unclear/not reported

3c. CSF sample input volume: specify; unclear/not reported

(For CSF, if WHO SOP not followed, briefly describe specimen processing in comments.)

WHO SOPs for CSF

If more than 5 mL of CSF is available for testing.

1. *Transfer all of the CSF specimen to a conical centrifuge tube and concentrate the specimen at 3000 × g for 15 minutes.*
2. *Resuspend the pellet to a final volume of 2 mL by adding Xpert® MTB/RIF Sample Reagent.*
3. *Transfer 2 mL of the resuspended CSF sample to the Xpert® MTB/RIF cartridge.*
4. *Load the cartridge into the GeneXpert instrument according to the manufacturer's instructions.*

If 1 mL to 5 mL of CSF is available.

1. *Add an equal volume of Sample Reagent to the CSF.*
2. *Mix the specimen and the Sample Reagent by vortexing as described above. After seven to eight minutes at room temperature, vortex the sample as above a second time.*
3. *Incubate for an additional seven to eight minutes (15 minutes total incubation) at room temperature.*
4. *Add 2 mL of the sample mixture directly to the Xpert® MTB/RIF cartridge.*
5. *Load the cartridge into the GeneXpert instrument according to the manufacturer's instructions.*

C. Body fluids, other than CSF

4a. Body fluid: specify; processed as per manufacturer for sputum

Yes/No/Unclear

4b. Body fluid: specify; sample input volume: specify; unclear/not reported

5a. Body fluid: specify; processed as per manufacturer for sputum (WHO followed)

Yes/No/Unclear

5b. Body fluid: specify; sample input volume: specify; unclear/not reported

(Add additional specimens as needed.)

(For body fluids other than CSF, if manufacturer's instructions not followed, briefly describe specimen processing in comments.)

Manufacturer's instructions for sputum

Raw specimen

1. *Pour or pipette (pipette not provided) approximately 2 times the volume of Sample Reagent into the specimen (2:1 dilution, Sample Reagent: specimen).*
2. *Shake vigorously 10 to 20 times or vortex for at least 10 seconds.*
3. *Incubate sample for a total of 15 minutes at 20°C to 30°C.*
4. *Between 5 and 10 minutes into the incubation period, shake vigorously 10 to 20 times or vortex for at least 10 seconds.*

Specimen sediment

Assay requires at least 0.5 mL of resuspended specimen sediment after digestion, decontamination, and concentration.

1. *Use the method of Kent and Kubica and resuspend the sediment in a 67 mM phosphate/H_2O buffer.*
2. *After resuspension, keep at least 0.5 mL of the resuspended sediment for the Xpert® MTB/RIF assay.*
3. *Add 1.5 mL of Sample Reagent to 0.5 mL of the resuspended sediment (3:1 dilution, Sample Reagent: specimen)*
4. *Follow steps 2 to 4 above.*

Comments on specimen processing.

VII. Specimen processing for culture

Specimen collected from sterile site: Yes/No/Unclear

Specimen processed for culture as per American Thoracic Society Diagnostic Standards? Yes/No/Unclear

(ATS guidelines: specimens collected from normally sterile sites may be placed directly into the culture medium.)

Note: All specimens such as CSF, pleura, lymph node aspirates and tissues, peritoneal fluid, pericardial fluid, bone or joint fluid and tissue, and urine are considered sterile.

VIII. Results

TB detection: number error or invalid or both Xpert® MTB/RIF results over total number of cultures performed. The denominator includes contaminated cultures and cultures that were uninterpretable.

Unclear/not reported.

RIF resistance: number indeterminate Xpert results (over total number of cultures performed).

Unclear/not reported.

Non-tuberculous mycobacteria (NTM): number of cultures with NTM (over total number of cultures performed).

Unclear/not reported.

IX. Tables

(Non-tuberculous mycobacteria (NTM) should be included as not TB.)

TB detection against culture reference standard (example; provide additional tables for other extrapulmonary specimens).

(Continued)

Xpert in lymph node fluid		Definite TB		
		Yes	No	Total
Xpert result	Positive			
	Negative			
	Total			
	Error/invalid			

By smear status (extrapulmonary specimens)

(Continued)

Xpert in smear-positives		Definite TB		
		Yes	No	Total
Xpert result	Positive			
	Negative			
	Total			
	Error/invalid			

(Continued)

Xpert in smear-negatives		Definite TB		
		Yes	No	Total
Xpert result	Positive			
	Negative			
	Total			
	Error/invalid			

Rifampicin resistance detection (for all culture-positive, extrapulmonary specimens)

Rifampicin resistance detection	Rifampicin resistance

(Continued)

		Yes	No	Total
Xpert result	**Positive**			
	Negative			
	Total			
	Indeterminate			

Appendix 3. Rules for QUADAS-2

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 scored "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 could not tell.

Signalling question 2: was a case-control design avoided?

We did not include in the review studies using a case-control design because this study design, especially when used to compare results in severely ill patients versus those in relatively healthy individuals, may lead to overestimation of accuracy in diagnostic studies. We scored "yes" for all studies.

Signalling question 3: did the study avoid inappropriate exclusions?

We scored "yes" if the study included both smear-positive and smear-negative specimens or included only smear-negative specimens. We judged "no" if the study included only smear-positive specimens or excluded specimens based on physical appearance (such as purulence) or a biochemical analysis (e.g. adenosine deaminase (ADA), cytology (cell analysis)). We scored "unclear" if we could not tell.

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

We were interested in how Xpert performed in patients presumed to have extrapulmonary TB who were evaluated as they would be in routine practice. We scored "low concern" if patients were evaluated at local hospitals or primary care centres. We scored "high concern" if patients were evaluated exclusively as inpatients at tertiary care centres. We scored "unclear concern" if the clinical setting was not reported or if information was insufficient to allow a decision. We also scored "unclear concern" if Xpert testing was done at a reference laboratory and the clinical setting was not reported for the following reason. It was difficult to tell if a given reference laboratory provided services mainly to very sick patients (inpatients in tertiary care) or to all patients, 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 results of the reference standard?

We answered this question "yes" for all studies because Xpert 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 pre-specified?

As the threshold is pre-specified in all versions of Xpert, we answered this question "yes" for all studies.

Applicability: are there concerns that the index test, its conduct, or its interpretation differ from the review question?

We note that variations in execution of the test might affect accuracy estimates. We judged "low concern" if the test was performed according to WHO standard operating procedures (WHO 2014a), or if the index test was performed as recommended by the manufacturer for sputum. We scored "high concern" if the test was performed in a way that deviated from these recommendations. We scored "unclear concern" if we could not tell. In studies that evaluated several different types of specimens, we used the following rule: if ≥ 75% of the specimen types were processed per WHO standard operating procedure (SOP) or as per the manufacturer's instructions, we judged "low concern"; if < 50% of the specimen types were processed per WHO SOP or as per the manufacturer's instructions, we scored "high concern";

and if at least 50% to 74% of the specimen types were processed per WHO SOP or as per the manufacturer's instructions, or if we could not tell, we scored "unclear concern".

Domain 3: reference standard

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

We considered this domain separately for the reference standard for detection of extrapulmonary TB and the reference standard for detection of rifampicin resistance.

Signalling question 1: is the reference standard likely to correctly classify the target condition?

For detection of extrapulmonary TB, culture is generally considered the best reference standard. For the diagnosis of all forms of extrapulmonary TB (except as noted for pleural TB below), culture is a criterion for inclusion in the review. However, limitations are associated with culture; bacterial load is usually low in extrapulmonary TB, leading to a reduction in the sensitivity of culture. Concerning the conduct of the reference standard (preparation of the specimen for culture), N-acetyl-L-cysteine-sodium hydroxide is routinely used to homogenize, decontaminate, and liquefy non-sterile specimens for TB culture (American Thoracic Society 2000). However, CSF, pleural fluid, and lymph node aspirates are usually considered sterile, and standards specify, "specimens collected from normally sterile sites may be placed directly into the culture medium" (American Thoracic Society 2000). Overly processing (sterile) specimens with N-acetyl-L-cysteine-sodium hydroxide may lead to a decrease in viable TB bacteria and consequently false-negative cultures. We scored "yes" if studies did not use N-acetyl-L-cysteine-sodium hydroxide for processing specimens and "unclear" if studies used N-acetyl-L-cysteine-sodium hydroxide. We discussed this further under Discussion and Strengths and weaknesses of the review.

For detection of pleural TB, use of culture or a composite reference standard was a criterion for inclusion in the review. We answered this question "yes" for all studies of pleural TB.

For detection of rifampicin resistance, culture-based drug susceptibility testing (DST, also called conventional phenotypic method) is considered to be the best reference standard. MTBDR*plus* is also a WHO-recommended test for rifampicin resistance. As we extracted data only for studies that used culture-based DST or MTBDR*plus*, we answered this question "yes" for all studies.

Signalling question 2: were the reference standard results interpreted without knowledge of results of the index test?

We scored "yes" if the reference test provided an automated result (e.g. MGIT 960), if blinding was explicitly stated, or if it was clear that the reference standard was performed at a separate laboratory and/or was performed by different people. We scored "no" if the study stated that the reference standard result was interpreted with knowledge of the Xpert® MTB/RIF test result. We scored "unclear" if we could not tell.

Signalling question 3: (rifampicin resistance) were the reference standard results interpreted without knowledge of results of the index test?

We added a signalling question for rifampicin resistance detection. We scored "yes" if the reference test provided an automated result (e.g. MGIT 960), if solid culture was performed followed by speciation, if blinding was explicitly stated, or if it was clear that the reference standard was performed at a separate laboratory or was performed by different people, or both. We scored "no" if the study stated that the reference standard result was interpreted with knowledge of the Xpert test result. We scored "unclear" if we could not tell. Not all studies evaluated detection of rifampicin resistance; therefore this question was not applicable to all studies.

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

We judged "high concern" if included studies did not speciate mycobacteria isolated in culture, "low concern" if speciation was performed, and "unclear concern" if we could not tell.

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 the reference standard?

In most included studies, we expected that specimens for Xpert and culture would be obtained at the same time, when patients were evaluated for presumptive extrapulmonary TB. However, even if there were a delay of several days between index test and reference standard, TB is a chronic disease, and we considered misclassification of disease status to be unlikely, as long as treatment was not initiated in the interim. We judged "yes" if the index test and the 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 could not tell.

Signalling question 2: did all patients receive the same reference standard?

For the diagnosis of any form of extrapulmonary TB, except pleural TB, we answered this question "yes" if all participants in the study or a subset of participants in the study (for whom we extracted data) received the acceptable reference standard (solid culture, liquid culture, or both), which we specified as a criterion for inclusion in the review. However, we acknowledge that it is possible that some specimens could undergo solid culture and others liquid culture as the reference standard. This could potentially result in variations in accuracy, but we think the variation would be minimal. For the diagnosis of pleural TB as measured against a composite reference standard, we answered this

question "yes" if all participants received the same reference standard, "no" if not all participants received the same reference standard, and "unclear" if we could not tell.

For rifampicin resistance detection, we answered "yes" if all participants received the same reference standard (either culture-based DST or MTBDR*plus*), "no" if not all participants received the same reference standard, and "unclear" if we could not 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 × 2 tables. We will answer "yes" if the numbers matched and "no" if there were patients enrolled in the study who were not included in the analysis. We will answer "unclear" if we cannot tell.

Judgements for overall 'Risk of bias' assessments.

- If we answered all signalling questions for a domain "yes", then we scored risk of bias as "low".
- If we answered all or most signalling questions for a domain "no", then we scored risk of bias as "high".
- If we answered only one signalling question for a domain "no", we discussed further the "risk of bias" judgement.
- If we answered all or most signalling questions for a domain "unclear", then we scored risk of bias as "unclear".
- If we answered only one signalling question for a domain "unclear", we discussed further the "risk of bias" judgement for the domain.
- In the Reference Standard Domain, if we answered "yes" for both signalling questions concerning detection of extrapulmonary TB, we scored risk of bias as "low", regardless of our judgement for blinding of the reference standard for detection of rifampicin resistance.

Appendix 4. OpenBugs

In this section we provide OpenBUGS models for the bivariate meta-analysis as well as the latent class meta-analysis. Any alternative prior distributions used are provided in the comments within each model.

BIVARIATE MODEL ASSUMING PERFECT CULTURE REFERENCE TEST

```
model {

for(i in 1:N) { # N is the number of studies

############################## LIKELIHOOD
logit(TPR[i]) <- l[i,1]
logit(FPR[i]) <- -l[i,2]
pos[i]<-TP[i]+FN[i]
neg[i]<-TN[i]+FP[i]
TP[i] ~ dbin(TPR[i],pos[i])
FP[i] ~ dbin(FPR[i],neg[i])
se[i] <- TPR[i]
sp[i] <- 1-FPR[i]
l[i,1:2] ~ dmnorm(mu[1:2], T[1:2,1:2])

}
############################## HYPER PRIOR DISTRIBUTIONS
mu[1] ~ dnorm(0,0.25) # replaced by mu[1] ~ dnorm(0, 0.01) in sensitivity analysis to check impact of less informative prior
mu[2] ~ dnorm(0,0.25) # replaced by mu[2] ~ dnorm(0, 0.01) in sensitivity analysis to check impact of less informative prior
T[1:2,1:2]<-inverse(TAU[1:2,1:2])

#### BETWEEN-STUDY VARIANCE-COVARIANCE MATRIX
TAU[1,1] <- tau[1]*tau[1]
TAU[2,2] <- tau[2]*tau[2]
TAU[1,2] <- rho*tau[1]*tau[2]
TAU[2,1] <- rho*tau[1]*tau[2]
tau[1]<-pow(prec[1],-0.5) # replaced by tau[1] ~ dunif(0,3) in sensitivity analysis to check impact of less informative prior
tau[2]<-pow(prec[2],-0.5) # replaced by tau[2] ~ dunif(0,3) in sensitivity analysis to check impact of less informative prior
sigma.sq[1] <- pow(tau[1], 2)
sigma.sq[2] <- pow(tau[2], 2)

#### prec = between-study precision in the logit(sensitivity) and logit(specificity)

prec[1] ~ dgamma(2,0.5) # replaced by prec[1] <- 1/pow(tau[1],-2) in sensitivity analysis to check impact of less informative prior
prec[2] ~ dgamma(2,0.5) # replaced by prec[2] <- 1/pow(tau[2],-2) in sensitivity analysis to check impact of less informative prior
```

```
rho ~ dunif(-1,1)

############################## OTHER PARAMETERS OF INTEREST

#### POOLED SENSITIVITY AND SPECIFICITY

Pooled_S<-1/(1+exp(-mu[1]))
Pooled_C<-1/(1+exp(-mu[2]))

#### POOLED POSITIVE AND NEGATIVE LIKELIHOOD RATIOS

PLR <- Pooled_S/(1-Pooled_C)
NLR <- (1-Pooled_S)/Pooled_C

#### PREDICTED SENSITIVITY AND SPECIFICITY IN A FUTURE STUDY

l.new[1:2] ~ dmnorm(mu[],T[,])
sens.new <- 1/(1+exp(-l.new[1]))
spec.new <- 1/(1+exp(-l.new[2]))

} #### END OF PROGRAM
```

LATENT CLASS META-ANALYSIS MODEL

```
# WinBUGS PROGRAM FOR ESTIMATING A BIVARIATE HIERARCHICAL META-ANALYSIS MODEL
# FOR SENSITIVITY AND SPECIFICITY ALLOWING FOR HETEROGENEITY BETWEEN STUDIES

model {

############################## ##############################

for(i in 1:N) {# N is the number of studies

############################## LIKELIHOOD
logit(p[1, i]) <- l[i,1]
logit(p[2, i]) <- -l[i,2]


prob[i,1] <- pi[i]*(p[1,i]* s2[i] + covp[i]) + (1-pi[i])*(p[2,i]*(1-c2[i]) + covn[i])
prob[i,2] <- pi[i]*(p[1,i]* (1-s2[i]) - covp[i]) + (1-pi[i])*(p[2,i]*c2[i] - covn[i])
prob[i,3] <- pi[i]*((1-p[1,i])*s2[i] - covp[i]) + (1-pi[i])*((1-p[2,i])*(1-c2[i]) - covn[i])
prob[i,4] <- pi[i]*((1-p[1,i])*(1-s2[i]) + covp[i]) + (1-pi[i])*((1-p[2,i])*c2[i] + covn[i])

n[i] <- sum(cell[i,1:4])
cell[i,1:4] ~ dmulti(prob[i,1:4],n[i])

pi[i] ~ dbeta(1,1)

se[i] <- p[1,i]
sp[i] <- 1-p[2,i]


l[i,1:2] ~ dmnorm(mu[1:2], T[1:2,1:2])

#=================================================================
# CONDITIONAL DEPENDENCE
#=================================================================

#=====================================
# upper limits of covariance parameters
#=====================================
us[i]<-min(se[i],s2[i])-(se[i]*s2[i]);
```

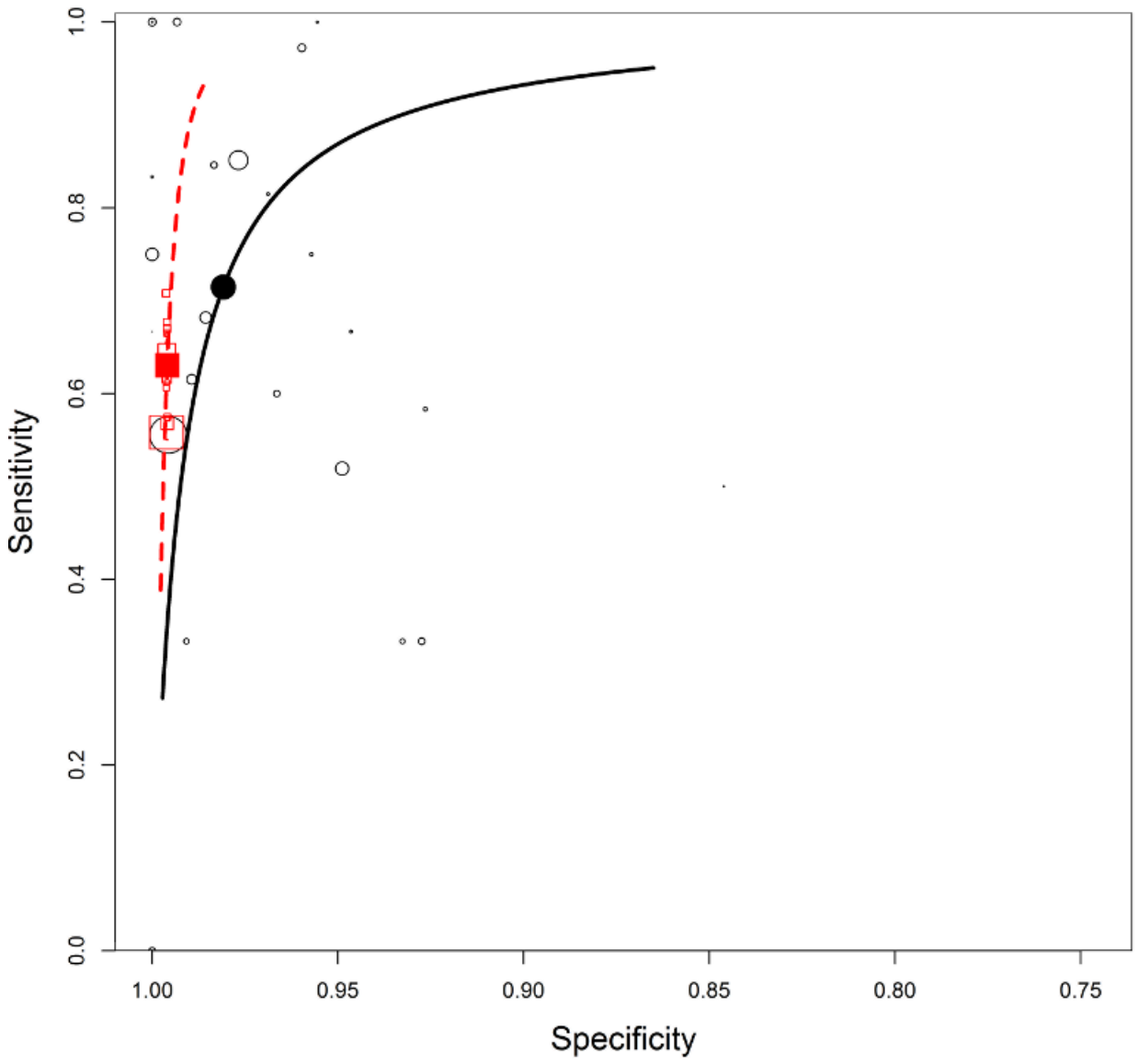
```
uc[i]<-min(sp[i],c2[i])-(sp[i]*c2[i]);

ls[i]<- -(1-se[i])*(1-s2[i])
lc[i]<- -(1-sp[i])*(1-c2[i])


#============================================================
# prior distribution of transformed covariances on (0,1) range
#============================================================
covp[i]~dunif(ls[i],us[i]);
covn[i]~dunif(lc[i],uc[i]);
#covn[i]<-0


}


# ==================================
# NON-INFORMATIVE HIERARCHICAL PRIOR DISTRIBUTION OVER REF STD PROPERTIES
# ==================================

for(j in 1:29) {

logit(s2[j]) <- l2[j,1]
logit(c2[j]) <- l2[j,2]
l2[j,1:2] ~ dmnorm(mu2[1:2], T2[1:2,1:2])


}

############################## HYPER PRIOR DISTRIBUTIONS ##############################

############################################################
############################################################
###
### XPERT TEST
###
############################################################
############################################################

mu[1] ~ dnorm(0,0.25)
mu[2] ~ dnorm(0,0.25) #dnorm(4.59512,10)
T[1:2,1:2]<-inverse(TAU[1:2,1:2])

#### BETWEEN-STUDY VARIANCE-COVARIANCE MATRIX
TAU[1,1] <- tau[1]*tau[1]
TAU[2,2] <- tau[2]*tau[2]
TAU[1,2] <- rho*tau[1]*tau[2]
TAU[2,1] <- rho*tau[1]*tau[2]

tau[1]<-pow(prec[1],-0.5)
tau[2]<-pow(prec[2],-0.5)

sigma.sq[1] <- pow(tau[1], 2)
sigma.sq[2] <- pow(tau[2], 2)

#### prec = between-study precision in the logit(sensitivity) and logit(specificity)
prec[1] ~ dgamma(2,0.5)
prec[2] ~ dgamma(2,0.5)
rho ~ dunif(-1,1)

############################## OTHER PARAMETERS OF INTEREST
```

```
#### POOLED SENSITIVITY AND SPECIFICITY OF XPERT
Pooled_S<-1/(1+exp(-mu[1]))
Pooled_C<-1/(1+exp(-mu[2]))

#### POOLED POSITIVE AND NEGATIVE LIKELIHOOD RATIOS
PLR <- Pooled_S/(1-Pooled_C)
NLR <- (1-Pooled_S)/Pooled_C

#### PREDICTED SENSITIVITY AND SPECIFICITY OF XPERT IN A FUTURE STUDY
l.new[1:2] ~ dmnorm(mu[],T[,])
sens.new <- 1/(1+exp(-l.new[1]))
spec.new <- 1/(1+exp(-l.new[2]))


############################################################
############################################################
###
### CULTURE TEST
###
############################################################
############################################################


mu2[1] ~ dnorm(0,0.25)
mu2[2] ~ dnorm(0,0.25)
T2[1:2,1:2]<-inverse(TAU2[1:2,1:2])

#### BETWEEN-STUDY VARIANCE-COVARIANCE MATRIX
TAU2[1,1] <- tau2[1]*tau2[1]
TAU2[2,2] <- tau2[2]*tau2[2]
TAU2[1,2] <- rho2*tau2[1]*tau2[2]
TAU2[2,1] <- rho2*tau2[1]*tau2[2]

tau2[1] <-pow(prec2[1],-0.5)
tau2[2] <-pow(prec2[2],-0.5)

sigma.sq2[1] <- pow(tau2[1], 2)
sigma.sq2[2] <- pow(tau2[2], 2)

#### prec = between-study precision in the logit(sensitivity) and logit(specificity)
prec2[1] ~ dgamma(2,0.5)
prec2[2] ~ dgamma(2,0.5)
rho2 ~ dunif(-1,1)



#### POOLED SENSITIVITY AND SPECIFICITY OF CULTURE
S2<-1/(1+exp(-mu2[1]))
C2<-1/(1+exp(-mu2[2]))

s2.new <- 1/(1+exp(-ls2.new))
c2.new <- 1/(1+exp(-lc2.new))
ls2.new ~ dnorm(mu2[1],prec2[1])
lc2.new ~ dnorm(mu2[2],prec2[2])

}
```

Appendix 5. Receiver operating characteristic plot for TB meningitis

Figure 10 displays the receiver operating characteristic plot for TB meningitis.

Figure 10. Receiver operating characteristic plot for TB meningitis. The black curve corresponds to the model that assumes culture is a perfect reference standard. The black emptied circles are plotted at co-ordinates corresponding to study sensitivity and specificity estimates obtained from the data. The filled black circle is the pooled estimate of sensitivity and specificity obtained from the bivariate model under the assumption that culture is a perfect reference standard. The red dashed line corresponds to the latent class meta-analysis. The red emptied squares are plotted at sensitivity and specificity co-ordinates corresponding to sensitivity and specificity estimates obtained from the latent class model. The filled red square has co-ordinates corresponding to pooled sensitivity and specificity estimates from the latent class model. The size of the emptied circles and squares is proportionate to the size of the studies.



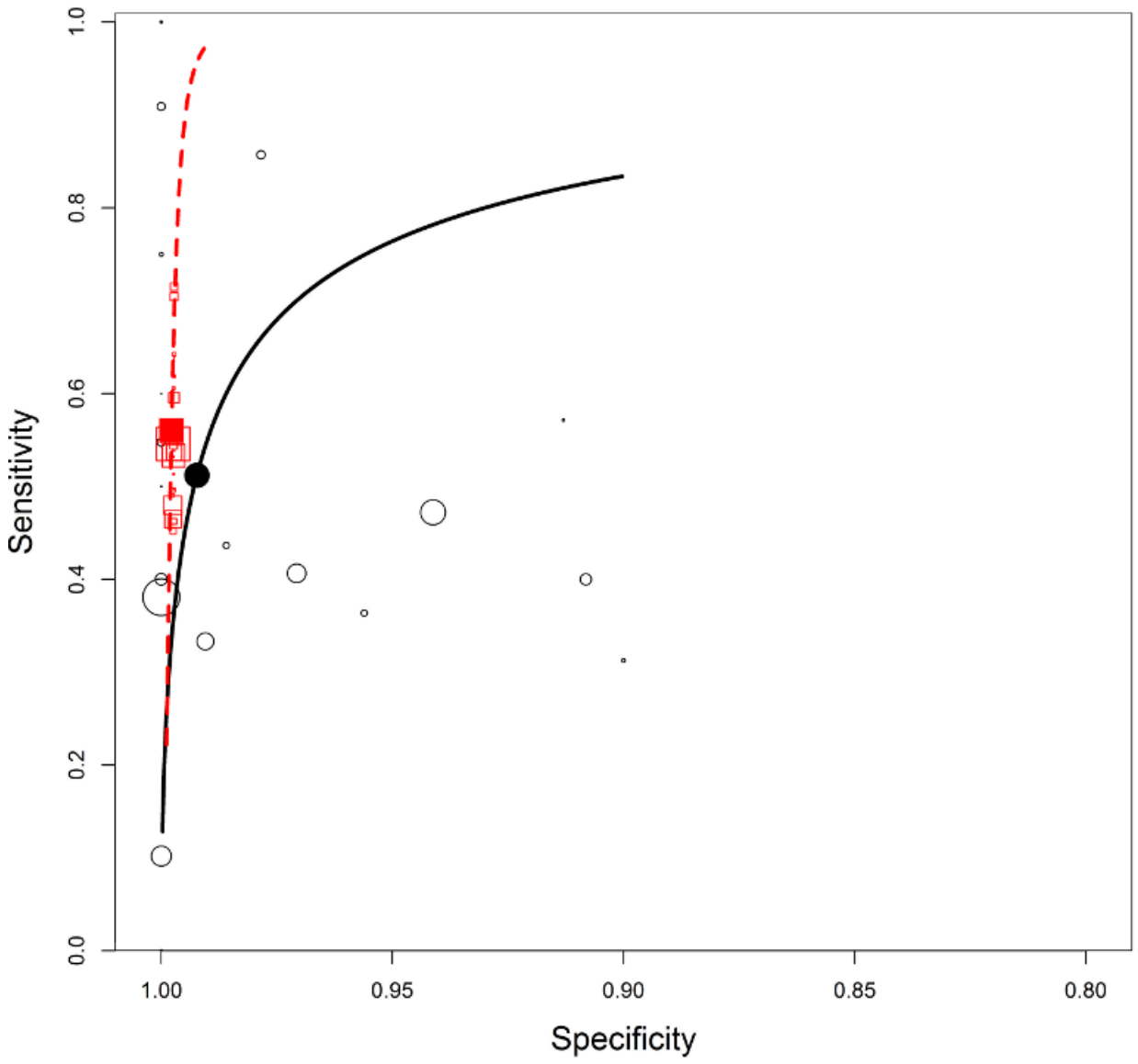
Appendix 6. Impact of concentrating cerebrospinal fluid on Xpert® MTB/RIF accuracy

Covariate (number of studies, participants)	Pooled sensitivity (95% credible interval)	Pooled specificity (95% credible interval)
Concentration step		
Concentrated specimen (15, 2758)	74.8% (63.1 to 84.4)	98.3% (97.1 to 99.1)
Unconcentrated specimen (12, 905)	66.2% (48.5 to 81.4)	97.7% (95.4 to 99.0)
Difference (concentrated minus unconcentrated)	8.5% (-9.9 to 27.7)	0.6% (-1.1 to 2.9)
Probability (concentrated minus unconcentrated)	0.825	0.754

Appendix 7. Receiver operating characteristic plot for pleural fluid

Figure 11 displays the receiver operating characteristic plot for pleural fluid.

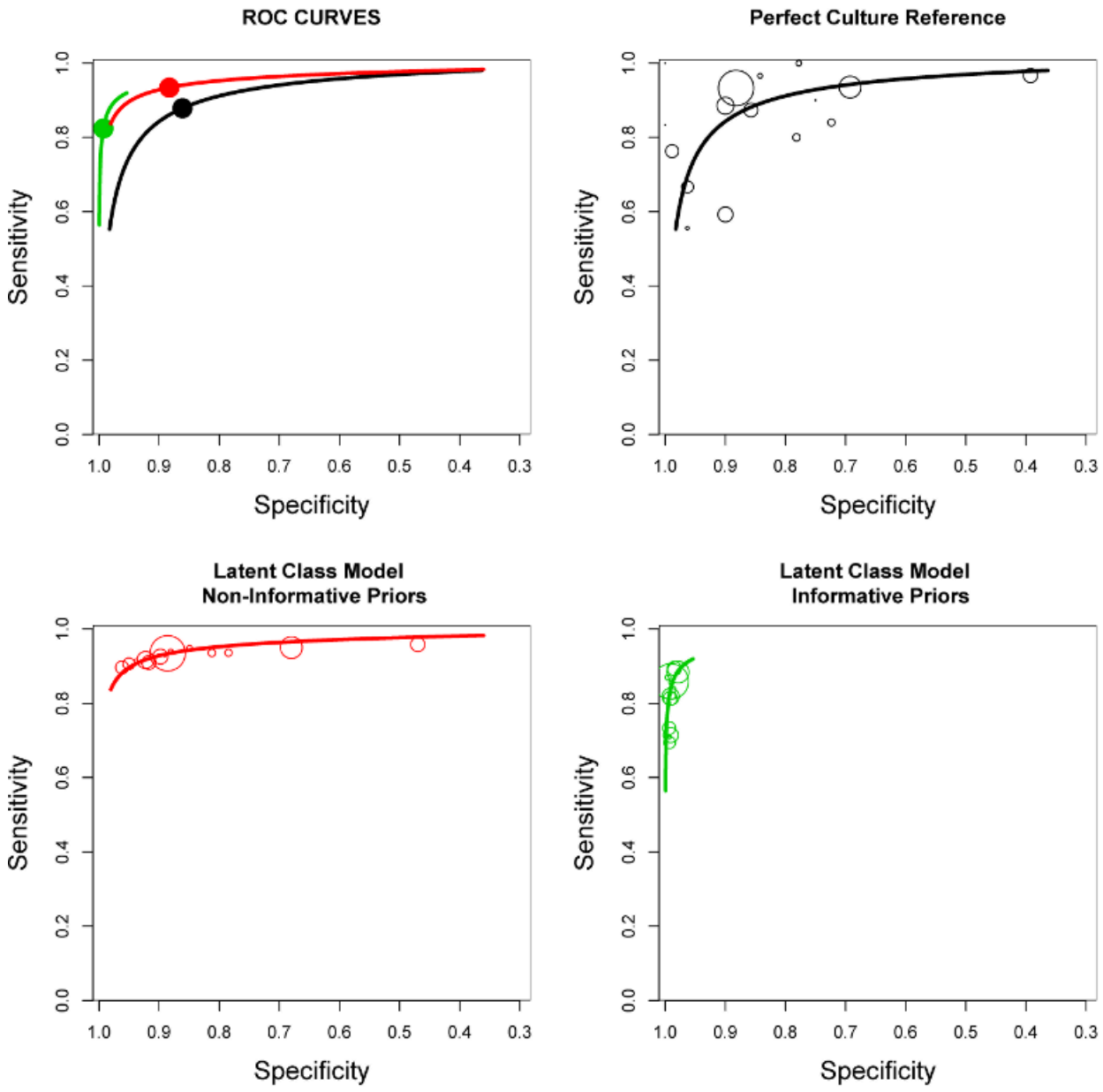
Figure 11. Receiver operating characteristic plot for pleural fluid. The black curve corresponds to the model that assumes culture is a perfect reference standard. The black emptied circles are plotted at co-ordinates corresponding to study sensitivity and specificity estimates obtained from the data. The filled black circle is the pooled estimate of sensitivity and specificity obtained from the bivariate model under the assumption that culture is a perfect reference standard. The red dashed line corresponds to the latent class meta-analysis. The red emptied squares are plotted at sensitivity and specificity co-ordinates corresponding to sensitivity and specificity estimates obtained from the latent class model. The filled red square has co-ordinates corresponding to pooled sensitivity and specificity estimates from the latent class model. The size of the emptied circles and squares is proportionate to the size of the studies.



Appendix 8. Receiver operating characteristic plot for lymph node aspirate

Figure 12 displays the receiver operating characteristic plot for lymph node aspirate.

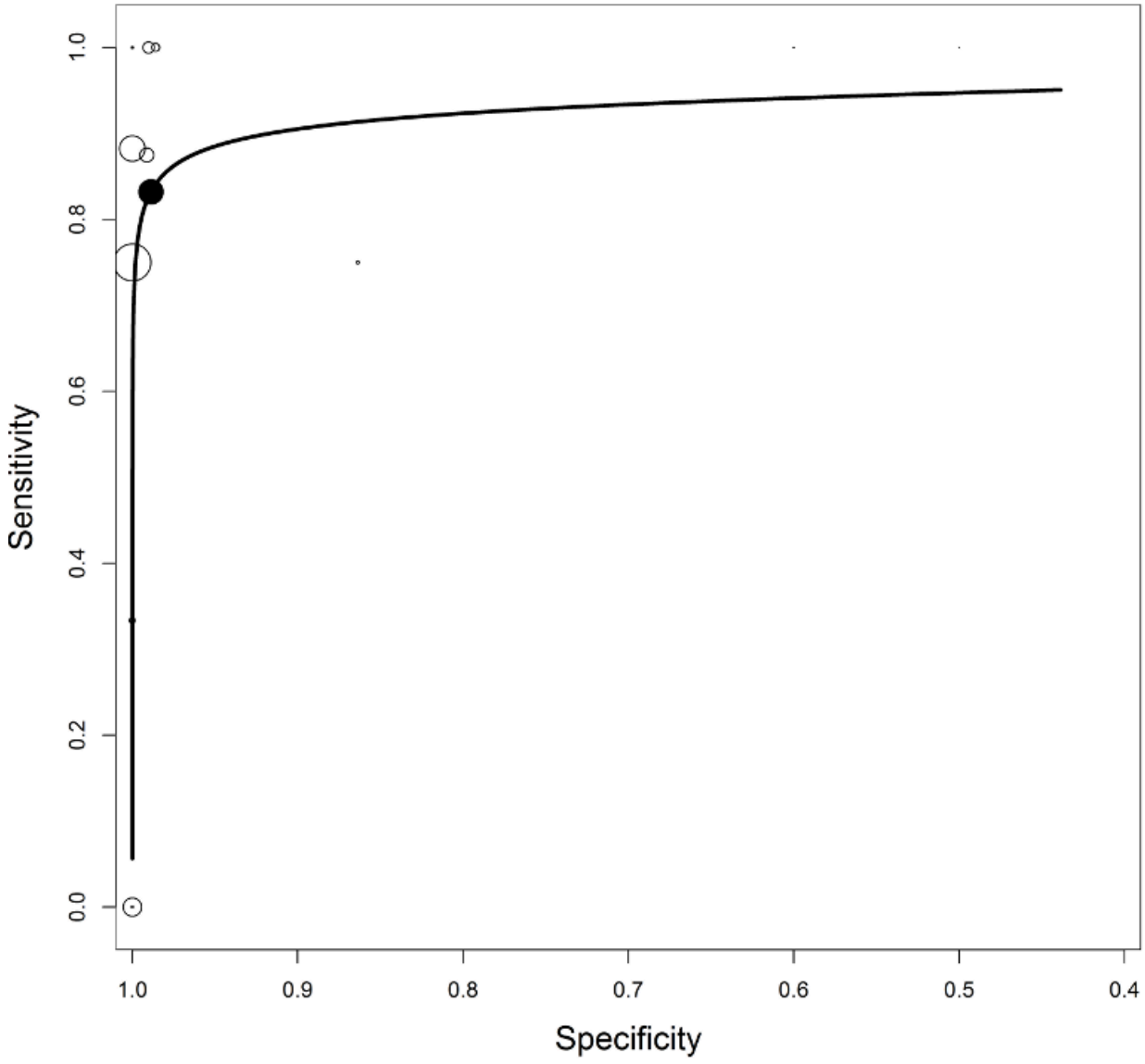
Figure 12. Receiver operating characteristic plot for lymph node aspirate. The black curve corresponds to the model that assumes culture is a perfect reference standard. The red curve corresponds to the latent class meta-analysis model with non-informative priors. The green curve corresponds to the latent class meta-analysis model with informative priors. The filled circles of each colour correspond to the pooled sensitivity and specificity of the respective model. The empty circles for each colour are plotted at sensitivity and specificity co-ordinates corresponding to sensitivity and specificity estimates obtained from the respective models. The size of the emptied circles is proportionate to the size of the studies.



Appendix 9. Receiver operating characteristic plot for urine

Figure 13 displays the receiver operating characteristic plot for urine.

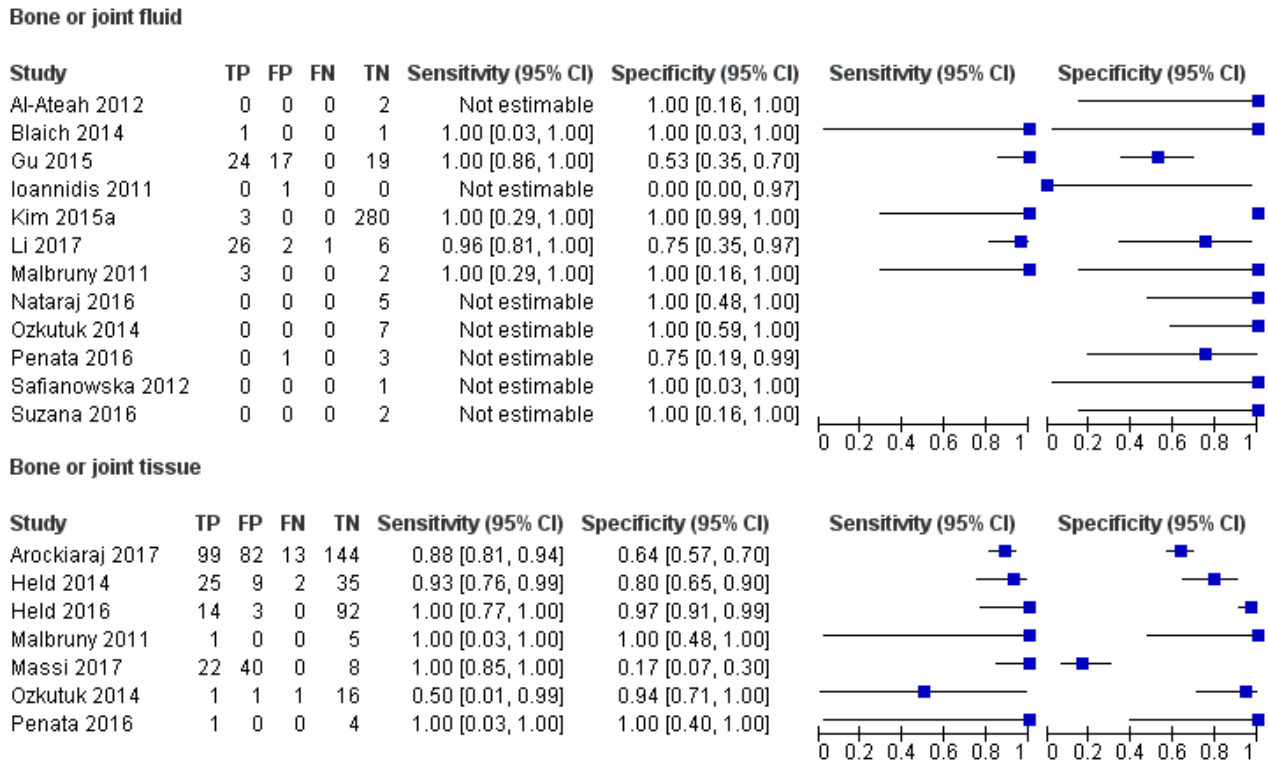
Figure 13. Receiver operating characteristic plot for urine. The black curve corresponds to the model that assumes culture is a perfect reference standard. The black emptied circles are plotted at co-ordinates corresponding to study sensitivity and specificity estimates obtained from the data. The filled black circle is the pooled estimate of sensitivity and specificity obtained from the bivariate model under the assumption that culture is a perfect reference standard. The size of the emptied circles and squares is proportionate to the size of the studies.



Appendix 10. Bone or joint TB

Figure 14 displays forest plots of Xpert sensitivity and specificity in bone or joint fluid and tissue.

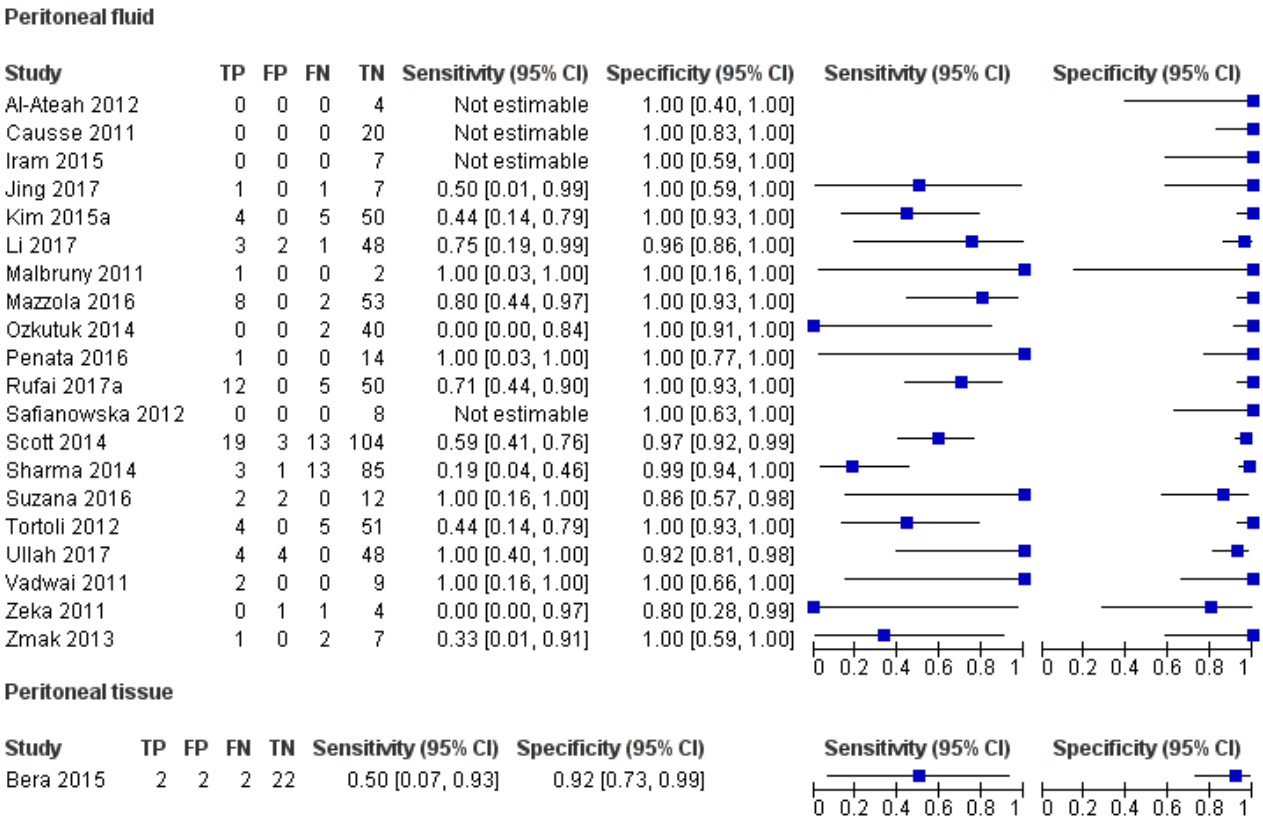
Figure 14. Forest plots of Xpert® MTB/RIF sensitivity and specificity for bone or joint TB (fluid and tissue) with respect to a culture reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive.



Appendix 11. Peritoneal TB

Figure 15 displays forest plots of Xpert sensitivity and specificity in peritoneal fluid and tissue.

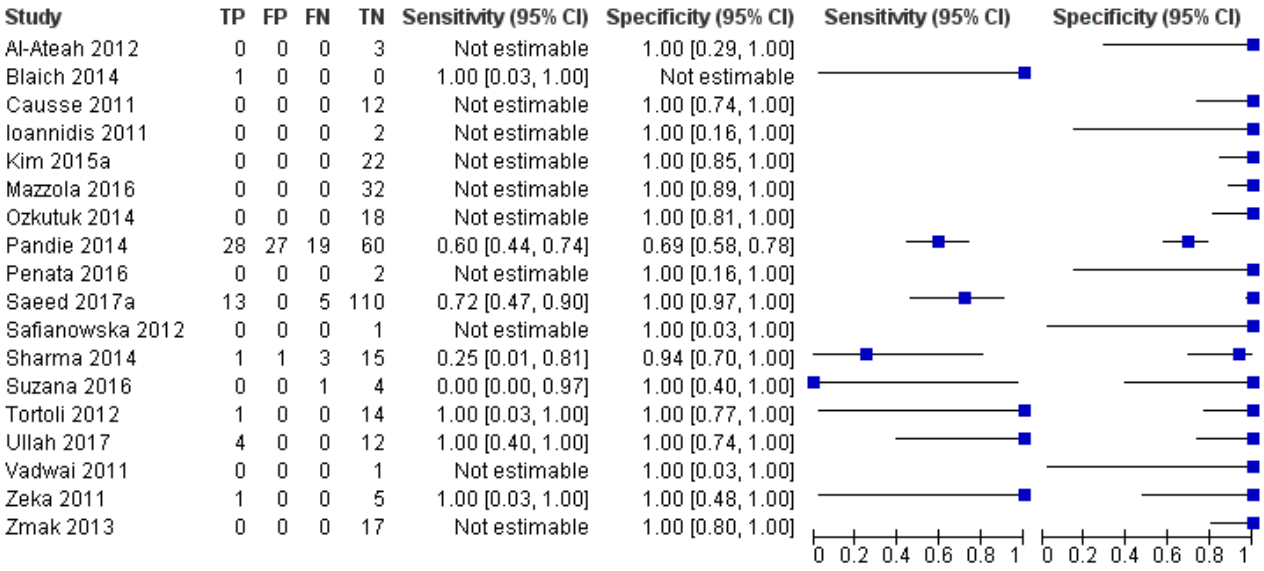
Figure 15. Forest plots of Xpert® MTB/RIF sensitivity and specificity for peritoneal TB (fluid and tissue) with respect to a culture reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive.



Appendix 12. Pericardial TB

Figure 16 displays forest plots of Xpert sensitivity and specificity in pericardial fluid.

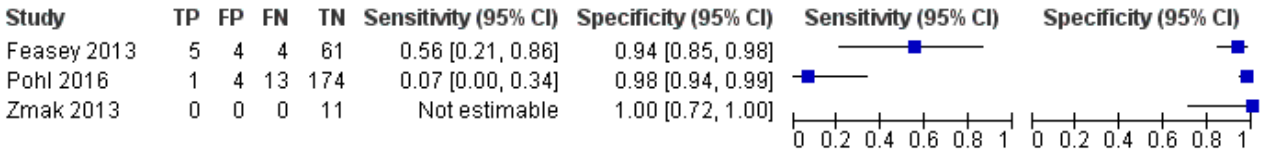
Figure 16. Forest plots of Xpert® MTB/RIF sensitivity and specificity in pericardial fluid with respect to a culture reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive.



Appendix 13. Disseminated TB

Figure 17 displays forest plots of Xpert sensitivity and specificity in blood.

Figure 17. Forest plots of Xpert® MTB/RIF sensitivity and specificity in blood with respect to a culture reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive.



CONTRIBUTIONS OF AUTHORS

MK and KRS wrote early drafts of the protocol. CMD and SGS contributed methodological advice. KD contributed clinical expertise. CMD and SGS tailored QUADAS-2 to the review. MK and KRS reviewed the studies and extracted accuracy data. MK and KRS assessed the methodological quality of included studies. IS and ND performed the statistical analyses. All review authors interpreted the findings. MK, ND, and KRS wrote the first draft of the review. MK and KRS prepared the 'Summary of findings' tables. All review authors contributed to the final manuscript.

DECLARATIONS OF INTEREST

We have no financial involvement with any organization or entity that has a financial interest in, or financial conflict with, the subject matter or materials discussed in the review apart from those disclosed.

CMD and SGS work for FIND. FIND is a non-for-profit foundation whose mission is to find diagnostic solutions to overcome diseases of poverty in low- and middle-income countries. FIND works closely with the private and public sectors and receives funding from donors and some of its industry partners. FIND has an independent Scientific Advisory Committee and organizational firewalls that protect it against any undue influences in its work or in publication of its findings. More information on FIND's policy and guidelines for working with private sector partners can be found at www.finddx.org/business-model.

KRS received financial support for the submitted work from the CIDG, and has received financial support for the preparation of systematic reviews and educational materials, consultancy fees from FIND (for the preparation of a systematic review), honoraria, and travel support to attend WHO guideline meetings.

ND received funding from the CIDG.

SOURCES OF SUPPORT

Internal sources

- Liverpool School of Tropical Medicine, UK.

External sources

- Department for International Development, UK.

  Project number 300342-104

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

QUADAS-2: we modified QUADAS-2 as follows. Reference standard domain: we clarified that CSF, pleural fluid, and lymph node aspirates are usually considered to be sterile, and standards specify that these specimens may be placed directly into the culture medium. Overly processing specimens may lead to false-negative cultures. We scored ‘yes' if studies did not use N-acetyl-L-cysteine-sodium hydroxide for processing sterile specimens and "unclear" if studies used N-acetyl-L-cysteine-sodium hydroxide.

Investigations of heterogeneity: for specimen volume, we restricted this analysis to CSF because it was most clinically meaningful. For other fluid specimen types, the manufacturer's instructions for sputum were usually followed requiring 2 mL of input fluid for the Xpert cartridge. In terms of the WHO standard operating procedure for lymph node tissue, we did not investigate this further because 80% (8/10) of the included studies followed WHO recommendations. In performing the review, it became clear that because a homogenization step is part of the WHO standard operating procedure for preparing tissue specimens, there was no need to perform an additional separate analysis to confirm the presence of a homogenization step. We removed condition of specimen (fresh or frozen) from the analysis because we identified only six studies in the current review that used frozen specimens, and we had already performed an analysis of this possible source of heterogeneity for the Cochrane Review on Xpert for pulmonary TB (Steingart 2014).

In the case of lymph node TB, for which we suspected a systematic bias in the performance of culture, we used informative prior distributions over the specificity of culture (ranging from 99% to 100%) and the specificity of Xpert (ranging from 98% to 100%).

We performed sensitivity analyses that limited inclusion to studies that reported one specimen per patient, and for lymph node aspirate limited inclusion to studies that involved only adults.

We have tried to eliminate stigmatizing language, for example, by changing ‘suspected TB' to ‘presumptive TB'.

INDEX TERMS

Medical Subject Headings (MeSH)

*Reagent Kits, Diagnostic; Antibiotics, Antitubercular [*therapeutic use]; Bacterial Proteins [*genetics]; DNA-Directed RNA Polymerases [*genetics]; Drug Resistance, Bacterial [*genetics]; False Negative Reactions; False Positive Reactions; Mycobacterium tuberculosis [drug effects] [*genetics] [isolation & purification]; Reference Standards; Rifampin [*therapeutic use]; Sensitivity and Specificity; Tuberculosis [cerebrospinal fluid] [*diagnosis] [drug therapy]; Tuberculosis, Meningeal [cerebrospinal fluid] [drug therapy]

MeSH check words

Humans