# Using the Xpert MTB/RIF assay to detect pulmonary and extrapulmonary tuberculosis and rifampicin resistance in adults and children




World Health Organization

Using the Xpert MTB/RIF assay to detect pulmonary and extrapulmonary tuberculosis and rifampicin resistance in adults and children

# Expert Group Meeting Report

2013

This report contains the collective views of an international group of experts, and does not necessarily represent the decisions or the stated policy of the World Health Organization. Mention of a technology in the report does not imply endorsement of any specific commercial product.

World Health Organization

**© World Health Organization 2013**

All rights reserved. Publications of the World Health Organization are available on the WHO web site (www.who.int) or can be purchased from WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland (tel.: +41 22 791 3264; fax: +41 22 791 4857; e-mail: bookorders@who.int).

Requests for permission to reproduce or translate WHO publications – whether for sale or for noncommercial distribution – should be addressed to WHO Press through the WHO web site (http://www.who.int/about/licensing/copyright_form/en/index.html).

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by the World Health Organization to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall the World Health Organization be liable for damages arising from its use.

Designed by GPS Publishing

WHO/HTM/TB/2013.14

## Contents


**1. EXECUTIVE SUMMARY** **X**

**2. BACKGROUND** **1**

**3. EVIDENCE BASE** **2**

**3.1. Evidence synthesis** **2**

**3.2. Meeting objectives** **2**

**3.3. GRADE evaluation** **3**

3.3.1. PICO questions for each review 4

3.3.2. Determining the relative importance of patients' outcomes 6

3.3.3. Assessment of study quality 6

**3.4. PROCEDURAL ISSUES** **6**

**4. RESULTS** **7**

**4.1. Using Xpert MTB/RIF to diagnose pulmonary TB** **7**

4.1.1. Characteristics of the studies 7

4.1.2. Quality of the studies 8

4.1.3. Using Xpert MTB/RIF as an initial test to replace smear microscopy 9

4.1.4. Investigations of heterogeneity: TB detection in smear-positive and smear-negative individuals suspected of having TB 11

4.1.5. Detecting TB in HIV-negative and HIV-positive individuals suspected of having pulmonary TB 13

4.1.6. Detecting TB among HIV-positive individuals by smear status 13

4.1.7. Effect of the condition of the specimen 14

4.1.8. Effect of specimen preparation 14

4.1.9. Effect of the proportion of culture-confirmed cases of TB on a study 15

4.1.10. Effect of a country's income status 15

4.1.11. Detecting rifampicin resistance 15

4.1.12. Investigations of heterogeneity: detecting rifampicin resistance 16

4.1.13. Summary of findings and GRADE evidence profiles 19

4.1.14. Strengths and limitations of the evidence base 19

4.1.15. GRADE evaluation and recommendations 20

**4.2. Using Xpert MTB/RIF to diagnose extrapulmonary TB and rifampicin resistance in adults and children** **30**

4.2.1. Characteristics of the studies 30

4.2.2. Quality of the studies 36

4.2.3. Using Xpert MTB/RIF to detect extrapulmonary TB 37
4.2.4. Detecting rifampicin resistance 44
4.2.5. Summary of findings and GRADE evidence profiles 44
4.2.6. Strengths and limitations of the evidence base 46
4.2.7. GRADE evaluations and recommendations 46
4.2.8. Further research needs 56
**4.3. Using Xpert MTB/RIF to diagnose pulmonary TB, peripheral lymph node TB, TB meningitis and rifampicin resistance in children 56**
4.3.1. Characteristics of the studies 56
4.3.2. Quality of the studies 57
4.3.3. Using Xpert MTB/RIF as the initial test to detect pulmonary TB 58
4.3.4. Xpert MTB/RIF compared with smear microscopy 60
4.3.5. Investigations of heterogeneity 62
4.3.6. Using Xpert MTB/RIF to detect peripheral lymph node TB in children 66
4.3.7. Using Xpert MTB/RIF to detect TB meningitis in children 67
4.3.8. Using Xpert MTB/RIF to detect rifampicin resistance in children 67
4.3.9. Summary of findings and GRADE evidence profiles 68
4.3.10. Strengths and limitations of the evidence base 68
4.3.11. GRADE evaluations and recommendations 69
**4.4. Affordability and cost effectiveness of using Xpert MTB/RIF to diagnose TB 82**
4.4.1. Summary of cost analyses 82
4.4.2. Summary of cost–effectiveness analyses 82
4.4.3. Summary of findings 83
4.4.4. Recommendations 83

Annex 1. Meeting participants 84
Annex 2. Meeting agenda 86
Annex 3. Declarations of Interests 91
Annex 4. Selection of studies evaluating the accuracy of Xpert MTB/RIF in diagnosing pulmonary TB and rifampicin resistance in adults 92
Annex 5. Selection of studies evaluating the accuracy of Xpert MTB/RIF in diagnosing extrapulmonary TB and rifampicin resistance in adults and children 102
Annex 6. Selection of studies evaluating the accuracy of Xpert MTB/RIF in diagnosing pulmonary and extrapulmonary TB, and rifampicin resistance in children 108
Annex 7. Literature review of affordability and cost effectiveness of Xpert MTB/RIF for the diagnosis of TB 112

## Tables


**Table 1.** Pooled sensitivity and specificity of the Xpert MTB/RIF assay for detecting pulmonary TB and rifampicin resistance 10

**Table 2.** Impact of covariates on the heterogeneity of the sensitivity and specificity of Xpert MTB/RIF in detecting pulmonary TB in 24 studies (33 study centres) 12

**Table 3.** GRADE evidence profile: accuracy of Xpert MTB/RIF in diagnosing pulmonary TB in adults 21

**Table 4.** GRADE evidence profile: accuracy of Xpert MTB/RIF in diagnosing pulmonary TB in sputum smear-positive adults 22

**Table 5.** GRADE evidence profile: accuracy of Xpert MTB/RIF in diagnosing pulmonary TB in sputum smear-negative adults 23

**Table 6.** GRADE evidence profile: accuracy of Xpert MTB/RIF in diagnosing pulmonary TB in adults living with HIV 24

**Table 7.** GRADE evidence profile: accuracy of Xpert MTB/RIF in diagnosing pulmonary TB in adults without HIV infection 25

**Table 8.** GRADE evidence profile: the incremental yield of Xpert MTB/RIF compared with microscopy in patients with culture-confirmed TB 26

**Table 9.** GRADE evidence profile: accuracy of Xpert MTB/RIF in diagnosing pulmonary TB in adults as an add-on test following negative sputum-smear microscopy 27

**Table 10.** Sensitivity of Xpert MTB/RIF in smear-negative culture-confirmed pulmonary TB in individuals, by HIV status 28

**Table 11.** GRADE evidence profile: additional yield of Xpert MTB/RIF over microscopy in smear-negative TB 28

**Table 12.** GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting rifampicin resistance, where Xpert MTB/RIF replaces phenotypic culture-based drug-susceptibility testing as the initial test 29

**Table 13.** Sample types and processing methods used in the 22 studies included in the review of the accuracy of Xpert MTB/RIF in diagnosing extrapulmonary TB and rifampicin resistance in adults and childrena 31

**Table 14.** Sensitivity and specificity of Xpert MTB/RIF in detecting extrapulmonary TB, by sample type 39

**Table 15.** Accuracy of Xpert MTB/RIF in detecting TB in lymph node fluid and tissue (A. Evidence profile, B. Summary of findings) 47

**Table 16.** Accuracy of Xpert MTB/RIF in detecting TB in pleural fluid (A. Evidence profile, B. Summary of findings) 49

**Table 17.** Accuracy of Xpert MTB/RIF in detecting TB in cerebrospinal fluid (A. Evidence profile, B. Summary of findings) 51

**Table 18.** GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting TB in gastric fluid 53

**Table 19.** GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting TB in tissue samples 54

**Table 20.** GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting rifampicin resistance in nonrespiratory specimens 55

**Table 21.** Meta-analysis of the estimated sensitivity and specificity of Xpert MTB/RIF in diagnosing pulmonary TB, peripheral lymph node TB and TB meningitis in children compared against culture as a reference standard (in published and unpublished studies) as well as against the clinical TB reference standard 60

**Table 22.** Meta-analysis of the estimated sensitivity and specificity of smear microscopy in diagnosing pulmonary TB, peripheral lymph node TB and TB meningitis in children compared against culture as a reference standard in published and unpublished studies 61

**Table 23.** Meta-analysis of the sensitivity and specificity of Xpert MTB/RIF in diagnosing pulmonary TB, peripheral lymph node TB and TB meningitis compared against culture as a reference standard in smear-negative and smear-positive children in published and unpublished studies 63

**Table 24.** Meta-analysis comparing Xpert MTB/RIF for diagnosing pulmonary TB, peripheral lymph node TB and TB meningitis using culture as a reference standard in HIV-positive and HIV-negative children, stratified by smear status 64

**Table 25.** Metaregression model for Xpert MTB/RIF using samples of expectorated or induced sputum from children, controlling for smear status and HIV status 64

**Table 26.** GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting TB in children compared with culture as a reference standard (A. Expectorated sputum and induced sputum, B. Gastric lavage or aspirate, C. Summary of findings) 70

**Table 27.** GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting TB in children compared with a clinical reference standard (A. Expectorated sputum and induced sputum, and gastric lavage and aspirate, B. Summary of findings) 73

**Table 28.** GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting TB in children following negative smear microscopy (A. Evidence profile, B. Summary of findings, C. Additional yield of Xpert MTB/RIF over microscopy) 75

**Table 29.** Incremental yield of Xpert MTB/RIF compared with smear microscopy in children with culture-confirmed TB (A. Expectorated sputum and induced sputum, B. Gastric lavage or aspirate) 78

**Table 30.** GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting rifampicin resistance in respiratory specimens from children 79

**Table 31.** GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting peripheral lymph node TB in children 80

**Table 32.** GRADE evidence profile and summary of findings: accuracy of Xpert MTB/RIF in detecting TB meningitis in children 81

## Figures


**Figure 1.** Risks of bias (%) and judgements about applicability (%) for each domain of the QUADAS-2 tool in studies assessing the use of Xpert MTB/RIF to diagnose pulmonary TB for 36 study centres included in 27 studies 9

**Figure 2.** Forest plot of the sensitivity and specificity of Xpert MTB/RIF for detecting pulmonary TB in 27 studies (36 study centres) 11

**Figure 3.** Forest plot of the sensitivity and specificity of Xpert MTB/RIF for detecting pulmonary TB in smear-negative individuals suspected of having TB in 24 studies (33 study centres) 11

**Figure 4.** Forest plots of the sensitivity and specificity of Xpert MTB/RIF for detecting pulmonary TB in HIV-negative individuals suspected of having TB (9 studies, 18 study centres) and HIV-positive individuals suspected of having TB (10 studies, 16 centres) 14

**Figure 5.** Forest plots of the sensitivity and specificity of Xpert MTB/RIF for detecting rifampicin resistance when Xpert MTB/RIF was used as an initial test replacing phenotypic culture-based drug-susceptibility testing in 24 studies (33 study centres) (Studies are presented in order of decreasing sensitivity and decreasing number of true-positive results) 16

**Figure 6.** Risks of bias (%) and judgements about applicability (%) for each domain of the QUADAS-2 tool in studies using culture as the reference standard for TB detection and using Xpert MTB/RIF to diagnose extrapulmonary TB and rifampicin resistance in adults and children 36

**Figure 7.** Risks of bias (%) and judgements about applicability (%) for each domain of the QUADAS-2 tool in studies using an author-defined composite reference standard for TB detection and Xpert MTB/RIF to diagnose extrapulmonary TB and rifampicin resistance in adults and children 36

**Figure 8.** Forest plot of the sensitivity of Xpert MTB/RIF in detecting extrapulmonary TB in the smear-positive subgroup of participants in 22 studies 38

**Figure 9.** Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting extrapulmonary TB in lymph node samples (tissue or aspirate) compared with culture as the reference standard 39

**Figure 10.** Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting extrapulmonary TB in lymph node samples (tissue or aspirate) compared with a composite reference standard 40

**Figure 11.** Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting TB using pleural fluid compared with culture as a reference standard 41

**Figure 12.** Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting TB in cerebrospinal fluid compared with culture as a reference standard 42

**Figure 13.** Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting TB in cerebrospinal fluid compared with a composite reference standard 42

**Figure 14.** Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting TB in gastric fluid compared with culture as a reference standard 43

**Figure 15.** Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting TB in tissue samples (other than from a lymph node) compared with culture as a reference standard 44

**Figure 16.** Summary estimates for sensitivity (A) and specificity (B) of Xpert MTB/RIF in detecting extrapulmonary TB in adults and children, by sample type 45

**Figure 17.** Risks of bias (%) and judgements about applicability (%) for each domain of the QUADAS-2 tool in 16 studies assessing the use of Xpert MTB/RIF to diagnose pulmonary TB, peripheral lymph node TB, TB meningitis and rifampicin resistance in children 58

**Figure 18.** Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting pulmonary TB, peripheral lymph node TB and TB meningitis in children compared against culture as a reference standard, by study and specimen type 59

**Figure 19.** Forest plot of the sensitivity and specificity of smear microscopy in detecting pulmonary TB, peripheral lymph node TB and TB meningitis in children compared against culture as a reference standard , by study and specimen type 61

**Figure 20.** Forest plot of the sensitivity of Xpert MTB/RIF in diagnosing pulmonary TB, peripheral lymph node TB and TB meningitis in smear-positive and smear-negative children, by study and specimen type 62

**Figure 21.** Forest plot of the sensitivity of Xpert MTB/RIF in detecting pulmonary TB, peripheral lymph node TB and TB meningitis in children aged 0–4 and 5–15 years, by study and specimen type 64

**Figure 22.** Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting pulmonary TB, peripheral lymph node TB and TB meningitis in HIV-positive and HIV-negative children, by study and specimen type 65

**Figure 23.** Forest plot of the sensitivity and specificity of Xpert MTB/RIF for detecting resistance to rifampicin, peripheral lymph node TB and TB meningitis in children, by study and specimen type 67

## Abbreviations

| | |
|---|---|
| CI | confidence interval |
| CrI | credible interval |
| CRS | composite reference standard |
| CSF | cerebrospinal fluid |
| DOI | Declaration of Interests |
| DST | drug-susceptibility testing |
| FIND | Foundation for Innovative New Diagnostics |
| FNA | fine needle aspiration |
| GRADE | Grading of Recommendations Assessment, Development and Evaluation |
| HIV | human immunodeficiency virus |
| MDR-TB | multidrug-resistant tuberculosis |
| MGIT | mycobacterial growth indicator tube |
| NAAT | nucleic acid amplification test |
| NTM | nontuberculous mycobacteria |
| PCR | polymerase chain reaction |
| PEPFAR | United States President's Emergency Plan for AIDS Relief |
| PRISMA | Preferred Reporting Items for Systematic Reviews and Meta-analyses |
| QUADAS | Quality Assessment of Diagnostic Accuracy Studies |
| *rpo*B | gene encoding for the ß-subunit of the DNA-dependent RNA polymerase of *Mycobacterium tuberculosis* |
| STAG-TB | Strategic and Technical Advisory Group for TB |
| TB | tuberculosis |
| WHO | World Health Organization |

# 1. Executive summary

## Background

The global priorities for tuberculosis (TB) care and control are to improve case-detection and to detect cases earlier, including cases of smear-negative disease which are often associated with coinfection with the human immunodeficiency virus (HIV) and young age, and to enhance the capacity to diagnose multidrug-resistant tuberculosis (MDR-TB). In September 2010, the World Health Organization (WHO) convened an Expert Group to review the evidence on the Xpert MTB/RIF assay (Cepheid, Sunnyvale, CA, United States) for the purpose of formulating recommendations to guide the use of the test. Policy recommendations on using the Xpert MTB/RIF assay were issued by WHO early in 2011[1], supported by an operational how-to document[2] and a checklist for implementation at the country level.[3]

WHO's current policy recommends that Xpert MTB/RIF be used as an initial diagnostic test in individuals suspected of having MDR-TB or HIV-associated TB (strong recommendation, moderate quality of evidence)[1].The guidance also provides a conditional recommendation that Xpert MTB/RIF be used as a follow-on test to smear microscopy in settings where MDR-TB or HIV are of lesser concern, especially for further testing of smear-negative specimens. In acknowledgement of the difficulties of obtaining microbiological confirmation of the diagnosis in children, this recommendation generalizes from data on adults to include the use of Xpert MTB/RIF in children.

Since 2010, more than 85 peer-reviewed research papers have been published on using Xpert MTB/RIF to diagnose pulmonary, extrapulmonary and paediatric TB, and studies continue to be performed. Extrapulmonary TB accounts for about 25% of all cases of TB, and an even higher percentage of cases in children and immunocompromised patients. Diagnosing extrapulmonary TB is often challenging, requiring the clinician to obtain specimens for microscopy, culture and histopathological examination from the suspected sites of involvement. However, the availability of these tests is limited and the need for alternatives to use to diagnose TB in nonrespiratory samples is great. In 2011, the global burden of TB in children was estimated at 500 000 cases, representing approximately 6% of all cases. However, in all likelihood this burden is an underestimate due to the difficulties in obtaining microbiological confirmation of the diagnosis of TB in children.

The Xpert MTB/RIF assay remains the only fully automated cartridge-based real-time DNA-based test that can detect both TB and resistance to rifampicin in less than 2 hours, and it is the only mature technology representing a new generation of automated platforms for molecular diagnosis.

Given the amount of additional data on Xpert MTB/RIF that have emerged since 2010, an update of WHO's policy guidance was warranted. WHO's Global TB Programme therefore commissioned three systematic reviews to update and revise the guidance; these reviews assessed the utility of Xpert MTB/RIF for diagnosing TB and rifampicin resistance in pulmonary, extrapulmonary and paediatric TB. Published studies on the affordability and cost effectiveness of Xpert MTB/RIF were also reviewed.

1 *Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF system. Policy statement.* Geneva, World Health Organization, 2011 (http://whqlibdoc.who.int/publications/2011/9789241501545_eng.pdf).

2 *Rapid implementation of the Xpert MTB/RIF diagnostic test. Technical and operational 'how-to': practical considerations.* Geneva, World Health Organization, 2011 (http://whqlibdoc.who.int/publications/2011/9789241501569_eng.pdf).

3 *Prerequisites to country implementation of Xpert MTB/RIF and key action points at country level: checklist.* Geneva, World Health Organization, 2011 (http://whqlibdoc.who.int/hq/2011/WHO_HTM_TB_2011.12_eng.pdf).

WHO convened an Expert Group to review the evidence at Les Pensierès, Veyrier-du-Lac, France, during 20–21 May 2013. The major findings and recommendations of this Expert Group are summarized below, and this detailed meeting report is available at: http://www.who.int/tb/laboratory/policy_statements/en/.

## Summary of results

### *Using Xpert MTB/RIF to diagnose pulmonary TB in adults*

Twenty seven unique studies involving 9558 participants were included in the review. Two of the 27 studies were multicentre, international studies (one with five distinct study centres and the other with six). Two of the 27 studies evaluated Xpert MTB/RIF in primary care clinics where the results could be used to begin treatment on the same day. Sixteen studies (59%) were performed in low-income or middle-income countries. The reference standards for detecting pulmonary TB were solid culture or liquid culture. For rifampicin resistance, the reference standard was phenotypic culture-based drug-susceptibility testing (DST).

When used as an initial diagnostic test to replace smear microscopy, Xpert MTB/RIF achieved an overall pooled sensitivity of 88% (95% credible interval [CrI], 84–92%)[4] and pooled specificity of 99% (95% CrI, 98–99%) (22 studies, 9008 participants). When used as an add-on test following a negative smear-microscopy result, Xpert MTB/RIF yielded a pooled sensitivity of 68% (95% CrI, 61–74%) and a pooled specificity of 99% (95% CrI, 98–99%) (23 studies, 7151 participants). For smear-positive culture-positive TB, the pooled sensitivity of Xpert MTB/RIF was 98% (95% CrI, 97–99%) (23 studies, 1952 participants); for smear-negative culture-positive TB it was 68% (95% CrI, 61–74%) (23 studies, 7151 participants).

For people living with HIV, the pooled sensitivity of Xpert MTB/RIF was 79% (95% CrI, 70–86%) (7 studies, 1789 participants); for people without HIV infection, the pooled sensitivity was 86% (95% CrI, 76–92%) (7 studies, 1470 participants).

When used to detect rifampicin resistance, Xpert MTB/RIF achieved a pooled sensitivity of 95% (95% CrI, 90–97%) (17 studies, 555/2624 total specimens) and a pooled specificity of 98% (95% CrI, 97–99%) (24 studies, 2414 specimens, true negatives and false positives).

### *Expert Group consensus*

The process of evidence synthesis confirmed there was a high-quality evidence base to support the widespread use of Xpert MTB/RIF to detect adult pulmonary TB and rifampicin resistance. Therefore:

- Xpert MTB/RIF should be used rather than conventional microscopy, culture and DST as the initial diagnostic test in adults suspected of having MDR-TB or HIV-associated TB (strong recommendation, high-quality evidence);
- Xpert MTB/RIF may be used as a follow-on test to microscopy in adults where MDR-TB and HIV are of lesser concern, especially for further testing of smear-negative specimens (conditional recommendation acknowledging resource implications, high-quality evidence);
- Xpert MTB/RIF may be used rather than conventional microscopy and culture as the initial diagnostic test in all adults suspected of having TB (conditional recommendation acknowledging resource implications, high-quality evidence).

4 The credible interval (CrI) is the Bayesian equivalent of the confidence interval (CI).

### *Using Xpert MTB/RIF to diagnose extrapulmonary TB in adults and children*

Fifteen published studies and 7 unpublished studies, involving 5922 samples, were included in the review. The majority of studies (59%) were performed in settings with a high burden of TB. Due to the heterogeneity of sample types included in the studies, prespecified subgroups of samples (pleural fluid, lymph node samples [biopsy and aspirate combined], other tissues and cerebrospinal fluid [CSF]) with a comparison against culture and against a composite reference standard (CRS) were included in the meta-analysis. In some cases the CRS included a nucleic acid amplification test other than Xpert MTB/RIF, histology, smear, culture, results of biochemical testing, presenting signs, or a response to treatment with anti-TB therapy, or a combination of these.

Using culture as the reference standard, the pooled sensitivity of Xpert MTB/RIF in lymph node tissues or aspirates was 84.9% (95% confidence interval [CI], 72.1–92.4%) and the pooled specificity was 92.5% (95% CI, 80.3–97.4%) (14 studies, 849 samples). Five studies (one unpublished) assessed Xpert MTB/RIF on lymph node samples compared against an author-defined CRS. The pooled sensitivity was estimated to be 83.7% (95% CI, 73.8–90.3%) and the pooled specificity to be 99.2% (95% CI, 88.4–100%).

In **CSF**, the pooled sensitivity of Xpert MTB/RIF compared against culture as a reference standard was 79.5% (95% CI, 62.0–90.2%) and the pooled specificity was 98.6% (95% CI, 95.8–99.6%) (16 studies, 709 samples). Comparing Xpert MTB/RIF on CSF against a CRS yielded a pooled sensitivity of 55.5% (95% CI, 44.2–66.3%) and a pooled specificity of 98.8% (95% CI, 94.5–99.8%) (6 studies, 512 samples).

Using culture as the reference standard, the pooled sensitivity of Xpert MTB/RIF in **gastric fluid** was 83.8% (95% CI, 65.9–93.2%) and the pooled specificity was 98.1% (95% CI, 92.3–99.5%) (12 studies, 1258 samples); and in other tissue samples (12 studies, 699 samples) the pooled sensitivity of Xpert MTB/RIF was 81.2% (95% CI, 67.7–89.9%) and the pooled specificity was 98.1% (95% CI, 87.0–99.8%).

In **pleural fluid**, the pooled sensitivity of Xpert MTB/RIF compared against culture was 43.7% (95% CI, 24.8– 64.7%) (17 studies, 1385 samples); compared against a CRS it was 17.0% (95% CI, 7.5–34.2%) (7 studies, 698 samples). Pooled specificity was high when compared against both culture as a reference standard and the CRS. The data for additional sample types (such as, ascitic fluid, pericardial fluid, urine, blood and stool) were limited and therefore not considered for analysis.

### *Expert Group consensus*

- **CSF:** The Expert Group recommends that Xpert MTB/RIF should be used in preference to conventional microscopy and culture as the initial diagnostic test when testing CSF from patients suspected of having TB meningitis (strong recommendation given the urgency of rapid diagnosis, very low-quality evidence). The Expert Group noted that a negative Xpert MTB/RIF result on CSF should be followed by other tests. The Expert Group also noted that concentration methods should be used to enhance yield when sufficient volumes of CSF are available. These recommendations apply to both children and adults.
- **Lymph node and tissue:** The Expert Group recommends that Xpert MTB/RIF may be used as a replacement test for usual practice (including conventional microscopy, culture and histology) in testing lymph nodes and tissues from patients suspected of having extrapulmonary TB (conditional recommendation, very low-quality evidence). The Expert Group noted that a negative result from

Xpert MTB/RIF should be followed by other tests. The Expert Group also noted that sample processing methods for lymph nodes and tissues need to be standardized to optimize yields. These recommendations apply to both adults and children.

- **Pleural fluid:** The Expert Group noted that pleural fluid is a suboptimal sample for diagnosing pleural TB, and that a pleural biopsy is the preferred sample for bacteriological confirmation (including by Xpert MTB/RIF). The Expert Group noted that if Xpert MTB/RIF is used in the diagnostic work-up of patients suspected of having pleural TB then a positive result from Xpert MTB/RIF is considered confirmatory (conditional recommendation, very low-quality evidence), and that a negative result from Xpert MTB/RIF should be followed by other tests. These recommendations apply to both adults and children.

### *Using Xpert MTB/RIF to diagnose pulmonary TB and rifampicin resistance in children*

Sixteen studies (12 published and 4 unpublished) were included in the review. All studies were performed at higher levels of care, and the children included in the studies were mainly inpatients. Thirteen studies were performed in low-income or middle-income countries.

Pulmonary TB was evaluated in 13 studies that included 2603 participants. The overall pooled sensitivity of Xpert MTB/RIF compared against culture as a reference standard in children suspected of having TB was 66% in 10 studies where expectorated sputum or induced sputum was used (95% CrI, 52–77%), and 66% in 7 studies where gastric lavage aspirates were used (95% CrI, 51–81%). The pooled specificity of Xpert MTB/RIF compared against culture as the reference standard was at least 98%, with narrow confidence intervals.

The pooled sensitivity of Xpert MTB/RIF in culture-negative specimens from children compared with clinical TB used as the reference standard was very low at 4% for expectorated sputum or induced sputum (8 studies), and 15% for gastric lavage aspirates (3 studies), both sensitivities had wide confidence intervals. It is likely that the apparently poor performance of Xpert MTB/RIF was the result of a reference standard for clinical TB that lacked specificity. The sensitivity of Xpert MTB/RIF to detect rifampicin resistance in specimens from children was 86% (95% CrI, 53–98%).

### *Expert Group consensus*

- The Expert Group recommends that Xpert MTB/RIF should be used rather than conventional microscopy, culture and DST as the initial diagnostic test in children suspected of having MDR-TB or HIV-associated TB (strong recommendation given the difficulties in diagnosing paediatric TB, very low-quality evidence).
- The Expert Group also recommends that Xpert MTB/RIF may be used rather than conventional microscopy and culture in all other children suspected of having pulmonary TB (conditional recommendation acknowledging resource implications, very low-quality evidence).
- The Expert Group noted that Xpert MTB/RIF should not be used as the only test in the diagnostic pathway of children suspected of having TB, and that a child in whom there is a high clinical suspicion of TB should be treated even if the result from Xpert MTB/RIF is negative or if the test is not available.

### *Affordability and cost effectiveness of using Xpert MTB/RIF to diagnose TB*

Twelve published papers were identified that compared the costs of current diagnostic algorithms for diagnosing TB and MDR-TB with the costs of using Xpert MTB/RIF as the initial diagnostic test or as a follow-on test to microscopy. The setting for the majority of analyses was South Africa; two

studies included other countries in sub-Saharan Africa (Botswana, Lesotho, Namibia, Swaziland and Uganda); one study included countries in the former Soviet Union; and one global analysis included all countries. Seven of the 12 studies analysed costs, and 5 were cost–effectiveness analyses. Wide variations in the methods used, the underlying assumptions, and the intended use of Xpert MTB/RIF made a systematic review impossible.

### *Expert Group consensus*

Although the use of Xpert MTB/RIF was found to be cost effective overall, more directly measured costing evidence is needed from more countries to improve analyses of cost effectiveness.

## 2. BACKGROUND

The global priorities for tuberculosis (TB) care and control are to improve case-detection and to detect cases earlier, including cases of smear-negative disease which are often associated with coinfection with the human immunodeficiency virus (HIV) and young age, and to enhance capacity to diagnose multidrug-resistant tuberculosis (MDR-TB). In September 2010, the World Health Organization (WHO) convened an Expert Group to review the evidence on the Xpert MTB/RIF assay (Cepheid, Sunnyvale, CA, United States) in order to formulate recommendations to guide the use of the test. Policy recommendations on using the Xpert MTB/RIF assay were issued by WHO early in 2011[5], supported by an operational how-to document[6] and a checklist for implementation at the country level [7].

In accordance with WHO's standards for assessing evidence when formulating policy recommendations, WHO engages in a systematic, transparent process using the GRADE approach (the Grading of Recommendations Assessment, Development and Evaluation, see http://www.gradeworkinggroup.org/). GRADE provides a structured framework for evaluating the accuracy of diagnostic tests, and the impact on patients and public health of new diagnostic tests. The Expert Group meeting assessed an updated systematic review on the accuracy of Xpert MTB/RIF for diagnosing pulmonary TB and rifampicin resistance in adults. Two additional systematic reviews had been performed, and were presented as separate reports: a review assessing the use of the Xpert MTB/RIF assay for diagnosing TB in nonrespiratory specimens (extrapulmonary TB) and the use of the Xpert MTB/RIF assay to diagnose TB and rifampicin resistance in children.

### Xpert MTB/RIF

Xpert MTB/RIF is an automated polymerase chain reaction (PCR) test (that is, a molecular test) utilizing the GeneXpert platform. Xpert MTB/RIF is a single test that can detect both *Mycobacterium tuberculosis* complex and rifampicin resistance within 2 hours of starting the test, with minimal hands-on technical time. Unlike conventional nucleic acid amplification tests (NAATs), Xpert MTB/RIF is unique because sample processing, and PCR amplification and detection, are integrated into a single self-enclosed test unit, the Xpert MTB/RIF cartridge. 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 properties (that is, it has the ability to kill TB bacteria), and this largely eliminates biosafety concerns during the test procedure. These features allow the technology to be taken out of a reference laboratory and used nearer to patients. However, Xpert MTB/RIF requires an uninterrupted and stable electrical power supply, temperature control and yearly calibration of the instrument's modules.

The test procedure may be used directly on clinical specimens, either raw sputum samples or sputum pellets (also called sputum sediment), which are created after decontaminating and concentrating the sputum. In both cases, the test material is combined with the reagent, mixed by hand or vortex, and incubated at room temperature for

5 *Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF system. Policy statement.* Geneva, World Health Organization, 2011 (http://whqlibdoc.who.int/publications/2011/9789241501545_eng.pdf).

6 *Rapid implementation of the Xpert MTB/RIF diagnostic test. Technical and operational 'how-to': practical considerations.* Geneva, World Health Organization, 2011 (http://whqlibdoc.who.int/publications/2011/9789241501569_eng.pdf).

7 *Prerequisites to country implementation of Xpert MTB/RIF and key action points at country level: checklist.* Geneva, World Health Organization, 2011 (http://whqlibdoc.who.int/hq/2011/WHO_HTM_TB_2011.12_eng.pdf).

15 minutes. After incubation, 2 ml of the treated sample are transferred to the cartridge, and the run is initiated. According to the manufacturer, Xpert MTB/RIF may be used with fresh sputum samples and freshly prepared sediments.

Xpert MTB/RIF 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 *rpo*B gene of TB. Five different coloured beacons are used, each covering a separate nucleic acid sequence within the amplified *rpo*B gene. When a beacon binds to the matching sequence, it fluoresces (or lights up), which indicates the presence of one of the gene sequences that is characteristic of rifampicin-susceptible TB. If a beacon fails to bind to the matching sequence or if binding is delayed, the sample is potentially resistant to rifampicin. The number of positive beacons and the timing of their detection (when the fluorescent signal rises above a predetermined baseline cycle threshold), as well as the results of sample processing controls, allow the test to distinguish among the following results: no TB; TB detected, rifampicin resistance detected; TB detected, no rifampicin resistance detected; and an invalid result.

## 3. EVIDENCE BASE

### 3.1. Evidence synthesis

In order to facilitate the development of policy guidance on the use of new diagnostic tools, new diagnostic methods, and novel approaches to diagnosis using existing tools, WHO has developed a structured evidence-based process. The first step involves systematically reviewing the data, using standard methods appropriate for studies assessing diagnostic accuracy. The second step involves convening an Expert Group to (1) evaluate the strength of the evidence base, (2) identify operational and logistical considerations relevant to mainstreaming the tools or approaches into national TB-control programmes, and (3) identify any gaps that need to be addressed by research. The third step involves presenting WHO's guidance on the use of these tools or approaches to WHO's Strategic and Technical Advisory Group for TB (STAG-TB) for endorsement; after endorsement, the guidance is disseminated to Member States for implementation.

This document presents the findings and recommendations from the meeting of the Expert Group on Xpert MTB/RIF convened by WHO at Les Pensierès, Veyrier-du-Lac, France in May 2013. The Expert Group (*Annex 1*) consisted of researchers, clinicians, epidemiologists, end-users (representatives of programmes and laboratories), a participant who has had TB, and an expert on evidence synthesis. The meeting followed a structured agenda (*Annex 2*) and was chaired by a clinical epidemiologist with expertise and extensive experience in evidence synthesis and guideline development.

### 3.2. Meeting objectives

The objectives of the meeting were:

- to review the evidence base and evaluate data from an updated systematic review on the accuracy of the Xpert MTB/RIF assay in diagnosing pulmonary TB and rifampicin resistance in adults;
- to review the evidence base and evaluate data from a systematic review on the accuracy of the Xpert MTB/RIF assay in diagnosing TB in nonrespiratory samples;

- to review the evidence base and evaluate data from a systematic review on the accuracy of the Xpert MTB/RIF assay in diagnosing TB and rifampicin resistance in children;
- to review the evidence on the cost effectiveness and affordability of the Xpert MTB/RIF assay in different epidemiological settings and differently resourced settings;
- to outline issues to be addressed by WHO in subsequent policy recommendations.

## 3.3. GRADE evaluation

To comply with current standards for assessing evidence when formulating policy recommendations, the GRADE system was used; this system has been adopted by WHO and is used for developing all policies and guidelines[8]. The GRADE approach assesses both the *quality of evidence and the strength of recommendations*, and aims to provide a comprehensive and transparent approach for developing policies and guidance.

Evaluations of the evidence used the GRADE system for grading the quality of the evidence and the strength of recommendations for diagnostic tests. The *quality of evidence* was evaluated according to the following six criteria.

- *Overall study design*
  Cross-sectional studies: select patients at risk or specimens randomly or consecutively (preferred);
  Case–control studies: select patients or specimens according to a reference standard.
- *Risk of bias or limitations in study design and execution* (as reflected by the Quality Assessment of Diagnostic Accuracy Studies [QUADAS-2] tool).[9]
- *Directness*
  Assessment of the presence of direct evidence of the impact on patient-important outcomes, and assessment of generalizability in relation to the population, the diagnostic test used, the comparator of the test, and whether tests were directly or indirectly compared.
- *Inconsistency*
  Assessment of unexplained inconsistency in estimates of sensitivity or specificity.
- *Imprecision*
  Assessment to determine to whether there were wide confidence intervals for pooled estimates of sensitivity or specificity.
- *Publication bias*
  Assessment of whether research was published based on its nature and outcome – for example, was there language bias in the studies published? Were studies showing poor performance not published?

As called for by GRADE, the Expert Group also considered the strength of the recommendations being made (which were classified as strong or conditional); this assessment was based on the balance of effects (that is, the advantages weighed against the disadvantages), patients' values and preferences, and costs.

Using the GRADE framework, sensitivity and specificity results were interpreted as proxy measures for patient-important outcomes based on the relative importance or impact of false-positive results and false-negative results. Poor sensitivity would result in false-negative results where patients with TB or MDR-TB would be missed, with negative consequences for morbidity, mortality and transmission of disease. Poor specificity would result in false-positive results where patients without TB or MDR-TB would be

8 *Handbook for guideline development*. Geneva, World Health Organization, 2012 (http://apps.who.int/iris/bitstream/10665/75146/1/9789241548441_eng.pdf).

9 Whiting PF et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 2011, 155:529–536.

prescribed unnecessary treatment, with negative consequences, such as serious adverse events related to the use of second-line anti-TB agents.

Rates for true positives, true negatives, false positives and false negatives were calculated based on pretest probabilities – that is, an assumed prevalence of TB of 2.5%, 5% and 10% among patients suspected of having TB who were being screened, and an assumed prevalence of rifampicin resistance of 5% and 15% (as a proxy for MDR-TB) among patients with confirmed TB.

The evaluation of the impact on patients was based on a balance among the following values:

- *true positives* – the benefit to patients from rapid diagnosis and treatment;
- *true negatives* – the benefit to patients who would be spared unnecessary treatment; the benefit of reassurance and an alternative diagnosis;
- *false positives* – the likelihood of anxiety and morbidity caused by additional testing or unnecessary treatment, or both; the chance that a false positive may halt further diagnostic evaluation;

false negatives – the increased risk of morbidity and mortality, and the continued risk of community transmission of TB.

### *3.3.1. PICO questions for each review*

The evaluation of the evidence used the GRADE system to assess the quality of evidence and provide information about the strength of the recommendations; these evaluations were based on a priori questions (the PICO questions) agreed by the Expert Group. PICO refers to the following four elements that should be included in questions that govern a systematic search of the evidence: the Population targeted by the action or intervention, the Intervention, the Comparator, and the Outcome. The PICO questions for each review are given in *Box 1*.

**Box 1.** PICO questions for the four systematic reviews evaluating the accuracy of the Xpert MTB/RIF assay in diagnosing TB

**Review 1. Updated systematic review: Xpert MTB/RIF for diagnosis of pulmonary tuberculosis and rifampicin resistance in adults**

1. What is the diagnostic accuracy of Xpert MTB/RIF for detection of pulmonary TB in adults, where Xpert MTB/RIF is used as a replacement test for smear microscopy?

2. What is the diagnostic accuracy of Xpert MTB/RIF for detection of pulmonary TB in adults, where Xpert MTB/RIF is used as an add-on test following a negative smear-microscopy result?

3. What is the diagnostic accuracy of Xpert MTB/RIF for detection of smear-positive pulmonary TB in adults?

4. What is the diagnostic accuracy of Xpert MTB/RIF for detection of smear-negative (culture-positive) pulmonary TB in adults?

5. What is the diagnostic accuracy of Xpert MTB/RIF for detection of pulmonary TB in people living with HIV (adults)?

6. What is the diagnostic accuracy of Xpert MTB/RIF for detection of pulmonary TB in adults without HIV infection?

7. What is the diagnostic accuracy of Xpert MTB/RIF for detection of rifampicin resistance, where Xpert MTB/RIF is used as an initial test replacing phenotypic culture-based drug-susceptibility testing?

**Review 2. Systematic review: Xpert MTB/RIF for diagnosis of tuberculosis and detection of rifampicin resistance on nonrespiratory samples (extrapulmonary TB)**

1. What is the diagnostic accuracy of Xpert MTB/RIF overall compared with culture for nonrespiratory specimens, where Xpert MTB/RIF is used as a replacement test for usual practice?[a]

2. What is the diagnostic accuracy of Xpert MTB/RIF overall compared with a combined clinical and laboratory reference standard for nonrespiratory specimens, where Xpert MTB/RIF is used as a replacement test for usual practice?

2a. What is the diagnostic accuracy of Xpert MTB/RIF for lymph node fluid and tissue, where Xpert MTB/RIF is used as a replacement test for usual practice?

2b. What is the diagnostic accuracy of Xpert MTB/RIF for pleural fluid, where Xpert MTB/RIF is used as a replacement test for usual practice?

2c. What is the diagnostic accuracy of Xpert MTB/RIF for cerebrospinal fluid, where Xpert MTB/RIF is used as a replacement test for usual practice?

2d. What is the diagnostic accuracy of Xpert MTB/RIF for gastric fluid, where Xpert MTB/RIF is used as a replacement test for usual practice?

2e. What is the diagnostic accuracy of Xpert MTB/RIF for tissue samples, where Xpert MTB/RIF is used as a replacement test for usual practice?

3.What is the diagnostic accuracy of Xpert MTB/RIF for detection of rifampicin resistance in nonrespiratory specimens, where Xpert MTB/RIF is used as an initial test replacing phenotypic culture-based drug-susceptibility testing?

**Review 3. Systematic review: Xpert MTB/RIF for diagnosis of tuberculosis and rifampicin resistance in children**

1. What is the diagnostic accuracy of Xpert MTB/RIF for detection of TB in children compared with culture, where Xpert MTB/RIF is used as a replacement test for usual practice?[b]

2. What is the diagnostic accuracy of Xpert MTB/RIF for detection of TB in children compared with a combined clinical and laboratory reference standard, where Xpert MTB/RIF is used as a replacement test for usual practice?

3. What is the diagnostic accuracy of Xpert MTB/RIF for detection of TB in children, where Xpert MTB/RIF is used as an add-on test following a negative smear-microscopy result?

4. What is the diagnostic accuracy of Xpert MTB/RIF compared with smear microscopy for detection of TB in children?

5. What is the diagnostic accuracy of Xpert MTB/RIF for detection of rifampicin resistance in children, where Xpert MTB/RIF is used as an initial test replacing phenotypic culture-based drug-susceptibility testing?

6. What is the diagnostic accuracy of Xpert MTB/RIF for detection of peripheral lymph node TB in children, where Xpert MTB/RIF is used as a replacement test for usual practice?

7. What is the diagnostic accuracy of Xpert MTB/RIF for detection of TB meningitis in children, where Xpert MTB/RIF is used as a replacement test for usual practice?

**Review 4. Systematic review: Affordability, cost effectiveness and resource implications for Xpert MTB/RIF scale up**

1. For which diagnostic and screening algorithms is the Xpert MTB/RIF assay an affordable and a cost-effective intervention?

a Most analyses were performed using two reference standards: culture (the current reference standard) and a combined clinical and laboratory reference standard chosen by the study's authors (given the technical limitations of using culture for diagnosis).
b Given the difficulties of diagnosing TB in children, usual practice refers to customary practice in the field, which may vary from setting to setting. The usual practice for children (aged 0–15 years) suspected of having intrathoracic TB (that is, pulmonary, pleural, and mediastinal or hilar lymph node TB) normally requires bacteriological confirmation through examination of sputum (obtained by expectoration, gastric washings, or induction) for smear microscopy and culture. In the event of negative bacteriological results, a diagnosis of TB may be based on the presence of abnormalities consistent with TB on chest radiography, a history of exposure to an infectious case, evidence of TB infection (that is, a positive tuberculin skin test or interferon-γ release assay) and clinical findings suggestive of TB. For children suspected of having extrapulmonary TB, appropriate specimens from the suspected sites of involvement may be obtained for microscopy, and for culture and histopathological examination.

### *3.3.2. Determining the relative importance of patients' outcomes*

PICO questions were drafted by the WHO Steering Group and were presented to the Expert Group for discussion and modification. The Steering Group also prepared an initial list of relevant outcomes, including desirable effects and undesirable effects, and requested that the Expert Group identify any other relevant outcomes.

A webinar was conducted with members of the Expert Group prior to the meeting to refine and finalize the proposed outcomes seen as important to patients, and to rate their relative importance. The following outcomes for each PICO question were determined, and the ratings of their importance were unanimously agreed by the Expert Group:

- critical outcomes – diagnostic accuracy as reflected by true-positive, true-negative, false-positive and false-negative results; time to diagnosis;
- important outcome – cost.

### *3.3.3. Assessment of study quality*

The appraisal of the studies included in the reviews used the QUADAS -2 tool[10], which consists of four domains: patient selection, index test, reference standard, and flow and timing.

## 3.4. Procedural issues

The meeting was chaired by an expert in evidence synthesis. Decisions were based on consensus. Concerns raised by members were noted and included in the final report of the meeting. A detailed report of the meeting was prepared by the WHO Steering Group; the report went through several iterations before being finally signed off by members of the Expert Group.

The systematic reviews and reports were made available to members of the Expert Group for scrutiny before the meeting, and full copies of the reviews were also available during the meeting.

As agreed, participation in the meeting was restricted to members who could attend the meeting in person, both for the discussion and for follow-up dialogue. Individuals were carefully selected to become members of the group to represent and balance perspectives deemed to be important for the process of formulating the recommendations. Therefore the Expert Group included technical experts, end-users, patients' representatives and specialists in evidence synthesis.

Members were asked to submit completed Declarations of Interests (DOI) prior to the meeting. The completed forms were evaluated by the WHO Steering Group before any formal invitations to the meetings were issued. Each

10 Whiting PF et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 2011, 155:529–536.

DOI was reviewed by the Steering Group to determine if an interest had been declared and, if so, whether it was insignificant or potentially significant. If the Steering Group determined that no relevant interest had been declared or such interest was insignificant or minimal, a letter was sent inviting participation.

When the Steering Group's review indicated that a declared interest was significant or potentially significant, WHO's Legal Department was consulted, and their advice on the meeting's procedures was followed. DOI statements were summarized by the Chairperson at the start of the meeting. Individuals who had been determined to have significant relevant interests were invited to attend as observers to provide technical input and answer technical questions. However, these individuals were not permitted to participate in formulating the recommendations. In addition, they were excluded from the process of developing the final report of the Guideline Development Group and from preparing WHO's final policy update. A summary of DOI statements appears in *Annex 3*.

Selected individuals with intellectual or research involvement in Xpert MTB/RIF, or both, were invited as observers to provide technical input and answer technical questions. These individuals did not participate in the GRADE evaluation process at the meeting nor during the final discussions when recommendations were developed. Also, they were not involved in developing the report of the Expert Group's meeting, nor in preparing documentation for STAG-TB or WHO's final policy update

# 4. Results

## 4.1. Using Xpert MTB/RIF to diagnose pulmonary TB

### *4.1.1. Characteristics of the studies*

Studies that assessed the accuracy of Xpert MTB/RIF in diagnosing pulmonary TB or rifampicin resistance, or both, typically were cross-sectional in design, comparing Xpert MTB/RIF to a reference standard (defined below). Studies were included in the review when true-positive, true-negative, false-positive and false-negative values could be determined. Participants in the studies were predominantly adult patients aged 15 years or older, who were suspected of having pulmonary TB or MDR-TB with or without HIV infection. Studies that assessed respiratory samples other than sputum, such as samples obtained by bronchoalveolar lavage, were included. Data on specimens obtained by gastric aspiration were excluded.

The reference standard for pulmonary TB was culture on solid media, or the use of a commercial liquid culture system, such as the BACTEC MGIT (mycobacterial growth indicator tube) 960 Mycobacterial Detection System, (Becton Dickinson, Franklin Lakes, NJ, United States). The reference for rifampicin resistance was WHO's recommended conventional phenotypic drug-susceptibility testing (DST) on solid media or liquid media.[11]

Two literature searches were performed, on 25 September 2011 and 15 December 2011, from which 18 studies were identified. These

11 *Policy guidance on drug-susceptibility testing (DST) of second-line antituberculosis drugs*. Geneva, World Health Organization, 2008 (WHO/HTM/TB/2008.392).

18 studies were included in the Cochrane review published 31 January 2013.[12] A third literature search was performed on 7 February 2013; this identified nine additional studies. *Annex 4* shows the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) diagram of the studies. From the 3 literature searches, 27 relevant studies (26 published studies and 1 unpublished study) involving 36 study centres were identified and evaluated. *Annex 4* lists the included and excluded studies, along with the reasons for exclusion.

A total of 27 studies of TB detection included 9558 participants. Of the 27 studies, 24 (33 study centres, 2969 participants) also provided data on the detection of rifampicin resistance. Three studies were not included: one presented combined results for pulmonary specimens and extrapulmonary specimens; one did not report information on rifampicin resistance; and one study did not use the defined reference standard. Seven of the 27 studies detected no rifampicin resistance with the reference standard.

### *4.1.2. Quality of the studies*

The quality of the studies included in the review was assessed with the QUADAS-2 tool. As recommended, all domains (patient selection, index test, reference standard, and flow and timing) were assessed in terms of their risk of bias, and the first domains were also assessed in terms of concerns about applicability.

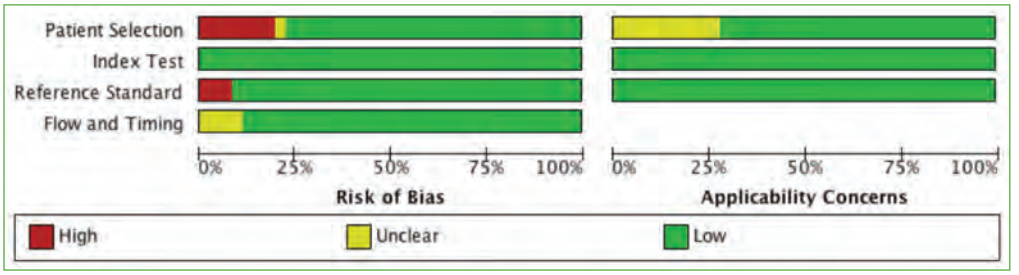
In the patient-selection domain, 28 of 36 study centres (78%) were considered to be at low risk of bias because participants were enrolled consecutively, and these centres avoided inappropriate exclusions. The remaining study centres were considered to be at high risk of bias because either (1) convenience sampling was used (five study centres) or the manner in which patients were selected was not stated (one study), or (2) the study preselected smear-positive patients (two study centres). With regard to the applicability of the study's findings in terms of patients' characteristics and settings, for 26 of the 36 study centres (72%) (corresponding to 18 of the 27 studies, 67%) applicability was judged to be of low concern because these centres performed Xpert MTB/RIF in intermediate-level or peripheral-level laboratories associated with primary care clinics. For the remaining centres, applicability was judged to be of unclear concern. These studies either did not provide any clinical information (one study) or ran Xpert MTB/RIF in central-level laboratories where culture (the reference standard) could also be performed (nine studies).

In the index test domain (that is, Xpert MTB/RIF), concerns about the risk of bias and applicability were considered to be low for all study centres. In the reference-standard domain, 33 study centres (92%) were deemed to be at low risk of bias for TB, and 34 study centres (94%) to be at low risk of bias for rifampicin resistance because the results of the reference standard were interpreted without knowledge of the results of the Xpert MTB/RIF assay.

In the reference-standard domain, applicability was considered to be of low concern for all studies. In the flow and timing domain, the risk of bias was considered to be of low concern for 32 centres (89%) because all patients were accounted for in the analysis, and information about uninterpretable results was provided. Inconclusive Xpert MTB/RIF results (that is, errors, invalid tests, or no results) were excluded from the analyses used to determine the sensitivity and specificity for both TB detection and rifampicin resistance. Although the pooled rate of inconclusive results reported in the studies in the meta-analysis was considered to be low, there were some concerns regarding potential bias, given observations from the field that rates might be higher (up to 8%). *Figure 1* shows the overall rating of the quality of the 36 centres where the studies were conducted.

12 Steingart KR et al. Xpert MTB/RIF® assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database of Systematic Reviews*, 2013, (1):CD009593.

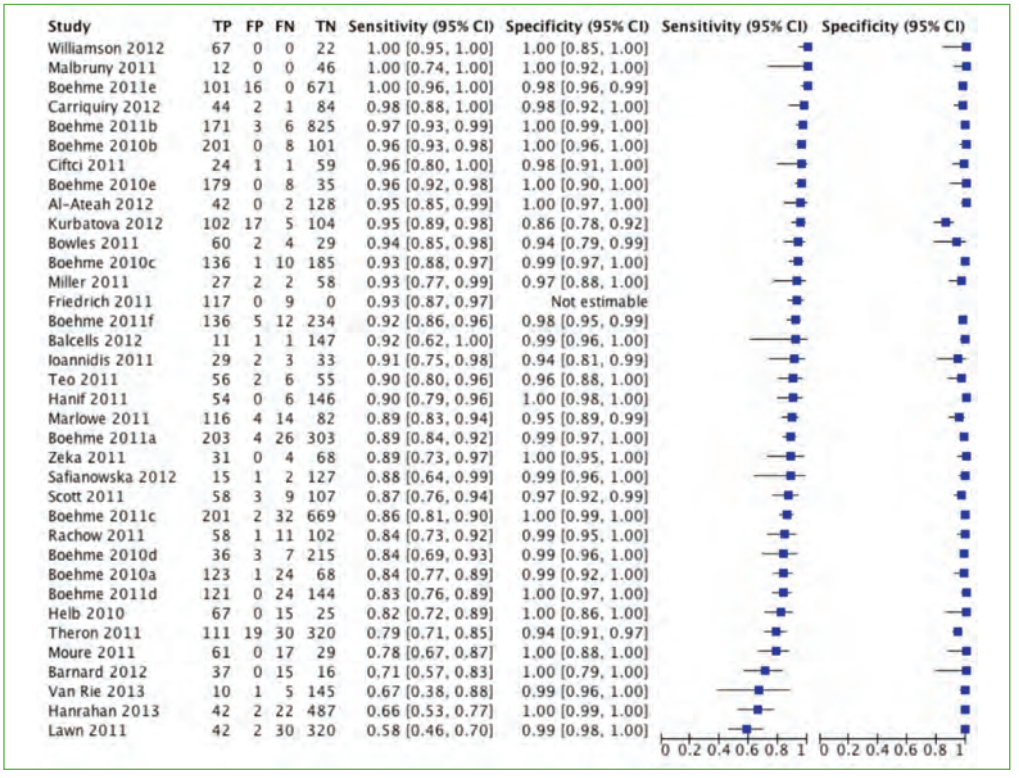
Figure 1. Risks of bias (%) and judgements about applicability (%) for each domain of the QUADAS-2 tool in studies assessing the use of Xpert MTB/RIF to diagnose pulmonary TB for 36 study centres included in 27 studies




### *4.1.3. Using Xpert MTB/RIF as an initial test to replace smear microscopy*

Forest plots of the sensitivity and specificity of Xpert MTB/RIF for detecting TB in the 27 studies (36 study centres) are presented in *Figure 2*. Sensitivity estimates varied from 58% to 100%; specificity estimates varied from 86% to 100%.

Figure 2. Forest plot of the sensitivity and specificity of Xpert MTB/RIF for detecting pulmonary TB in 27 studies (36 study centres)[a]




TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

Twenty two of the 27 studies (9008 participants) were included in this meta-analysis. Five studies that enrolled primarily smear-positive or smear-negative patients were excluded. The pooled sensitivity of Xpert MTB/RIF for detecting pulmonary TB was 88% (95% credible interval [CrI], 84–92%); the pooled specificity was 99% (95% CrI, 98–99%) (*Table 1*).

Twenty-one studies (8880 participants) provided data from which to compare the sensitivity of Xpert MTB/RIF and smear microscopy. For smear microscopy, the pooled sensitivity was 65% (95% CrI, 57–72%); for Xpert MTB/RIF, the pooled sensitivity was 88% (95% CrI, 84–92%). Therefore, in comparison with smear microscopy, Xpert MTB/RIF increased TB detection among culture-confirmed cases by 23% (95% CrI, 15–32%). When Xpert MTB/RIF was used as an add-on test following a negative smear-microscopy result (23 studies, 7151 participants), the pooled sensitivity was 68% (95% CrI, 61–74%); the pooled specificity was 99% (95% CrI, 98–99%) (*Table 1*). In other words, 68% of smear-negative culture-confirmed TB cases were detected using Xpert MTB/RIF following smear microscopy, increasing case detection by 68% (95% CrI, 61–74%) in this group.

Table 1. **Pooled sensitivity and specificity of the Xpert MTB/RIF assay for detecting pulmonary TB and rifampicin resistance**

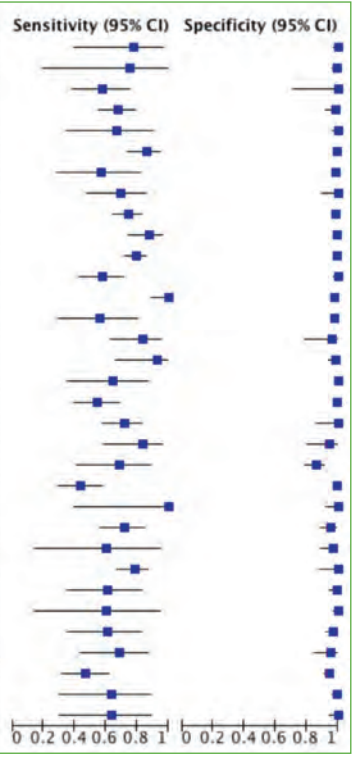
| Type of analysis (No. of studies, No. of participants) | Median (%) pooled sensitivity (95% CrI) | Median (%) pooled specificity (95% CrI) |
|---|---|---|
| Xpert MTB/RIF used as an initial test for TB detection replacing microscopy (22, 9008) | 88 (84–92) | 99 (98–99) |
| Xpert MTB/RIF used as an add-on test for TB detection following a negative smear-microscopy result (23, 7151) | 68 (61–74) | 99 (98–99) |
| Xpert MTB/RIF used as an initial test for detecting rifampicin resistance replacing conventional drug-susceptibility testing as the initial test[a] | 95 (90–97) | 98 (97–99) |

CrI, credible interval; the CrI is the Bayesian equivalent of the confidence interval.

a The pooled sensitivity estimates and specificity estimates for detecting rifampicin resistance were determined separately by univariate analyses. The pooled sensitivity analysis included 17 studies (555 participants); the pooled specificity analysis included 24 studies (2414 participants).

Figure 3. Forest plot of the sensitivity and specificity of Xpert MTB/RIF for detecting pulmonary TB in smear-negative individuals suspected of having TB in 24 studies (33 study centres)[a]

| Study | TP | FP | FN | TN | Sensitivity (95% CI) | Specificity (95% CI) |
|---|---|---|---|---|---|---|
| Al-Ateah 2012 | 7 | 0 | 2 | 127 | 0.78 [0.40, 0.97] | 1.00 [0.97, 1.00] |
| Balcells 2012 | 3 | 1 | 1 | 145 | 0.75 [0.19, 0.99] | 0.99 [0.96, 1.00] |
| Barnard 2012 | 18 | 0 | 13 | 11 | 0.58 [0.39, 0.75] | 1.00 [0.72, 1.00] |
| Boehme 2010a | 46 | 1 | 22 | 68 | 0.68 [0.55, 0.78] | 0.99 [0.92, 1.00] |
| Boehme 2010b | 8 | 0 | 4 | 101 | 0.67 [0.35, 0.90] | 1.00 [0.96, 1.00] |
| Boehme 2010c | 44 | 1 | 7 | 185 | 0.86 [0.74, 0.94] | 0.99 [0.97, 1.00] |
| Boehme 2010d | 8 | 3 | 6 | 215 | 0.57 [0.29, 0.82] | 0.99 [0.96, 1.00] |
| Boehme 2010e | 18 | 0 | 8 | 35 | 0.69 [0.48, 0.86] | 1.00 [0.90, 1.00] |
| Boehme 2011a | 68 | 4 | 23 | 303 | 0.75 [0.65, 0.83] | 0.99 [0.97, 1.00] |
| Boehme 2011b | 37 | 3 | 5 | 825 | 0.88 [0.74, 0.96] | 1.00 [0.99, 1.00] |
| Boehme 2011c | 121 | 2 | 32 | 669 | 0.79 [0.72, 0.85] | 1.00 [0.99, 1.00] |
| Boehme 2011d | 30 | 0 | 22 | 144 | 0.58 [0.43, 0.71] | 1.00 [0.97, 1.00] |
| Boehme 2011e | 31 | 16 | 0 | 671 | 1.00 [0.89, 1.00] | 0.98 [0.96, 0.99] |
| Boehme 2011f | 9 | 5 | 7 | 234 | 0.56 [0.30, 0.80] | 0.98 [0.95, 0.99] |
| Bowles 2011 | 20 | 1 | 4 | 23 | 0.83 [0.63, 0.95] | 0.96 [0.79, 1.00] |
| Carriquiry 2012 | 13 | 1 | 1 | 82 | 0.93 [0.66, 1.00] | 0.99 [0.93, 1.00] |
| Hanif 2011 | 9 | 0 | 5 | 146 | 0.64 [0.35, 0.87] | 1.00 [0.98, 1.00] |
| Hanrahan 2013 | 26 | 2 | 22 | 478 | 0.54 [0.39, 0.69] | 1.00 [0.99, 1.00] |
| Helb 2010 | 38 | 0 | 15 | 25 | 0.72 [0.58, 0.83] | 1.00 [0.86, 1.00] |
| Ioannidis 2011 | 15 | 2 | 3 | 32 | 0.83 [0.59, 0.96] | 0.94 [0.80, 0.99] |
| Kurbatova 2012 | 11 | 17 | 5 | 104 | 0.69 [0.41, 0.89] | 0.86 [0.78, 0.92] |
| Lawn 2011 | 23 | 2 | 30 | 320 | 0.43 [0.30, 0.58] | 0.99 [0.98, 1.00] |
| Malbruny 2011 | 4 | 0 | 0 | 45 | 1.00 [0.40, 1.00] | 1.00 [0.92, 1.00] |
| Marlowe 2011 | 31 | 4 | 12 | 82 | 0.72 [0.56, 0.85] | 0.95 [0.89, 0.99] |
| Miller 2011 | 3 | 2 | 2 | 58 | 0.60 [0.15, 0.95] | 0.97 [0.88, 1.00] |
| Moure 2011 | 61 | 0 | 17 | 29 | 0.78 [0.67, 0.87] | 1.00 [0.88, 1.00] |
| Rachow 2011 | 11 | 1 | 7 | 102 | 0.61 [0.36, 0.83] | 0.99 [0.95, 1.00] |
| Safianowska 2012 | 3 | 0 | 2 | 120 | 0.60 [0.15, 0.95] | 1.00 [0.97, 1.00] |
| Scott 2011 | 11 | 3 | 7 | 104 | 0.61 [0.36, 0.83] | 0.97 [0.92, 0.99] |
| Teo 2011 | 13 | 2 | 6 | 42 | 0.68 [0.43, 0.87] | 0.95 [0.85, 0.99] |
| Theron 2011 | 22 | 19 | 25 | 320 | 0.47 [0.32, 0.62] | 0.94 [0.91, 0.97] |
| Van Rie 2013 | 7 | 1 | 4 | 142 | 0.64 [0.31, 0.89] | 0.99 [0.96, 1.00] |
| Zeka 2011 | 7 | 0 | 4 | 68 | 0.64 [0.31, 0.89] | 1.00 [0.95, 1.00] |




TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

### *4.1.4. Investigations of heterogeneity: TB detection in smear-positive and smear-negative individuals suspected of having TB*

#### *4.1.4.1. Smear-positive TB*

There was little heterogeneity in the sensitivity estimates (range, 95–100%) for studies reporting data on smear-positive participants (24 studies, 33 study centres, 2071 participants). In the meta-analysis, the pooled sensitivity for smear-positive culture-positive TB was high at 98% (95% CrI, 97–99%) (23 studies, 1952 participants). Estimates of pooled specificity for Xpert MTB/RIF were not performed for studies of the smear-positive subgroup because almost all participants were considered to be true positives for TB.

#### *4.1.4.2. Smear-negative TB*

*Figure 3* displays the forest plots for studies reporting data on smear-negative participants (24 studies, 33 study centres, 7247 participants). There was considerable variability in the sensitivity estimates (range, 43–100%). Specificity estimates showed less variation (range, 86–100%). The meta-analysis included 23 studies that made direct comparisons between smear-positive subgroups and smear-negative subgroups. The pooled sensitivity estimate for smear-negative culture-positive TB was 68% (95% CrI, 61–74%), considerably lower than the pooled sensitivity estimate for smear-positive culture-positive TB, which was 98% (95% CrI, 97–99%) (*Table 2*). The 95% credible interval for the difference in the sensitivity of Xpert MTB/RIF for the smear-positive and smear-negative subgroups

did not cross 0, which suggests that this finding is statistically significant (*Table 2*).

Table 2. Impact of covariates on the heterogeneity of the sensitivity and specificity of Xpert MTB/RIF in detecting pulmonary TB in 24 studies (33 study centres)

| Covariate (No. of studies) | Median (%) pooled sensitivity (95% CrI) | Median (%) pooled specificity (95% CrI) |
|---|---|---|
| **Smear status** | | |
| Smear positive (23) | 98 (97–99) | [b] |
| Smear negative (23) | 68 (61–74) | 99 (98–99) |
| Difference (smear positive minus smear negative) | 31 (24–37) | [b] |
| P value (smear positive > smear negative) | 1.00 | [b] |
| **HIV status** | | |
| HIV negative (7) | 86 (76–92) | 99 (98–100) |
| HIV positive (7) | 79 (70–86) | 98 (96–99) |
| Difference (HIV negative minus HIV positive) | 7 (–5–18) | 1 (–1–3) |
| *P* value (HIV negative > HIV positive) | 0.90 | 0.85 |
| **Condition of specimen** | | |
| Fresh (13) | 89 (83–93) | 99 (98–100) |
| Frozen (6) | 85 (75–92) | 98 (95–99) |
| Difference (fresh minus frozen) | 3 (–5–14) | 1 (–0.4–4) |
| *P* value (fresh > frozen) | 0.78 | 0.91 |
| **Specimen preparation** | | |
| Unprocessed (11) | 90 (84–93) | 98 (97–99) |
| Processed (11) | 88 (81–93) | 99 (98–100) |
| Difference (unprocessed minus processed) | 2 (–5–10) | –1 (–2–1) |
| P value (unprocessed > processed) | 0.71 | 0.20 |
| **Proportion of TB cases in the study** | | |
| > 30% (12)[a] | 90 (85–93) | 98 (96–99) |
| ≤ 30% (10)[a] | 86 (79–92) | 99 (98–100) |
| Difference (> 30% minus ≤ 30%) | 3 (–4–11) | –1 (–3–0.3) |
| P value (> 30% minus ≤ 30%) | 0.81 | 0.06 |
| **Country income level** | | |
| High income (8) | 92 (87–96) | 99% (97–99) |
| Low income and middle income (14) | 86 (81–91) | 99% (97–99) |
| Difference (high income minus low income and middle income) | 6 (–1–12) | 0.1% (–2–2) |
| *P* value (high income > low income and middle income) | 0.96 | 0.56 |

CrI, credible interval; the CrI is the Bayesian equivalent of the confidence interval.

a The cut off of 30% was selected based on the median proportion in the studies included in the meta-analysis.
b Values could not be determined.

### *4.1.5. Detecting TB in HIV-negative and HIV-positive individuals suspected of having pulmonary TB*

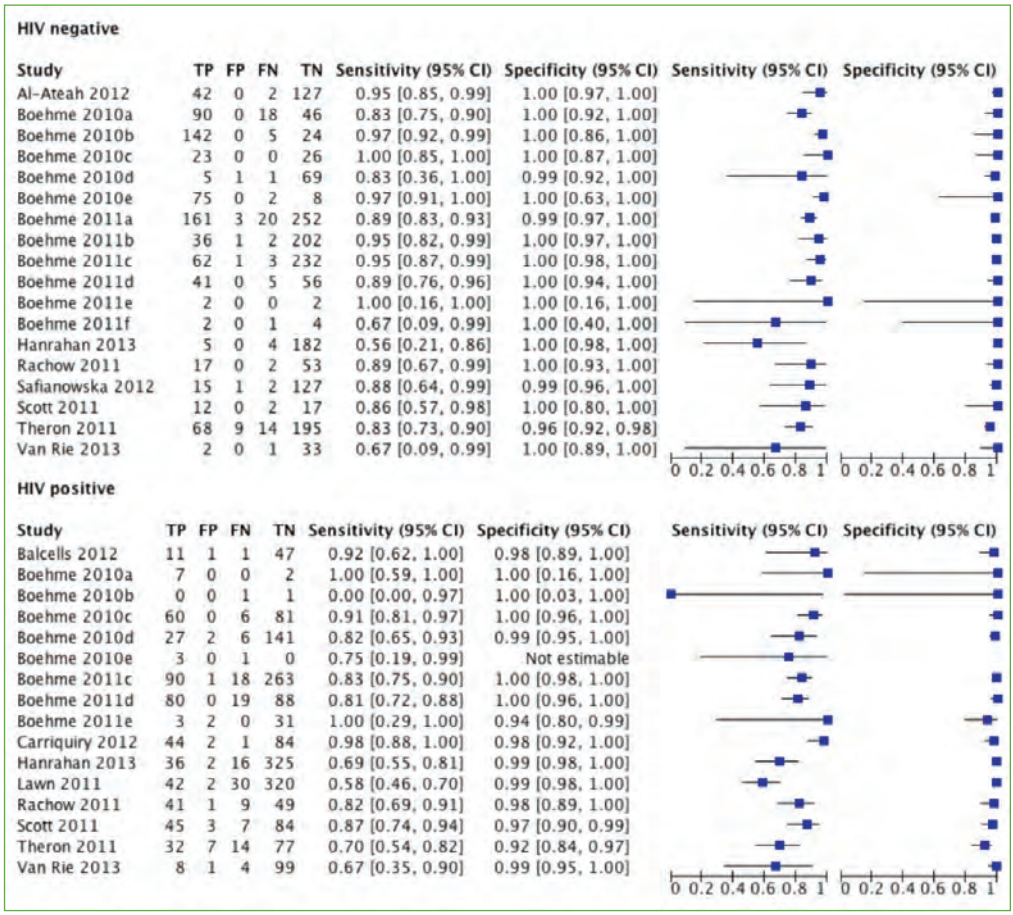
*Figure 4* displays the forest plots for studies reporting data for HIV-negative individuals (9 studies, 18 study centres, 2555 participants) and HIV-positive individuals (10 studies, 16 study centres, 2378 participants). There was variability in the sensitivity of Xpert MTB/RIF in both the HIV-negative subgroup (range, 56–100%) and the HIV-positive subgroup (range, 0–100%). The small number of participants in several studies may have contributed to some of the variability. Specificity varied less than sensitivity in both subgroups: 96–100% in the HIV-negative subgroup and 92–100% in the HIV-positive subgroup.

The meta-analysis included seven studies that provided data for both HIV-negative individuals (1470 participants) and HIV-positive individuals (1789 participants). The pooled sensitivity was 86% (95% CrI, 76–92%) in the HIV-negative subgroup and 79% (95% CrI, 70–86%) in the HIV-positive subgroup (*Table 2*). The corresponding pooled specificities were similar: in the HIV-negative subgroup the pooled specificity was 99% (95% CrI, 98–100%); in the HIV-positive subgroup it was 98% (95% CrI, 96–99%). When adjusting for the percentage of smear-positive patients in each study, the impact of the HIV covariate decreased, which suggests that some of the differences between the HIV-positive and HIV-negative subgroups could be attributed to differences in smear status.

### *4.1.6. Detecting TB among HIV-positive individuals by smear status*

Five studies reported data from which it was possible to assess the accuracy of Xpert MTB/RIF in HIV-positive individuals who had smear-negative culture-positive TB. The sensitivity of Xpert MTB/RIF ranged from 43% to 93% for smear-negative culture-positive TB; sensitivity ranged from 91% to 100% for smear-positive culture-positive TB. Data were sufficient to perform a univariate meta-analysis to assess the sensitivity of Xpert MTB/RIF. Among people living with HIV, the pooled sensitivity of Xpert MTB/RIF was 61% (95% CrI, 42–79%) for smear-negative culture-positive TB compared with 97% (95% CrI, 91–99%) for smear-positive culture-positive TB, which was a statistically significant result (data not shown). Hence, among people coinfected with HIV and TB, those with smear-positive disease were more likely to be diagnosed with TB using Xpert MTB/RIF than those with smear-negative disease.

Figure 4. Forest plots of the sensitivity and specificity of Xpert MTB/RIF for detecting pulmonary TB in HIV-negative individuals suspected of having TB (9 studies, 18 study centres) and HIV-positive individuals suspected of having TB (10 studies, 16 centres)[a]




TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

### *4.1.7. Effect of the condition of the specimen*

Although the manufacturer of the Xpert MTB/RIF assay recommends using fresh specimens, some studies have been conducted using frozen specimens. The effect of the condition of the specimen on the performance of Xpert MTB/RIF was explored. The pooled sensitivity for Xpert MTB/RIF using fresh specimens (13 studies) was 89% (95% CrI, 83–93%); this was slightly higher than the pooled sensitivity for frozen specimens (6 studies), which was 85% (95% CrI, 75–92%) (*Table 2*). The pooled specificity for fresh specimens was 99% (95% CrI, 98–100%); for frozen specimens it was 98% (95% CrI, 95–99%). The difference in the sensitivity and specificity of Xpert MTB/RIF for fresh and frozen specimens was not significantly different from 0 (*Table 2*). Metaregression modelling suggested that once smear-status was taken into account, there was no conclusive evidence supporting the impact of

the condition of the specimen on the sensitivity of Xpert MTB/RIF.

#### *4.1.8. Effect of specimen preparation*

The pooled sensitivity estimate for studies using unprocessed specimens was 90% (95% CrI, 84–93%) (11 studies), which was higher than the pooled sensitivity estimate for studies using processed specimens (88%; 95% CrI, 81–93%) (11 studies) (Table 2). The pooled specificity for unprocessed specimens was 98% (95% CrI, 97–99%), which was similar to the pooled specificity for processed specimens (99%; 95% CrI, 98–100%). The difference in the sensitivity and specificity of Xpert MTB/RIF for processed and unprocessed specimens was not significantly different from 0.

#### *4.1.9. Effect of the proportion of culture-confirmed cases of TB on a study*

For this analysis, a cut-off value of 30% for culture-confirmed cases of TB in a study was used because 30% was around the median proportion of such cases in the studies included in the review. A total of 12 studies were found to have more than 30% culture-confirmed cases of TB; 10 studies had 30% or less. The pooled sensitivity for studies with more than 30% culture-confirmed cases was 90% (95% CrI, 85–93%); the pooled sensitivity for studies with 30% or less culture-confirmed cases was 86% (95% CrI, 79–92%) (*Table 2*). The corresponding pooled specificity estimates were similar: 98% for studies with more than 30% culture-confirmed cases (95% CrI, 96–99%) and 99% for studies with less than 30% culture-confirmed cases (95% CrI, 98–100%) (*Table 2*). After adjusting for smear status, the probability of any differences in accuracy was further decreased, which suggests there was no conclusive evidence of the impact of TB prevalence on the performance of Xpert MTB/RIF.

#### *4.1.10. Effect of a country's income status*

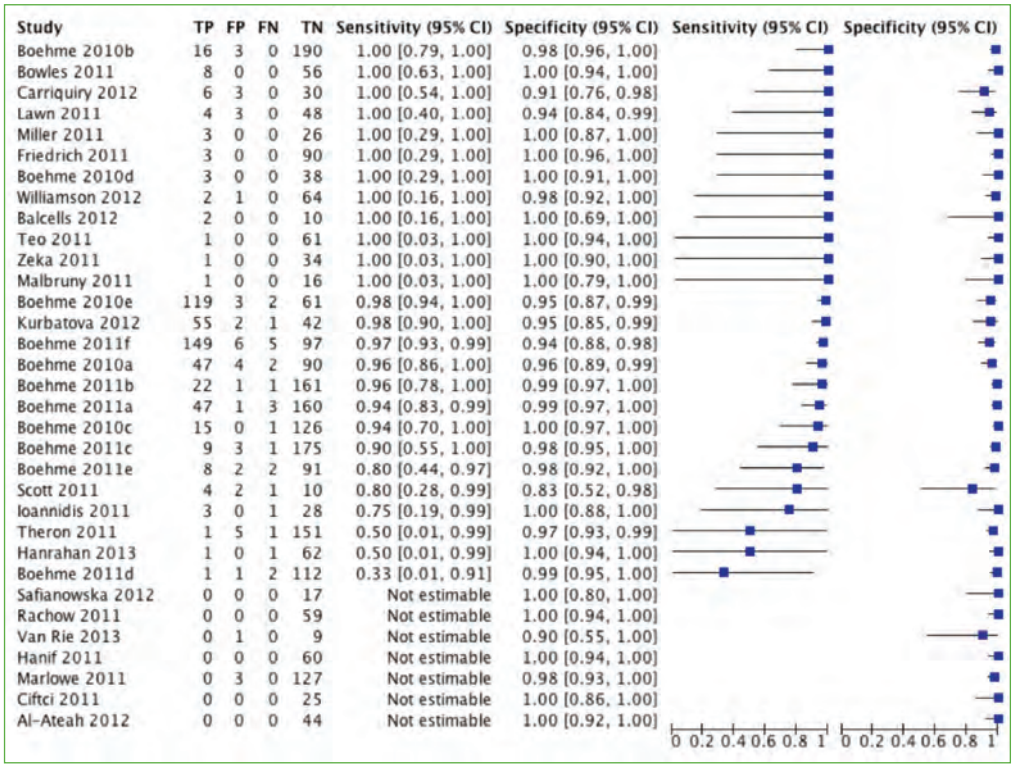
The pooled sensitivity for the 8 studies in high-income countries was 92% (95% CrI, 87–96%), which was higher than the pooled sensitivity for the 14 studies in low-income and middle-income countries (86%; 95% CrI, 81–91%) (*Table 2*). However, after adjusting for smear status, there was no conclusive evidence supporting the impact of a country's income status on the sensitivity of Xpert MTB/RIF.

#### *4.1.11. Detecting rifampicin resistance*

##### *4.1.11.1. Xpert MTB/RIF used as an initial test replacing conventional DST*

The 24 studies (33 study centres) in the analysis of using Xpert MTB/RIF as an initial test to replace conventional DST included 555 rifampicin-resistant specimens. *Figure 5* shows the forest plots of the sensitivity and specificity found in this analysis. The individual study centres in the plots are presented in order of decreasing sensitivity. Although there was heterogeneity in the estimates of sensitivity (range, 33–100%), in general there was less variability among study centres with a higher number of rifampicin-resistant specimens. Specificity showed less variability than sensitivity (range, 83–100%). The pooled sensitivity using univariate analysis was 95% (95% CrI, 90–97%); the pooled specificity using univariate analysis was 98% (95% CrI, 97– 99%) (*Table 1*). For the subset of studies that provided data for both sensitivity and specificity (17 studies, 2624 participants), the pooled sensitivity and specificity were the same by bivariate analysis.

Figure 5. Forest plots of the sensitivity and specificity of Xpert MTB/RIF for detecting rifampicin resistance when Xpert MTB/RIF was used as an initial test replacing phenotypic culture-based drug-susceptibility testing in 24 studies (33 study centres).[a]




TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

### *4.1.12. Investigations of heterogeneity: detecting rifampicin resistance*

#### *4.1.12.1. Effect of the version of the Xpert MTB/RIF assay*

The basis in the Xpert MTB/RIF system for detecting rifampicin resistance is the difference between the first *M. tuberculosis*-specific beacon or probe (the early-cycle threshold) and the last beacon (the late-cycle threshold). This difference is referred to as the delta-cycle threshold. The original Xpert MTB/RIF system configuration reported rifampicin resistance when the delta-cycle threshold was higher than 3.5 cycles; rifampicin sensitivity was reported when the delta-cycle threshold was 3.5 cycles or lower (using the Xpert MTB/RIF G1 cartridge). After May 2010, the manufacturer modified the delta-cycle threshold cut-off value to improve the specificity of Xpert MTB/RIF in detecting rifampicin resistance (using the G2 and G3 cartridges). Another modification was implemented in late 2011 (using the G4 cartridge) which changed the molecular beacon sequence of Probe B to improve detection of rifampicin resistance when there were fluctuations in annealing temperatures. Fluidic changes and software changes virtually eliminated the signal-loss detection error (known as a 5011 error), and allowed high sensitivity and specificity to be maintained when detecting TB and rifampicin resistance. These enhancements to the assays were considered to be part of a routine process

of product improvement. Cepheid, the Foundation for Innovative New Diagnostics (FIND) and the University of Medicine and Dentistry of New Jersey will continue to monitor the clinical performance of the Xpert MTB/RIF test. From 2013, the G4 cartridges were the only type of cartridge available.

The effect of the version of Xpert MTB/RIF on the sensitivity and specificity for detecting rifampicin resistance was investigated. The pooled sensitivity for studies using Xpert MTB/RIF G2, G3 or G4 cartridges (13 studies) was 93% (95% CrI, 87–97%); for studies using the Xpert MTB/RIF G1 cartridge (4 studies) it was 97% (95% CrI, 91–99%). The pooled specificity for studies using Xpert MTB/RIF G2, G3 or G4 cartridges (15 studies) was 98% (95% CrI, 96–99%); for studies using the Xpert MTB/RIF G1 cartridge (4 studies) it was 99% (95% CrI, 98–100%). The overlapping credible intervals indicate that there was no statistically significant difference in either the sensitivity or specificity estimates for the Xpert MTB/RIF G1 cartridge when compared with later versions of the assay.

#### *4.1.12.2. Accuracy of the Xpert MTB/RIF G4 cartridge*

Two studies used the Xpert MTB/RIF G4 cartridge and provided data for specificity determinations. One study observed a specificity of 100% (10/10 tests) (95% confidence interval [CI], 69–100%); the second study reported a specificity of 95% (42/44 tests) (95% CI, 85–99%) (*Figure 5*).

FIND evaluated the diagnostic accuracy of the G4 cartridge[13] in a study that involved testing 233 archived sputum specimens that had been stored in Borstel, Germany, and were from individuals suspected of having TB; additionally, there were 184 frozen sediments from Lima, Peru, that were positive for acid-fast bacilli (AFB), as well as frozen sputum specimens from a further 231 patients who had been consecutively enrolled from Baku, Azerbaijan. All of the samples were shipped to and tested in Germany using the G4 cartridge. Fresh sputum samples from 30 patients were tested using both the G3 cartridge and the G4 cartridge in Kampala, Uganda; a further 218 specimens were evaluated using both the G3 cartridge and the G4 cartridge in Cape Town, South Africa.

The reference standard used across all sites included at least one Löwenstein–Jensen culture and at least one BACTEC MGIT 960 culture, with *M. tuberculosis species* confirmed using Capilia TB-Neo (Tauns Laboratories, Shizuoka, Japan), GenoType MTBDR*plus* (Hain Lifescience, Nehren, Germany) or GenoType Mycobacterium CM/AS (Hain Lifescience, Nehren, Germany). Conventional testing for rifampicin resistance was performed using either the Löwenstein–Jensen proportion method or BACTEC MGIT 960 and, in a few cases, using only the Genotype MTBDR*plus* assay. Genetic sequencing was performed on results discordant between Xpert MTB/RIF and conventional DST. Six patients (smear-negative and culture-negative) were started on anti-TB treatment and excluded from the analysis. Genetic sequencing was used to resolve discordant results to determine sensitivity and specificity.

The overall sensitivity for rifampicin resistance was 98.9% (87/88 tests) (95% CI, 93.8–99.8%); the overall specificity for rifampicin-sensitive TB was 99.8% (433/434 tests) (95% CI, 98.7–100.0%). For four cases in which results were discordant (Xpert MTB/RIF identified samples as rifampicin sensitive but DST identified as resistant), the *rpo*B region was sequenced; the discordant results resolved in three of these cases in favour of Xpert MTB/RIF. For nine cases in which results were discordant and Xpert MTB/RIF identified the specimens as rifampicin resistant but DST identified as rifampicin sensitive, sequencing of the *rpo*B region was performed; discordant results resolved in eight of these cases in favour of Xpert MTB/RIF.

13 *Report: performance of Xpert MTB/RIF version G4 assay*. Geneva, Foundation for Innovative New Diagnostics, 2011 (available at: http://www.stoptb.org/wg/gli/assets/documents/map/findg4cartridge.pdf).

#### *4.1.12.3. Accuracy of the reference standards used*

Culture is regarded as the best reference standard for active TB, and was the reference standard used for TB in this review. Phenotypic culture-based DST methods, using WHO's recommended critical concentrations, were the reference standards for rifampicin resistance.[14]

Three recent studies have raised concerns about using phenotypic DST to detect rifampicin resistance, in particular the automated BACTEC MGIT 960 system, when the recommended critical concentrations are used. Van Deun and colleagues reported that the BACTEC 460 and the BACTEC MGIT 960 systems missed certain strains associated with low-level rifampicin resistance[15]. Furthermore, using Xpert MTB/RIF and gene sequencing, Williamson and colleagues identified four patients (three with clinical information available) whose TB isolates contained mutations to the *rpo*B gene but whose results from the BACTEC MGIT 960 system indicated that the isolates were rifampicin susceptible. In this study, 2/49 (4.1%) patients whose isolates did not have apparent *rpo*B gene mutations experienced treatment failure compared with 3/3 (100%) patients whose isolates did have *rpo*B gene mutations and had been deemed rifampicin susceptible using phenotypic methods[16]. In a study involving retreatment patients, Van Deun and colleagues found that disputed *rpo*B mutations conferring low-grade resistance were often missed by rapid phenotypic DST, particularly with the BACTEC MGIT 960 system, but to a lesser extent also by conventional slow DST[17]. The authors suggested this may be the reason for the perceived insufficient specificity of molecular DST for rifampicin. Although the study involved retreatment patients, the results also appear to be similar for individuals newly diagnosed with TB (A. Van Deun, personal communication, 2013). Specifically, the determination of the specificity of a molecular DST method based on phenotypic DST alone may underestimate the specificity of molecular DST. In light of these findings, it is unclear whether and to what extent Xpert MTB/RIF might outperform phenotypic DST methods for detecting rifampicin resistance.

#### *4.1.12.4. Effect of the proportion of rifampicin-resistant samples on a study*

For this analysis, we used a cut-off value of 15% for the proportion of rifampicin-resistant samples in a study. The pooled sensitivity for studies in which more than 15% of samples were rifampicin resistant (4 studies) was 96% (95% CrI, 91–98%), which was higher than the pooled sensitivity for studies in which 15% or less of the samples were rifampicin resistant (7 studies) (91%; 95% CrI, 79–97%). The pooled specificity for studies in which more than 15% of samples were rifampicin resistant was 97% (95% CrI, 94–99%); for studies in which 15% or less of the samples were rifampicin resistant the pooled specificity was 99% (95% CrI, 98–99%). The differences in sensitivity and specificity for Xpert MTB/RIF were not significantly different from 0.

#### *4.1.12.5. Sensitivity analyses*

Sensitivity analyses for detecting TB were undertaken by limiting inclusion in the meta-analysis to (1) studies that provided age data that met the inclusion criterion for adults (that is, people aged 15 years or older), (2) studies that used consecutive sampling, (3) studies where a single specimen yielded a single Xpert MTB/RIF result for a given individual, and (4) studies that explicitly tested individuals presumed to have

14 *Policy guidance on drug-susceptibility testing (DST) of second-line antituberculosis drugs*. Geneva, World Health Organization, 2008 (WHO/HTM/TB/2008.392).
15 Van Deun A et al. Mycobacterium tuberculosis strains with highly discordant rifampin susceptibility test results. *Journal of Clinical Microbiology*, 2009, 47:3501–3506.
16 Williamson DA et al. An evaluation of the Xpert MTB/RIF assay and detection of false-positive rifampicin resistance in Mycobacterium tuberculosis. *Diagnostic Microbiology and Infectious Disease*, 2012, 74:207–209.
17 Van Deun A et al. Rifampicin drug resistance tests for tuberculosis: challenging the gold standard. *Journal of Clinical Microbiology*, 2013, 51:2633–2640.

TB. A sensitivity analysis was also performed by excluding from the meta-analysis the two large multicentre studies. These sensitivity analyses made no difference to any of the findings.

#### *4.1.12.6. Nontuberculous mycobacteria*

Fourteen studies (2626 participants) provided data on a variety of nontuberculous mycobacteria (NTM) that grew from specimens tested to look for evidence of cross-reactivity. Among these 14 studies comprising 180 NTM, Xpert MTB/RIF was positive for only one specimen (0.6%) that grew NTM.

### *4.1.13. Summary of findings and GRADE evidence profiles*

In adults (with presumptive) TB with or without HIV infection, Xpert MTB/RIF is sensitive and specific. In comparison with smear microscopy, Xpert MTB/RIF substantially increases TB detection among culture-confirmed cases. Xpert MTB/RIF has higher sensitivity for detecting TB in smear-positive patients than in smear-negative patients. Nonetheless, Xpert MTB/RIF may also be valuable as an add-on test following smear microscopy in patients who have previously been found to be smear negative. In adults suspected of having TB or MDR-TB, Xpert MTB/RIF achieves high sensitivity and high specificity for detecting rifampicin resistance and can allow rapid initiation of treatment for MDR-TB.

- When used as an initial test to replace smear microscopy, Xpert MTB/RIF detected 88% of TB cases with high specificity (99%).
- When used as an add-on test following smear microscopy, Xpert MTB/RIF detected 68% of TB cases with high specificity (99%).
- The sensitivity of Xpert MTB/RIF for smear-positive culture-positive TB was 98%.
- The sensitivity of Xpert MTB/RIF for smear-negative culture-positive TB was 68%.
- Xpert MTB/RIF detected 79% of pulmonary cases of TB in people living with HIV, and 86% of cases of pulmonary TB in people without HIV infection. However, after adjusting for smear status, there was no evidence of a difference between the HIV-positive and HIV-negative subgroups.
- When used as an initial test to replace phenotypic culture-based DST, Xpert MTB/RIF detected 95% of rifampicin-resistant TB cases with a specificity of 98%.
- Phenotypic DST is an imperfect reference standard. Hence, determining the specificity of a molecular DST method by using phenotypic DST alone may underestimate the specificity of the molecular test.
- When genetic sequencing was used to resolve discordant results, the specificity of Xpert MTB/RIF for detecting rifampicin-sensitive TB was 99.8% (433/434 tests) (95% CI, 98.7–100.0%).

### *4.1.14. Strengths and limitations of the evidence base*

The review's findings are based on the use of comprehensive searches, strict inclusion criteria and standardized methods for data extraction. The strength of this review is that it allows an assessment to be made of the diagnostic accuracy of Xpert MTB/RIF in detecting TB when Xpert MTB/RIF is used as a replacement test for smear microscopy or as an add-on test following smear microscopy. In addition, the review allows for a determination of the accuracy of Xpert MTB/RIF in detecting rifampicin resistance when Xpert MTB/RIF is used as an initial test to replace conventional DST.

This data set involved comprehensive searching and correspondence with experts in the field and the test's manufacturer to identify additional studies, as well as repeated correspondence with the authors of the studies identified to obtain additional data and information that were missing from the literature. The search included studies published in all languages.

The majority of studies selected participants consecutively and interpreted the results of the

reference standard without knowledge of the results from Xpert MTB/RIF. Xpert MTB/RIF results are generated automatically, and do not require subjective interpretation. In the majority of studies, Xpert MTB/RIF tests were performed in intermediate-level and peripheral-level laboratories, settings that matched the review question. In general, studies were fairly well reported, though we corresponded with almost all authors to obtain additional data and missing information.

A major limitation in determining the specificity of Xpert MTB/RIF for detecting rifampicin resistance is the lack of a perfect reference standard. Phenotypic DST fails to detect some strains of *M. tuberculosis* with certain mutations in the *rpo*B gene; this means that the specificity of Xpert MTB/RIF for detecting rifampicin resistance is underestimated, a situation which incorrectly inflates the number of expected false-positive results.

### *4.1.15. GRADE evaluation and recommendations*

GRADE evidence profiles are provided in *Tables 3 to 12*. The GRADE process confirmed there was a solid evidence base to support the widespread and decentralized use of Xpert MTB/RIF to detect TB and rifampicin resistance. The Expert Group therefore concluded that:

- Xpert MTB/RIF should be used rather than conventional microscopy, culture and DST as the initial diagnostic test in individuals suspected of having MDR-TB or HIV-associated TB (strong recommendation, high-quality evidence);
- Xpert MTB/RIF may be used as a follow-on test to microscopy in adults where MDR-TB and HIV are of lesser concern, especially in further testing of smear-negative specimens (conditional recommendation acknowledging resource implications, high-quality evidence);
- Xpert MTB/RIF may be used rather than conventional microscopy and culture as the initial diagnostic test in all adults suspected of having TB (conditional recommendation acknowledging resource implications, high-quality evidence).

Table 3. GRADE evidence profile: accuracy of Xpert MTB/RIF in diagnosing pulmonary TB in adults

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for detection of pulmonary TB in adults, where Xpert MTB/RIF is used as a replacement test for smear microscopy?
Participants: Adults suspected of having pulmonary TB
Setting: Mainly intermediate-level laboratories and primary health-care facilities
Target condition: Pulmonary TB
Reference standard: Solid culture or liquid culture
Number of studies (number of participants): 22 (9008)
Pooled sensitivity: 88% (95% CrI, 84–92%); pooled specificity: 99% (95% CrI, 98– 99%)

| Outcome | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence | Number of results/1000 individuals tested (95% CrI)[a] | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | | Prevalence 25/1000[b] | Prevalence 50/1000[b] | Prevalence 100/1000[b] | Prevalence 300/1000[b] |
| True positives (individuals with TB) | Cross-sectional | None[c] | None[d] | None | None | Undetected[e] | High ⊕⊕⊕⊕ | 22 (21–23) | 44 (42–46) | 88 (84–92) | 264 (252–276) |
| False negatives (individuals incorrectly classified as not having TB) | Cross-sectional | None[c] | None[d] | None | None | Undetected[e] | High ⊕⊕⊕⊕ | 3 (2–4) | 6 (4–8) | 12 (8–16) | 36 (24–48) |
| False positives (individuals incorrectly classified as having TB | Cross-sectional | None[c] | None[d] | None | None | Undetected[e] | High ⊕⊕⊕⊕ | 10 (10–20) | 10 (10–19) | 9 (9–18) | 7 (7–14) |
| True negatives (individuals without TB) | Cross-sectional | None[c] | None[d] | None | None | Undetected[e] | High ⊕⊕⊕⊕ | 965 (956–965) | 941 (931–941) | 891 (882–891) | 693 (686–693) |

CrI, credible interval.

a The expected number of Xpert MTB/RIF results was based on the sensitivity and specificity estimates calculated from a comparison with culture for different prevalences of TB.
b The estimates of TB prevalence were provided by the WHO Steering Group.
c The QUADAS-2 tool was used to assess the risk of bias. The majority of studies enrolled individuals consecutively, and assessed the result of the reference standard blinded to the result from Xpert MTB/RIF. The Xpert MTB/RIF result is automated and was considered blinded in all studies.
d The rating of the quality of the evidence may be lowered if there are important differences in the tests studied and in the expertise of those applying the tests in the studies compared with the settings for which the recommendations are intended. The majority of studies (15/22; 68%) evaluated Xpert MTB/RIF in settings where it was intended to be used. Although studies of diagnostic accuracy may not provide direct evidence about patient-important outcomes, two studies provided information about the time until treatment initiation. In Boehme 2011, for smear-negative culture-positive TB, the median delay until beginning treatment before Xpert MTB/RIF was introduced was 56 days (interquartile range [IQR], 39–81 days) compared with 5 days (IQR, 2–8 days) after Xpert MTB/RIF was introduced. In Van Rie 2013, for smear-negative culture-positive TB patients with positive results from Xpert MTB/RIF, treatment was begun on the same day compared with after 13 days for patients diagnosed by other methods.
e One unpublished study was included in the analysis. No formal assessment of publication bias was conducted using methods such as funnel plots or regression tests because such techniques have not been found to be helpful in studies of diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

Table 4. GRADE evidence profile: accuracy of Xpert MTB/RIF in diagnosing pulmonary TB in sputum smear-positive adults

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for detection of pulmonary TB in smear-positive individuals?
Participants: Adults who are smear positive and culture positive
Setting: Mainly intermediate-level laboratories and primary health-care facilities
Target condition: Pulmonary TB
Reference standard: Solid culture or liquid culture
Number of studies (number of participants): 23 (1952)
Pooled sensitivity: 98% (95% CrI, 97–99%); pooled specificity: Not estimated

| Outcome | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence | Number of results/1000 individuals tested (95% CrI) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | | Prevalence 25/1000 | Prevalence 50/1000 | Prevalence 100/1000 |
| True positives (individuals with TB) | Cross-sectional | None[a] | None[b] | None[c] | None | Undetected[d] | High ⊕⊕⊕⊕ | 25 (24-25) | 49 (49-50) | 98 (97-99) |
| False negatives (individuals incorrectly classified as not having TB) | Cross-sectional | None[a] | None[b] | None[c] | None | Undetected[d] | High ⊕⊕⊕⊕ | 1 (0-1) | 1 (1-2) | 2 (1-3) |
| False positives[e] (individuals incorrectly classified as having TB | – | – | – | – | – | – | – | – | – | – |
| True negatives[e] (individuals without TB) | – | – | – | – | – | – | – | – | – | – |

CrI, credible interval.

a The QUADAS-2 tool was used to assess the risk of bias. The majority of studies enrolled individuals consecutively, and assessed the result of the reference standard blinded to the result from Xpert MTB/RIF. The Xpert MTB/RIF result is automated and was considered blinded in all studies.
b The majority of studies (16/23; 70%) evaluated Xpert MTB/RIF in settings where it was intended to be used. Although studies of diagnostic accuracy may not provide direct evidence about patient-important outcomes, the quality of the evidence was not downgraded.
c The estimates of sensitivity were highly consistent.
d One unpublished study was included in the analysis. No formal assessment of publication bias was conducted using methods such as funnel plots or regression tests because such techniques have not been found to be helpful for studies of diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.
e The pooled specificity for Xpert MTB/RIF was not estimated in these studies because almost all participants were considered to be true positives for TB.

Table 5. GRADE evidence profile: accuracy of Xpert MTB/RIF in diagnosing pulmonary TB in sputum smear-negative adults

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for detection of culture-confirmed pulmonary TB in smear-negative individuals?
Participants: Adults who were smear-negative and suspected of having pulmonary TB
Setting: Mainly intermediate-level laboratories and primary health-care facilities
Target condition: Pulmonary TB
Reference standard: Solid culture or liquid culture
Number of studies (number of participants): 23 (7151)
Pooled sensitivity: 68% (95% CrI, 61–74%); pooled specificity: 99% (95% CrI, 98–99%)

| Outcome | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence | Number of results/1000 smear-negative individuals tested (95% CrI) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | | Prevalence 25/1000 | Prevalence 50/1000 | Prevalence 100/1000 |
| True positives (individuals with TB) | Cross-sectional | None[a] | None[b] | Serious (–1)[c] | None[d] | Undetected[e] | Moderate ⊕⊕⊕○ | 17 (15-19) | 34 (31-37) | 68 (61-74) |
| False negatives (individuals incorrectly classified as not having TB) | Cross-sectional | None[a] | None[b] | Serious (–1)[c] | None[d] | Undetected[e] | Moderate ⊕⊕⊕○ | 8 (7-10) | 16 (13-20) | 32 (26-39) |
| False positives (individuals incorrectly classified as having TB | Cross-sectional | None[a] | None[b] | None | None | Undetected[e] | High ⊕⊕⊕⊕ | 10 (10-20) | 10 (10-19) | 9 (9-18) |
| True negatives (individuals without TB) | Cross-sectional | None[a] | None[b] | None | None | Undetected[e] | High ⊕⊕⊕⊕ | 965 (956-965) | 941 (931-941) | 891 (882-891) |

CrI, credible interval.

a The QUADAS-2 tool was used to assess the risk of bias. The majority of studies enrolled individuals consecutively, and assessed the results of the reference standard blinded to the result from Xpert MTB/RIF. The Xpert MTB/RIF result is automated and was considered blinded in all studies.
b The majority of studies (16/23; 70%) evaluated Xpert MTB/RIF in settings where it was intended to be used. Although studies of diagnostic accuracy may not provide direct evidence about patient-important outcomes, the quality of the evidence was not downgraded.
c There was some variability in sensitivity estimates across studies. This heterogeneity could not be explained by study quality or by removing the study by Lawn 2011 from the analysis. The study by Lawn, which found the lowest sensitivity, evaluated the use of Xpert MTB/RIF to screen HIV-positive patients with advanced immunodeficiency, regardless of their symptoms, who were enrolling in antiretroviral therapy. The observed sensitivity may have varied among studies as a result of characteristics associated with the participants, but there were insufficient data to investigate this possibility. The quality of the evidence was downgraded by one point.
d At a pretest probability of 10%, the confidence intervals for true positives and false negatives were relatively narrow.
e One unpublished study was included in the analysis. No formal assessment of publication bias was conducted using methods such as funnel plots or regression tests because such techniques have not been found to be helpful in studies of diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

Table 6. GRADE evidence profile: accuracy of Xpert MTB/RIF in diagnosing pulmonary TB in adults living with HIV

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for detection of pulmonary TB in people living with HIV?
Participants: Adults living with HIV and suspected of having pulmonary TB
Setting: Mainly intermediate-level laboratories and primary health-care facilities
Target condition: Pulmonary TB
Reference standard: Solid culture or liquid culture
Number of studies (number of participants): 7 (1789)
Pooled sensitivity: 79% (95% CrI, 70–86%); pooled specificity: 98% (95% CrI, 96–99%)

| Outcome | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence | Number of results/1000 individuals tested (95% CrI) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | | Prevalence 25/1000 | Prevalence 50/1000 | Prevalence 100/1000 |
| True positives (individuals with TB) | Cross-sectional | None[a] | None[b] | Serious (–1)[c] | None[d] | Undetected[e] | Moderate ⊕⊕⊕◯ | 20 (18-22) | 40 (35-43) | 79 (70-86) |
| False negatives (individuals incorrectly classified as not having TB) | Cross-sectional | None[a] | None[b] | Serious (–1)[c] | None[d] | Undetected[e] | Moderate ⊕⊕⊕◯ | 5 (4-8) | 11 (7-15) | 21 (14-30) |
| False positives (individuals incorrectly classified as having TB | Cross-sectional | None[a] | None[b] | None | None | Undetected[e] | High ⊕⊕⊕⊕ | 20 (10-39) | 19 (10-38) | 18 (9-36) |
| True negatives (individuals without TB) | Cross-sectional | None[a] | None[b] | None | None | Undetected[e] | High ⊕⊕⊕⊕ | 956 (936-965) | 931 (912-941) | 882 (864-891) |

CrI, credible interval.

a The QUADAS-2 tool was used to assess the risk of bias. All studies enrolled individuals consecutively, and assessed the result of the reference standard blinded to the result from Xpert MTB/RIF. The Xpert MTB/RIF result is automated and was considered blinded in all studies.
b The rating of the quality of the evidence may be lowered if there are important differences in the tests studied and the expertise of those applying the tests in the studies compared with the settings for which the recommendations are intended. The majority of studies (6/7; 86%) evaluated Xpert MTB/RIF in settings where it was intended to be used. Although studies of diagnostic accuracy may not provide direct evidence about patient-important outcomes, the quality of the evidence was not downgraded.
c There was some variability in sensitivity estimates across studies but this heterogeneity could not be explained by study quality. The observed sensitivity may have varied among studies as a result of characteristics associated with the participants, but there were insufficient data to investigate this possibility. The quality of evidence was downgraded by one point.
d At a pretest probability of 10%, the confidence intervals for true positives and false negatives were relatively narrow.
e No formal assessment of publication bias was conducted using methods such as funnel plots or regression tests because such techniques have not been found to be helpful in studies of diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

Table 7. GRADE evidence profile: accuracy of Xpert MTB/RIF in diagnosing pulmonary TB in adults without HIV infection

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for detection of pulmonary TB in adults without HIV infection?
Participants: Adults who were HIV negative and suspected of having pulmonary TB
Setting: Mainly intermediate-level laboratories and primary health-care facilities
Target condition: Pulmonary TB
Reference standard: Solid culture or liquid culture
Number of studies (number of participants): 7 (1470)
Pooled sensitivity: 86% (95% CrI, 76–92%); pooled specificity: 99% (95% CrI, 98–100%)

| Outcome | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence | Number of results/1000 individuals tested (95% CrI) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | | Prevalence 25/1000 | Prevalence 50/1000 | Prevalence 100/1000 |
| True positives (individuals with TB) | Cross-sectional | None[a] | None[b] | Serious (–1)[c] | None[d] | Undetected[e] | Moderate ⊕⊕⊕◯ | 22 (19-23) | 43 (38-46) | 86 (76-92) |
| False negatives (individuals incorrectly classified as not having TB) | Cross-sectional | None[a] | None[b] | Serious (–1)[c] | None[d] | Undetected[e] | Moderate ⊕⊕⊕◯ | 4 (2-6) | 7 (4-12) | 14 (8-24) |
| False positives (individuals incorrectly classified as having TB | Cross-sectional | None[a] | None[b] | None | None | Undetected[e] | High ⊕⊕⊕⊕ | 10 (0-20) | 10 (0-19) | 9 (0-18) |
| True negatives (individuals without TB) | Cross-sectional | None[a] | None[b] | None | None | Undetected[e] | High ⊕⊕⊕⊕ | 956 (956-975) | 941 (931-950) | 891 (882-900) |

CrI, credible interval.

a The QUADAS-2 tool was used to assess the risk of bias. All studies enrolled individuals consecutively, and assessed the result from the reference standard blinded to the result from Xpert MTB/RIF. The Xpert MTB/RIF result is automated and was considered blinded in all studies.
b The rating of the quality of the evidence may be lowered if there are important differences in the tests studied and the expertise of those applying the tests in the studies compared with the settings for which the recommendations are intended. The majority of studies (6/7; 86%) evaluated Xpert MTB/RIF in settings where it was intended to be used. Although studies of diagnostic accuracy may not provide direct evidence about patient-important outcomes, the quality of the evidence was not downgraded.
c There was some variability in sensitivity estimates across studies but this heterogeneity could not be explained by study quality. The observed sensitivity may have varied among studies as a result of characteristics associated with the participants, but there were insufficient data to investigate this possibility. The quality of the evidence was downgraded by one point.
d At a pretest probability of 10%, the confidence intervals for true positives and false negatives were relatively narrow.
e No formal assessment of publication bias was conducted using methods such as funnel plots or regression tests because such techniques have not been found to be helpful in studies of diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

Table 8. GRADE evidence profile: the incremental yield of Xpert MTB/RIF compared with microscopy in patients with culture-confirmed TB

PICO question: What is the incremental yield of Xpert MTB/RIF compared with microscopy in patients with culture-confirmed TB?
Participants: Adults with culture-confirmed TB
Setting: Mainly intermediate-level laboratories and primary health-care facilities
Target condition: Pulmonary TB
Reference standard: Solid culture or liquid culture
Number of studies (number of participants): 21 (8880)
Pooled sensitivity for smear microscopy : 65% (95% CrI, 57–72%); pooled sensitivity for Xpert MTB/RIF: 88% (95% CrI, 84–92%)

| Outcome | Number of results/1000 individuals tested (95% CrI)[a] | | | | | | | | Quality of evidence[b] |
|---|---|---|---|---|---|---|---|---|---|
| | Prevalence 25/1000 | | Prevalence 50/1000 | | Prevalence 100/1000 | | Prevalence 300/1000 | | |
| | Smear microscopy | Xpert MTB/RIF | Smear microscopy | Xpert MTB/RIF | Smear microscopy | Xpert MTB/RIF | Smear microscopy | Xpert MTB/RIF | |
| True positives (individuals with TB) | 16 (14–18) | 22 (21–23) | 33 (29–36) | 44 (42–46) | 65 (57–72) | 88 (84–92) | 195 (171–216) | 264 (252–276) | High ⊕⊕⊕⊕ |
| True positives (**absolute difference**) | 6 more | | 11 more | | 23 more | | 69 more | | |
| False negatives (individuals incorrectly classified as not having TB) | 9 (7–11) | 3 (2–4) | 18 (14–22) | 6 (4–8) | 35 (28–43) | 12 (8–16) | 105 (84–129) | 36 (24–48) | High ⊕⊕⊕⊕ |
| False negatives (**absolute difference**) | 6 fewer | | 12 fewer | | 23 fewer | | 69 fewer | | |

CrI, credible interval.

a The sensitivity results were taken from bivariate analyses (including both sensitivity and specificity) to obtain the values for true positives and false negatives.
b The GRADE framework was used to assess the quality of the evidence. For microscopy, the sensitivity estimates for individual studies were variable (range, 29–83%), and the pooled sensitivity estimate was imprecise. The main reason for heterogeneity and imprecision in the sensitivity estimate was considered to be the variability in smear-positive status across studies. Several additional factors may have contributed to this heterogeneity, including type of microscopy, method of specimen processing and HIV status. The quality of evidence was not downgraded.

Table 9. GRADE evidence profile: accuracy of Xpert MTB/RIF in diagnosing pulmonary TB in adults as an add-on test following negative sputum-smear microscopy

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for detection of pulmonary TB in adults, where Xpert MTB/RIF is used as an add-on test following a negative smear-microscopy result?
Participants: Adults who are smear negative and suspected of having pulmonary TB
Setting: Mainly intermediate-level laboratories and primary health-care facilities
Target condition: Pulmonary TB
Reference standard: Solid culture or liquid culture
Number of studies (number of participants): 23 (7151)
Pooled sensitivity: 68% (95% CrI, 61–74%); pooled specificity: 99% (95% CrI, 98–99%)

| Outcome | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence | Number of results/1000 smear-negative individuals tested (95% CrI) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | | Prevalence 25/1000 | Prevalence 50/1000 | Prevalence 100/1000 |
| True positives (individuals with TB) | Cross-sectional | None[a] | None[b] | Serious (–1)[c] | None[d] | Undetected[e] | Moderate ⊕⊕⊕○ | 17 (15-19) | 34 (31-37) | 68 (61-74) |
| False negatives (individuals incorrectly classified as not having TB) | Cross-sectional | None[a] | None[b] | Serious (–1)[c] | None[d] | Undetected[e] | Moderate ⊕⊕⊕○ | 8 (7-10) | 16 (13-20) | 32 (26-39) |
| False positives (individuals incorrectly classified as having TB | Cross-sectional | None[a] | None[b] | None | None | Undetected[e] | High ⊕⊕⊕⊕ | 10 (10-20) | 10 (10-19) | 9 (9-18) |
| True negatives (individuals without TB) | Cross-sectional | None[a] | None[b] | None | None | Undetected[e] | High ⊕⊕⊕⊕ | 965 (956-965) | 941 (931-941) | 891 (882-891) |

CrI, credible interval.

a The QUADAS-2 tool was used to assess the risk of bias. The majority of studies enrolled individuals consecutively, and assessed the result from the reference standard blinded to the result from Xpert MTB/RIF. The Xpert MTB/RIF result is automated and was considered blinded in all studies.
b The majority of studies (16/23; 70%) evaluated Xpert MTB/RIF in settings where it was intended to be used. Although studies of diagnostic accuracy may not provide direct evidence about patient-important outcomes, the quality of the evidence was not downgraded.
c There was some variability in sensitivity estimates across studies but this heterogeneity could not be explained by study quality or by removing the study by Lawn 2011 from the analysis. The study by Lawn, which found the lowest sensitivity, evaluated the use of Xpert MTB/RIF to screen HIV-positive patients with advanced immunodeficiency, regardless of their symptoms, who were enrolling in antiretroviral therapy. The observed sensitivity may have varied among studies as a result of characteristics associated with the participants, but there were insufficient data to investigate this possibility. The quality of the evidence was downgraded by one point.
d At a pretest probability of 10%, the confidence intervals for true positives and false negatives were relatively narrow.
e One unpublished study was included in the analysis. No formal assessment of publication bias was conducted using methods such as funnel plots or regression tests because such techniques have not been found to be helpful in studies of diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

Table 10. **Sensitivity of Xpert MTB/RIF in smear-negative culture-confirmed pulmonary TB in individuals, by HIV status**

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for detection of pulmonary TB in adults, where Xpert MTB/RIF is used as an add-on test following a negative smear microscopy result, stratified by HIV status?
Participants: Adults with smear-negative culture-confirmed pulmonary TB
Setting: One intermediate-level laboratory and one primary health-care clinic
Reference standard: Phenotypic culture using solid media or liquid media
Number of studies (number of participants): 2 (91)

| Study | HIV-positive participants (n = 33) | HIV-negative participants (n = 58) |
|---|---|---|
| | Sensitivity (%) (95% CI) | Sensitivity (%) (95% CI) |
| Theron 2011 | 48 (27–69) | 45 (25–67) |
| Van Rie 2013 | 60 (27–86) | 67 (13–98) |

CI, confidence interval.

Table 11. **GRADE evidence profile: additional yield of Xpert MTB/RIF over microscopy in smear-negative TB**

PICO question: What is the additional yield of Xpert MTB/RIF over microscopy in smear-negative TB?
Participants: Adults who are smear negative and culture positive
Setting: Mainly intermediate-level laboratories and primary health-care facilities
Target condition: Pulmonary TB
Reference standard: Solid culture or liquid culture
Number of studies (number of participants): 23 (7151)
Polled sensitivity for smear microscopy: 0%; pooled sensitivity for Xpert MTB/RIF : 68% (95% CrI, 61–74%)

| Outcome | Number of results/1000 individuals tested (95% CrI)[a] | | | | | | Quality of evidence[b] |
|---|---|---|---|---|---|---|---|
| | Prevalence 25/1000 | | Prevalence 50/1000 | | Prevalence 100/1000 | | |
| | Smear microscopy | Xpert MTB/RIF | Smear microscopy | Xpert MTB/RIF | Smear microscopy | Xpert MTB/RIF | |
| True positives (individuals with TB) | 0 | 17 (15–19) | 0 | 34 (31–37) | 0 | 68 (61–74) | Moderate ⊕⊕⊕○ |
| True positives (**absolute difference**) | 17 more | | 34 more | | 68 more | | |
| False negatives (individuals incorrectly classified as not having TB) | 25 | 8 (7–10) | 50 | 16 (13–20) | 100 | 32 (26–29) | Moderate ⊕⊕⊕○ |
| False negatives (**absolute difference**) | 17 fewer | | 34 fewer | | 68 fewer | | |

CrI, credible interval.

a The sensitivity results were taken from the bivariate analyses (including both sensitivity and specificity) to obtain the values for true positives and false negatives.
b The GRADE framework was used to assess the quality of the evidence. The quality of the evidence was downgraded one point for inconsistency/imprecision.

Table 12. GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting rifampicin resistance, where Xpert MTB/RIF replaces phenotypic culture-based drug-susceptibility testing as the initial test

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for detection of rifampicin resistance, where Xpert MTB/RIF is used as an initial test replacing phenotypic culture-based drug-susceptibility testing?
Participants: Adults with confirmed TB. Setting: Mainly intermediate-level laboratories and primary health-care facilities.
Target condition: Rifampicin resistance. Reference standard: Phenotypic culture-based drug-susceptibility testing[a].
Number of studies (number of participants) pooled sensitivity: 17 (555). Number of studies (number of participants) pooled specificity: 24 (2414).
Pooled sensitivity: 95% (95% CrI, 90–97%); Pooled specificity: 98% (95% CrI, 97–99%)[b]

| Outcome | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence | Number of results/1000 individuals tested (95% CrI) | |
|---|---|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | | Prevalence 50/1000[c] | Prevalence 100/1000[c] |
| True positives (individuals with TB and rifampicin resistance) | Cross-sectional | None[d] | None[e] | None[f] | None | Undetected[g] | High ⊕⊕⊕⊕ | 84 (45-49) | 143 (135-146) |
| False negatives (individuals incorrectly classified as having TB that is rifampicin susceptible) | Cross-sectional | None[d] | None[e] | None[f] | None | Undetected[g] | High ⊕⊕⊕⊕ | 3 (2-5) | 5 (8-15) |
| False positives (individuals incorrectly classified as having TB with rifampicin resistance) | Cross-sectional | None[d] | None[e] | None[f] | None | Undetected[g] | High ⊕⊕⊕⊕ | 19 (10-29) | 17 (9-26) |
| True negatives (individuals with TB that is rifampicin susceptible) | Cross-sectional | None[d] | None[e] | None[f] | None | Undetected[g] | High ⊕⊕⊕⊕ | 931 (922-941) | 833 (825-842) |

CrI, credible interval.

a Phenotypic drug-susceptibility testing is considered an imperfect reference standard but its use should not downgrade the rating of the quality of evidence.
b Estimates of sensitivity and specificity were determined separately using univariate analyses.
c Prevalence estimates were provided by the WHO Steering Group. The upper limit for the prevalence of rifampicin resistance in new cases was estimated to be 5% (50/1000 cases); the lower limit for the prevalence of rifampicin resistance in previously treated cases was estimated to be 15% (150/1000 cases).
d The QUADAS-2 tool was used to assess the risk of bias. The majority of studies enrolled individuals consecutively, and assessed the result from the reference standard blinded to the result from Xpert MTB/RIF. The Xpert MTB/RIF result is automated and was considered blinded in all studies.
e The rating of the quality of the evidence may be lowered if there are important differences in the tests studied and the expertise of those applying the tests in the studies compared with the settings for which the recommendations are intended. The majority of studies (12/17, 71% for sensitivity; and 16/24, 67% for specificity) evaluated Xpert MTB/RIF in settings where it was intended to be used. Only two studies evaluated Xpert MTB/RIF with the G4 cartridge. In 2013, the G4 cartridges were the only type of cartridge available. It is possible that the performance of Xpert MTB/RIF using the G4 cartridge will be different. Although studies of diagnostic accuracy may not provide direct evidence about patient-important outcomes, the quality of the evidence was not downgraded.
f Estimates of sensitivity and specificity were consistent. Several studies have suggested that determining the specificity of a molecular method of drug-susceptibility testing, such as Xpert MTB/RIF, using phenotypic drug-susceptibility testing alone may underestimate the specificity of the molecular test. Gene sequencing of discordant results obtained using Xpert MTB/RIF and phenotypic drug-susceptibility testing (discrepant analysis) is often resolved in favour of Xpert MTB/RIF, suggesting that Xpert MTB/RIF has higher specificity. However, discrepant analyses may introduce bias.
g One unpublished study was included in the analysis. No formal assessment of publication bias was conducted using methods such as funnel plots or regression tests because such techniques have not been found to be helpful in studies of diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

## 4.2. Using Xpert MTB/RIF to diagnose extrapulmonary TB and rifampicin resistance in adults and children

### *4.2.1. Characteristics of the studies*

*Annex 5* shows the PRISMA diagram of the studies. From the literature search, 194 citations and 51 full-text articles were identified. Twenty two studies met the eligibility criteria for inclusion in the review. *Annex 5* lists the included and excluded studies along with the reasons for exclusion.

The 22 studies (5922 samples) included in the analysis of using Xpert MTB/RIF to diagnose extrapulmonary TB and rifampicin resistance in adults and children had a TB prevalence (based on culture) that ranged from 7% to 81%. Only one study (unpublished) was written in a language other than English (Portuguese). Thirteen studies (59%) were conducted in low-income or middle-income countries. All studies were performed in tertiary care centres or reference laboratories. In 19 studies, the HIV status of the participants was known; 5 studies did not include any HIV-positive patients. Fourteen studies included HIV-positive patients; their percentage of the study population ranged from 1% to 87%. Of the three studies in which participants' HIV status was unknown, two were done in settings where the incidence of HIV is low (France and Germany) and one was done where the incidence is high (South Africa). Two studies (one by Bates 2012 and one by Walters 2012) included only children; nine studies included no children at all. In the remaining 11 studies, the percentages of children in the study population ranged from 2% to 34%. The median number of samples per study was 145 (interquartile range, 67–342). Three published studies and four unpublished studies included only one type of sample (for example, only pleural fluid). The remainder of the studies included different sample types in varying percentages. Twelve studies reported on only one sample per patient; the other studies either reported on multiple samples per patient or did not report the number of samples per patient. Six studies used archived samples (frozen); 15 studies used fresh samples; one study used both fresh and frozen samples.

Studies varied widely in how specimens were processed (*Table 13*). Only four studies used the protocol recommended by the manufacturer for unprocessed respiratory samples. Thirteen studies (59%) reported using a mechanical homogenization step for nonliquid samples. Twelve studies (55%) reported using a solution of N-acetyl-L-cysteine and sodium hydroxide (NaOH) for specimen digestion and decontamination; one study used only NaOH. Some studies did not consistently decontaminate all specimens but used decontamination only when bacterial contamination was identified. Most of the studies that included a mechanical homogenization step also performed a decontamination procedure. Fourteen studies reported a concentration step; and 10 had a resuspension step using varying volumes for the two steps. Data on sample processing were extracted on a study level, although some steps (for example, homogenization) might apply only to certain types of samples.

The ratio of the volume of sample reagent to the volume of the sample also varied. Seven studies used a ratio of reagent to sample of 3:1; 15 studies used a ratio of 2:1. Of the 12 studies that used digestion and decontamination, and a concentration step, 6 used a ratio of sample reagent to sample of 2:1. The manufacturer recommends a ratio of reagent to sample of 3:1 for samples processed by the Kent and Kubica protocol[18] and 2:1 for unprocessed sputum.

18 Kent PA, Kubica GP. *Public health mycobacteriology: a guide for the level III laboratory.* Atlanta, GA, United States Department of Health and Human Services, 1985.

Table 13. Sample types and processing methods used in the 22 studies included in the review of the accuracy of Xpert MTB/RIF in diagnosing extrapulmonary TB and rifampicin resistance in adults and children[a]

| No. | Study (year) | Setting | Specimen condition | Homogenization | Digestion or decontamination | Concentration (volume) | Resuspension | Ratio of sample reagent to sample | Comments on sample processing | Input volume | Total No. of samples | Specimen type and No. of samples |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Al-Ateah (2012) | Tertiary care centre | Fresh | Mechanical | NALC–NaOH | Yes | Yes (2.5 ml) | 2:1 | – | 1 | 67 | Pleural fluid: 13<br>Lymph node: 8<br>CSF: 14<br>Tissue: 16 |
| 2 | Armand (2011) | Tertiary care centre | Frozen | Mechanical | NALC–NaOH | Yes (variable) | No | 3:1 | – | 2 | 56 | Pleural fluid: 8<br>Lymph node: 20<br>CSF: 5<br>Gastric fluid: 12<br>Tissue: 5 |
| 3 | Bates (2012) | Tertiary care centre | Fresh | Mechanical | NALC–NaOH | Y (variable) | Yes (2 ml) | 3:1 | – | 2 | 800 | Gastric fluid: 800 |
| 4 | Causse (2011) | Reference laboratory | Fresh | No | NALC–NaOH | Yes (variable) | No | 2:1 | – | 1 | 340 | Pleural fluid: 34<br>Lymph node: 80<br>CSF: 50<br>Gastric fluid: 54<br>Tissue: 18 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | Friedrich (2011) | Tertiary care centre | Fresh | No | No | Yes (50 ml) | Yes (2 ml) | 2:1 | – | >2 | 24 | Pleural fluid: 25 |
| 6 | Hanif (2011) | Reference laboratory | Fresh | No | No | No | No | 2:1 | – | 2 | 29 | Pleural fluid: 11<br>Lymph node: 9<br>CSF: 5<br>Gastric fluid: 6 |
| 7 | Hillemann (2011) | Reference laboratory | Fresh | Mechanical (for tissue) | NALC–NaOH | Yes | Yes (1.5 ml) | 3:1 | – | 2 | 495 | Pleural fluid: 111<br>Lymph node: 65<br>CSF: 19<br>Gastric fluid: 30<br>Tissue: 171 |
| 8 | Lighthelm (2011) | Tertiary care centre | Fresh | No | No | No | No | 2:1 | If sample with <0.7 ml, PBS added to increase volume to 0.7 ml | 2 | 48 | Lymph node: 48 |
| 9 | Malbruny (2011) | Tertiary care centre | Fresh and frozen | Mechanical | NALC–NaOH | Yes (variable) | No | 3:1 | All samples were concentrated except for CSF | 0.5-1 | 124 | Pleural fluid: 12<br>Lymph node: 23<br>CSF: 15<br>Gastric fluid: 33<br>Tissue: 17 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | Moure (2012) | Tertiary care centre | Frozen | Mechanical | NALC–NaOH | Yes (variable) | Yes (2 ml) | 2:1 | 3 ml went into cartridge: 1 ml sediment and 2 ml sample reagent | 3 | 149 | Pleural fluid: 34<br>Lymph node: 38<br>CSF: 14<br>Gastric fluid: 8<br>Tissue: 29 |
| 11 | Safi-anowska (2012) | Tertiary care centre | Fresh | Mechanical | NALC–NaOH | Yes (see comment) | Yes (2.5 ml) | 3:1 | Centrifugation then additional washing step with 50 ml distilled water and repeat centrifugation | 2 | 68 | Pleural fluid: 32<br>Lymph node: 2<br>CSF: 6<br>Tissue: 17 |
| 12 | Tortoli (2012) | Tertiary care centre | Frozen | Mechanical | NALC–NaOH | Yes (variable) | Yes (2 ml) | 2:1 | Decontamination done only for non-sterile samples | 2 | 1491 | Pleural fluid: 330<br>Lymph node: 118<br>CSF: 138<br>Gastric fluid: 224<br>Tissue: 250 |
| 13 | Vadwai (2011) | Tertiary care centre | Fresh | Mechanical (for tissue) | No | No | No | 2:1 | Ratio of reagent to sample was 2:1 except for CSF which was usually <1 ml and was raised to 2 ml with sample reagent | 2 | 342 | Pleural fluid: 29<br>Lymph node: 188<br>CSF: 19<br>Tissue: 113 |
| 14 | Walters (2012) | Tertiary care centre | Fresh | Mechanical | NALC–NaOH | Yes (see comment) | Yes (1.5 ml) | 2:1 | PBS added up to volume of 40 ml then centrifuged | 2.1 | 37 | Gastric fluid: 20 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 15 | Zeka (2011) | Tertiary care centre | Frozen | No | No | No | No | 3:1 | – | 2 | 175 | Pleural fluid: 56<br>Lymph node: 26<br>CSF: 31<br>Gastric fluid: 6<br>Tissue: 20 |
| 16 | Caws | Tertiary care centre | Fresh | No | No | Yes (5–8 ml) | Yes (0.7ml) | 2:1 | 200 ul deposit was taken and 500 ul PBS added | 2 | 141 | CSF: 141 |
| *17* [redacted] | | *Tertiary Care Center* | *Fresh* | *Mechanical (for tissue)* | *No* | *No* | *No* | *2:1* | – | *2* | *170* | *Pleural fluid: 84*<br>*Lymph node: 29*<br>*CSF: 40* |
| *18* [redacted] | | *Tertiary Care Center* | *Fresh* | *No* | *No* | *No* | *No* | *2:1* | – | *2* | *76* | *Pleural fluid: 76* |
| 19 | Oliveira | Tertiary care centre | Frozen | No | NaOH | No | No | 2:1 | – | 2 | 90 | Pleural fluid: 90 |
| 20 | Patel | Tertiary care centre | Frozen | No | No | No | No | 2:1 | – | 2 | 151 | CSF: 151 |
| *21* [redacted] | | *Tertiary Care Center* | *Fresh* | *Mechanical (for tissue)* | *NALC–NaOH* | *Yes* | *Yes (5 ml)* | *3:1* | *Decontaminated only if contaminated. If sample >0.5 ml, then ratio of sample reagent to sample was 2:1* | *2* | *427* | *Pleural fluid: 383*<br>*CSF: 27*<br>*Gastric fluid: 1*<br>*Tissue: 9* |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *22[redacted]h* | *Tertiary Care Center* | *Fresh* | *Mechanical (for tissue)* | *NALC–NaOH* | *Yes* | *No* | *2:1* | *NALC–NaOH used for all samples except CSF; 2–3 ml deposit was used* | *2* | *629* | *Pleural fluid: 70*<br>*Lymph node: 193*<br>*CSF: 42*<br>*Gastric fluid: 87*<br>*Tissue: 38* |

NALC– NaOH, N-acetyl-L-cysteine and sodium hydroxide; CSF, cerebrospinal fluid; PBS, phosphate-buffered saline.

a Data from unpublished studies appear in italics, and the names of unpublished studies containing confidential data have been obscured. Where no year has been given, the data are from 2013.

### *4.2.2. Quality of the studies*

The methodological quality of each study included in the review was assessed separately for culture as the reference standard and for the CRS. Most analyses were performed using two reference standards: culture (the current reference standard) and a combined clinical and laboratory reference standard chosen by the study's authors (given the technical limitations of diagnosis by culture alone).The quality assessment was conducted similarly for both culture and CRS except for the assessment of flow and timing, and the blinding to the result of the reference standard. The overall quality of all studies included in the review is summarized in *Figure 6* for those using culture as the reference standard and in Figure 7 for those using a CRS.

*Figure 6* shows the risks of bias and judgements about applicability as judged by the review's authors for each QUADAS-2 domain for the 22 studies included in the review.

Figure 6. **Risks of bias (%) and judgements about applicability (%) for each domain of the QUADAS-2 tool in studies using culture as the reference standard for TB detection and using Xpert MTB/RIF to diagnose extrapulmonary TB and rifampicin resistance in adults and children**

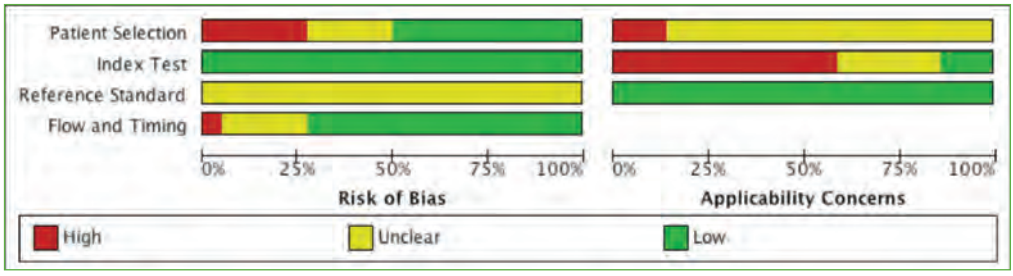
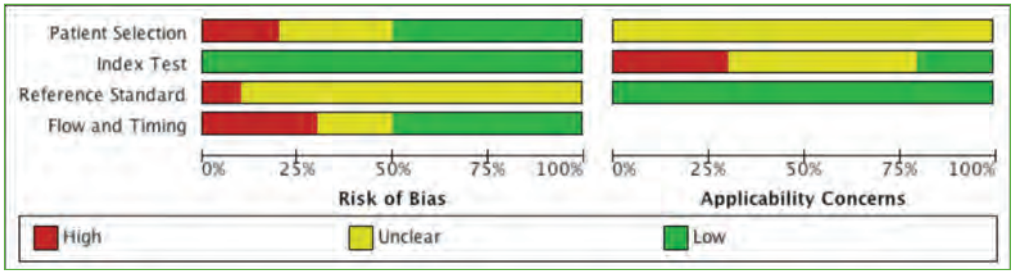



*Figure 7* shows the risks of bias and judgements about applicability as judged by the review's authors for each QUADAS-2 domain for the 22 studies included in the review.

Figure 7. **Risks of bias (%) and judgements about applicability (%) for each domain of the QUADAS-2 tool in studies using an author-defined composite reference standard for TB detection and Xpert MTB/RIF to diagnose extrapulmonary TB and rifampicin resistance in adults and children**




In the patient-selection domain, six studies were judged to have a high risk of bias because they were either case–control studies or used convenience sampling to select participants for enrolment. The majority of studies collected data prospectively (17 out of 22; 77%). Concerns about applicability in respect to the intended site of use for Xpert MTB/RIF (as described for pulmonary TB – that is, in district or subdistrict health-care settings) were judged to be high if the

test was used in reference laboratories; concerns about applicability were judged to be unclear if Xpert MTB/RIF was used in a hospital laboratory.

The results from the index test were considered blinded with respect to the results from the reference standard since interpretation of the result from Xpert MTB/RIF does not require human judgement. Similarly, the threshold for positivity for TB detection is fixed by the manufacturer and thus, by definition, it is prespecified. In the index-test domain, all studies were considered to have a low risk of bias.

With respect to the applicability domain, variations in the use of the test were considered to possibly affect estimates of the diagnostic accuracy of Xpert MTB/RIF to different degrees. There was low concern in 3 out of 22 studies (14%) for sputum samples if the samples were unprocessed and the test was done according to the manufacturer's recommendations. There was high concern in 13 out of 22 studies (59%) that the test's performance could be altered by adding a mechanical homogenization step because it was unclear whether the homogenization would be sufficient, and what quantity of sample particles would ultimately be included in the sample input volume. It was also considered conceivable that the particles could clog the valves and result in a higher rate of test results being classified as indeterminate. In 6 of the 22 studies (27%) the processing protocols were unclear.

The reference standards were considered to introduce bias due to the possible misclassification of participants, and hence all studies in this domain were rated as unclear. Blinding to the result from the Xpert MTB/RIF test was considered to be relevant for culture only if species identification was not done (using the Capilia test, NAAT or sequencing). Blinding was more important for the CRS, particularly if the CRS included smear tests without species identification, or if a clinical evaluation was included. The risk of bias resulting from clinical evaluation was a concern in one study, but the remaining studies showed a low risk of bias.

In the domain of flow and timing, an interval of several days between performing the index test and the reference standard was not considered to be problematic for diagnosing TB in presumptive cases (that is, in patients not being treated for TB): TB is a chronic disease and test results are unlikely to change within a few days, therefore the misclassification of disease status is unlikely. All patients across all studies were included in the analysis of culture as the reference standard; therefore partial verification bias was not considered to be a problem. For the CRS, three studies indicated that some patients did not receive culture as part of the CRS. These studies were rated as being of high concern for a risk of bias.

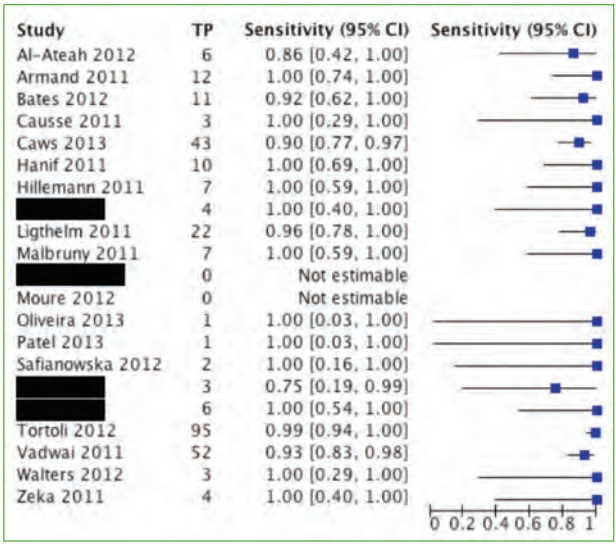
### *4.2.3. Using Xpert MTB/RIF to detect extrapulmonary TB*

#### *4.2.3.1. Detecting extrapulmonary TB: all sample types*

The studies that looked at using Xpert MTB/RIB to detect extrapulmonary TB with all types of samples were diverse with respect both to the different types of samples tested and the relative percentages of each type of sample in each study. The heterogeneity in performance characteristics across the different sample types was substantial, primarily in sensitivity. Therefore, combining these studies to obtain an overall estimate of the accuracy of Xpert MTB/RIF in extrapulmonary TB would not be meaningful.

However, when all studies were assessed by smear status, the heterogeneity was restricted primarily to the smear-negative samples. For all smear-positive samples across the studies (390 samples), heterogeneity was limited (*Figure 8*). A univariate analysis was done only for sensitivity (97.6%; 95% CI, 95.2–99.9%) since data were too limited for to allow for an estimation of specificity.

Figure 8. Forest plot of the sensitivity of Xpert MTB/RIF in detecting extrapulmonary TB in the smear-positive subgroup of participants in 22 studies[a]




TP, true positive; FN, false negative; CI, confidence interval.

a The specificity of Xpert MTB/RIF for detecting extrapulmonary TB was not estimated because only limited data were available. The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test results are the number of each type of result (true positive or false negative).

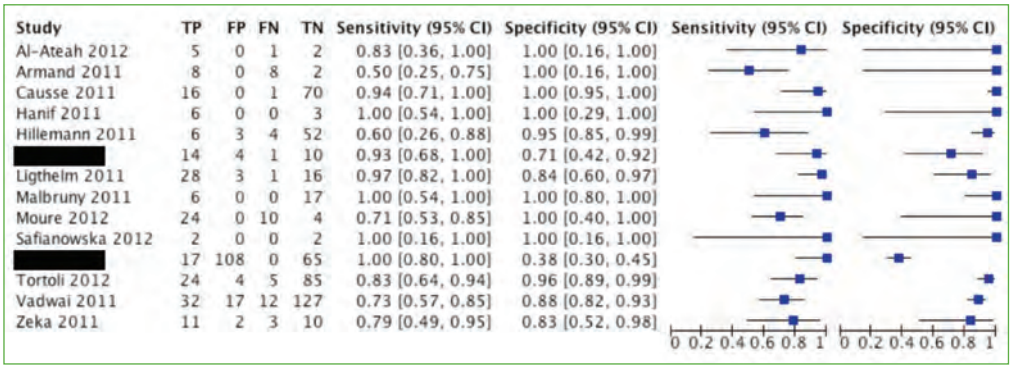
An analysis of predefined subgroups of sample types (for example, pleural fluid, lymph node aspirate or tissue, cerebrospinal fluid [CSF], gastric fluid, and tissue other than lymph node) was undertaken to account for the heterogeneity among studies. Data on the smear status of samples were not available for the individual types of samples. Therefore, samples included in the subgroups were either smear-positive, smear-negative or of unknown smear status.

### *4.2.3.2. Detecting lymph node TB in samples from biopsy or fine-needle aspiration*

Fourteen studies were identified that tested the accuracy of Xpert MTB/RIF on samples from lymph node biopsies or fine-needle aspiration (FNA) and compared the results against culture as a reference standard (*Figure 9*). A meta-analysis was performed for each sample type if at least 4 studies had at least 10 samples in each study. For the 11 studies with more than 10 samples (total, 849 samples), estimates for sensitivity ranged from 50% to 100%. The pooled sensitivity across studies was 84.9% (95% CI, 72.1–92.4%); and the pooled specificity was 92.5% (95% CI, 80.3–97.4%). Only two studies reported any indeterminate results from Xpert MTB/RIF testing: Armand 2011 reported 10% indeterminate results (2/20) and one unpublished study reported 1.6% (3/193).

One unpublished study had a much lower specificity for Xpert MTB/RIF than the other studies (38%, compared with 71–100% in the other studies). If the subset of published studies was analysed separately, the pooled specificity improved slightly to 94.4% (95% CI, 88.2–97.4%), with the lower bound of the 95% confidence interval shifted upward by 8%. In the analysis of the same subset of published studies, the sensitivity decreased slightly to 80.8% (95% CI, 67.9–89.4%). In the subset of studies that performed consecutive sampling of participants, the pooled sensitivity was slightly increased to 89.4% (95% CI, 74.1–96.1%), and the specificity was decreased to 86.9% (95% CI, 67.5–95.5%), but the precision of these estimates also decreased. The overall estimates did not differ significantly if the two case–control studies (Armand 2011 and Moure 2012) were removed (*Table 14*).

Figure 9. Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting extrapulmonary TB in lymph node samples (tissue or aspirate) compared with culture as the reference standard[a]




TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

Table 14. Sensitivity and specificity of Xpert MTB/RIF in detecting extrapulmonary TB, by sample type

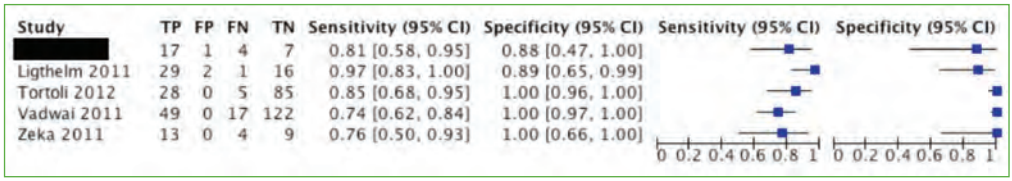
| Sample type | Pooled sensitivity (%) (95% CI) | Pooled specificity (%) (95% CI) |
|---|---|---|
| **Lymph node (overall)** | 85 (72–92) | 93 (80–97) |
| Published studies | 81 (68–89) | 94 (88–97) |
| Studies with consecutive sampling | 89 (74–96) | 87 (68–96) |
| Excluding case–control studies | 89 (77–95) | 90, (76–97) |
| **Pleural fluid (overall)** | 44 (25–65) | 98 (95–99) |
| Published studies | 48 (22–75) | 99 (98–100) |
| Studies with consecutive sampling | 54 (39–68) | 98 (93–99) |
| Excluding case–control studies | 46 (23–72) | 98 (95–99) |
| **Cerebrospinal fluid (overall)** | 80 (62–90) | 99 (96–100) |
| Published studies | 77 (48–100) | 99 (97–100) |
| Studies with consecutive sampling | 70 (47–94) | 99 (98–100) |
| Excluding case–control studies | 74 (56–93) | 99 (98–100) |
| **Gastric fluid (overall)** | 84 (66–93) | 98 (92–100) |
| Published studies | 84 (73–94) | 99 (99–100) |
| Studies with consecutive sampling | 89 (72–100) | 91 (81–100) |
| Excluding case–control studies | 89 (77–100) | 96 (92–100) |
| **Tissue (overall)** | 81 (68–90) | 98 (87–100) |
| Published studies | 80 (66–89) | 99 (89–100) |
| Studies with consecutive sampling | 80 (69–88) | 98 (77–100) |
| Excluding case–control studies | 84 (76–90) | 98 (86–100) |

CI, confidence interval.

Five studies (one unpublished) assessed Xpert MTB/RIF using lymph node samples, and compared the results against an author-defined CRS (*Figure 10*). In the different studies the CRS included some combination of NAAT other than Xpert MTB/RIF, histology, smear, culture, biochemical testing, presenting signs, or a response to treatment with anti-TB therapy. The pooled sensitivity was estimated to be 83.7% (95% CI, 73.8–90.3%); the pooled specificity was estimated to be 99.2% (95% CI, 88.4–100%).

Studies that used fresh samples showed a slightly higher sensitivity and a lower specificity than those that used frozen samples; however, the precision of these estimates was low because data were limited. Only nine studies provided information on the prevalence of HIV, and only two studies included more than 10% HIV-positive patients. Accuracy estimates for these studies did not differ substantially from those that included fewer HIV-positive patients (*Figure 10*). Given the limited amount of data for the group that had a prevalence of HIV greater than 10%, a summary estimate was not determined.

Figure 10. Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting extrapulmonary TB in lymph node samples (tissue or aspirate) compared with a composite reference standard[a]




TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

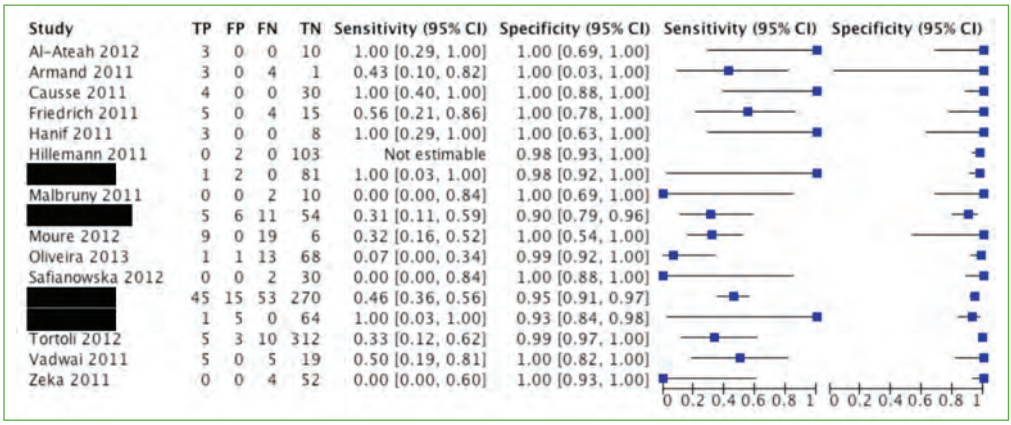
### *4.2.3.3. Detecting pleural TB in pleural fluid*

Seventeen studies (1385 samples, 217 culture-positive) provided data that could be used to estimate the sensitivity and specificity of Xpert MTB/RIF for testing pleural fluid. Results from the assessment of the accuracy of Xpert MTB/RIF using samples from pleural biopsy were integrated into the assessment of Xpert MTB/RIF for testing tissue biopsies of all kinds other than lymph node (see section 4.2.3.6).

The sensitivity of Xpert MTB/RIF for testing pleural fluid varied from 0% to 100% among the studies (*Figure 11*). The outliers at the lower end of the range and the upper end were studies with few culture-confirmed cases of TB. For the meta-analysis, studies that did not contribute to either sensitivity or specificity, and those that included fewer than 10 pleural fluid specimens, were excluded. The pooled sensitivity was low at 43.7%, with a wide 95% confidence interval (24.8–64.7%); the pooled specificity was high at 98.1% (95% CI, 95.3–99.2%).

Seven studies (4 published and 3 unpublished) with 698 samples (188 culture-positive) evaluated Xpert MTB/RIF for testing pleural fluid compared with a CRS. In the different studies the CRS included some combination of NAAT other than Xpert MTB/RIF, histology, smear, culture, biochemical testing, presenting signs, or a response to treatment with anti-TB therapy. Compared with studies that used culture as the reference standard, the CRS subgroup yielded an even lower pooled sensitivity (17.0%; 95% CI, 7.5–34.2%) with a high specificity (99.9%; 95% CI, 93.7–100.0%).

Figure 11. Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting TB using pleural fluid compared with culture as a reference standard[a]




TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

Sensitivity was increased in studies in which participants had a low rate of coinfection with HIV (48% compared with 31% in studies in which more than 10% of participants also had HIV), however the confidence intervals were wide and overlapped. There was also no difference in the results of studies that used a concentration step. In assessing the condition of the specimen, an improved sensitivity was observed for fresh samples (50%; 95% CI, 36–64%) compared with frozen samples (26%; 95% CI, 14–40%), but specificity was lower for fresh samples (95%; 95% CI, 93–98%) versus frozen samples (99%; 95% CI, 97–100%).

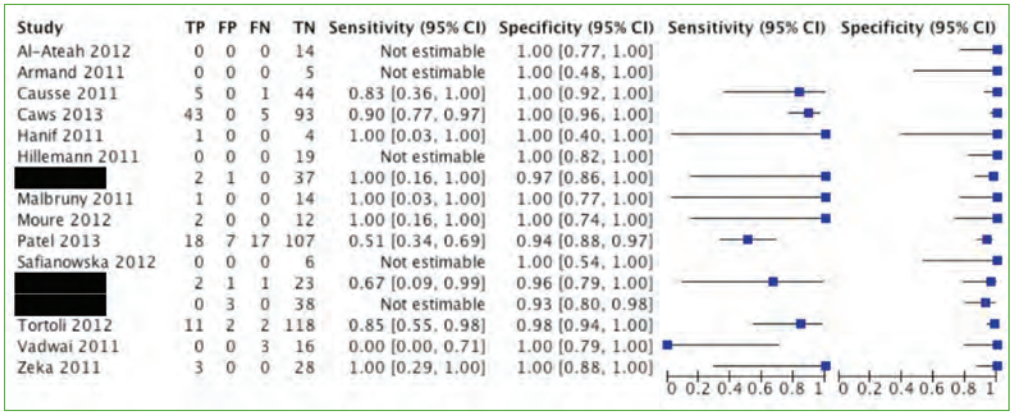
#### *4.2.3.4. Detecting TB in samples of cerebrospinal fluid*

In total, 709 CSF samples were tested with Xpert MTB/RIF and compared against culture as a reference standard in 16 studies (13 studies had more than 10 samples, and 10 of 13 provided information on both sensitivity and specificity). Only 117 culture-confirmed cases of TB were found. Estimates of sensitivity varied widely and ranged from 51% to 100%: one study with 19 samples (3 false negatives) was an outlier at 0% (*Figure 12*). The pooled sensitivity across studies was 79.5% (95% CI, 62.0–90.2%) and pooled specificity was 98.6% (95% CI, 95.8–99.6%), suggesting good performance of Xpert MTB/RIF in detecting TB in CSF when tested against culture as a reference standard.

For the subset of eight published studies only a univariate analysis was feasible; sensitivity and specificity were largely unchanged from the overall estimate (*Table 14*). Sensitivity in the 10 studies that used consecutive sampling rather than convenience sampling was decreased, at 70.1% (95% CI, 46.6–93.7%); specificity remained largely the same. All estimates from these sensitivity analyses were accompanied by wide and overlapping confidence intervals.

Only 6 studies (3 unpublished) assessed the accuracy of using Xpert MTB/RIF on CSF samples compared against an author-defined CRS; sensitivity estimates ranged from 20% to 86% (*Figure 13*). The pooled sensitivity was estimated to be 55.5% (95% CI, 44.2–66.3%) and the pooled specificity was estimated to be 98.8% (95% CI, 94.5–99.8%). The reduced sensitivity of Xpert MTB/RIF compared with the CRS versus culture as a reference standard suggests that either the CRS was too broad or that culture as the single reference standard is inadequate.

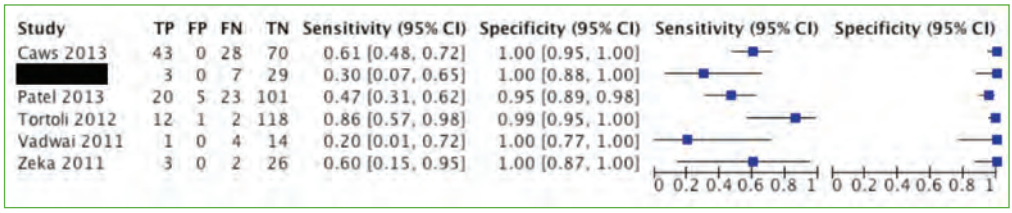
Figure 12. Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting TB in cerebrospinal fluid compared with culture as a reference standard[a]




TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

Figure 13. Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting TB in cerebrospinal fluid compared with a composite reference standard[a]




TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

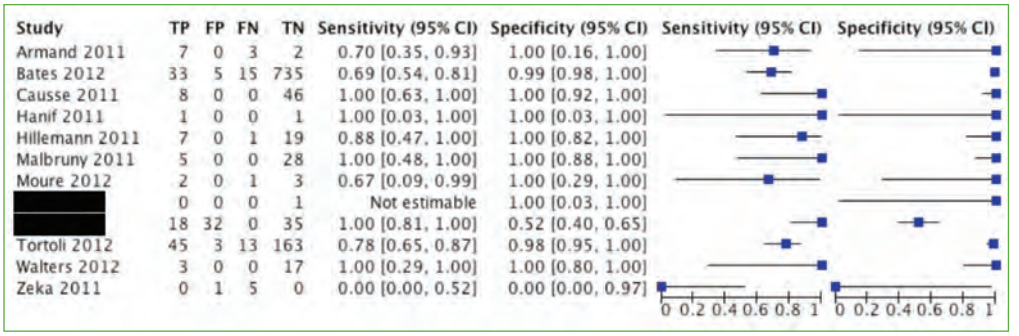
The prevalence of HIV and the condition of the specimen did not have an effect on the performance estimates for Xpert MTB/RIF used to test CSF. Ten of the sixteen studies comparing Xpert MTB/RIF with culture as the reference standard used a concentration step in processing the sample. Six studies did not use a concentration step. A concentration step appeared to increase the sensitivity of the Xpert MTB/RIF test (82%, 95% CI, 71–93% for concentrated samples versus 56%, 95% CI, 36–77% for unconcentrated samples), although the confidence intervals overlapped. A concentration step did not affect the specificity.

#### *4.2.3.5. Detecting TB in gastric fluid*

A total of 12 studies (1258 samples) examined the performance of Xpert MTB/RIF in samples of gastric fluid and compared the results against culture as a reference standard (8 studies had more than 10 samples). Two studies included only children. The remaining 10 studies included

adults and children (the proportion of children included across sample types ranged from 0% to 33.5%). One study with 788 samples that had valid results using Xpert MTB/RIF accounted for 62.6% of all samples of gastric fluid specimens. The estimates of sensitivity varied from 69% to 100%; specificity varied from 98% to 100%, with one estimate at 52% (*Figure 14*). The pooled sensitivity across the studies was 83.8% (95% CI, 65.9–93.2%) and the pooled specificity was 98.1% (95% CI, 92.3–99.5%), both of which suggest good accuracy for Xpert MTB/RIF in detecting TB in gastric fluid. Indeterminate results from Xpert MTB/RIF were reported for 2 studies (Bates 2012, 1.5% of samples, and one unpublished study, 2.3%).

Figure 14. Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting TB in gastric fluid compared with culture as a reference standard[a]




TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

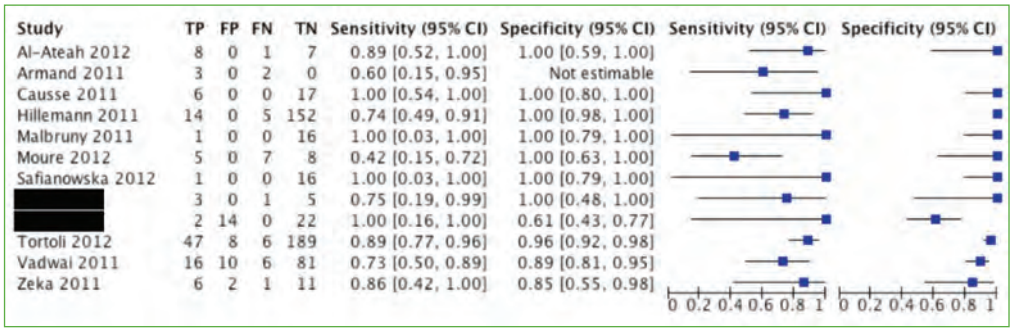
The pooled estimates for the seven published studies were analysed by univariate analysis; the estimates for both sensitivity and specificity were largely unchanged. Univariate sensitivity for studies that used consecutive sampling (4 studies) was slightly increased (89.1%; 95% CI, 72.1–100%); specificity was decreased at 90.5% (95% CI, 81–100%). Both estimates had wide confidence intervals, which highlights the lack of precision caused by the limited amount of data.

All studies included a concentration step, therefore this was not considered a source of heterogeneity within the data. The condition of the gastric-fluid specimen (fresh versus frozen) did not appear to have a strong effect on the performance of Xpert MTB/RIF.

#### *4.2.3.6. Detecting TB in tissue samples*

There were 12 studies (699 samples) that tested Xpert MTB/RIF using tissue samples (from any site other than a lymph node) and compared the results against culture as a reference standard (10 studies had more than 10 samples). The estimates of sensitivity varied widely and ranged from 42% to 100% (*Figure 15*). The estimate of pooled sensitivity was calculated as 81.2% (95% CI, 67.7–89.9%). The pooled specificity was 98.1% (95% CI, 87.0–99.8%). One study represented an outlier with respect to specificity, having an estimate of 61.1%; specificity in other studies ranged from 84.6% to 100%. Indeterminate results from Xpert MTB/RIF were present in the studies by Hilleman 2011 (6/176 samples, 3.4%) and Tortoli 2012 (4/254, 1.6%). The condition of the specimen did not have an impact on the performance of Xpert MTB/RIF in the tissue samples.

Figure 15. Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting TB in tissue samples (other than from a lymph node) compared with culture as a reference standard[a]




TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

### *4.2.4. Detecting rifampicin resistance*

Data on detecting rifampicin resistance was used only from published studies because data collection was incomplete in some of the unpublished studies. Furthermore, data from a study were included only if DST had been done for all samples that were positive by culture and Xpert MTB/RIF because the selective confirmation of results could have introduced bias.

In total, data on resistance testing were available for 566 samples from 13 studies. Forty one samples were confirmed to be rifampicin resistant by phenotypic DST. Given the limited amount of data, no summary estimate was calculated. The average prevalence of rifampicin resistance across the studies was 5.4%, with the highest prevalence reported from India (25.6%). Xpert MTB/RIF did not identify 2 of the 41 phenotypically rifampicin-resistant samples. Six of the 41 samples that were identified as rifampicin resistant by Xpert MTB/RIF were found to be susceptible by phenotypic DST. Five of these six samples underwent sequencing of the *rpo*B gene, and four were found to have a mutation in the same region of the *rpo*B gene at codon 533. Hence, Xpert MTB/RIF detected four additional rifampicin-resistant strains that would have been missed by phenotypic DST alone.

### *4.2.5. Summary of findings and GRADE evidence profiles*

The review of the diagnostic accuracy of Xpert MTB/RIF in nonrespiratory samples identified 15 published studies and 7 unpublished studies. An analysis was performed using either culture as a reference standard or an author-defined CRS (that included culture in all studies).

Good performance of Xpert MTB/RIF in was observed in smear-positive samples across sample types (sensitivity, 97.6%; 95% CI, 95.2–99.9%), and the low number of indeterminate results (1.4%) supports the use of this technique in specific nonrespiratory samples. However, the variability in performance, particularly in sensitivity, among nonrespiratory sample types suggests that a different approach might be needed for different types of samples. Although Xpert MTB/RIF can be considered a diagnostic tool for evaluating TB in tissue samples and lymph node samples, gastric fluid and CSF, the benefit in testing pleural fluid is limited. Furthermore, these recommendations do not apply to stool, urine or blood, given that data on the utility of Xpert MTB/RIF for these specimens are limited.

In respect to testing for rifampicin resistance, the data were limited, and did not allow pooled

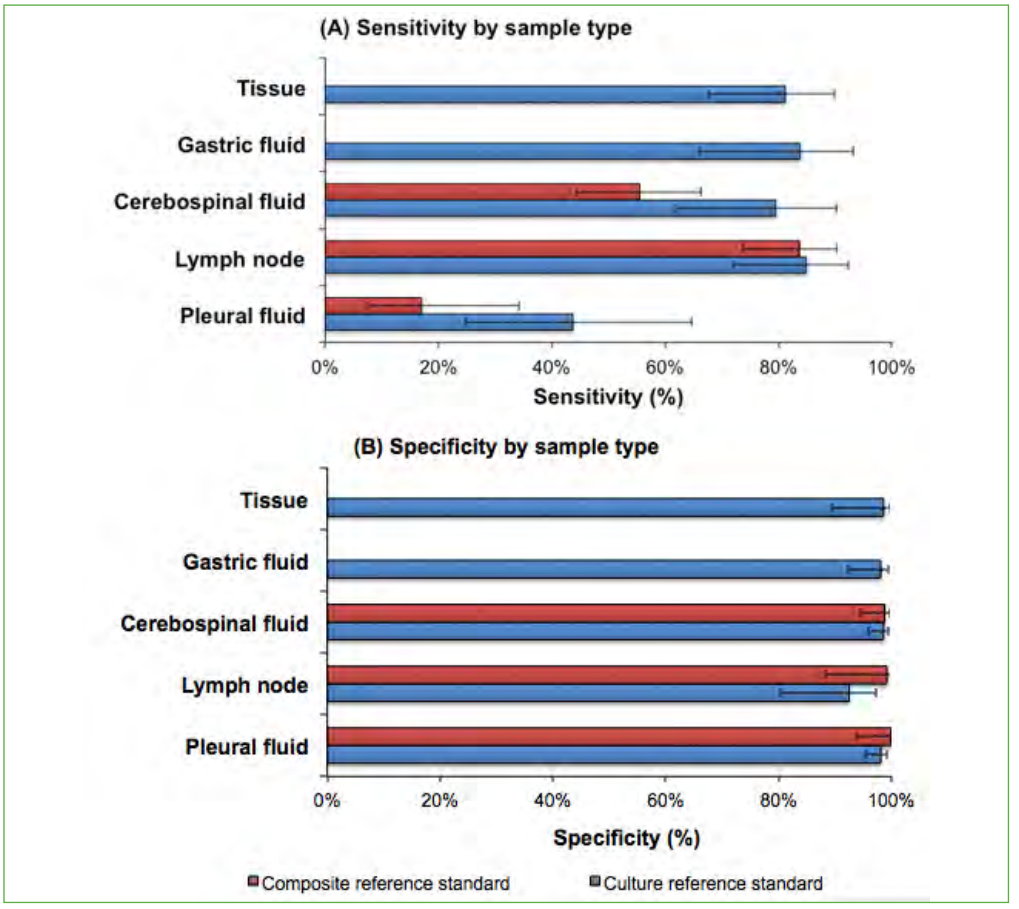
estimates to be calculated. However, given the mechanism of detection used with Xpert MTB/RIF it is unlikely that accuracy will differ from that estimated for respiratory samples.

The results demonstrated that for detecting TB in nonrespiratory samples:

- the sensitivity of Xpert MTB/RIF varied widely across different sample types if the smear result was negative;
- Xpert MTB/RIF had good sensitivity when used to test smear-positive samples;
- Xpert MTB/RIF had good sensitivity when compared with using culture to test lymph node tissues or aspirates, gastric fluid, CSF and other tissue samples;
- Xpert MTB/RIF had good sensitivity when compared with an author-defined CRS for testing lymph node tissues or aspirates;
- Xpert MTB/RIF had poor sensitivity for testing pleural fluid;
- the proportion of indeterminate results was low; and
- there was substantial heterogeneity even within subgroups classified by sample type, and therefore the pooled estimates must be interpreted with caution.

*Figure 16* shows the summary estimates for different sample types.

Figure 16. Summary estimates for sensitivity (A) and specificity (B) of Xpert MTB/RIF in detecting extrapulmonary TB in adults and children, by sample type



### *4.2.6. Strengths and limitations of the evidence base*

The strengths of the review include the use of a standard protocol to assess the studies, strict inclusion criteria, standardized methods of data extraction, independent reviewers, a bivariate model for meta-analysis and prespecified subgroups to account for heterogeneity.

Compiling this data set involved comprehensive searching and correspondence with experts in the field and the test's manufacturer to identify additional published and unpublished studies, as well as repeated correspondence with the authors of studies to obtain additional data and information that was missing from the literature. The search strategy included studies published in all languages. The majority of studies selected participants consecutively, and the results from the reference standard were interpreted without knowledge of the results from the Xpert MTB/RIF test. Xpert MTB/RIF results are generated automatically, and do not require subjective interpretation.

However, the review also had several limitations. The meta-analysis was limited by the small number of studies used to assess Xpert MTB/RIF for different types of samples; this was particularly true for those studies using a CRS. Also, low event rates (that is, the number of confirmed TB cases) limited the precision of the estimates of sensitivity. Furthermore, the way in which samples were processed varied greatly across studies and within studies.

The sensitivity and specificity estimates in the meta-analysis might be overly optimistic for the following reasons: ( 1) some studies lacked a representative spectrum of patients, which could result in exaggerated estimates of the accuracy of the test, and (2) all of the studies were performed in tertiary care centres or reference laboratories, where performance characteristics might be better and where patients might present later in their disease process.

### *4.2.7. GRADE evaluations and recommendations*

GRADE evidence profiles are provided in *Tables 15 to 20*. The GRADE evaluation supports the use of Xpert MTB/RIF to diagnose extrapulmonary TB and rifampicin resistance in adults and children. The Expert Group therefore concluded that:

- Xpert MTB/RIF should be used in preference to conventional microscopy and culture as the initial diagnostic test for CSF from patients suspected of having TB meningitis (strong recommendation given the urgency for rapid diagnosis, very low-quality evidence);
  *Note: The Expert Group noted that a negative result from CSF tested using Xpert MTB/RIF should be followed by other tests. The Expert Group also noted that concentration methods should be used to enhance yield when a sufficient volume of CSF is available. These recommendations apply to samples from both children and adults.*
- Xpert MTB/RIF may be used as a replacement test for usual practice, including conventional microscopy and culture, for testing lymph nodes and tissues from patients suspected of having extrapulmonary TB (conditional recommendation, very low-quality evidence);
  *Note: The Expert Group noted that in these cases a negative result from Xpert MTB/RIF should be followed by other tests. The Expert Group also noted that processing methods for samples from lymph nodes and tissues need to be standardized to optimize yield. These recommendations apply to samples from both adults and children.*
- Xpert MTB/RIF should not be used in the diagnostic work-up of patients suspected of having pleural TB (conditional recommendation, very low-quality evidence).
  *Note: The Expert Group noted that pleural fluid is a suboptimal sample for the diagnosis of pleural TB in general, and that a pleural biopsy is the preferred sample for bacteriological confirmation. These recommendations apply to samples from both adults and children.*

Table 15. Accuracy of Xpert MTB/RIF in detecting TB in lymph node fluid and tissue (A. Evidence profile, B. Summary of findings)

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for TB detection in lymph node fluid and tissue, where Xpert MTB/RIF is used as a replacement test for usual practice?

**A. Evidence profile**

| Outcome[a] | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence |
|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | |
| True positives (individuals with TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ |
| False negatives (individuals without TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ |
| False positives (individuals incorrectly classified as having TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ |
| False negatives (patients incorrectly classified as not having TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ |

a For each outcome, the rating of the quality of evidence started as high when the study designs were randomized controlled trials or high-quality observational studies (such as cross-sectional studies with diagnostic uncertainty and direct comparison of the results of the index test with a reference standard), or low when there were case–control studies. The evidence was downgraded one point when a serious issue was identified; it was downgraded two points when a very serious issue was identified in any of the other five factors that may decrease the quality of evidence: limitations, indirectness, inconsistency, imprecision or publication bias.
b The QUADAS-2 tool was used to assess the risk of bias. The majority of studies enrolled patients consecutively, and assessed the result from the reference standard blinded to the result from Xpert MTB/RIF. The Xpert MTB/RIF result is automated and was considered blinded in all studies. However, the variable use of the composite reference standard across studies was a concern for its possibility of introducing bias. The evidence was downgraded one point.
c The majority of studies used Xpert MTB/RIF in tertiary care centres or reference laboratories. Because obtaining a sample requires an invasive procedure, the test will likely be performed at higher levels of care than are feasible for diagnosing pulmonary TB. Thus, the study populations are probably representative of the population that will receive the test. The evidence was not downgraded.
d Unexplained heterogeneity in the findings might originate from differences in sample processing, the condition of samples, and differences in study populations (for example, varying prevalences of TB or HIV). The evidence was downgraded one point.
e Confidence intervals are wide, likely due in part to unexplained heterogeneity as described in footnote d and in part due to verification bias because of the use of an imperfect reference standard. The evidence was downgraded one point.
f Unpublished studies were included. Data did not allow for formal assessment of publication bias using methods such as funnel plots or regression tests because such techniques have not been found to be helpful in studies of diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

**B. Summary of findings**

Reference standard: Culture
Number of studies (number of samples): 14 (849); 11 studies had more than 10 samples
Pooled sensitivity: 85% (95% CI, 72–92); pooled specificity: 93% (95% CI, 80–97)
Reference standard: Composite reference standard
Number of studies (number of samples): 5 (409)
Pooled sensitivity: 84% (95% CI, 74–90%); pooled specificity: 99% (95% CI, 88–100%)

| Outcome | Number of results/1000 individuals tested (95% CrI) | | | | | | Quality of evidence |
|---|---|---|---|---|---|---|---|
| | Prevalence 2.5% | | Prevalence 5% | | Prevalence 10% | | |
| | Culture | CRS | Culture | CRS | Culture | CRS | |
| True positives (individuals with TB) | 21 (18–23) | 21 (19–23) | 43 (36–46) | 42 (37–45) | 85 (72–92) | 84 (74–90) | Very low ⊕○○○ |
| True negatives (individuals without TB) | 907 (780–946) | 965 (858–975) | 884 (760–922) | 941 (836–950) | 837 (720–873) | 891 (792–900) | Very low ⊕○○○ |
| False positives (individuals incorrectly classified as having TB) | 68 (29–195) | 10 (0–117) | 67 (29–190) | 10 (0–114) | 63 (27–180) | 9 (0–108) | Very low ⊕○○○ |
| False negatives (individuals incorrectly classified as not having TB) | 4 (2–7) | 3 (4–7) | 8 (4–14) | 8 (5–13) | 15 (8–28) | 16 (10–26) | Very low ⊕○○○ |

CI, confidence interval; CRS, composite reference standard.

Table 16. Accuracy of Xpert MTB/RIF in detecting TB in pleural fluid (A. Evidence profile, B. Summary of findings)

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for TB detection in pleural fluid, where Xpert MTB/RIF is used as a replacement test for usual practice?

**A. Evidence profile**

| Outcome[a] | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence |
|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | |
| True positives (individuals with TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ |
| False negatives (individuals without TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | None | Undetected[f] | Low ⊕⊕○○ |
| False positives (individuals incorrectly classified as having TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | None | Undetected[f] | Low ⊕⊕○○ |
| False negatives (patients incorrectly classified as not having TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ |

a For each outcome, the rating of the quality of evidence started as high when the study designs were randomized controlled trials or high-quality observational studies (such as cross-sectional studies with diagnostic uncertainty and direct comparison of the results of the index test with a reference standard) , or low when there were case–control studies. The evidence was downgraded one point when a serious issue was identified; it was downgraded two points when a very serious issue was identified in any of the other five factors that may decrease the quality of evidence: limitations, indirectness, inconsistency, imprecision or publication bias.
b The QUADAS-2 tool was used to assess the risk of bias. The majority of studies enrolled patients consecutively, and assessed the result from the reference standard blinded to the result from Xpert MTB/RIF. The Xpert MTB/RIF result is automated and was considered blinded in all studies. However, the variable use of the composite reference standard across studies was a concern for its possibility of introducing bias. The evidence was downgraded one point.
c The majority of studies used Xpert MTB/RIF in tertiary care centres or reference laboratories. Because obtaining a sample requires an invasive procedure, the test will likely be performed at higher levels of care than are feasible for diagnosing pulmonary TB. Thus, the study populations are probably representative of the population that will receive the test. The evidence was not downgraded.
d Unexplained heterogeneity in the findings might originate from differences in sample processing, the condition of the sample, and differences in study populations (for example, varying prevalences of TB or HIV). The evidence was downgraded one point.
e Confidence intervals are wide, likely due in part to unexplained heterogeneity as described in footnote d and in part due to verification bias because of the use of an imperfect reference standard. The evidence was downgraded one point.
f Unpublished studies were included. Data did not allow for formal assessment of publication bias using methods such as funnel plots or regression tests because such techniques have not been found to be helpful in studies of diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

**B. Summary of findings**

Reference standard: Culture
Number of studies (number of samples): 17 (1384), 16 studies had more than 10 samples
Pooled sensitivity: 44% (95% CI, 25–65%); pooled specificity: 98% (95% CI, 95–99%)
Reference standard: Composite reference standard
Number of studies (number of samples): 7 (698)
Pooled sensitivity: 17% (95% CI, 8–34%); pooled specificity: 100% (95% CI, 94–100%)

| Outcome | Number of results/1000 individuals tested (95% CrI) | | | | | | Quality of evidence |
|---|---|---|---|---|---|---|---|
| | Prevalence 2.5% | | Prevalence 5% | | Prevalence 10% | | |
| | Culture | CRS | Culture | CRS | Culture | CRS | |
| True positives (individuals with TB) | 11 (6–16) | 4 (2–9) | 22 (13–33) | 9 (4–17) | 44 (25–65) | 17 (8–34) | Very low ⊕○○○ |
| True negatives (individuals without TB) | 956 (926–965) | 975 (917–975) | 931 (903–941) | 950 (893–950) | 882 (855–891) | 900 (846–900) | Low ⊕⊕○○ |
| False positives (individuals incorrectly classified as having TB) | 20 (10–49) | 0 (0–59) | 19 (10–48) | 0 (0–57) | 18 (9–45) | 0 (0–54) | Low ⊕⊕○○ |
| False negatives (individuals incorrectly classified as not having TB) | 14 (9–19) | 21 (17–23) | 28 (18–38) | 42 (33–46) | 56 (35–75) | 83 (66–92) | Very low ⊕○○○ |

CI, confidence interval; CRS, composite reference standard.

Table 17. Accuracy of Xpert MTB/RIF in detecting TB in cerebrospinal fluid (A. Evidence profile, B. Summary of findings)

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for TB detection in cerebrospinal fluid, where Xpert MTB/RIF is used as a replacement test for usual practice?

**A. Evidence profile**

| Outcome[a] | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence |
|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | |
| True positives (individuals with TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ |
| False negatives (individuals without TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | None | Undetected[f] | Low ⊕⊕○○ |
| False positives (individuals incorrectly classified as having TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | None | Undetected[f] | Low ⊕⊕○○ |
| False negatives (patients incorrectly classified as not having TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ |

a For each outcome, the rating of the quality of evidence started as high when the study designs were randomized controlled trials or high-quality observational studies (such as cross-sectional studies with diagnostic uncertainty and direct comparison of the results of the index test with a reference standard), or low when there were case–control studies. The evidence was downgraded one point when a serious issue was identified; it was downgraded two points when a very serious issue was identified in any of the other five factors that may decrease the quality of evidence: limitations, indirectness, inconsistency, imprecision or publication bias.
b The QUADAS-2 tool was used to assess the risk of bias. The majority of studies enrolled patients consecutively, and assessed results from the reference standard blinded to the result from Xpert MTB/RIF. The Xpert MTB/RIF result is automated and was considered blinded in all studies. However, the variable use of the composite reference standard across studies was a concern for its possibility of introducing bias. The evidence was downgraded one point.
c The majority of studies used Xpert MTB/RIF in tertiary care centres or reference laboratories. Because obtaining a sample requires an invasive procedure, the test will likely be performed at higher levels of care than are feasible for diagnosing pulmonary TB. Thus, the study populations are probably representative of the population that will receive the test. The evidence was not downgraded.
d Unexplained heterogeneity in the findings might originate from differences in sample processing, the condition of the samples, and differences in study populations (for example, varying prevalences of TB or HIV). The evidence was downgraded one point.
e Confidence intervals are wide, likely due in part to unexplained heterogeneity as described in footnote d and in part due to verification bias because of the use of an imperfect reference standard. The evidence was downgraded one point.
f Unpublished studies were included. Data did not allow for formal assessment of publication bias using methods such as funnel plots or regression tests because such techniques have not been found to be helpful in studies of diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

**B. Summary of findings**

Reference standard: Culture
Number of studies (number of samples): 16 (709); 13 studies had more than 10 samples
Pooled sensitivity: 80% (95% CI, 62–90%); pooled specificity: 99% (95% CI, 96–100%)
Reference standard: Composite reference standard
Number of studies (number of samples): 7 (698)
Pooled sensitivity: 56% (95% CI, 44–66%); pooled specificity: 99% (95% CI, 95–100%)

| Outcome | Number of results/1000 individuals tested (95% CrI) | | | | | | Quality of evidence |
|---|---|---|---|---|---|---|---|
| | Prevalence 2.5% | | Prevalence 5% | | Prevalence 10% | | |
| | Culture | CRS | Culture | CRS | Culture | CRS | |
| True positives (individuals with TB) | 20 (16–23) | 14 (11–17) | 40 (31–45) | 28 (22–33) | 80 (62–90) | 56 (44–66) | Very low ⊕○○○ |
| True negatives (individuals without TB) | 965 (936–975) | 965 (926–975) | 941 (912–950) | 941 (903–950) | 891 (864–900) | 891 (855–900) | Low ⊕⊕○○ |
| False positives (individuals incorrectly classified as having TB) | 10 (0–39) | 10 (0–49) | 10 (0–38) | 10 (0–48) | 9 (0-36) | 9 (0–45) | Low ⊕⊕○○ |
| False negatives (individuals incorrectly classified as not having TB) | 5 (3-10) | 9 (11-14) | 10 (5-19) | 22 (17-28) | 20 (10-38) | 56 (44-66) | Very low ⊕○○○ |

CI, confidence interval; CRS, composite reference standard.

Table 18. GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting TB in gastric fluid

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for TB detection in gastric fluid, where Xpert MTB/RIF is used as a replacement test for usual practice?
Reference standard: Culture
Number of studies (number of samples): 12 (1258 samples); 8 studies had more than 10 samples
Pooled sensitivity: 84% (95% CI, 66– 93%); pooled specificity: 98% (95% CI, 92–100%)

| Outcome[a] | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence | Number of results/1000 individuals tested (95% CrI) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | | Prevalence 2.5% | Prevalence 5% | Prevalence 10% |
| True positives (individuals with TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ | 21 (17-23) | 42 (33-47) | 84 (66-93) |
| False negatives (individuals without TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ | 956 (897-975) | 931 (874-950) | 882 (828-900) |
| False positives (individuals incorrectly classified as having TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ | 20 (0-78) | 19 (0-76) | 18 (0-72) |
| False negatives (individuals incorrectly classified as not having TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ | 4 (2-9) | 8 (4-17) | 16 (7-34) |

a For each outcome, the rating of the quality of evidence started as high when the study designs were randomized controlled trials or high-quality observational studies (such as cross-sectional studies with diagnostic uncertainty and direct comparison of index test results with a reference standard), or low when there were case–control studies. The evidence was downgraded one point when a serious issue was identified; it was downgraded two points when a very serious issue was identified in any of the other five factors that may decrease the quality of evidence: limitations, indirectness, inconsistency, imprecision or publication bias.
b The QUADAS-2 tool was used to assess the risk of bias. The majority of studies enrolled patients consecutively, and assessed results from the reference standard blinded to the results from Xpert MTB/RIF. The Xpert MTB/RIF result is automated and was considered blinded in all studies. However, the variable use of the composite reference standard across studies was a concern for its possibility of introducing bias. The evidence was downgraded one point.
c The majority of studies used Xpert MTB/RIF in tertiary care centres or reference laboratories. Because obtaining a sample requires an invasive procedure, the test will likely be performed at higher levels of care than are feasible for diagnosing pulmonary TB. Thus, the study populations are probably representative of the population that will receive the test. The evidence was not downgraded.
d Unexplained heterogeneity in the findings might originate from differences in sample processing, the condition of the samples, and differences in study populations (for example, varying prevalences of TB or HIV). The evidence was downgraded one point.
e Confidence intervals are wide, likely due in part to unexplained heterogeneity as described in footnote d and in part due to verification bias because of the use of an imperfect reference standard. The evidence was downgraded one point.
f Unpublished studies were included. Data did not allow for formal assessment of publication bias using methods such as funnel plots or regression tests because such techniques have not been found to be helpful in studies of diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

Table 19. GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting TB in tissue samples

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for TB detection in tissue samples, where Xpert MTB/RIF is used as a replacement test for usual practice?
Reference standard: Culture
Number of studies (number of samples): 12 (699); 10 studies had more than 10 samples
Pooled sensitivity: 81% (95% CI, 68– 90%); pooled specificity: 98% (95% CI, 87–100%)

| Outcome[a] | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence | Number of results/1000 individuals tested (95% CrI) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | | Prevalence 2.5% | Prevalence 5% | Prevalence 10% |
| True positives (individuals with TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ | 20 (17-23) | 41 (34-45) | 81 (68-90) |
| False negatives (individuals without TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ | 956 (848-975) | 931 (827-950) | 882 (783-900) |
| False positives (individuals incorrectly classified as having TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ | 20 (0-127) | 19 (0-124) | 18 (0-117) |
| False negatives (individuals incorrectly classified as not having TB) | Majority cross-sectional | Serious (–1)[b] | None[c] | Serious (–1)[d] | Serious (–1)[e] | Undetected[f] | Very low ⊕○○○ | 5 (3-8) | 10 (5-16) | 19 (10-32) |

CI, confidence interval.

a For each outcome, the rating of the quality of evidence started as high when the study designs were randomized controlled trials or high-quality observational studies (such as cross-sectional studies with diagnostic uncertainty and direct comparison of index test results with a reference standard), or low when there were case–control studies. The evidence was downgraded one point when a serious issue was identified; it was downgraded two points when a very serious issue was identified in any of the other five factors that may decrease the quality of evidence: limitations, indirectness, inconsistency, imprecision or publication bias.

b The QUADAS-2 tool was used to assess the risk of bias. The majority of studies enrolled patients consecutively, and assessed the result from the reference standard blinded to the result from Xpert MTB/RIF. The Xpert MTB/RIF result is automated and was considered blinded in all studies. However, the variable use of the composite reference standard across studies was a concern for its possibility of introducing bias. The evidence was downgraded one point.

c The majority of studies used Xpert MTB/RIF in tertiary care centres or reference laboratories. Because obtaining a sample requires an invasive procedure, the test will likely be performed at higher levels of care than are feasible for diagnosing pulmonary TB. Thus, the study populations are probably representative of the population that will receive the test. The evidence was not downgraded.

d Unexplained heterogeneity in the findings might originate from differences in sample processing, the condition of the samples, and differences in study populations (for example, varying prevalences of TB or HIV). The evidence was downgraded one point.

e Confidence intervals are wide, likely due in part to unexplained heterogeneity as described in footnote d and in part due to verification bias because of the use of an imperfect reference standard. The evidence was downgraded one point.

f Unpublished studies were included. Data did not allow for formal assessment of publication bias using methods such as funnel plots or regression tests because such techniques have not been found to be helpful in studies of diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

Table 20. GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting rifampicin resistance in nonrespiratory specimens

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for rifampicin-resistance detection in nonrespiratory specimens, where Xpert MTB/RIF is used as an initial test replacing phenotypic culture-based drug-susceptibility testing?
Number of studies (number of samples): 13 (566)

| Outcome[a,b] | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence |
|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | |
| True positives (individuals with TB) | Majority cross-sectional | None[c] | Serious (−1)[d] | Serious (−1)[e] | Serious (−1)[f] | Undetected[g] | Very low ⊕○○○ |
| False negatives (individuals without TB) | Majority cross-sectional | None[c] | Serious (−1)[d] | Serious (−1)[e] | Serious (−1)[f] | Undetected[g] | Low ⊕⊕○○ |
| False positives (individuals incorrectly classified as having TB) | Majority cross-sectional | None[c] | Serious (−1)[d] | Serious (−1)[e] | Serious (−1)[f] | Undetected[g] | Low ⊕⊕○○ |
| False negatives (patients incorrectly classified as not having TB) | Majority cross-sectional | None[c] | Serious (−1)[d] | Serious (−1)[e] | Serious (−1)[f] | Undetected[g] | Very low ⊕○○○ |

a Given the limited amount of data, we did not calculate summary estimates. Six studies reported a sensitivity of 100% (34 rifampicin-resistant cases); 1 study had 50% sensitivity (2 rifampicin-resistant cases); 1 study had 0% sensitivity (1 rifampicin-resistant case).
b For each outcome, the rating of the quality of evidence started as high when the study designs were randomized controlled trials or high-quality observational studies (such as cross-sectional studies with diagnostic uncertainty and direct comparison of index test results with a reference standard), or low when there were case–control studies. The evidence was downgraded one point when a serious issue was identified; it was downgraded two points when a very serious issue was identified in any of the other five factors that may decrease the quality of evidence: limitations, indirectness, inconsistency, imprecision or publication bias.
c The QUADAS-2 tool was used to assess the risk of bias. The majority of studies enrolled patients consecutively, and assessed the result from the reference standard blinded to the result from Xpert MTB/RIF. The Xpert MTB/RIF result is automated and was considered blinded in all studies. The reference standard of phenotypic drug-susceptibility testing was used in all studies. The evidence was not downgraded. However, there are concerns that phenotypic drug-susceptibility testing is not a perfect reference standard for detecting rifampicin resistance since sequencing data have shown that Xpert MTB/RIF might correctly identify resistance that is not identified by phenotypic drug-susceptibility testing. Across all studies included here, six false-positive results from Xpert MTB/RIF testing were identified. Of those, five were tested with sequencing, and four out of five were found to have a mutation in codon 533.
d The majority of studies used Xpert MTB/RIF in tertiary care centres or reference laboratories. Because obtaining a sample requires an invasive procedure, the test will likely be performed at higher levels of care than are feasible for diagnosing pulmonary TB. Thus, the study populations are probably representative of the population that will receive the test. The evidence was not downgraded.
e Unexplained heterogeneity in the findings might originate from differences in the types of specimen tested and in sample processing. The evidence was downgraded one point.
f There were insufficient data to obtain a summary estimate for rifampicin resistance. The results from individual studies varied widely; therefore, the evidence was downgraded one point.
g Unpublished studies were not included in the assessment of using Xpert MTB/RIF to detect rifampicin resistance because data collection on drug-susceptibility testing was not complete for some of the studies. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

#### *4.2.8. Further research needs*

##### *4.2.8.1. Optimizing sample processing*

There was evidence that concentration methods enhanced the sensitivity of Xpert MTB/RIF in detecting TB in CSF specimens, although a detailed investigation of optimized sample processing was not possible within this review. As a priority, the Expert Group highlighted the need to develop standardized protocols for processing samples that can be used in subsequent studies assessing Xpert MTB/RIF for nonrespiratory specimens.

The Expert Group recommended that processing protocols focus not only on one the type of sample but also assess differences in performance when processing is modified in the individual steps (such as homogenization, concentration, decontamination) and when changes are made in the ratio of sample reagent to sample. Optimization of sample preparation, DNA extraction and purification might further enhance accuracy. The specific steps necessary to optimize processing are likely to vary by sample type.

##### *4.2.8.2. Sample types for testing*

All of the published studies examined samples obtained from routine-care services in hospitals or reference laboratories. Further research is needed to assess differences in the performance of the Xpert MTB/RIF test after samples have undergone a freeze–thaw cycle, since implementing centralized testing and specimen transport would require that specimens be preserved by freezing.

##### *4.2.8.3. Reference standards for extrapulmonary TB*

Additional research would be beneficial to elucidate the performance of Xpert MTB/RIF in detecting extrapulmonary TB in nonrespiratory samples from subgroups at high risk of this type of TB (for example, HIV-positive people). Efforts should also be undertaken to optimize the reference standard so that the accuracy of Xpert MTB/RIF and other new diagnostics can be evaluated for use in diagnosing extrapulmonary TB. An optimized and standardized case definition for the diagnosis of extrapulmonary TB would facilitate comparisons across studies. A case definition has already been published for TB meningitis, however there are no definitions for pleural TB, musculoskeletal TB or TB lymphadenitis.

### 4.3. Using Xpert MTB/RIF to diagnose pulmonary TB, peripheral lymph node TB, TB meningitis and rifampicin resistance in children

#### *4.3.1. Characteristics of the studies*

An electronic literature search was performed initially on 24 January 2013. In total, 39 articles were identified. Of these, 26 were excluded based on a review of their title and abstract. Of the remaining 13 articles that underwent a full-text review, 10 were evaluated in this review. An additional 2 published studies were included that had been identified during a final electronic search conducted on 3 April 2013. An additional four unpublished studies were included that had been identified by querying networks of people working in childhood TB and contacting authors as previously described. In total, 16 studies were included (*Annex 6*).

Studies that assessed the accuracy of Xpert MTB/RIF in diagnosing pulmonary TB, peripheral lymph node TB, TB meningitis and rifampicin resistance in children aged 0–15 years with presumptive TB were included in the review. All published articles, articles that were in press, and unpublished studies were shared confidentially with the Expert Group members with the authors' agreement.

Cross-sectional studies, cohort studies and randomized controlled trials were included if they compared Xpert MTB/RIF with an acceptable reference standard (see section 4.3.3.1). Case–control studies, case reports, as well as studies presented only as abstracts were excluded.

Studies had to include children aged 0–15 years who were suspected of having TB, but the studies

could not use predefined criteria for the diagnosis of presumptive TB. Studies that recruited children from both inpatient and outpatient settings were considered, as well as studies performed at any level of the health-care system or in research laboratories.

Studies that included and excluded HIV-infected populations as well as children with other comorbidities, such as malnutrition, in order to improve generalizability were also included. Authors of studies that included both children and adults were contacted if the paediatric data were not presented separately. Studies that used respiratory samples, rifampicin-resistance testing of respiratory samples, and nonrespiratory samples (for example, CSF or samples from lymph node biopsies or aspirates) were included.

The reference standard for pulmonary TB, peripheral lymph node TB and TB meningitis was TB confirmed by at least one positive culture on solid media or a commercial liquid culture system, such as the BACTEC MGIT 960 system. The reference for rifampicin resistance was WHO's recommended conventional phenotypic DST on solid media or liquid media[19], or molecular line probe assays. Recognizing the limitations of mycobacterial culture in children (culture sensitivity is approximately 30% to 60%), a second reference standard, clinical TB, was applied only in children who were culture negative. Children were categorized as positive for the clinical TB reference standard if they were culture negative and had started anti-TB therapy based on a clinical diagnosis of TB.

### *4.3.2. Quality of the studies*

The quality of the studies included in the review was assessed using the QUADAS-2 tool. QUADAS-2 consists of four domains: patient selection, index test, reference standard, and flow and timing.

Within the patient-selection domain, 75% of studies were considered to have a low risk of bias since the studies recruited children consecutively and avoided inappropriate exclusions. The remaining studies were considered to have a high risk of bias because patient recruitment was either by convenience sampling or the methods were unclear. There was high concern for applicability in this domain because (1) most studies were performed among inpatients and (2) all studies were performed at higher levels of care (such as university hospitals or research laboratories), or were laboratory-based studies that had unclear criteria for patient selection and provided limited clinical information (three studies).

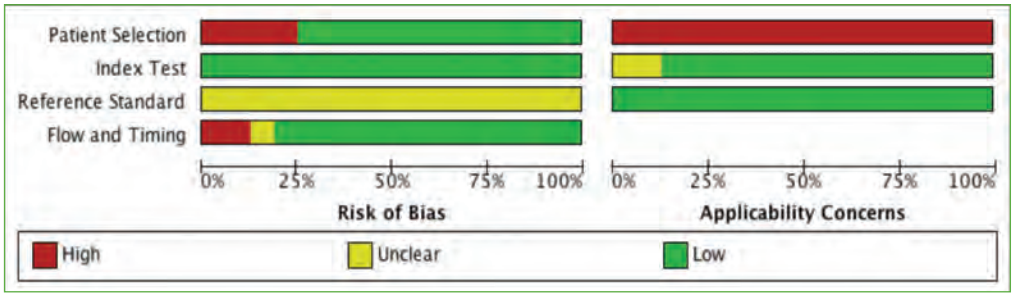
The risk of bias for the index test was judged to be of low concern for all studies since Xpert MTB/RIF is fully automated. The applicability of the index test was rated as being of low concern for the majority of studies. Two studies were rated as unclear regarding the applicability of the index test since the preparation of specimens was not clearly described.

The risk of bias for the reference standard was rated as unclear for all studies because mycobacterial culture as well as clinical case definitions are considered to be imperfect standards for diagnosing childhood TB. However, the applicability of the reference standard was considered to be of low concern for all studies.

The risk of bias for the flow and timing domain was considered to be high in two studies. In the first study, 99 children were excluded from the analysis because they were lost to follow up. In the second study, children with a clinical diagnosis of TB and negative culture results were excluded from participation. In contrast, culture-positive children with TB were included. *Figure 17* shows the concerns regarding the risks of bias and applicability for the studies included in the review.

19 *Policy guidance on drug-susceptibility testing (DST) of second-line antituberculosis drugs*. Geneva, World Health Organization, 2008 (WHO/HTM/TB/2008.392).

Figure 17. Risks of bias (%) and judgements about applicability (%) for each domain of the QUADAS-2 tool in 16 studies assessing the use of Xpert MTB/RIF to diagnose pulmonary TB, peripheral lymph node TB, TB meningitis and rifampicin resistance in children




### *4.3.3. Using Xpert MTB/RIF as the initial test to detect pulmonary TB*

Pulmonary TB was evaluated in 13 studies that included 2603 participants. Studies either collected the same specimen type from all children or different types of specimens from different subgroups of children (for example, samples of expectorated sputum were collected from older children; samples of induced sputum or gastric lavage or aspirate were collected from younger children). In three studies different types of specimens were collected from each child. As a result, a total of 3347 specimens were assessed (median number of specimens per study, 69; range, 3–788): expectorated sputum (4 studies, 270 children), induced sputum (7 studies, 1279 children), nasopharyngeal aspirate (1 study, 474 children), gastric lavage or aspirate (6 studies, 1324 children).

The individual sensitivities and specificities of Xpert MTB/RIF for each type of specimen compared against culture as a reference standard are expressed as Forest plots by study in *Figure 18*.

Sensitivities varied from 55% to 90% for expectorated sputum, 40% to 100% for induced sputum, and 40% to 100% for gastric lavage or aspirate. Confidence intervals overlapped for each specimen type, suggesting that no specimen type was superior. Specificities for all studies and specimen types ranged from 93% to 100%.

One study examined the yield of nasopharyngeal specimens. The sensitivity of Xpert MTB/RIF in these specimens was 44% (95% CI, 33–55%). In the same group of children, the sensitivity in samples from induced sputum was 60% (95% CI, 49–70).

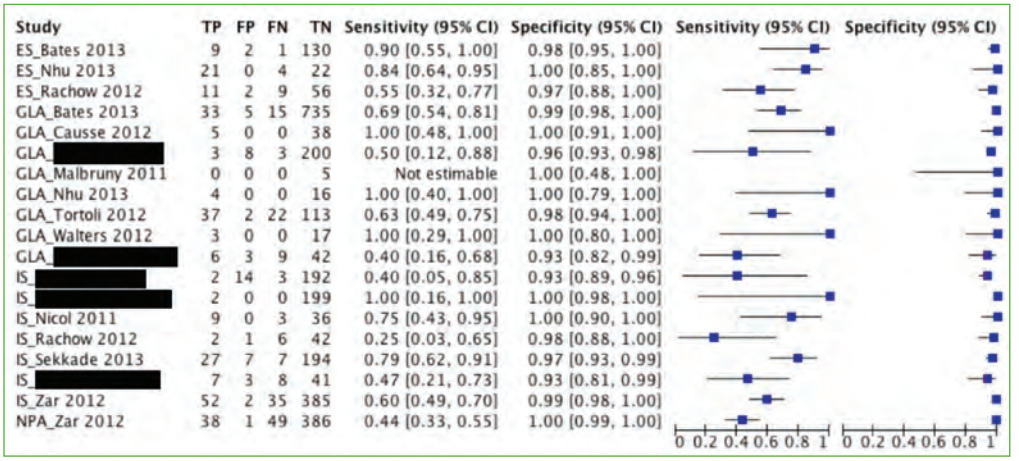
The sensitivity and specificity of Xpert MTB/RIF were estimated using mycobacterial culture as the reference standard as well as clinical TB.

#### *4.3.3.1 Differences in the reference standards used*

In all the studies included in the review, 13.2% of children had culture-confirmed TB. The proportion of children with culture-confirmed TB varied by study and specimen type (range, 0–54.2%). In the majority of studies (9/13; 69%), multiple cultures were performed on samples from single participants. Hence, the definition of culture positive was based on the presence of at least one positive culture result out of as many as six cultures performed. The average bacteriological yield in studies using multiple cultures was increased compared with the group of four studies that defined a participant as culture positive based only on one culture result. Studies also used different culture techniques, but the impact of this potential source of bias was not evaluated.

Children were categorized as positive using the clinical TB reference standard if they were culture negative and had started anti-TB therapy

Figure 18. Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting pulmonary TB, peripheral lymph node TB and TB meningitis in children compared against culture as a reference standard, by study and specimen type[a]




TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval; ES, expectorated sputum; GLA, gastric lavage or aspirate; IS, induced sputum; NPA, nasopharyngeal aspirate.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

used and the clinical definitions expected in the studies. Children assigned to the group "clinical not TB" (that is, negative according to the clinical TB reference standard) either (1) did not have another diagnosis assigned or (2) did not start anti-TB treatment but nonetheless improved, or their condition did not worsen after at least 1 month of follow-up after enrolment.

#### *4.3.3.2 Accuracy of Xpert MTB/RIF compared with culture*

The pooled sensitivity of Xpert MTB/RIF compared against culture was 66% for samples of expectorated or induced sputum (95% CrI, 52–77%) and 66% for samples of gastric lavage or aspirate (95% CrI, 51–81%). The width of the confidence intervals indicated a high level of heterogeneity among studies. The specificity values for Xpert MTB/RIF compared against culture as the reference standard were all at least 98%, with narrow confidence intervals.

#### *4.3.3.3 Accuracy of Xpert MTB/RIF compared with clinical TB as a reference standard*

The sensitivity of Xpert MTB/RIF in culture-negative samples from paediatric patients compared against clinical TB as the reference standard was 4% (95% CrI, 1–12%) for expectorated or induced sputum and 15% (95% CrI, 5-31%) for gastric lavage or aspirate, with all confidence intervals being wide and therefore indicating a high level of heterogeneity. It is likely that the apparently poor performance of Xpert MTB/RIF was the result of a clinical TB reference standard that lacked specificity. The specificity values of Xpert MTB/RIF compared against the clinical TB reference standard were at least 99%, with narrow confidence intervals.

The estimated sensitivities and specificities for Xpert MTB/RIF compared against culture as the reference standard (in published and unpublished studies) as well as against the clinical TB reference

standard are given in *Table 21*. The exclusion of unpublished studies increased sensitivity to 69% (95% CrI, 55–81%) for expectorated sputum and induced sputum, and to 75% (95% CrI, 59-90%) for gastric aspiration or lavage, but the confidence intervals remained wide and overlapped between the two estimates, suggesting that heterogeneity was retained.

Table 21. **Meta-analysis of the estimated sensitivity and specificity of Xpert MTB/RIF in diagnosing pulmonary TB, peripheral lymph node TB and TB meningitis in children compared against culture as a reference standard (in published and unpublished studies) as well as against the clinical TB reference standard**

| Comparison | Specimen type (No. of studies, No. of children) | Median (%) pooled sensitivity (pooled 95% CrI) | Median (%) pooled specificity (pooled 95% CrI) |
|---|---|---|---|
| Xpert MTB/RIF compared against culture as a reference standard (published and unpublished studies) | Expectorated sputum or induced sputum (10, 1546)[a] | 66 (5277) | 98 (96–99) |
| | Gastric lavage or aspiration (7, 1319)[b] | 66 (51–81) | 98 (96–99) |
| Xpert MTB/RIF compared against culture as a reference standard (published studies only) | Expectorated sputum or induced sputum (7, 1075) | 69 (55–81) | 98 (97–99) |
| | Gastric lavage or aspiration (5, 1045) | 75 (59–90) | 99 (97–100) |
| Xpert MTB/RIF compared against clinical TB as a reference standard (published and unpublished studies) | Expectorated sputum or induced sputum (8, 995)[c] | 4 (01–12) | 100 (99–100) |
| | Gastric lavage or aspiration (3, 269)[d] | 15 (05–31) | 99 (96–100) |

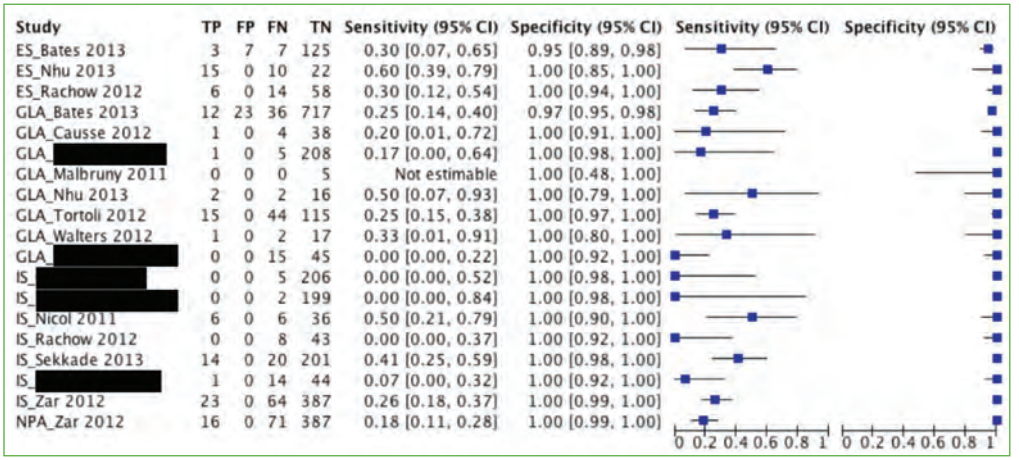
CrI, credible interval; the CrI is the Bayesian equivalent of the confidence interval.

a The studies included were Bates 2013, Chisti unpublished, LaCourse unpublished, Nhu 2013, Nicol 2011, Rachow 2012 (ES and IS cohorts included as separate studies), Sekadde 2013, Walters unpublished, Zar 2012
b Studies included Bates 2013, Causse 2011, Chisti unpublished, Nhu 2013, Tortoli 2012, Walters 2012, Walters unpublished
c Studies included Chisti unpublished, LaCourse unpublished, Nhu 2013, Nicol 2011, Rachow 2012 (ES and IS cohorts included as separate studies), Walters unpublished, Zar 2012
d Studies included: Chisti unpublished, Nhu 2013, Walters unpublished

### *4.3.4. Xpert MTB/RIF compared with smear microscopy*

The diagnostic accuracy of smear microscopy was calculated against culture as a reference standard for the same studies and specimen types that were used to calculate the accuracy of Xpert MTB/RIF (see section 4.3.3.2). The forest plot in *Figure 19* shows the estimates for individual studies. Sensitivity varied from 30% to 60% for expectorated sputum, from 0% to 50% for induced sputum, from 0% to 50% for gastric lavage or aspirate, and was 18% in the one cohort using nasopharyngeal aspirate. Confidence intervals were wide and overlapping. The specificity was high (greater than 93%) for all studies and specimen types.

Figure 19. Forest plot of the sensitivity and specificity of smear microscopy in detecting pulmonary TB, peripheral lymph node TB and TB meningitis in children compared against culture as a reference standard , by study and specimen type[a]




TP, true positive; FP, false positive; FN, true negative; CI, confidence interval; ES, expectorated sputum; GLA, gastric lavage or aspirate; IS, induced sputum; NPA, nasopharyngeal aspirate.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

For expectorated or induced sputum the pooled sensitivity was 29% (95% CrI, 16–42%); for gastric lavage or aspirate it was 22% (95% CrI, 12–35%). The sensitivities for both expectorated or induced sputum and for gastric lavage or aspirate had wide pooled credible intervals, indicating a high level of heterogeneity. Similar to the analyses of Xpert MTB/RIF, the specificity of microscopy was greater than 99% in both comparison groups (*Table 22*).

Table 22. Meta-analysis of the estimated sensitivity and specificity of smear microscopy in diagnosing pulmonary TB, peripheral lymph node TB and TB meningitis in children compared against culture as a reference standard in published and unpublished studies

| Comparison | Specimen type (No. of studies, No. of children) | Median (%) pooled sensitivity (pooled 95% CrI) | Median (%) pooled specificity (pooled 95% CrI) |
|---|---|---|---|
| Smear microscopy compared against culture as a reference standard (published and unpublished studies) | Expectorated sputum and induced sputum (10, 1546) | 29 (16–42) | 100 (99–100) |
| | Gastric lavage or aspirate (7, 1319) | 22 (12–35) | 99 (97–100) |

CrI, credible interval; the CrI is the Bayesian equivalent of the confidence interval.

These data suggest that in comparison with smear microscopy, the sensitivity of Xpert MTB/RIF was 37% higher if performed on samples of expectorated or induced sputum, and 44% higher if performed on samples of gastric lavage or aspirate.

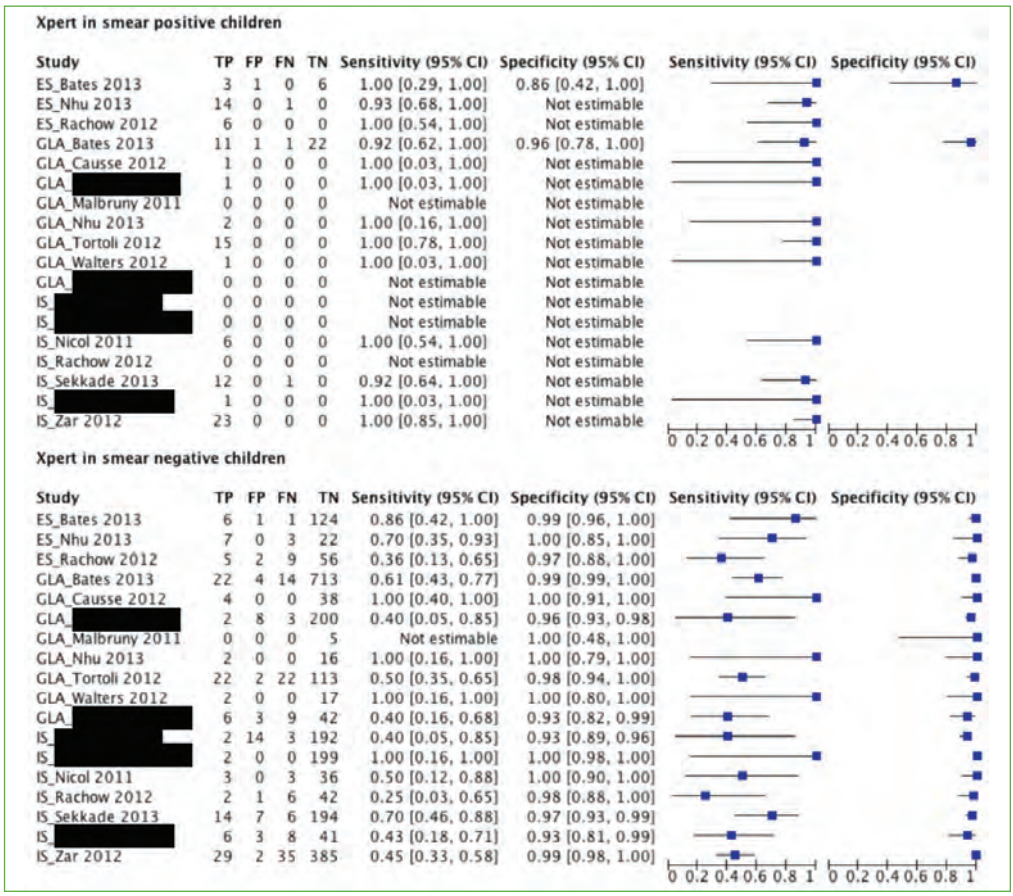
### *4.3.5. Investigations of heterogeneity*

Factors that might cause heterogeneity in results and that might be associated with smear status, HIV infection, age, as well as the approach used to confirm TB, were investigated.

#### *4.3.5.1. Performance of Xpert MTB/RIF in smear-positive and smear-negative children*

A total of 7 studies, collectively containing data from 1083 children, were included in the analysis of the accuracy of Xpert MTB/RIF in samples of expectorated or induced sputum. Six studies with data from 1259 children were included in the analysis of using Xpert MTB/RIF to test samples from gastric lavage or aspiration. All studies that were included reported results for both smear-positive and smear-negative children (*Figure 20*).

Figure 20. Forest plot of the sensitivity of Xpert MTB/RIF in diagnosing pulmonary TB, peripheral lymph node TB and TB meningitis in smear-positive and smear-negative children, by study and specimen type[a]




TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval; ES, expectorated sputum; GLA, gastric lavage or aspirate; IS, induced sputum.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

The sensitivity of Xpert MTB/RIF on samples from expectorated or induced sputum from children with smear-negative results ranged from 25% to 86%. In contrast, the sensitivity of Xpert MTB/RIF among children with smear-positive results ranged from 92% to 100%. The pooled estimate of sensitivity in smear-positive children was 96% (95% CrI, 90–99%) and in smear-negative children it was 55% (95% CrI, 41–69%). The estimates were similar for Xpert MTB/RIF used for samples of gastric lavage or aspirate, with an overall sensitivity of 95% (95% CrI, 83–99%) among smear-positive children; for smear-negative children overall sensitivity was 62% (95% CrI, 44–80%). The credible intervals were wide (indicating variability), but did not overlap (*Table 23*).

Table 23. **Meta-analysis of the sensitivity and specificity of Xpert MTB/RIF in diagnosing pulmonary TB, peripheral lymph node TB and TB meningitis compared against culture as a reference standard in smear-negative and smear-positive children in published and unpublished studies**

| Comparison | Specimen type (No. of studies, No. of children) | Median (%) pooled sensitivity (pooled 95% CrI) | Median (%) pooled specificity (pooled 95% CrI) |
|---|---|---|---|
| Xpert MTB/RIF in smear-positive children | Expectorated sputum or induced sputum (7, 68)[a] | 96 (90–99) | [c] |
| | Gastric lavage or aspirate (6, 32)[b] | 95 (83–99) | [c] |
| Xpert MTB/RIF in smear-negative children | Expectorated sputum or induced sputum (7, 1008)[a] | 55 (41–69) | 98 (96–99) |
| | Gastric lavage or aspirate (6, 1204)[b] | 62 (44–80) | 99 (97–99) |

CrI, credible interval; the CrI is the Bayesian equivalent of the confidence interval.

a This analysis included studies by Bates 2013, Nhu 2013, Nicol 2011, Rachow 2012 (expectorated sputum only), Sekadde 2013, an unpublished study by Walters and Zar 2012.
b This analysis included studies by Bates 2013, Causse 2011, an unpublished study by Chisti, Nhu 2013, Tortoli 2012 and Walters 2012.
c There were not enough data to calculate specificity.

As expected, smear status was associated with the performance of Xpert MTB/RIF, indicating that the sensitivity of Xpert MTB/RIF was greater in children who had a higher mycobacterial burden than in those with paucibacillary disease. Xpert MTB/RIF detected 55% of smear-negative culture-positive children from samples of expectorated or induced sputum, and 62% of smear-negative culture-positive children from samples of gastric lavage or aspirate. This finding suggests that Xpert MTB/RIF has the potential to effectively contribute to a diagnostic algorithm as an add-on test following negative results from smear microscopy.

The data indicated that smear status was associated with the performance of Xpert MTB/RIF. This finding suggests that the sensitivity of Xpert MTB/RIF was greater in children who had a higher mycobacterial burden than in those with paucibacillary disease.

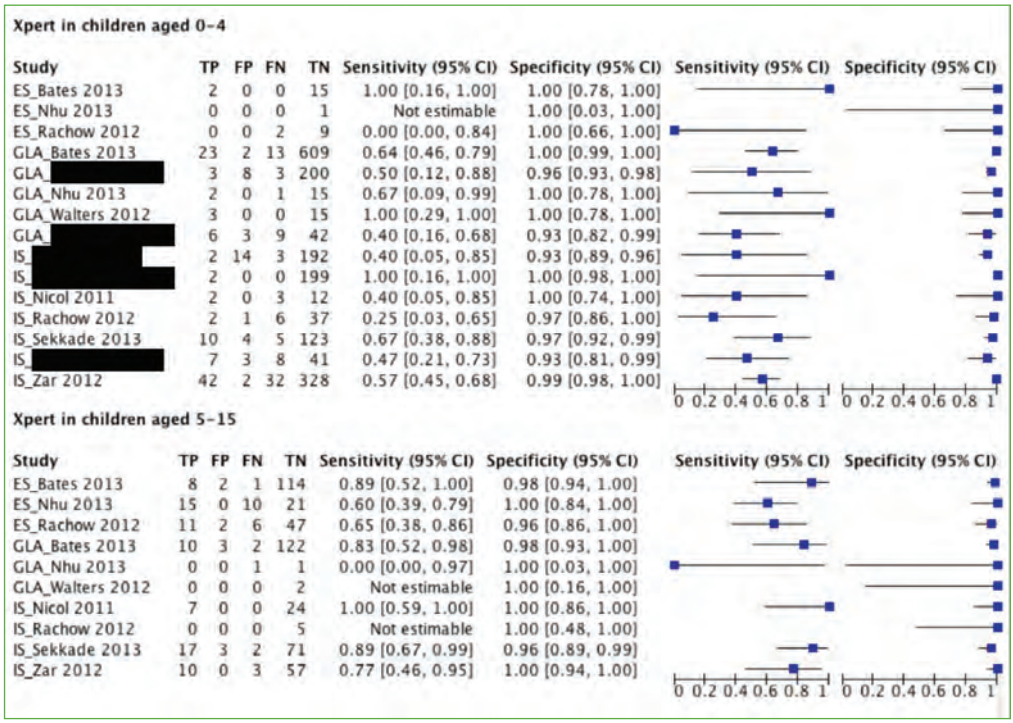
#### *4.3.5.2. Xpert MTB/RIF in children aged 0–4 years and 5–15 years*

Five of the seven studies reported results for children aged 0–4 years and 5–15 years. Data from 976 children were included in the analysis of using Xpert MTB/RIF on samples from

expectorated or induced sputum (*Figure 21*). The estimated accuracy of Xpert MTB/RIF using samples from gastric lavage or aspiration was reported only for children aged 0–4 years (5 studies, 957 children).

The sensitivity of Xpert MTB/RIF in both age groups ranged from 0% to 100%. The pooled sensitivity among children aged 0–4 years was (57%; 95%CrI, 36-75) for Xpert MTB/RIF used on samples of expectorated or induced sputum and gastric lavage or aspirate. The pooled sensitivity for samples of expectorated or induced sputum was higher in children aged 5–15 years (83%; 95% CrI, 68–92%). The pooled specificity was at least 98% for all groups assessed, with relatively narrow confidence intervals.

Figure 21. Forest plot of the sensitivity of Xpert MTB/RIF in detecting pulmonary TB, peripheral lymph node TB and TB meningitis in children aged 0–4 and 5–15 years, by study and specimen type[a]




TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval; ES, expectorated sputum; GLA, gastric lavage or aspirate; IS, induced sputum.

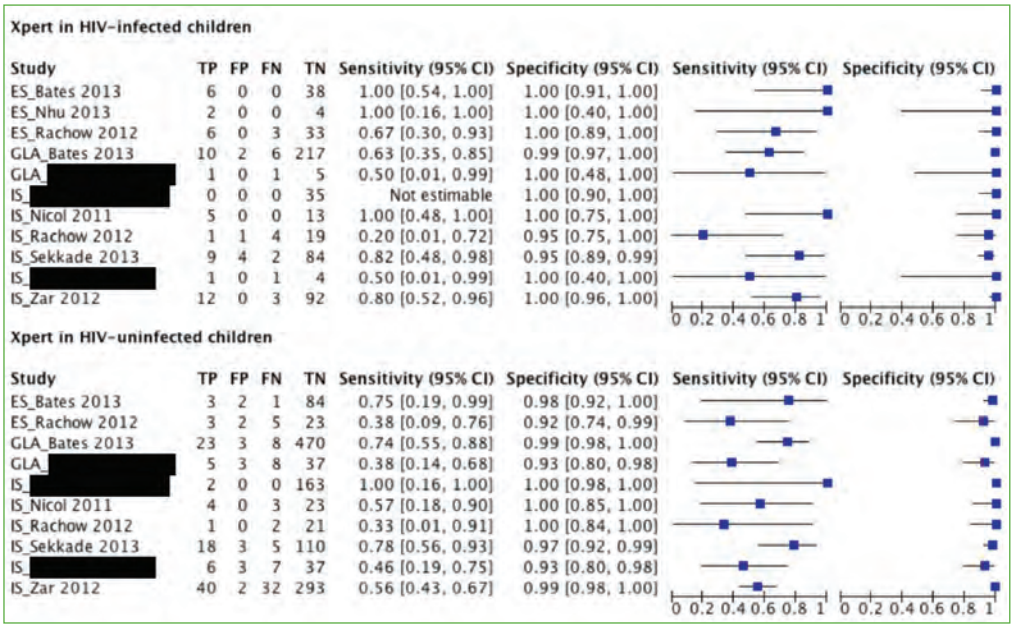
a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

### *4.3.5.3. Xpert MTB/RIF in HIV-positive and HIV-negative children*

A total of 7 studies with data from 1074 children were included in the analysis of the accuracy of Xpert MTB/RIF using samples of expectorated and induced sputum. All studies reported results for both HIV-positive and HIV-negative children.

The sensitivity of the test among HIV-positive children ranged from 20% to 100%; among HIV-negative children it ranged from 33% to 100% (*Figure 22*).

Figure 22. Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting pulmonary TB, peripheral lymph node TB and TB meningitis in HIV-positive and HIV-negative children, by study and specimen type[a]




TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval; ES, expectorated sputum; GLA, gastric lavage or aspirate; IS, induced sputum.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

The pooled sensitivity among HIV-positive children (75%; 95% CrI, 57–88%) was higher than the sensitivity for HIV-negative children (57%; 95% CrI, 41–71%); however, the credible intervals were wide and overlapping. The pooled specificity was 98% for both groups.

An analysis of the performance of Xpert MTB/RIF stratified by smear status and HIV status demonstrated that the test had high sensitivity among smear-positive children regardless of their HIV status. The sensitivity of Xpert MTB/RIF was lowest among HIV-negative children, although the credible intervals were wide and overlapping (*Table 24*).

A metaregression model that simultaneously controlled for smear status and HIV status using Xpert MTB/RIF on samples of expectorated or induced sputum showed that the odds of test positivity were fourfold greater in smear-positive children than in smear-negative children. The odds of Xpert MTB/RIF positivity were not statistically significant for HIV-positive children compared with HIV-negative children (*Table 25*).

Table 24. Meta-analysis comparing Xpert MTB/RIF for diagnosing pulmonary TB, peripheral lymph node TB and TB meningitis using culture as a reference standard in HIV-positive and HIV-negative children, stratified by smear status

| Category | Specimen type (No. of studies, No. of children) | Pooled sensitivity (%) (95% CrI) |
|---|---|---|
| HIV positive and smear positive | Expectorated sputum or induced sputum (5, 21)[a] | 97 (85–100) |
| HIV positive and smear negative | Expectorated sputum or induced sputum (5, 25)[a] | 69 (46–87) |
| HIV negative and smear positive | Expectorated sputum or induced sputum (5, 29)[a] | 94 (81–99) |
| HIV negative and smear negative | Expectorated sputum or induced sputum (5, 85)[a] | 48 (29–67) |

CrI, credible interval; the CrI is the Bayesian equivalent of the confidence interval.

a The studies included in this analysis were Bates 2013, Nicol 2011, Rachow 2012 (expectorated sputum), Sekadde 2013 and Zar 2012.

Table 25. Metaregression model for Xpert MTB/RIF using samples of expectorated or induced sputum from children, controlling for smear status and HIV status

| Node | | Mean | SD | MC error | 2.5% | Median | 97.5% |
|---|---|---|---|---|---|---|---|
| Beta 0 | | 0.06201 | 0.385 | 0.0054 | –0.8159 | -0.064 | 0.7059 |
| Beta 1 (HIV) | | 0.5863 | 0.5551 | 0.008862 | –0.4919 | 0.5789 | 1.705 |
| Beta 2 (Smear) | | 3.98 | 1.076 | 0.02855 | 2.159 | 3.878 | 6.399 |
| Pooled sensitivity | Smear negative and HIV negative | 0.485 | 0.09264 | 0.0013 | 0.3066 | 0.484 | 0.6695 |
| | Smear positive and HIV negative | 0.9694 | 0.03031 | 6.402 | 0.8873 | 0.9785 | 0.9983 |
| | Smear negative and HIV positive | 0.6213 | 0.1101 | 0.001197 | 0.3944 | 0.6257 | 0.8216 |
| | Smear positive and HIV positive | 0.988 | 0.01977 | 3.951 | 0.9284 | 0.9879 | 0.9991 |

SD, standard deviation.

There were insufficient data to conduct a meta-analysis comparing the performance of Xpert MTB/RIF when using samples of gastric lavage or aspirate from HIV-positive and HIV-negative children.

### *4.3.6. Using Xpert MTB/RIF to detect peripheral lymph node TB in children*

The use of FNA or biopsies of lymph nodes to diagnose peripheral lymph node TB was evaluated in five studies. Two studies were

excluded from the meta-analysis because the sample included fewer than five participants. Therefore, the analysis included 3 studies with data from 172 children (*Figure 23*). The pooled sensitivity of Xpert MTB/RIF compared with culture as a reference standard was 86% (95% CrI, 65–96%); the pooled specificity was 81% (95% CrI, 54–93%). Confidence intervals were wide for both sensitivity and specificity, indicating heterogeneity among the studies.

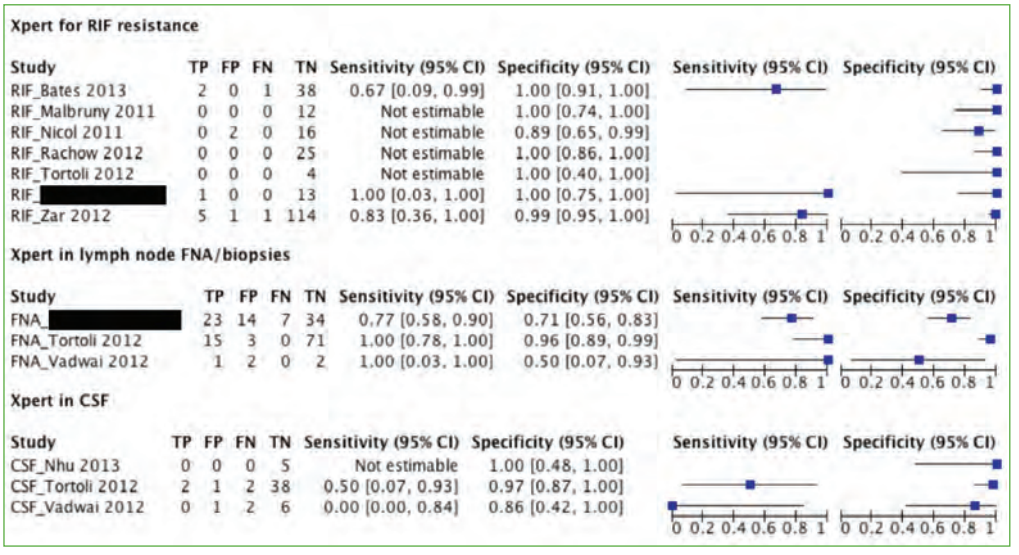
### *4.3.7. Using Xpert MTB/RIF to detect TB meningitis in children*

The use of CSF to diagnose TB meningitis was evaluated in five studies that included 61 children. In total 7/61 (11.5%) children had TB meningitis confirmed by CSF culture. Of these, 3 children were positive by Xpert MTB/RIF (3/61; 4.9%). Two studies were excluded from the meta-analysis because they did not include any culture-positive children. One of the remaining studies had a subgroup sample size that included fewer than 5 children. Hence, there were insufficient data to calculate sensitivity from the two remaining studies, in which 2/6 culture-positive children had positive results from Xpert MTB/RIF (*Figure 23*). The pooled specificity, which included 3 studies and 51 children, was 95% (95% CrI, 81–99%), with relatively wide credible intervals.

### *4.3.8. Using Xpert MTB/RIF to detect rifampicin resistance in children*

In total, seven studies provided data on using Xpert MTB/RIF to detect rifampicin resistance (*Figure 23*). Four studies used conventional phenotypic DST, and three used line probe assays. A meta-analysis of 3 studies that included 176 participants showed a pooled sensitivity of 86% (95% CrI, 53–98%) and a pooled specificity of 98% (95% CrI, 94–100%).

Figure 23. Forest plot of the sensitivity and specificity of Xpert MTB/RIF for detecting resistance to rifampicin, peripheral lymph node TB and TB meningitis in children, by study and specimen type[a]




RIF, rifampicin; TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval; FNA, fine needle aspiration; CSF, cerebrospinal fluid.

a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

### *4.3.9. Summary of findings and GRADE evidence profiles*

- Xpert MTB/RIF shows moderate sensitivity (66%) in detecting pulmonary TB in children using samples of expectorated or induced sputum and gastric lavage and aspirate when compared against culture as a reference standard.
- Xpert MTB/RIF shows high sensitivity in smear-positive children.
- The specificity of Xpert MTB/RIF was consistently high (greater than 93%). Few additional clinically confirmed culture-negative TB cases can be detected by Xpert MTB/RIF, but the challenge of collecting good quality specimens from children may explain the low yield in culture-negative children.
- The sensitivity of Xpert MTB/RIF was poor when compared against a clinical TB reference standard, which highlights the need for a universal composite reference standard that can be used to evaluate tests for diagnosing TB in children.
- The performance of Xpert MTB/RIF is superior to smear microscopy when both are compared against culture as a reference standard. Hence, Xpert MTB/RIF identifies additional cases of TB if it is used as an add-on test in children who have negative smear results.
- The performance of Xpert MTB/RIF appears to be similar in HIV-positive children and HIV-negative children, and it may have higher sensitivity as disease severity increases since disease severity is associated with smear positivity.
- The sensitivity of Xpert MTB/RIF was higher among children aged 5–15 years than among children aged 0–4 years, and sensitivity is affected by smear status.
- Xpert MTB/RIF detected 86% of children with culture-confirmed peripheral lymph node TB using specimens from FNA.
- There were insufficient data to calculate the pooled sensitivity of Xpert MTB/RIF using CSF to detect TB meningitis in children. However, Xpert MTB/RIF showed good sensitivity in detecting TB meningitis in adults.
- Xpert MTB/RIF detected 86% of culture-confirmed resistance to rifampicin with high specificity, and has the potential to increase children's access to DST.

### *4.3.10. Strengths and limitations of the evidence base*

The findings of this review are based on comprehensive searching, and the use of strict inclusion criteria and standardized methods of data extraction. Despite limitations in the number of studies and participants included in the review, as well as the heterogeneous methodological approaches used by these studies, this review for the first time provides data on the accuracy of Xpert MTB/RIF in diagnosing pulmonary TB, rifampicin resistance, peripheral lymph node TB and TB meningitis in children.

Compiling the data set involved comprehensive searching and correspondence with experts in the field to identify additional studies, as well as repeated correspondence with the authors of studies to obtain additional data and information that were not provided in the literature. The search strategy included studies published in all languages. However, despite the comprehensive search strategy, some studies may have been missed, particularly those that are continuing or articles that are in press.

Culture is regarded as the best reference standard for active TB, and was the reference standard used for TB in this review. Yet its accuracy is suboptimal in children. The limitations and implications of using culture as the reference standard are described and discussed throughout the review. A second reference standard, clinical TB, which is also suboptimal, was applied only

to children who were culture negative. Although this approach mimics clinical practice, it is methodologically flawed. Ideally, studies would apply each reference standard to all included children.

The majority of studies selected participants consecutively.

Xpert MTB/RIF results are generated automatically, and do not require subjective interpretation. In the majority of studies Xpert MTB/RIF was used at higher levels of care and among inpatients. In general, studies were fairly well reported, though we corresponded with the authors of all studies to collect additional data and missing information.

### *4.3.11. GRADE evaluations and recommendations*

GRADE evidence profiles are provided in Tables 26 to 32. The GRADE evaluation supports the use of Xpert MTB/RIF to diagnose pulmonary and extrapulmonary TB, and rifampicin resistance in children. The Expert Group therefore concluded that:

- Xpert MTB/RIF should be used rather than conventional microscopy, culture and DST as the initial diagnostic test in children suspected of having MDR-TB or HIV-associated TB (strong recommendation given the difficulties in diagnosing paediatric TB, very low-quality evidence);
- Xpert MTB/RIF may be used rather than conventional microscopy and culture in all other children suspected of having pulmonary TB (conditional recommendation acknowledging resource implications, very low-quality evidence);

  *Note: The Expert Group noted that Xpert MTB/RIF should not be used as the only test in the diagnostic pathway of children suspected of having TB, and that a child in whom there is a high clinical suspicion of TB should be treated even if the result from Xpert MTB/RIF is negative or if the test is not available.*

- Xpert MTB/RIF should be used in preference to conventional microscopy and culture as the initial diagnostic test when using CSF from patients suspected of having TB meningitis (strong recommendation given the urgency for rapid diagnosis, very low-quality evidence);

  *Note: The Expert Group noted that a negative result from Xpert MTB/RIF should be followed by other tests, and that a child in whom there is a high clinical suspicion of TB should be treated even if the result from Xpert MTB/RIF is negative or if the test is not available. The Expert Group also noted that concentration methods should be used to enhance yield when a sufficient volume of CSF is available.*

- Xpert MTB/RIF may be used as a replacement test for usual practice, including conventional microscopy and culture, in testing lymph node fluid and tissue from children suspected of having peripheral lymph node TB (conditional recommendation, very low-quality evidence).

  *Note: The Expert Group noted that a negative result from Xpert MTB/RIF should be followed by other tests, and that a child in whom there is a high clinical suspicion of TB should be treated even if the result from Xpert MTB/RIF is negative or if the test is not available. The Expert Group also noted that processing methods for samples from lymph nodes and tissues need to be standardized to optimize yield.*

Table 26. GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting TB in children compared with culture as a reference standard (A. Expectorated sputum and induced sputum, B. Gastric lavage or aspirate, C. Summary of findings)

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for the detection of TB in children against culture, where Xpert MTB/RIF is used as a replacement test for usual practice?

**A. Expectorated sputum and induced sputum**

Participants: Children aged 0–15 years suspected of having TB. Setting: Mainly tertiary care referral hospitals and university hospitals
Target condition: Pulmonary TB. Reference test: Solid culture or liquid culture. Number of studies (number of participants): 10 (1546)[a].
Pooled sensitivity: 66% (95% CrI, 52–77%); pooled specificity: 98% (95% CrI, 96–99%)

| Outcome | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence |
|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | |
| True positives (individuals with TB) | Cross-sectional | None[b] | Serious (–1)[c] | Serious (–1)[d] | None[e] | Undetected[f] | Low ⊕⊕◯◯ |
| False negatives (individuals incorrectly classified as not having TB) | Cross-sectional | None[b] | Serious (–1)[c] | Serious (–1)[d] | None[e] | Undetected[f] | Low ⊕⊕◯◯ |
| False positives (individuals incorrectly classified as having TB) | Cross-sectional | None[b] | Serious (–1)[c] | Serious (–1)[d] | None[e] | Undetected[f] | Moderate ⊕⊕⊕◯ |
| True negatives (individuals without TB) | Cross-sectional | None[b] | Serious (–1)[c] | Serious (–1)[d] | None[e] | Undetected[f] | Moderate ⊕⊕⊕◯ |
| Time to diagnosis[g] | Only descriptive data available | | | | | | Very low ⊕◯◯◯ |
| Cost | No data | | | | | | Not rated |

CrI, credible interval.

a Rachow 2012 reported testing one cohort using expectorated sputum and one using induced sputum; this was counted as two studies.

b The QUADAS-2 tool was used to assess the risk of bias. The majority of studies enrolled patients consecutively. Results from Xpert MTB/RIF are automated and were considered blinded in all studies. The exclusion of enrolled patients from analyses were well explained in most studies; however, in one study (LaCourse, unpublished) 99/300 children were excluded because they were lost to follow up. Another study (Bates 2013) enrolled a number of children who had already initiated anti-TB treatment (77/1037 with samples of expectorated or induced sputum, or gastric lavage or aspirate ): 66 of these children were culture negative and were excluded from the analysis; 11 were culture positive and were included. Additionally, the following considerations were not taken into account in judgements about the limitations of the studies but they are important: culture is an imperfect reference standard. In children, culture sensitivity varies from 20% to 70%, depending on factors such as the child's smear status, severity of disease, the child's age, the culture method used, and specimen type. In order to improve the diagnostic yield, 9/10 studies performed at least two cultures for each child, but some performed as many as six cultures per child. The sensitivity of Xpert MTB/RIF may be underestimated when it is compared against culture as a reference standard. The evidence was not downgraded.

c The rating of the quality of evidence may be lowered if there are important differences in the populations studied and the tests studied, and in the expertise of those using the tests in studies compared with the settings for which the recommendations are intended. All studies were performed at higher levels of care – that is, higher-level referral hospitals and university hospitals. Six of the 10 studies (60%) included only inpatients; 4/10 studies included inpatients and outpatients. In the paediatric population, inpatients may present with more severe disease and have a higher likelihood of being smear positive; they are one group that might benefit from Xpert MTB/RIF. Outpatients may present differently, with less severe disease and so be less likely to be smear positive or culture positive. Outpatients are not represented by the studies included in this review. Only one study (Rachow 2012) included a broader group of children: approximately 20% of children suspected of having TB had been identified through contact tracing. Two unpublished studies (Chisti and LaCourse) included children who were severely malnourished, and TB was one differential diagnosis. (Children in the Chisti study additionally had to have cough lasting longer than 2 weeks or pneumonia on X-ray, or both.) There were serious concerns about indirectness in the majority of studies performed in higher-level referral hospitals. The evidence was downgraded by one point.

d The point estimates for sensitivity ranged from 25% to 100%, indicating a high level of heterogeneity. However, the 95% confidence intervals overlapped. Subgroup analyses suggested that there was an effect of smear status on overall sensitivity, with a pooled estimate for sensitivity among smear-positive children of 96% (95% CrI, 90–99%) and for smear-negative children of 55% (95% CrI, 41–69%). There was no relevant heterogeneity among smear-positive children (pooled estimates of sensitivity range, 92–100%), but there was considerable heterogeneity among smear-negative children (range, 25–100%). Subgroup analyses by age group (for those aged 0–4 years the pooled sensitivity estimate was 57% with 95% CrI, 36–74%; for children aged 5–15 years the pooled sensitivity estimate was 83% with 95% CrI, 68–92%) and HIV status (for HIV-positive children the pooled sensitivity estimate was 75% with 95% CrI, 57–88%; for HIV-negative children the pooled sensitivity estimate was 57% with 95% CrI, 41–71%) indicated that there were differences among these groups, but these estimates are heterogeneous and seem to be affected by smear status. Overall, heterogeneity cannot be fully explained by the subgroup analyses; therefore, inconsistency is rated as a serious concern for the sensitivity estimates. The evidence was downgraded by one point. Estimates of specificity did not show serious heterogeneity: point estimates ranged from 93% to 100%. There was no concern about specificity.

e Five of the 10 studies included more than 100 children, and the yield of culture among all studies varied from 2% to 52%. The confidence interval for sensitivity was wide (crossing +/– 10%), which is of concern. The evidence was not downgraded further given the downgrading for inconsistency. There was no concern about specificity.

f Three unpublished studies were included. A few other studies were in progress, but data were not available for analysis. The data that were included in the review did not allow for formal assessment of publication bias using methods such as funnel plots or regression tests; such techniques have not been found to be helpful in studies evaluating diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

g The time to diagnosis was mainly reported for the turnaround time in the laboratory rather than from the day of specimen collection: it ranged from 6 days to 7 weeks for culture (7 studies) and from 2 hours to 2.5 hours for Xpert MTB/RIF. Three studies described the time from enrolment or specimen collection to diagnosis using Xpert MTB/RIF (1 day) and culture (21 days using the BACTEC MGIT 960, and 30 days using Löwenstein–Jensen medium).

**B. Gastric lavage or aspirate**

Number of studies (number of participants): 7 (1319)

Pooled sensitivity: 66% (95% CrI, 51–81%); pooled specificity: 98% (95% CrI, 96–99%)

| Outcome | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence |
|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | |
| True positives (individuals with TB) | Cross-sectional | None[a] | Serious (–1)[b] | Serious (–1)[c] | None (serious)[d] | Undetected[e] | Low ⊕⊕◯◯ |
| False negatives (individuals incorrectly classified as not having TB) | Cross-sectional | None[a] | Serious (–1)[b] | Serious (–1)[c] | None (serious)[d] | Undetected[e] | Low ⊕⊕◯◯ |
| False positives (individuals incorrectly classified as having TB) | Cross-sectional | None[a] | Serious (–1)[b] | Serious (–1)[c] | None[d] | Undetected[e] | Moderate ⊕⊕⊕◯ |
| True negatives (individuals without TB) | Cross-sectional | None[a] | Serious (–1)[b] | Serious (–1)[c] | None[d] | Undetected[e] | Moderate ⊕⊕⊕◯ |

CrI, credible interval.

a The QUADAS-2 tool was used to assess the risk of bias. Three of seven studies enrolled individuals consecutively. Results from Xpert MTB/RIF are automated and were considered blinded in all studies. The exclusion of enrolled patients from analyses were well explained in most studies; however in one study (Bates 2013) a number of children who had already initiated anti-TB treatment were enrolled (77/1037, for specimens from expectorated or induced sputum and of gastric lavage and aspirate ). A total of 66 of these children were culture negative and were excluded; 11 were culture positive and were included in the analysis. Additionally, the following considerations were not taken into account in judgements about the limitations of the studies but they are important: culture is an imperfect reference standard. In children, culture sensitivity varies from 20% to 70%, depending on factors such as the severity of disease, the age of the child, the culture methods used, and specimen types. In order to improve the diagnostic yield, 9/10 studies performed at least two cultures for each child, but some performed as many as six cultures per child. The sensitivity of Xpert MTB/RIF may be underestimated when it is compared against culture as a reference standard. The evidence was not downgraded.
b The rating of the quality of evidence may be lowered if there are important differences in the populations studied and the tests studied, and in the expertise of those using the tests in the studies compared with the settings for which the recommendations are intended. All studies were performed at higher levels of care – that is, higher-level referral hospitals and university hospitals. Four of the seven studies included only inpatients; one included inpatients and outpatients; two were laboratory studies that did not provide information on the patient cohort. In the paediatric population, inpatients may present with more severe disease and have a higher likelihood of being smear positive; they are one group that might benefit from Xpert MTB/RIF. Outpatients may present differently, with less severe disease and so be less likely to be smear positive or culture positive. Outpatients are not represented by the studies included in this review. One unpublished study (Chisti) included severely malnourished children who had cough lasting longer than 2 weeks or pneumonia on X-ray, or both; TB was one differential diagnosis. There were serious concerns about indirectness. The evidence was downgraded by one point.
c The point estimates for sensitivity ranged from 40% to 100%, indicating a high level of heterogeneity. However, the 95% confidence intervals overlapped. Subgroup analyses suggested that there was an effect of smear status on overall sensitivity, with a pooled estimate for sensitivity among smear-positive children of 95% (95% CrI, 83–99%) and for smear-negative children of 62% (9%% CrI, 44– 80%). There was no relevant heterogeneity among smear–positive children (pooled estimates of sensitivity range, 92–100%), but there was considerable heterogeneity among smear-negative children (range, 36–100%). Analyses of subgroups by age was possible only for children aged 0–4 years (pooled estimate of sensitivity was 57% with 95% CrI, 38–75%). Three studies included only children aged 0–5 years; the number of gastric lavage and aspirations performed in children aged 5–15 in the remaining studies was small. Subgroup analysis for HIV status was not possible for specimens from gastric lavage or aspiration. Overall, heterogeneity could not be fully explained by the subgroup analyses; therefore, inconsistency was rated as a serious concern for the sensitivity estimates. The evidence was downgraded by one point. Estimates of specificity did not show serious heterogeneity: point estimates ranged from 93% to 100%. There was no concern about specificity.
d Three of 7 studies included more than 100 children, and the yield of culture among all studies varied from 2.8% to 32.7%. The confidence interval for sensitivity was wide (crossing +/– 10%), which is of concern. The evidence was not downgraded further given the downgrading for inconsistency. There was no concern about specificity.
e Two unpublished studies were included. A few other studies were in progress, but data were not available for analysis. The data that were included did not allow for formal assessment of publication bias using methods such as funnel plots or regression tests; such techniques have not been found to be helpful in studies evaluating diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

## C. Summary of findings

| Outcome | Xpert MTB/RIF using specimens of expectorate or induced sputum | | | | Xpert MTB/RIF using specimens of gastric lavage or aspirate | | | |
|---|---|---|---|---|---|---|---|---|
| | Number of results/1000 individuals tested (95% CrI)[a] | | | Quality of evidence | Number of results/1000 individuals tested (95% CrI)[a] | | | Quality of evidence |
| | Prevalence 10/1000[b] | Prevalence 50/1000[b] | Prevalence 100/1000[b] | | Prevalence 10/1000[b] | Prevalence 50/1000[b] | Prevalence 100/1000[b] | |
| True positives (individuals with TB) | 7 (5-8) | 33 (26-39) | 66 (52-77) | Low ⊕⊕○○ | 7 (5-8) | 33 (26-41) | 66 (51-81) | Low ⊕⊕○○ |
| False negatives (individuals incorrectly classified as not having TB) | 3 (2-5) | 17 (12-24) | 34 (23-48) | Low ⊕⊕○○ | 3 (2-5) | 17 (10-25) | 34 (19-49) | Low ⊕⊕○○ |
| False positives (individuals incorrectly classified as having TB) | 20 (10-40) | 19 (10-38) | 18 (9-36) | Very low ⊕○○○ | 20 (10-40) | 19 (10-38) | 18 (9-36) | Moderate ⊕⊕⊕○ |
| False negatives (individuals incorrectly classified as not having TB) | 970 (950-980) | 931 (912-941) | 882 (864-891) | Very low ⊕○○○ | 970 (950-980) | 931 (912-941) | 882 (864-891) | Moderate ⊕⊕⊕○ |

CrI, credible interval.

a The expected number of Xpert MTB/RIF results was based on the sensitivity and specificity estimates calculated from a comparison with culture for different prevalences of TB.
b The estimates of TB prevalence were provided by the WHO Steering Group.

Table 27. GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting TB in children compared with a clinical reference standard (A. Expectorated sputum and induced sputum, and gastric lavage and aspirate, B. Summary of findings)

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for detection of TB in children compared with a clinical reference standard, where Xpert MTB/RIF is used as a replacement test for usual practice?

**A. Expectorated and induced sputum, and gastric lavage and aspirate**

Participants: Culture negative children aged 0-15 years suspected of having TB. Setting: Mainly tertiary care referral hospitals, university hospitals.
Target condition: Pulmonary TB. Reference test: Clinical TB. Expectorated sputum and induced sputum combined.
Number of studies (number of participants): 8 (995)[a]. Pooled sensitivity: 4% (95% CrI, 1–12%); pooled specificity 100% (95% CrI, 99–100%).
Gastric lavage and aspirate. Number of studies (number of participants): 3 (269). Pooled sensitivity: 15% (95% CrI, 5–31%); pooled specificity: 99% (95% CrI, 96–100%)

| Outcome | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence |
|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | |
| True positives (individuals with TB) | Cross-sectional | Serious (–1)[b] | Serious (–1)[c] | Serious (–1)[d] | None[e] | Undetected[f] | Very low ⊕○○○ |
| False negatives (individuals incorrectly classified as not having TB) | Cross-sectional | Serious (–1)[b] | Serious (–1)[c] | Serious (–1)[d] | None[e] | Undetected[f] | Very low ⊕○○○ |
| False positives (individuals incorrectly classified as having TB) | Cross-sectional | Serious (–1)[b] | Serious (–1)[c] | Serious (–1)[d] | None[e] | Undetected[f] | Low ⊕⊕○○ |
| True negatives (individuals without TB) | Cross-sectional | Serious (–1)[b] | Serious (–1)[c] | Serious (–1)[d] | None[e] | Undetected[f] | Low ⊕⊕○○ |

CrI, credible interval.

a Rachow 2012 reported testing one cohort using expectorated sputum and one using induced sputum; this was counted as two studies.

b The QUADAS-2 tool was used to assess the risk of bias. Six of seven studies included in the meta-analysis enrolled individuals consecutively (seven studies used specimens of expectorated or induced sputum; three used specimens from gastric lavage or aspiration). Results from Xpert MTB/RIF are automated and were considered blinded in all studies. Clinical TB is an imperfect reference standard. In this review it was defined on the basis of whether culture-negative children suspected of having TB were initiated on anti-TB treatment. Data are lacking on how well this clinical reference standard identifies true cases, both in terms of overdiagnosis and underdiagnosis. The risk of bias for the reference standard was rated as unclear for all studies. There were very serious concerns about the limitations of the studies. The evidence was downgraded by one point.

c The rating of the quality of evidence may be lowered if there are important differences in the tests studied and the expertise of those using the tests in the studies compared with the settings for which the recommendations are intended. All studies were performed at higher levels of care – that is, higher-level referral hospitals and university hospitals. Five of the seven studies (71%) studies included only inpatients; two studies included inpatients and outpatients. In the paediatric population, inpatients may present with more severe disease and have a higher likelihood of being smear positive. In one study (Rachow 2012), approximately 20% of children suspected of having TB were identified through contact tracing; this study is most representative of the paediatric population that might benefit from Xpert MTB/RIF Two unpublished studies (Chisti and LaCourse) included children who were severely malnourished, and TB was one differential diagnosis. Additionally, children in the Chisti study had to have had cough lasting longer than 2 weeks or pneumonia on X-ray, or both. There were serious concerns about indirectness. The evidence was downgraded by one point.

d The point estimates for sensitivity ranged from 0% to 35% for testing specimens of expectorated or induced sputum; for testing specimens from gastric lavage or aspiration, the point estimates ranged from 0% to 20%; these results indicate moderate heterogeneity. However, the 95% confidence intervals overlapped. No subgroup analysis was performed for Xpert MTB/RIF compared against a clinical reference standard. Heterogeneity may partly be the result of differences in patient populations (for example, caused by the use of different inclusion criteria). Inconsistency was rated as a serious concern for the estimates of sensitivity . The evidence was downgraded by one point. The estimates of specificity did not show heterogeneity: point estimates ranged from 99% to 100%. There was no concern about specificity.

e The majority of cohorts included in the review had a sample size greater than 60 (5/8 studies using expectorated or induced sputum, and 2/3 using gastric lavage or aspirate ); the yield of culture among all cohorts varied from 2.4% to 54.2%. The confidence interval for sensitivity for expectorated or induced sputum was within the +/-10% margin; however, for gastric lavage and aspirate it was wide. There were no concerns for expectorated or induced sputum; there were serious concerns about gastric lavage and aspirate. The evidence was not downgraded further given the downgrading for inconsistency. There was no concern about the evidence for specificity.

f Two unpublished studies were included. A few other studies were in progress but data were not available for analysis. The data that were included do not allow for formal assessment of publication bias using methods such as funnel plots or regression tests; such techniques have not been found to be helpful in studies evaluating diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

## B. Summary of findings

| Outcome | Xpert MTB/RIF using specimens of expectorate or induced sputum | | | | Xpert MTB/RIF using specimens of gastric lavage or aspirate | | | |
|---|---|---|---|---|---|---|---|---|
| | Number of results/1000 individuals tested (95% CrI)[a] | | | Quality of evidence | Number of results/1000 individuals tested (95% CrI)[a] | | | Quality of evidence |
| | Prevalence 10/1000[b] | Prevalence 50/1000[b] | Prevalence 100/1000[b] | | Prevalence 10/1000[b] | Prevalence 50/1000[b] | Prevalence 100/1000[b] | |
| True positives (individuals with TB) | 0 (0-1) | 2 (1-6) | 4 (1-12) | Low ⊕⊕◯◯ | 4 (1-8) | 8 (3-16) | 15 (5-31) | Very low ⊕◯◯◯ |
| False negatives (individuals incorrectly classified as not having TB) | 10 (9-10) | 48 (44-50) | 96 (88-99) | Low ⊕⊕◯◯ | 21 (17-24) | 43 (35-48) | 85 (69-95) | Very low ⊕◯◯◯ |
| False positives (individuals incorrectly classified as having TB) | 0 (1-10) | 0 (0-10) | 0 (1-9) | Moderate ⊕⊕⊕◯ | 10 (0-39) | 10 (0-38) | 9 (0-9) | Moderate ⊕⊕⊕◯ |
| False negatives (individuals incorrectly classified as not having TB) | 0 (0-1) | 2 (1-6) | 4 (1-12) | Moderate ⊕⊕⊕◯ | 965 (936-975) | 941 (912-950) | 891 (864-900) | Moderate ⊕⊕⊕◯ |

CrI, credible interval.

a The expected number of Xpert MTB/RIF results was based on the sensitivity and specificity estimates calculated from a comparison with culture for different prevalences of TB.
b The estimates of TB prevalence were provided by the WHO Steering Group.

**Table 28.** GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting TB in children following negative smear microscopy (A. Evidence profile, B. Summary of findings, C. Additional yield of Xpert MTB/RIF over microscopy)

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for detection of TB in children, where Xpert MTB/RIF is used as an add-on test following a negative smear-microscopy result?

**A. Evidence profile**

Participants: Smear negative children aged 0-15 years suspected of having TB. Setting: Mainly tertiary care referral hospitals and university hospitals. Target condition: Pulmonary TB. Reference test: Solid culture or liquid culture. Expectorated sputum and induced sputum combined. Number of studies (number of participants): 7 (1008). Pooled sensitivity 44% (95% CrI, 41–69%); pooled specificity: 98% (95% CrI, 96–99%). Gastric lavage or aspirate. Number of studies (number of participants): 6 (1204). Pooled sensitivity: 62% (95% CrI, 44–80%); pooled specificity: 99% (95% CrI, 97–99%)

| Outcome | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence |
|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | |
| True positives (individuals with TB) | Cross-sectional | None[a] | Serious (–1)[b] | Serious (–1)[c] | Expectorated or induced sputum: none[d] Gastric lavage or aspirate: serious (–1)[d] | Undetected[e] | Expectorated or induced sputum: Low ⊕⊕◯◯ Gastric lavage or aspirate: Very low ⊕◯◯◯ |
| False negatives (individuals incorrectly classified as not having TB) | Cross-sectional | None[a] | Serious (–1)[b] | Serious (–1)[c] | Expectorated or induced sputum: none[d] Gastric lavage or aspirate: serious (–1)[d] | Undetected[e] | Expectorated or induced sputum: Low ⊕⊕◯◯ Gastric lavage or aspirate: Very low ⊕◯◯◯ |
| False positives (individuals incorrectly classified as having TB) | Cross-sectional | None[a] | Serious (–1)[b] | None[c] | None[d] | Undetected[e] | Moderate ⊕⊕⊕◯ |
| True negatives (individuals without TB) | Cross-sectional | None[a] | Serious (–1)[b] | None[c] | None[d] | Undetected[e] | Moderate ⊕⊕⊕◯ |

CrI, credible interval.

a The QUADAS-2 tool was used to assess the risk of bias for all 11 studies that were included in this estimate (7 cohorts contributing samples of expectorated or induced sputum; 4 cohorts with specimens from gastric lavage or aspiration). The majority of studies enrolled individuals consecutively. Results from Xpert MTB/RIF result are automated and were considered blinded in all studies. The exclusion of enrolled patients from analyses were well explained in most studies; however in one study (Bates 2013) that used specimens from gastric lavage or aspiration, a number of children who had already initiated anti-TB treatment were enrolled (77/1037 samples of expectorated or induced sputum, and samples from gastric lavage or aspiration); 66 of these children were culture negative and were excluded from the analysis; 11 were culture positive and were included. Additionally, the following considerations were not taken into account in judgements about the limitations of the studies but they are important: culture is an imperfect reference standard. In children, culture sensitivity varies from 20% to 70%, depending on factors such as the severity of disease, the age of the child, the culture methods used, and specimen types. In order to improve the diagnostic yield, 2/11 studies (both using specimens from gastric lavage or aspiration performed at least two cultures for each child, but some performed as many as six cultures per child. The sensitivity of Xpert MTB/RIF may be underestimated when it is compared against culture as a reference standard. The evidence was not downgraded.

b All studies were performed at higher levels of care – that is, at higher-level referral hospitals and university hospitals. Eight studies included only inpatients; two studies included inpatients and outpatients (two studies tested specimens of expectorated or induced sputum; one study tested specimens from gastric lavage or aspiration); two studies were laboratory-based and provided little clinical information (these tested specimens from gastric lavage or aspiration). In the paediatric population, inpatients may present with more severe disease and have a higher likelihood of being smear positive; they are one group that might benefit from Xpert MTB/RIF. Outpatients may present differently, with less severe disease and so be less likely to be smear positive or culture positive. Outpatients are not represented by the studies included in this review. Only one study (Rachow 2012, using specimens from expectorated sputum) represented a broader group of children: approximately 20% of children suspected of having TB had been identified through contact tracing. One unpublished study (Chisti) included severely malnourished children with cough lasting longer than 2 weeks or pneumonia on X-ray, or both; in these cases, TB was one differential diagnosis. There were serious concerns about indirectness in the majority of studies performed in higher-level referral hospitals. The evidence was downgraded by one point.

c No study directly addressed the question of inconsistency by performing prior testing with microscopy and then subsequently performing Xpert MTB/RIF. Estimates of sensitivity were variable across studies and specimen types, and ranged from 36% to 86% for expectorated sputum, from 43% to 70% for induced sputum, and from 40% to 100% for specimens from gastric lavage or aspiration. The heterogeneity could not be explained; the evidence was therefore downgraded by one point. Estimates of specificity did not show serious heterogeneity: point estimates ranged from 93% to 100%. There was no concern about specificity.

d The pooled estimate of sensitivity for specimens from expectorated or induced sputum had a wide 95% confidence interval (greater than +/-10% of the point estimate); the estimate for gastric lavage or aspiration had a wide 95% confidence interval +/-20% of the point estimate). The concerns about expectorated or induced sputum were serious; the concerns about samples from gastric lavage or aspiration were very serious. The evidence was not downgraded further given the downgrading for inconsistency. There was no concern about specificity.

e Two unpublished studies were included. The data that were included in the review did not allow for formal assessment of publication bias using methods such as funnel plots or regression tests; such techniques have not been found to be helpful in studies evaluating diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

## B. Summary of findings

| Outcome | Xpert MTB/RIF using specimens of expectorate or induced sputum | | | | Xpert MTB/RIF using specimens of gastric lavage or aspirate | | | |
|---|---|---|---|---|---|---|---|---|
| | Number of results/1000 individuals tested (95% CrI)[a] | | | Quality of evidence | Number of results/1000 individuals tested (95% CrI)[a] | | | Quality of evidence |
| | Prevalence 10/1000[b] | Prevalence 50/1000[b] | Prevalence 100/1000[b] | | Prevalence 10/1000[b] | Prevalence 50/1000[b] | Prevalence 100/1000[b] | |
| True positives (individuals with TB) | 6 (4-7) | 28 (21-35) | 55 (41-69) | Low ⊕⊕○○ | 6 (4-8) | 31 (22-40) | 62 (44-80) | Very low ⊕○○○ |
| False negatives (individuals incorrectly classified as not having TB) | 5 (3-6) | 23 (16-30) | 45 (31-59) | Low ⊕⊕○○ | 4 (2-6) | 19 (10-28) | 38 (20-56) | Very low ⊕○○○ |
| False positives (individuals incorrectly classified as having TB) | 20 (10-40) | 19 (10-38) | 18 (9-36) | Moderate ⊕⊕⊕○ | 10 (10-30) | 10 (10-29) | 9 (9-27) | Moderate ⊕⊕⊕○ |
| False negatives (individuals incorrectly classified as not having TB) | 970 (950-980) | 931 (912-941) | 882 (864-891) | Moderate ⊕⊕⊕○ | 980 (960-941) | 941 (922-941) | 891 (873-891) | Moderate ⊕⊕⊕○ |

CrI, credible interval.

a The expected number of Xpert MTB/RIF results was based on the sensitivity and specificity estimates calculated from a comparison with culture for different prevalences of TB.
b The estimates of TB prevalence were provided by the WHO Steering Group.

**C. What is the additional yield of Xpert MTB/RIF compared with microscopy in children with smear-negative culture-positive TB?**

| Outcome | Number of results/1000 smear-negative culture-positive children tested (95% CrI) | | | | | | Quality of evidence |
|---|---|---|---|---|---|---|---|
| | Prevalence 10/1000 | | Prevalence 50/1000 | | Prevalence 100/1000 | | |
| | Smear microscopy | Xpert MTB/RIF | Smear microscopy | Xpert MTB/RIF | Smear microscopy | Xpert MTB/RIF | |
| Specimens from expectorated or induced sputum | | | | | | | |
| True positives (patients with TB) | 0 | 6 (4–7) | 0 | 28 (21–35) | 0 | 55 (41–69) | Low ⊕⊕◯◯ |
| True positives (absolute difference) | 6 more | | 28 more | | 55 more | | |
| False negatives (patients incorrectly classified as not having TB) | 10 | 5 (3–6) | 50 | 23 (16–30) | 100 | 45 (31–59) | Low ⊕⊕◯◯ |
| False negatives (absolute difference) | 5 fewer | | 27 fewer | | 55 fewer | | |
| Specimens from gastric lavage or aspiration | | | | | | | |
| True positives (patients with TB) | 0 | 6 (4–8) | 0 | 31 (22–40) | 0 | 62 (44–80) | Low ⊕⊕◯◯ |
| True positives (absolute difference) | 6 more | | 31 more | | 62 more | | |
| False negatives (patients incorrectly classified as not having TB) | 10 | 4 (2–6) | 50 | 19 (10–28) | 100 | 38 (20–56) | Low ⊕⊕◯◯ |
| False negatives (absolute difference) | 6 fewer | | 31 fewer | | 62 fewer | | |

CrI, credible interval.

Table 29. Incremental yield of Xpert MTB/RIF compared with smear microscopy in children with culture-confirmed TB (A. Expectorated sputum and induced sputum, B. Gastric lavage or aspirate)

PICO question: What is the incremental yield of Xpert MTB/RIF compared with smear microscopy in children with culture-confirmed TB?

**A. Expectorated sputum and induced sputum**

Participants: Children aged 0-15 years with culture-confirmed TB
Setting: Mainly tertiary care referral hospitals and university hospitals
Target condition: Pulmonary TB. Reference standard: Solid culture or liquid culture.
Number of studies (number of participants): 10 (1546)[a]
Microscopy
Pooled sensitivity: 29% (95% CrI, 16–42%); pooled specificity: 100% (95% CrI, 99–100%)
Xpert MTB/RIF
Pooled sensitivity: 66% (95% CrI, 52–77%); pooled specificity: 98% (95% CrI, 96–99%)

| Outcome | Number of results/1000 culture-positive children tested (95% CrI) | | | | | | Quality of evidence |
|---|---|---|---|---|---|---|---|
| | Prevalence 10/1000 | | Prevalence 50/1000 | | Prevalence 100/1000 | | |
| | Smear microscopy | Xpert MTB/RIF | Smear microscopy | Xpert MTB/RIF | Smear microscopy | Xpert MTB/RIF | |
| True positives (patients with TB) | 3 (2–4) | 7 (5–8) | 15 (8–21) | 33 (26–39) | 29 (16–42) | 66 (52–77) | Low ⊕⊕○○ |
| True positives (**absolute difference**) | 4 more | | 18 more | | 37 more | | |
| False negatives (patients incorrectly classified as not having TB) | 7 (6–8) | 3 (2–5) | 36 (29–42) | 17 (12–24) | 71 (58–84) | 34 (23–48) | Low ⊕⊕○○ |
| False negatives (**absolute difference**) | 4 fewer | | 19 fewer | | 37 fewer | | |

CrI, credible interval.
a These are the same studies as those assessed in Table 24 A. Rachow 2012 reported testing one cohort using expectorated sputum and one using induced sputum; this was counted as two studies.

**B. Gastric lavage or aspirate**

Number of studies (number of participants): 7 (1319)
Microscopy
Pooled sensitivity: 11% (95% CrI, 12–35%); pooled specificity: 99% (95% CrI, 97–100%)
Xpert MTB/RIF
Pooled sensitivity: 66% (95% CrI, 51–81%); pooled specificity: 98% (95% CrI, 96–99%)

| Outcome | Number of results/1000 culture-positive children tested (95% CrI)[a] | | | | | | Quality of evidence |
|---|---|---|---|---|---|---|---|
| | Prevalence 10/1000 | | Prevalence 50/1000 | | Prevalence 100/1000 | | |
| | Smear microscopy | Xpert MTB/RIF | Smear microscopy | Xpert MTB/RIF | Smear microscopy | Xpert MTB/RIF | |
| True positives (patients with TB) | 2 (1–4) | 7 (5–8) | 11 (6–18) | 33 (26–41) | 22 (12–35) | 66 (51–81) | Low ⊕⊕○○ |
| True positives (**absolute difference**) | 5 more | | 22 more | | 44 more | | |
| False negatives (patients incorrectly classified as not having TB) | 8 (7–9) | 3 (2–5) | 39 (33–44) | 17 (10–25) | 78 (65–88) | 34 (19–49) | Low ⊕⊕○○ |
| False negatives (**absolute difference**) | 5 fewer | | 22 fewer | | 44 fewer | | |

CrI, credible interval.

Table 30. GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting rifampicin resistance in respiratory specimens from children

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for detection of rifampicin resistance in respiratory specimens from children?
Participants: Children aged 0–15 years suspected of having TB or multidrug-resistant TB
Setting: Mainly tertiary care referral hospitals and university hospitals
Target condition: Rifampicin resistance
Reference test: Culture and culture-based phenotypic drug-susceptibility testing (1 study) or Hain Lifescience's Genotype MTBDR*plus* (2 studies)
Number of studies (number of participants): 3 (176)
Pooled sensitivity: 86% (95% CrI, 53–98%); pooled specificity: 98% (95% CrI, 94–100%)

| Outcome | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence | Number of results/1000 individuals tested (95% CrI) | |
|---|---|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | | Prevalence 50/1000 | Prevalence 100/1000 |
| True positives (individuals with TB and rifampicin resistance) | Cross-sectional | None[a] | None[b] | None[c] | Very serious (–2)[d] | Undetected[e] | Low ⊕⊕○○ | 43 (27-49) | 129 (80-147) |
| False negatives (individuals incorrectly classified as having TB that is rifampicin susceptible) | Cross-sectional | None[a] | None[b] | None | Very serious (–2)[d] | Undetected[e] | Low ⊕⊕○○ | 7 (1-24) | 21 (3-71) |
| False positives (individuals incorrectly classified as having TB with rifampicin resistance) | Cross-sectional | None[a] | None[b] | None | None | Undetected[e] | High ⊕⊕⊕⊕ | 19 (0-57) | 17 (0-51) |
| True negatives (individuals with TB that is rifampicin susceptible) | Cross-sectional | None[a] | None[b] | None | None | Undetected[e] | High ⊕⊕⊕⊕ | 931 (893-950) | 833 (799-850) |

CrI, credible interval; DST, drug-susceptibility testing.

a The QUADAS-2 tool was used to assess the risk of bias. All three studies enrolled individuals consecutively. The result from Xpert MTB/RIF is automated and was considered blinded in all studies. There were no concerns about risk of bias.
b The rating of the quality of evidence may be lowered if there are important differences in the populations and tests studied, and in the expertise of those using the tests in the studies compared with the settings for which the recommendations are intended. All studies were performed at tertiary care referral hospitals or university hospitals, and all children included in the studies were inpatients.
c For the 3 studies, the point estimates for sensitivity were 67%, 83% and 100%, indicating some heterogeneity. The 95% confidence intervals were wide and overlapping. The estimates of specificity did not show serious heterogeneity: the point estimates ranged from 94% to 100%. There were no concerns about specificity.
d Only 3 studies with a total of 176 children were included in the review; 121 children were enrolled in one study. The confidence interval for sensitivity was wide (crossing +/– 20%). There were very serious concerns about this evidence. The evidence was downgraded by two points. There were no concerns about specificity.
e One unpublished study was included in the review. A few additional studies were in progress but data were not available for analysis. The data included in the review did not allow for formal assessment of publication bias using methods such as funnel plots or regression tests; such techniques have not been found to be helpful in studies evaluating diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

Table 31. GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting peripheral lymph node TB in children

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for detection of peripheral lymph node TB in children, where Xpert MTB/RIF is used as a replacement test for usual practice?
Participants: Children aged 0–15 years suspected of having peripheral lymph node TB
Setting: Mainly tertiary care referral hospitals and university hospitals
Target condition: Peripheral lymph node TB
Reference test: Solid culture or liquid culture of samples from fine needle aspiration or biopsy
Number of studies (number of participants): 3 (172)
Pooled sensitivity: 86% (95% CrI, 65–96%); pooled specificity: 81% (95% CrI, 54–93%)

| Outcome | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence | Number of results/1000 individuals tested (95% CrI) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | | Prevalence 10/1000 | Prevalence 50/1000 | Prevalence 100/1000 |
| True positives (individuals with TB) | Cross-sectional | None[a] | None[b] | Serious (–1)[c] | Serious (–1)[d] | Undetected[e] | Low ⊕⊕◯◯ | 22 (16-24) | 43 (33-48) | 86 (65-96) |
| False negatives (individuals incorrectly classified as not having TB) | Cross-sectional | None[a] | None[b] | Serious (–1)[c] | Serious (–1)[d] | Undetected[e] | Low ⊕⊕◯◯ | 4 (1-9) | 7 (2-18) | 14 (4-35) |
| False positives (individuals incorrectly classified as having TB) | Cross-sectional | None[a] | None[b] | Serious (–1)[c] | Serious (–1)[d] | Undetected[e] | Low ⊕⊕◯◯ | 185 (68-449) | 181 (67-437) | 171 (63-414) |
| True negatives (individuals without TB) | Cross-sectional | None[a] | None[b] | Serious (–1)[c] | Serious (–1)[d] | Undetected[e] | Low ⊕⊕◯◯ | 790 (527-907) | 770 (513-884) | 729 (486-837) |

CrI, credible interval.

a The QUADAS-2 tool was used to assess the risk of bias. Two of three studies enrolled children consecutively; one was a laboratory-based study and provided no clinical information. The result from Xpert MTB/RIF is automated and was considered blinded in all studies. Culture is an imperfect reference standard. There were no concerns about limitations.
b The rating of the quality of evidence may be lowered if there are important differences in the populations and tests studied, and in the expertise of those using the tests in the studies compared with the settings for which the recommendations are intended. All studies were performed at higher-level care facilities. The patient population that received the index test probably reflects the population of intended use. There were no concerns about indirectness.
c For the 3 studies, the point estimates for sensitivity were 77%, 100% and 100%, indicating some heterogeneity. The 95% confidence intervals overlapped. The estimates for specificity were 50%, 71% and 96%. There is no explanation for the heterogeneity. There were serious concerns about this evidence. The evidence was downgraded by one point.
d. Only 3 studies with a total of 172 children were included; 1 study had only 5 participants. The confidence intervals for both sensitivity and specificity were wide (crossing +/– 20%). There were very serious concerns about this evidence. However, the evidence was not downgraded further given the downgrading for inconsistency.
e One unpublished study was included. A few other studies were in progress but data were not available for analysis. The data included in the review did not allow for formal assessment of publication bias using methods such as funnel plots or regression tests; such techniques have not been found to be helpful in studies evaluating diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

Table 32. GRADE evidence profile and summary of findings: accuracy of Xpert MTB/RIF in detecting TB meningitis in children

PICO question: What is the diagnostic accuracy of Xpert MTB/RIF for detection of TB meningitis in children, where Xpert MTB/RIF is used as a replacement test for usual practice?
Participants: Children aged 0–15 years suspected of having TB meningitis
Setting: Mainly tertiary care referral hospitals and university hospitals
Target condition: TB meningitis
Reference test: Solid culture or liquid culture of cerebrospinal fluid
Number of studies (number of participants): 3 (51)
Pooled sensitivity: not enough data; pooled specificity: 95% (95% CrI, 81–99%)

| Outcome | Study design | Factors that may decrease the quality of evidence | | | | | Quality of evidence | Number of results/1000 individuals tested (95% CrI) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Limitations | Indirectness | Inconsistency | Imprecision | Publication bias | | Prevalence 10/1000 | Prevalence 50/1000 | Prevalence 100/1000 |
| True positives (individuals with TB) | Cross-sectional | Insufficient data[a] | | | | | | No data | No data | No data |
| False negatives (individuals incorrectly classified as not having TB) | Cross-sectional | Insufficient data[a] | | | | | | No data | No data | No data |
| False positives (individuals incorrectly classified as having TB) | Cross-sectional | None[a] | None[b] | None[c] | Serious (–1)[d] | Undetected[e] | Moderate ⊕⊕⊕○ | | | |
| True negatives (individuals without TB) | Cross-sectional | None[a] | None[b] | None[c] | Serious (–1)[d] | Undetected[e] | Moderate ⊕⊕⊕○ | | | |

CrI, credible interval.

a Three studies provided data about using Xpert MTB/RIF on cerebrospinal fluid to diagnose TB meningitis. There were insufficient data to calculate the sensitivity of Xpert MTB/RIF. In two studies, two out of six culture-positive children were also found to be positive using Xpert MTB/RIF. One study did not have any positive results for culture or Xpert MTB/RIF.
b The studies were all performed at higher levels of care (such as referral hospitals), which reflects the population that would benefit from using Xpert MTB/RIF on cerebrospinal fluid. There were no concern for indirectness.
c For the 3 studies, the point estimates for specificity were 86%, 97% and 100%, indicating little heterogeneity. The 95% confidence intervals overlapped. There was no concern about inconsistency.
d Only 3 studies with a total of 51 children were included in the review. The confidence intervals for both sensitivity and specificity were wide (crossing +/– 20%). There were very serious concerns about imprecision. The evidence was not downgraded further given the downgrading for inconsistency.
e One unpublished study was included. A few other studies were in progress but data were not available for analysis. The data included in the review did not allow for formal assessment of publication bias using methods such as funnel plots or regression tests; such techniques have not been found to be helpful in studies evaluating diagnostic accuracy. However, reporting bias was considered to be minimal since Xpert MTB/RIF is a new test that has had considerable attention and scrutiny.

## 4.4. Affordability and cost effectiveness of using Xpert MTB/RIF to diagnose TB

Twelve published papers were identified that compared the costs of using Xpert MTB/RIF and follow-on tests to diagnose TB and MDR-TB with the current diagnostic algorithm for diagnosing TB and MDR-TB. The setting for most of these analyses was South Africa (10 studies); 2 of these studies also included other countries in sub-Saharan Africa (Botswana, Lesotho, Namibia, Swaziland and Uganda); 1 study included countries in the Former Soviet Union; and 1 global analysis included all countries. A list of the studies included in the review is given in *Annex 7*.

Seven of the 12 studies were cost analyses, and 5 were cost-effectiveness analyses.

### *4.4.1. Summary of cost analyses*

Of the seven cost analyses, six were from South Africa and one was a global analysis. One study adopted a societal perspective. Five focused on a screening algorithm that used Xpert MTB/RIF for all individuals with signs and symptoms of TB (that is, all people suspected of having TB); and one study focused on using Xpert MTB/RIF for HIV-positive individuals suspected of having TB who initially had a negative result from Xpert MTB/RIF. Two of these studies assessed the budgetary impact for South Africa of scaling up at the national level the use of Xpert MTB/RIF for all people suspected to have TB. All seven studies used non-empirical methods of data collection; the main source for information about the costs of implementing Xpert MTB/RIF were the South African National Health Laboratory Service and WHO's manual on rapid implementation.[20]

The main conclusions of these analyses were:

- the cost of using Xpert MTB/RIF added 35% to the budget for TB in South Africa, but this cost represented only 2% of the public national health budget;
- the cost of using Xpert MTB/RIF at the level of point-of-treatment (the clinic level) was 51% more expensive than if it were used at subdistrict laboratories;
- the cost of using Xpert MTB/RIF for smear-negative individuals suspected of having TB was lower than using Xpert MTB/RIF for all individuals suspected of having TB;
- using Xpert MTB/RIF for smear-negative individuals suspected of having TB can be cost-saving for patients;
- the cost of using Xpert MTB/RIF for individuals who initially had a negative result from Xpert MTB/RIF was less than using culture;
- at the global level, the cost of using Xpert MTB/RIF to diagnose MDR-TB and TB in HIV-positive individuals was less than the cost of conventional diagnostics. On the other hand, the cost of using Xpert MTB/RIF to diagnose TB in all individuals suspected of having TB was almost five times higher than the cost of using conventional diagnostics.

### *4.4.2. Summary of cost–effectiveness analyses*

The screening algorithm for three of these studies included all individuals with signs and symptoms of TB; the two other studies focused on HIV-positive individuals about to start antiretroviral therapy. All five adopted a provider-cost perspective. Only one study (Vassall 2011) empirically collected cost data from sites piloting the use of Xpert MTB/RIF; the other four used the costs of Xpert MTB/RIF found by Vassall 2011. All five cost–effectiveness analyses found that using Xpert MTB/RIF to diagnose TB and as a follow-on test to diagnose MDR-TB was cost effective compared with the current approach in the settings where the studies were conducted. The current approach differed in each setting, and the costing methods were different in each study. Therefore, although all five

20 *Rapid implementation of the Xpert MTB/RIF diagnostic test. Technical and operational 'how-to': practical considerations.* Geneva, World Health Organization, 2011 (http://whqlibdoc.who.int/publications/2011/9789241501569_eng.pdf).

studies of cost effectiveness included incremental cost–effectiveness ratios, it is challenging to compare their data.

#### *4.4.2.1. Cost per Xpert MTB/RIF test*

The only study that directly measured cost data (using full-site costing) found that the overall cost per Xpert MTB/RIF test was between US$ 23.00 and US$ 28.00, with a cartridge price of US$ 19.40; when the cartridge price was US$ 11.70, the cost per test was between US$ 15.00 and US$ 20.00. The other three cost-analysis studies used market prices to estimate the cost of each Xpert MTB/RIF test; these found that the cost per test was between US$ 22.00 and US$ 39.00, the range reflecting different prices for cartridges and where the machines were placed.

#### *4.4.2.2. Affordability*

Two studies analysed affordability: one in South Africa and the other one at the global level, focusing on the 36 countries with a high burden of TB and MDR-TB. In South Africa, the increased annual costs including national coverage of the test, represented only 2% of the public health budget.

Worldwide, the cost of using Xpert MTB/RIF for all individuals suspected of having TB is about five times higher than it is for WHO's recommended diagnostics. But for countries in Europe that have a high burden of MDR-TB, as well as for Brazil and South Africa, the cost of Xpert MTB/RIF is less than 10% of the available funding for TB, and less than 0.2% of national health spending.

Worldwide, Xpert MTB/RIF costs less than WHO's recommended diagnostic tests for TB in people living with HIV. In countries in Africa that have a high burden of TB, the cost of using Xpert MTB/RIF is equivalent to only 1–3% of the funding approved for the operations of the United States President's Emergency Plan for AIDS Relief (known as PEPFAR). Worldwide, the cost of using Xpert MTB/RIF (with follow-on tests) for diagnosing MDR-TB is less than that for WHO's recommended diagnostics for MDR-TB. In countries with a high burden of MDR-TB, the cost of using Xpert MTB/RIF is equivalent to only 4% of the funding available for TB.

### *4.4.3. Summary of findings*

- Using Xpert MTB/RIF to diagnose TB and MDR-TB was found to be cost effective when compared with current practices for all individuals suspected of having TB and for HIV-positive individuals suspected of having TB.
- The majority of published studies refer only to the situation in South Africa.
- Only one of the five cost–effectiveness analyses directly measured the costs of laboratory resources.
- The cost of each Xpert MTB/RIF test was around US$ 15.00–39.00, depending on the cost of the cartridge, and where the machines were placed.
- The use of Xpert MTB/RIF could save costs for TB patients.
- Using Xpert MTB/RIF was more costly than following current practices, but the increased costs represented only a small share of the funding available for TB, and an even smaller share of the funding approved for PEPFAR's operations, and an even a smaller share of national health spending.

### *4.4.4. Recommendations*

- More directly measured costing evidence is needed to improve cost–effectiveness analyses and the recommendations made as a result of their findings.
- These analyses need to enlarge the sample of countries where studies are done, and they need to include low-income countries.
- Assessments of the impacts on budgets and analyses of affordability are needed when studies are done at the country level.
- The design of costing and cost–effectiveness studies should be standardized to facilitate comparisons of their findings.

## Annex 1. Meeting participants

**Sevim Ahmedov**
Senior Tuberculosis Technical Advisor
United States Agency for International Development
1300 Pennsylvania Ave NW
Washington, DC 20523
E-mail: sahmedov@usaid.gov

**Lucia Barrera**
Servicio Micobacterias (Mycobacteria Laboratory)
Velez Sarsfield 563
1281 Buenos Aires
Argentina
E-mail: lubarrera2000@yahoo.com.ar

**Catharina Boehme**
Senior Medical Officer
Foundation for Innovative New Diagnostics
Avenue de Budé 16
1202 Geneva
Switzerland
E-Mail: catharina.boehme@finddiagnostics.org

**Lucy Cheshire**
Tuberculosis Advocacy Consortium
Adalyn Court
Suite C6
Ngong Road
Nairobi
Kenya
E-mail: lucy@tbadvocacy.org

**Gavin Churchyard**
The Aurum Institute NPC
The Ridge
29 Queens Road
Parktown
Johannesburg
South Africa
E-mail: gchurchyard@auruminstitute.org

**Daniela Maria Cirillo**
Head
Emerging Bacterial Pathogens Unit
San Raffaele Scientific Institute
Via Olgettina 58
20132 Milan
Italy
E-mail: cirillo.daniela@hsr.it

**Frank Cobelens**
Amsterdam Institute of Global Health and Development &
Department of Global Health
Academic Medical Center
Amsterdam
The Netherlands
E-mail: f.cobelens@aighd.org

**Bill Coggin**
Department of State/PEPFAR
Office of the US Global AIDS Coordinator
2100 Pennsylvania Ave NW
Washington, DC 20522
United States
E-mail: CogginWL@state.gov

**Colleen Daniels**
Director TB/HIV
Treatment Action Group
New York, NY 10016
United States
E-mail:
colleen.daniels@treatmentactiongroup.org

**Claudia Denkinger**
McGill University
4075 Rue Cartier
Montreal, Quebec
Canada
E-mail: cdenking@bidmc.harvard.edu

**Anne Detjen**
Technical Consultant
North America Office
The International Union Against Tuberculosis and Lung Disease
61 Broadway, Suite 1720
New York, NY 10006
United States
E-mail: adetjen@theunion.org

**Mildred Fernando**
Management Sciences for Health, Inc.
181 Floro Subdivision
Malhacan, Meycauayan
Bulacan
Philippines
E-mail: mildspirit@gmail.com

**Nazir Ismail**
Head of Centre for Tuberculosis
National Institute for Communicable Diseases
1 Modderfontein Road
Sandringham
Johannesburg
South Africa
E-mail: naziri@nicd.ac.za

**Moses Joloba**
Department of Medical Microbiology
Makerere University College of Health Sciences
2nd Floor Pathology/Microbiology BL
Room C31, Upper Mulago Hill Road
256 Kampala
Uganda
E-mail: moses.joloba@case.edu

**Anna Mandalakas**
Director
Global Tuberculosis and Mycobacteriology Program
The Tuberculosis Initiative
Texas Children's Hospital
1102 Bates St, FC-630
Houston, Texas 77030
United States
E-mail: anna.mandalakas@bcm.edu

**Andrea Pantoja**
Ankenhofstrasse 23
Oberengstringen 8102
Switzerland
E-mail: andreagpantoja@gmail.com

**Holger Schünemann**
Chair
Department of Clinical Epidemiology & Biostatistics
McMaster University Health Sciences Centre
Room 2C10B
1200 Main Street West
Hamilton, Ontario ON L8N 3Z5
Canada
E-mail: schuneh@mcmaster.ca

**Moorine Penniah Sekadde**
National Tuberculosis and Leprosy Programme
Kampala
Uganda
E-mail: moorine.sekadde@gmail.com

**Thomas M Shinnick**
Associate Director for Global Laboratory Activities
Centers for Disease Control and Prevention
1600 Clifton Road
MS-G35, NE
Atlanta, GA 30333
United States
E-mail: tms1@cdc.gov

**Karen Steingart**
Editor
Cochrane Infectious Diseases Group
United States
E-mail: karen.steingart@gmail.com

**Sabira Tahseen**
National coordinator
National TB Control Programme
Islamabad
Pakistan
E-mail: sabira.tahseen@gmail.com

**Maarten van Cleeff**
KNCV Tuberculosis Foundation
Parkstraat 17
2514 JD The Hague
The Netherlands
E-mail: vancleeffm@kncvtbc.nl

**Francis Varaine**
Médecins sans Frontières
8 rue St Sabin
75011 Paris
France
E-mail: francis.varaine@paris.msf.org

**WHO staff members**
**Haileyesus Getahun**
E-mail: getahunh@who.int

**Christopher Gilpin**
E-mail: gilpinc@who.int

**Jean Iragena**
E-mail: iragenaj@who.int

**Knut Lönnroth**
Email: lonnrothk@who.int

**Lisa Nelson**
Email: nelsonl@who.int

**Fuad Mirzayev**
E-mail: mirzayevf@who.int

**Wayne van Gemert**
Email: vangemertw@who.int

**Karin Weyer**
E-mail: weyerk@who.int

**Matteo Zignol**
E-mail: zignolm@who.int

## Annex 2. Meeting agenda

**WHO Policy Guidance**

**on the utility of the Xpert MTB/RIF assay for the diagnosis of pulmonary and extra-pulmonary TB in adults and children**

**EXPERT GROUP MEETING**

Date: 20-21 May 2012
Venue: Les Pensierès, Veyrier-du-Lac, France

### Background

The World Health Organization (WHO) in 2011 issued a policy statement recommending the Xpert MTB/RIF assay (Cepheid, Sunnyvale, CA, USA) for the diagnosis of tuberculosis (TB) and rifampicin (RIF) resistance as a proxy for multidrug-resistant TB (MDR-TB). Xpert MTB/RIF should be used as an initial diagnostic test in individuals suspected of MDR or HIV-associated TB. It should be used as an add-on test to smear microscopy in settings where MDR-TB or HIV are of lesser concern, especially in smear-negative specimens. Generalizing from adult data, the recommendation includes the use of Xpert MTB/RIF in children, acknowledging the difficulties in the microbiological diagnosis of childhood TB.

Since that time, additional studies on Xpert MTB/RIF for both pulmonary and extrapulmonary TB have been published or are being performed. An update of the 2011 Xpert MTB/RIF policy guidance is planned for 2013.

In adults, Xpert MTB/RIF has shown a high sensitivity (99%) in smear- and culture-positive individuals with pulmonary TB. HIV coinfection did not significantly affect Xpert MTB/RIF performance to detect smear-negative/culture-positive patients (sensitivity 86%). The ability of Xpert MTB/RIF to detect TB in smear-negative individuals is encouraging and suggests that it may be a valuable TB diagnostic tool in children. There are emerging data available on the utility of Xpert MTB/RIF in children. Some of these studies have shown a better performance of Xpert MTB/RIF compared with microscopy, and similar performance when compared to liquid culture.

WHO has commissioned three systematic reviews that will include available data regarding the use of Xpert MTB/RIF for the diagnosis of pulmonary and extrapulmonary TB in adults and children, and will provide evidence to inform future policy recommendations.

In accordance with current WHO standards for evidence assessment in the formulation of policy recommendations, WHO engages in a systematic, transparent process using the GRADE approach (http://www.gradeworkinggroup.org). GRADE provides a structured framework for evaluating diagnostic test accuracy and the patient/public health impact of new diagnostic tests.

## Meeting objectives

- To review the evidence base and evaluate data from an updated systematic review to assess the accuracy of Xpert MTB/RIF assay for the diagnosis of pulmonary TB and rifampicin resistance in adults;
- To review the evidence base and evaluate data from a systematic review on the accuracy of the Xpert MTB/RIF assay for the diagnosis of tuberculosis on nonrespiratory samples;
- To review the evidence base and evaluate data from a systematic review on the accuracy of the Xpert MTB/RIF assay for the diagnosis of tuberculosis and rifampicin resistance in children;
- To review the evidence on the cost effectiveness and affordability of the Xpert MTB/RIF in different epidemiological and resource settings;
- To outline issues to be addressed by WHO in subsequent policy recommendations.

## Expected outcomes

- Evidence-based recommendations on the accuracy of the Xpert MTB/RIF assay for the diagnosis of pulmonary TB and rifampicin resistance in adults;
- Evidence-based recommendations on the accuracy of the Xpert MTB/RIF assay for the diagnosis of tuberculosis on nonrespiratory samples;
- Evidence-based recommendations on the accuracy of the Xpert MTB/RIF assay for the diagnosis of tuberculosis and rifampicin resistance in children;
- Evidence-based recommendations on the cost effectiveness and affordability of Xpert MTB/RIF in different epidemiological and resource settings;
- Consensus on issues to be addressed in development of subsequent WHO policy recommendations.

## Provisional agenda

**Monday, 20 May 2013: Expert Group Meeting on Xpert MTB/RIF DAY 1**
Chairperson: H Schünemann
WHO Secretariat: C Gilpin
Rapporteur: W van Gemert

| | | |
|---|---|---|
| **Opening session** | | |
| **09:00 – 09:10** | Welcome | K Weyer |
| **09:10 – 09:20** | Meeting scope and objectives | C Gilpin |
| **09:20 – 09:30** | Declarations of Interests by Expert Group members | C Gilpin |
| **09:30 – 10:00** | Grading the quality of evidence and strength of recommendations: Brief overview of GRADE | H Schünemann |
| **10:00 – 10:15** | Questions | |
| **10:15 – 10:45** | BREAK | |
| **Session 1: Xpert MTB/RIF for diagnosis of pulmonary TB** | | |
| **10:45 – 11:00** | Current WHO policy guidance on Xpert MTB/RIF | Wayne van Gemert |
| **11:00 – 11:15** | PICO questions and rating outcomes | Chris Gilpin |
| **11:15 – 12:00** | Updated systematic review: Xpert MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults | Karen Steingart |
| **12:00 – 13:00** | Discussion<br>1. What is the diagnostic accuracy of Xpert MTB/RIF for detection of pulmonary TB in adults, where Xpert MTB/RIF is used as a replacement test for smear microscopy?<br>2. What is the diagnostic accuracy of Xpert MTB/RIF for detection of pulmonary TB in adults, where Xpert MTB/RIF is used as an add-on test following a negative smear microscopy result?<br>3. What is the diagnostic accuracy of Xpert MTB/RIF for detection of smear-positive pulmonary TB in adults?<br>4. What is the diagnostic accuracy of Xpert MTB/RIF for detection of smear-negative (culture-positive) pulmonary TB in adults?<br>5. What is the diagnostic accuracy of Xpert MTB/RIF for detection of pulmonary TB in people living with HIV (adults)?<br>6. What is the diagnostic accuracy of Xpert MTB/RIF for detection of pulmonary TB in adults without HIV infection?<br>7. What is the diagnostic accuracy of Xpert MTB/RIF for detection of rifampicin resistance, where Xpert MTB/RIF is used as an initial test replacing phenotypic culture-based drug-susceptibility testing? | All |
| **13:00 – 14:00** | LUNCH | |
| **14:00 – 15:00** | Draft recommendations:<br>Xpert MTB/RIF in pulmonary TB | H Schünemann |

| Session 2: Xpert MTB/RIF for diagnosis of extrapulmonary TB | | |
|---|---|---|
| **15:00 – 15:45** | Systematic review: Xpert MTB/RIF assay for tuberculosis on nonrespiratory samples | Claudia Denkinger |
| **15:45 – 16:00** | BREAK | |
| **16:00– 17:00** | Discussion<br>1. What is the diagnostic accuracy of Xpert MTB/RIF for TB detection in lymph node fluid and tissue, where Xpert MTB/RIF is used as a replacement test for usual practice?<br>2. What is the diagnostic accuracy of Xpert MTB/RIF for TB detection in pleural fluid, where Xpert MTB/RIF is used as a replacement test for usual practice?<br>3. What is the diagnostic accuracy of Xpert MTB/RIF for TB detection in cerebrospinal fluid (CSF), where Xpert MTB/RIF is used as a replacement test for usual practice?<br>4. What is the diagnostic accuracy of Xpert MTB/RIF for TB detection in gastric fluid, where Xpert MTB/RIF is used as a replacement test for usual practice?<br>5. What is the diagnostic accuracy of Xpert MTB/RIF for TB detection in tissue samples, where Xpert MTB/RIF is used as a replacement test for usual practice?<br>6. What is the diagnostic accuracy of Xpert MTB/RIF for rifampicin resistance detection in nonrespiratory specimens, where Xpert MTB/RIF is used as an initial test replacing phenotypic culture-based drug-susceptibility testing? | All |
| **17:00 – 18:00** | Draft recommendations:<br>Xpert MTB/RIF in extrapulmonary TB | H Schünemann |
| **19:00** | Dinner reception | |

**Tuesday, 21 May 2013: Expert Group Meeting on Xpert MTB/RIF DAY 2**
Chairperson: H Schünemann
WHO Secretariat: C Gilpin
Rapporteur: W van Gemert

| | | |
|---|---|---|
| **Session 3: Xpert MTB/RIF in paediatric TB** | | |
| **09:00 – 9:45** | Systematic review: Xpert MTB/RIF assay for tuberculosis and rifampicin resistance in children | A Detjen |
| **9:45 – 10:45** | Discussion<br>1. What is the diagnostic accuracy of Xpert MTB/RIF for detection of TB in children compared with culture, where Xpert MTB/RIF is used as a replacement test for usual practice?<br>2. What is the diagnostic accuracy of Xpert MTB/RIF for detection of TB in children compared with a combined clinical and laboratory reference standard, where Xpert MTB/RIF is used as a replacement test for usual practice?<br>3. What is the diagnostic accuracy of Xpert MTB/RIF for detection of TB in children, where Xpert MTB/RIF is used as an add-on test following a negative smear-microscopy result?<br>4. What is the diagnostic accuracy of Xpert MTB/RIF compared with smear microscopy for detection of TB in children?<br>5. What is the diagnostic accuracy of Xpert MTB/RIF for detection of rifampicin resistance in children, where Xpert MTB/RIF is used as an initial test replacing phenotypic culture-based drug-susceptibility testing?<br>6. What is the diagnostic accuracy of Xpert MTB/RIF for detection of peripheral lymph node TB in children, where Xpert MTB/RIF is used as a replacement test for usual practice?<br>7. What is the diagnostic accuracy of Xpert MTB/RIF for detection of TB meningitis in children, where Xpert MTB/RIF is used as a replacement test for usual practice? | All |
| **10:45 – 11:00** | BREAK | |
| **11:00 – 12:00** | Draft recommendations:<br>Xpert MTB/RIF in paediatric TB | H Schünemann |
| **12:00 – 13:00** | LUNCH | |
| **Session 4: Cost effectiveness and affordability of Xpert MTB/RIF** | | |
| **13:00 – 13:45** | Review: Cost-effectiveness and resource implications for Xpert MTB/RIF implementation | A Pantoja |
| **13:45 – 14:45** | Discussion | All |
| **14:45 – 15:30** | Draft recommendations | H Schünemann |
| **15:30 – 15:45** | BREAK | |

| Session 5: Review of GRADE summaries and formulation of final recommendations | | |
|---|---|---|
| **15:45 – 16:15** | Discussion of issues to be addressed by WHO in subsequent policy recommendations on Xpert MTB/RIF | All |
| **16:15 – 16:45** | Review of GRADE process and summaries | H Schünemann |
| **16:45 – 17:30** | Final recommendations | H Schünemann |

## Annex 3. Declarations of Interests

**None declared**

Lucia Barrera, Lucy Cheshire, Colleen Daniels, Holger Schünemann (Chairperson),

Moorine Penniah Sekadde, Sabira Tahseen, Maarten van Cleeff

**Declared, insignificant**

Daniela Maria Cirillo: meeting participation sponsored by Cepheid; research grant support from FIND and ST microelectronics.

Mildred Fernando: Management Sciences for Health contributed approximately US$ 80 to enable her to participate in the Expert Group meeting; the Xpert MTB/RIF assay may have been used to determine if her TB had relapsed.

Thomas M Shinnick: received funding from the United States Government to participate in the meeting.

Francis Varaine: leader of the Médecins sans Frontières working group on TB; required to defend positions related to TB diagnostics.

**Declared, significant (observer status)**

Sevim Ahmedov: United States government employee; the US government has contributed funds to enable reductions in the end-user price of cartridges used for the Xpert MTB/RIF test.

Catharina Boehme: employed by FIND, which had a cofunding agreement with Cepheid supporting the development of Xpert MTB/RIF.

Gavin Churchyard: employed by the Aurum Institute, which received funding from the Bill and Melinda Gates Foundation for trials using the Xpert MTB/RIF assay to diagnose TB in South Africa, and to evaluate the impact of the test and its cost effectiveness during routine roll-out of the test.

Frank Cobelens: consultant to the Bill and Melinda Gates Foundation for research on rolling out and scaling up the use of Xpert MTB/RIF; also conducted a cost-effectiveness analysis of Xpert MTB/RIF for FIND.

Bill Coggin: United States government employee; the US government has contributed funds to enable reductions in the end-user price of cartridges used in the Xpert MTB/RIB test.

Claudia Denkinger: conducted the systematic review of using Xpert MTB/RIF to diagnose extrapulmonary TB

Anne Detjen: conducted the systematic review of using Xpert MTB/RIF to diagnose paediatric TB

Nazir Ismail: employed by the National Health Laboratory Service of South Africa and involved in the roll out of Xpert MTB/RIF testing in South Africa; South Africa purchased more than 1 million cartridges; Nazir Ismail received no financial gain from this process.

Moses Joloba: laboratory manger of the National TB Reference Laboratory in Kampala which was used by FIND as one testing site for the initial Xpert MTB/RIF demonstration studies.

Anna Mandalakas: conducted the systematic review of using Xpert MTB/RIF to diagnose paediatric TB.

Andrea Pantoja: conducted the review of the affordability and cost effectiveness of Xpert MTB/RIF.

Karen Steingart: conducted the systematic review of using Xpert MTB/RIF to diagnose pulmonary TB.

## Annex 4. Selection of studies evaluating the accuracy of Xpert MTB/RIF in diagnosing pulmonary TB and rifampicin resistance in adults

Figure 1. Flow diagram of studies identified by the initial literature searches

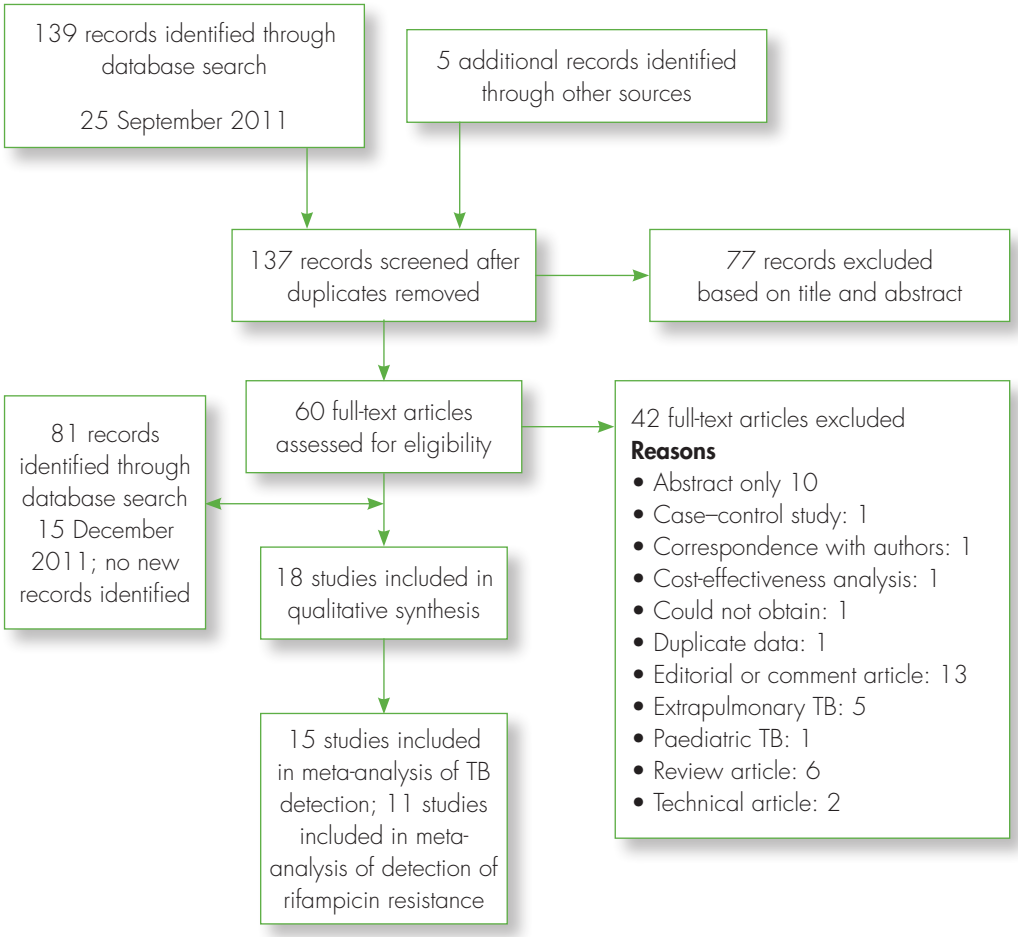
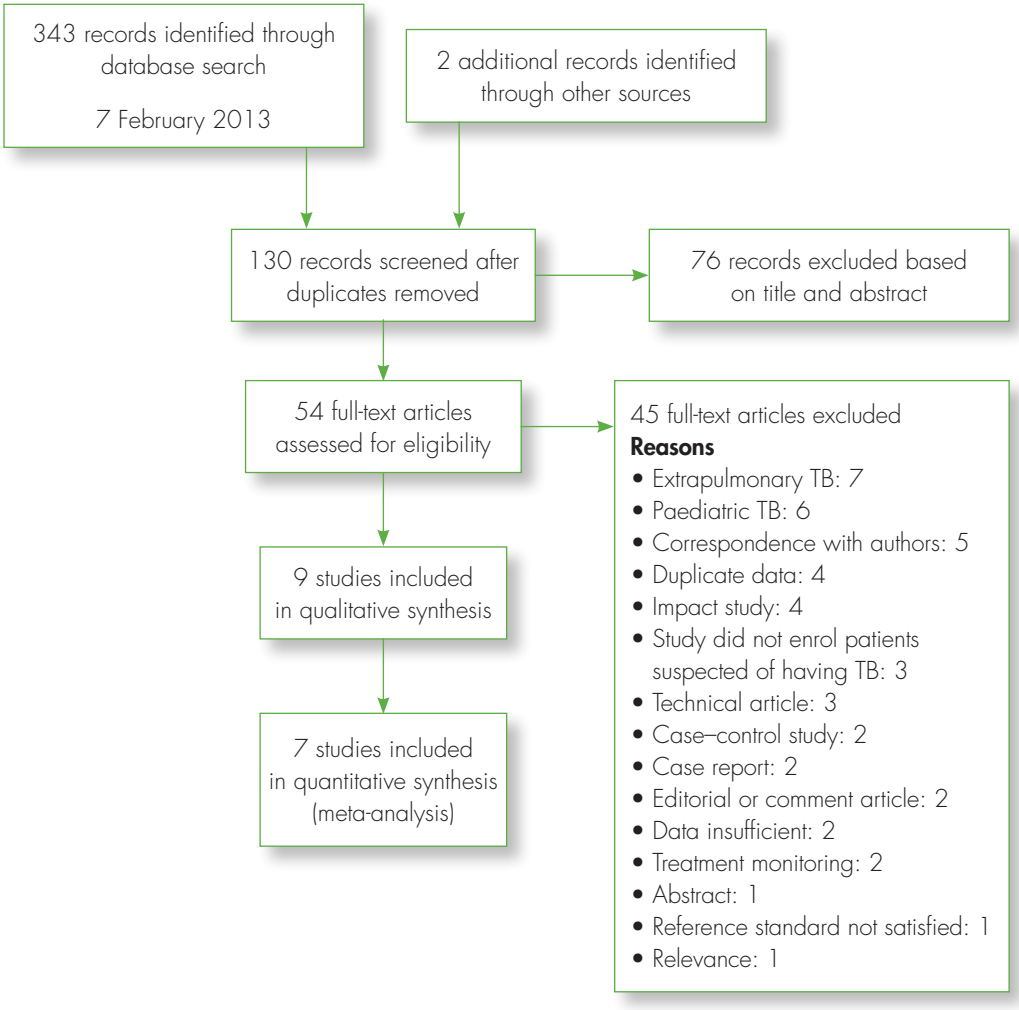


Figure 2. Flow diagram of studies identified by the updated literature search




## Included studies

### *Published studies*

1. 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 MTB/RIF system. *Saudi Medical Journal*, 2012, 33:1100–1105.

2. Balcells ME et al. Rapid molecular detection of pulmonary tuberculosis in HIV-infected patients in Santiago, Chile. *International Journal of Tuberculosis and Lung Disease*, 2012, 16:1349–1353.

3. Barnard M et al. The diagnostic performance of the GenoType MTBDRplus version 2 line probe assay is equivalent to that of the Xpert MTB/RIF assay. *Journal of Clinical Microbiology*, 2012, 50:3712–3716.

4. Boehme CC et al. Rapid molecular detection of tuberculosis and rifampin resistance. *New England Journal of Medicine*, 2010, 363:1005–1015.

5. Boehme CC et al. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *Lancet*, 2011, 377:1495–1505.

6. Bowles EC et al. Xpert MTB/RIF, a novel automated polymerase chain reaction-based tool for the diagnosis of tuberculosis. *International Journal of Tuberculosis and Lung Disease* 2011, 15:988–989.

7. Carriquiry G et al. A diagnostic accuracy study of Xpert MTB/RIF in HIV-positive patients with high clinical suspicion of pulmonary tuberculosis in Lima, Peru. *PLoS One*, 2012, 7:e44626 (doi:10.1371/journal.pone.0044626).

8. Ciftci IH, Aslan MH, Asik G. [Evaluation of Xpert MTB/RIF results for the detection of Mycobacterium tuberculosis in clinical samples]. *Mikrobiyoloji Bülteni*, 2011, 45:43–47.

9. Friedrich SO et al. Suitability of Xpert MTB/RIF and genotype MTBDRplus for patient selection for a tuberculosis clinical trial. *Journal of Clinical Microbiology*, 2011, 49:2827–2831.

10. Hanif SNM et al. GeneXpert MTB/RIF for rapid detection of Mycobacterium tuberculosis in pulmonary and extra-pulmonary samples. *International Journal of Tuberculosis and Lung Disease*, 2011,15:1274–1275.

11. Helb D et al. Rapid detection of Mycobacterium tuberculosis and rifampin resistance by use of on-demand, near-patient technology. *Journal of Clinical Microbiology*, 2010, 48:229–237.

12. Ioannidis P 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, 49:3068–3070.

13. Kurbatova EV et al. Performance of Cepheid Xpert MTB/RIF and TB-Biochip MDR in two regions of Russia with a high prevalence of drug-resistant tuberculosis. European Journal of Clinical Microbiology and Infectious Diseases, 2013, 32:735–743.

14. Lawn SD et al. Screening for HIV-associated tuberculosis and rifampicin resistance before antiretroviral therapy using the Xpert MTB/RIF assay: a prospective study. *PLoS Medicine*, 2011, 8:e1001067 doi:10.1371/journal.pmed.1001067).

15. Malbruny B et al. Rapid and efficient detection of Mycobacterium tuberculosis in respiratory and non-respiratory samples. *International Journal of Tuberculosis and Lung Disease*, 2011, 15:553–555.

16. Marlowe EM et al. Evaluation of the Cepheid Xpert MTB/RIF assay for direct detection of Mycobacterium tuberculosis complex in respiratory specimens. *Journal of Clinical Microbiology*, 2011, 49:1621–1623.

17. Miller MB et al. Performance of Xpert MTB/RIF RUO assay and IS6110 real-time PCR for Mycobacterium tuberculosis detection in clinical samples. *Journal of Clinical Microbiology*, 2011, 49:3458–3462.

18. Moure R et al. 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, 49:1137–1139.

19. Rachow A et al. Rapid and accurate detection of Mycobacterium tuberculosis in sputum samples by Cepheid Xpert MTB/RIF MTB/RIF assay–a clinical validation study. *PLoS One*, 2011, 6:e20458 (doi: 10.1371/journal.pone.0020458).

20. Safianowska A et al. [Two selected commercially based nucleic acid amplification tests for the diagnosis of tuberculosis]. *Pneumonologia i Alergologia Polska*, 2012, 80:6–12.

21. Scott LE et al. Comparison of Xpert MTB/RIF with other nucleic acid technologies for diagnosing pulmonary tuberculosis in a high HIV prevalence setting: a prospective study. *PLoS Medicine*, 2011, 8:e1001061 (DOI: 10.1371/journal.pmed.1001061).

22. Teo J et al. Comparison of two nucleic acid amplification assays, the Xpert MTB/RIF and the Amplified Mycobacterium Tuberculosis Direct (MTD) assay, for the detection of Mycobacterium tuberculosis in respiratory and non-respiratory specimens. *Journal of Clinical Microbiology,* 2011, 49:3659–3662.

23. Theron G et al. Evaluation of the Xpert MTB/RIF assay for the diagnosis of pulmonary tuberculosis in a high HIV prevalence setting. *American Journal of Respiratory and Critical Care Medicine*, 2011, 184:132–140.

24.Van Rie A et al. Point-of-care Xpert MTB/RIF® for smear-negative tuberculosis suspects at a primary care clinic in South Africa. *International Journal of Tuberculosis and Lung Disease*, 2013, 17:368–372.

25. Williamson DA et al. An evaluation of the Xpert MTB/RIF assay and detection of false-positive rifampicin resistance in Mycobacterium tuberculosis. *Diagnostic Microbiology and Infectious Disease*, 2012, 74:207–9.

26. Zeka AN, Tasbakan S, Cavusoglu C. Evaluation of the GeneXpert MTB/RIF assay for the rapid diagnosis of tuberculosis and detection of rifampin resistance in pulmonary and extrapulmonary specimens. *Journal of Clinical Microbiology*, 2011, 49:4138–4141.

### *Unpublished studies*

1. Hanrahan CF et al. Xpert MTB/RIF at point of care. 2013 (submitted).

## Excluded studies with reasons for exlusion

**Excluded studies were identified during electronic searches conducted on 25 September 2011, 15 December 2011 and 7 February 2013; the studies may have appeared in more than one search.**

1. Alvarez-Uria G et al. 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,* 2012:932862 (http://dx.doi.org/10.1155/2012/932862) . *Reference standard not satisfied.*

2. Andersen AB et al. [Treatment and diagnostics of tuberculosis: moving slowly forward]. Ugeskrift for Laeger, 2011, 173:897–899. *Editorial and comment only.*

3. Armand S et al. 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, 49:1772–1776. *This was a case–control study that compared Xpert MTB/RIF assay with an in-house IS6110-based real-time PCR using TaqMan probes (IS6110-TaqMan assay) to detect TB.*

4. Banada PP 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, 48:3551–3557. *Technical paper only.*

5. Bates M et al. Assessment of the Xpert MTB/RIF assay for diagnosis of tuberculosis with gastric lavage aspirates in children in sub-Saharan Africa: a prospective descriptive study. *Lancet Infectious Diseases*, 2013, 13:36–42. *This study evaluated the accuracy of using Xpert MTB/RIF to diagnose TB in children.*

6. Blakemore R et al. Evaluation of the analytical performance of the Xpert MTB/RIF assay. *Journal of Clinical Microbiology*, 2010, 48:2495–2501. Technical paper only.

7. Blakemore R et al. A multi-site assessment of the quantitative capabilities of the Xpert MTB/RIF assay. *American Journal of Respiratory Critical Care Medicine*, 2011, 184:1076–1084. *This was a technical paper that compared bacterial load quantitation determined by Xpert MTB/RIF with the load determined by conventional quantitative methods.*

8. Bodmer T, Strhle A. Diagnosing pulmonary tuberculosis in a low prevalence setting – the Xpert MTB/RIF test. Special Issue: Abstracts of the 20th ECCMID, Vienna, Austria, 10–13 April 2010. *Clinical Microbiology and Infection*, 2010, 16(Suppl. S2):S609–S610. *Abstract only.*

9. Causse M et al. Comparison of two molecular methods for rapid diagnosis of extrapulmonary tuberculosis. *Journal of Clinical Microbiology,* 2011, 49:3065–3067. *This study evaluated using Xpert MTB/RIF to diagnose extrapulmonary TB.*

10. Cavusolu C, Nazli A. Evaluation of the GeneXpert MTB/RIF assay for the rapid detection of rifampicin resistance in Mycobacterium tuberculosis isolates. Special Issue: Abstracts of the 20th ECCMID, Vienna, Austria, 10–13 April 2010. *Clinical Microbiology and Infection*, 2010, 16 (Suppl. S2):S53. *Abstract only.*

11. Chee CBE. Recent advancements in diagnosis of TB. *International Journal of Infectious Diseases*, 2011, 15 (Suppl.):S9–S10 (http://download.journals.elsevierhealth.com/pdfs/journals/1201-9712/PIIS1201971211600398.pdf). *Abstract only.*

12. Chegou NN et al. Tuberculosis assays: past, present and future. *Expert Review of Anti-Infective Therapy*, 2011, 9:457–469. Narrative review.

13. Clouse K et al. Implementation of Xpert MTB/RIF for routine point-of-care diagnosis of tuberculosis at the primary care level. *South African Medical Journal*, 2012, 102:805–807. Study on patient impact.

14. Cuevas LE. The urgent need for new diagnostics for symptomatic tuberculosis in children. *Indian Journal of Pediatrics*, 2011, 78:449–455. *Narrative review.*

15. Deforges L et al. Application of the rapid detection system for M. tuberculosis complex and rifampicin resistance Xpert MTB/RIF in decontaminated respiratory specimens, non-respiratory specimens and cultures. Special Issue: Abstracts of the 20th ECCMID, Vienna, Austria, 10–13 April 2010. *Clinical Microbiology and Infection*, 2010, 16(Suppl. S2):S53. *Abstract only.*

16. Dorjee K, Salvo F, Dierberg KL. Xpert MTB/RIF diagnosed disseminated smear-negative MDR-TB in a sub-district hospital in India. *International Journal of Tuberculosis and Lung Disease,* 2012, 16:1560–1561. *Case report.*

17. Dorman SE et al. Performance characteristics of the Cepheid Xpert MTB/RIF test in a tuberculosis prevalence survey. *PLoS One* 2012, 7:e43307 (doi:10.1371/journal.pone.004330). *Prevalence survey.*

18. Dowdy DW et al. Is scale-up worth it? Challenges in economic analysis of diagnostic tests for tuberculosis. *PLoS Medicine*, 2011, 8:e1001063 (doi:10.1371/journal.pmed.1001063). *Cost-effectiveness analysis.*

19. Evans C A. GeneXpert MTB/RIF – a game-changer for tuberculosis control? *PLoS Medicine*, 2011, ;8:e1001064 (doi:10.1371/journal.pmed.1001064). *Editorial and comment.*

20. Victorino Farga C. [New challenges in tuberculosis]. *Revista chilena de enfermedades respiratorias,* 2011, 27:161–168. *Narrative review.*

21. Fenner L, Boulle A, Egger M. In reply to 'Pre-screening with GeneXpert MTB/RIF may increase use of isoniazid preventive therapy in antiretroviral programmes'. 2011, 15:1273–1274. *Editorial and comment.*

22. Ferrara G et al. Xpert MTB/RIF test for tuberculosis. *Lancet,* 2011, 378:482; *author reply* 2011, 378:482–483. *Editorial and comment.*

23. Friedrich SO, von Groote-Bidlingmaier F, Diacon AH. Xpert MTB/RIF assay for the diagnosis of pleural tuberculosis. *Journal of Clinical Microbiology,* 2011, 49:4341–4342. *This study evaluated using Xpert MTB/RIF to diagnose of pleural TB, a form of extrapulmonary TB.*

24. Gotuzzo E. Xpert MTB/RIF MTB/RIF for diagnosis of pulmonary tuberculosis. *Lancet Infectious Diseases*, 2011, 11:802–803. *Editorial and comment.*

25. Hesseling AC, Graham SM, Cuevas LE. Rapid molecular detection of tuberculosis. *New England Journal of Medicine*, 2011, 364:183–184; author reply 2011, 364:184–185. *Editorial and comment.*

26. Hillemann D et al. Rapid molecular detection of extrapulmonary tuberculosis by the automated GeneXpert MTB/RIF system. *Journal of Clinical Microbiology*, 2011, 49:1202–1205. *This study evaluated using Xpert MTB/RIF to diagnose extrapulmonary TB.*

27. Hoek KG et al. Detecting drug-resistant tuberculosis: the importance of rapid testing. *Molecular Diagnosis & Therapy*, 2011, 15:189–194. Narrative review.

28. Ioannidis P 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.] *Acta Microbiologica Hellenica*, 2010, 55:175–182. *Could not obtain a copy of the study.*

29. Jones M et al. Rapid detection of Mycobacterium tuberculosis and rifampin-resistance from sputum samples with an easy-to-use *PCR test system with near-patient capability*. Clinical Microbiology and Infection, 2009, 15(Suppl. 4):S41. *Abstract only.*

30. Kim SY et al. The Xpert MTB/RIF assay evaluation in South Korea, a country with an intermediate tuberculosis burden. *International Journal of Tuberculosis and Lung Disease*, 2012, 16:1471–1476. *Case–control study.*

31. Kirwan DE, Cardenas MK, Gilman RH. Rapid implementation of new TB diagnostic tests: is it too soon for a global roll-out of Xpert MTB/RIF? *American Journal of Tropical Medicine and Hygiene*, 2012, 87:197–201. *Editorial and comment.*

32. Lawn SD. Pre-screening with GeneXpert MTB/RIF may increase use of isoniazid preventive therapy in antiretroviral programmes. *International Journal of Tuberculosis and Lung Disease*, 2011, 15:1272–1273. *Editorial and comment.*

33. Lawn SD, Nicol MP. Xpert MTB/RIF® assay: development, evaluation and implementation of a new rapid molecular diagnostic for tuberculosis and rifampicin resistance. *Future Microbiology*, 2011, 6:1067–1082. *Narrative review article.*

34. Lawn SD et al. Characteristics and early outcomes of patients with Xpert MTB/RIF-negative pulmonary tuberculosis diagnosed during screening before antiretroviral therapy. *Clinical Infectious Diseases*, 2012, 54:1071–1079. *Study of impacts on patient.*

35. Lawn SD et al. 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, 60:289–294. *Data insufficient for 2x2 table.*

36. Lawn SD, Kerkhoff, AD, Wood R. Characteristics of Xpert MTB/RIF -negative patients with pulmonary tuberculosis: reply. *Clinical Infectious Diseases*, 2012, 55:473–474. *Correspondence.*

37. Lawn SD et al. Clinical significance of lipoarabinomannan detection in urine using a low-cost point-of-care diagnostic assay for HIV-associated tuberculosis. *AIDS*, 2012, 26:1635–1643. *Primarily a study of lipoarabinomannan detection.*

38. Lawn SD et al. Diagnostic accuracy of a low-cost, urine antigen, point-of-care screening assay for HIV-associated pulmonary tuberculosis before antiretroviral therapy: a descriptive study. *Lancet Infectious Diseases*, 2012, 12:201–209. *Duplicate data.*

39. Ligthelm LJ et al. Xpert MTB/RIF for rapid diagnosis of tuberculous lymphadenitis from fine-needle-aspiration biopsy specimens. *Journal of Clinical Microbiology*, 2011, 49:3967–3970. *This study evaluated Xpert MTB/RIF for the diagnosis of TB lymphadenitis, a form of extrapulmonary TB.*

40. Lin JS et al. Evaluation of Xpert MTB/RIF assay and amplified Mycobacterium tuberculosis direct test in direct detection of pulmonary M. tuberculosis complex. Special Issue: Abstracts of the 20th ECCMID, Vienna, Austria, 10–13 April 2010. *Clinical Microbiology and Infection, 2010, 16(Suppl. S2):S609. Abstract only.*

41. Malbruny B et al. Rapid and efficient detection of Mycobacterium tuberculosis in respiratory and non-respiratory samples . *International Journal of Tuberculosis and Lung Disease*, 2011, 15:553–555. *Abstract only.*

42. Marlowe EM, Novak-Weekley SM, Cumpio JV. Comparison of three molecular methods for the direct detection of Mycobacterium tuberculosis complex from respiratory & non-respiratory specimens. Journal of Molecular Diagnostics 2010, Conference: Association for Molecular Pathology Annual Meeting San Jose, CA United States. Conference Start: 20101118 Conference End: 20101120. Conference Publication: (var.pagings). 12 (6):885. *Abstract only.*

43. Melzer M. An automated molecular test for Mycobacterium tuberculosis and resistance to rifampin (Xpert MTB/RIF) is sensitive and can be carried out in less than 2 h. *Evidence Based Medicine*, 2011, 16:19. *Editorial and comment.*

44. Miotto P et al. Early tuberculosis treatment monitoring by Xpert MTB/RIF. *European Respiratory Journal,* 2012, 39:1269–1271. *Treatment monitoring.*

45. Morris K. Xpert MTB/RIF TB diagnostic highlights gap in point-of-care pipeline. *Lancet Infectious Diseases*, 2010, 10:742–743. *Editorial and comment.*

46. Morris K. The new face of tuberculosis diagnosis. *Lancet Infectious Diseases*, 2011, 11:736–737. *Editorial and comment.*

47. 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, 50:513–515. *This study evaluated Xpert MTB/RIF for the diagnosis of extrapulmonary TB.*

48. Munoz L et al. GeneXpert MTB/RIF for smear-negative pulmonary tuberculosis: does it play a role in low-burden countries? *Diagnostic Microbiology and Infectious Disease*, 2012, 325–326. *Study of impacts on patients.*

49. Naidoo S. Evaluation of GeneXpert MTB/RIF assay on pulmonary and extra-pulmonary samples in a high-throughput laboratory. Special Issue: Abstracts of the 20th ECCMID, Vienna, Austria, 10–13 April 2010. *Clinical Microbiology and Infection*, 2010, 16(Suppl. S2):S604–605. *Abstract only.*

50. Narasimooloo R, Ross A. Delay in commencing treatment for MDR TB at a specialised TB treatment centre in KwaZulu-Natal. *South African Medical Journal,* 2012, 102:360–362. *Study of impacts on patients.*

51. Nhu NT 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). *This study evaluated Xpert MTB/RIF for the diagnosis of TB in children.*

52. Nicol MP et al. Accuracy of the Xpert MTB/RIF test for the diagnosis of pulmonary tuberculosis in children admitted to hospital in Cape Town, South Africa: a descriptive study. *Lancet Infectious Diseases*, 2011, 11:819–824. *This study evaluated Xpert MTB/RIF for the diagnosis of TB in children.*

53. Ntinginya EN et al. Performance of the Xpert MTB/RIF assay in an active case-finding strategy: a pilot study from Tanzania. *International Journal of Tuberculosis and Lung Disease,* 2012, 16:1468–1470. *This study included both adults and children. The study used an active case-finding strategy*

*involving previously known TB cases, and identified five additional culture-confirmed TB cases (5/219). Xpert MTB/RIF showed a positive result in all five culture-confirmed TB cases (sensitivity, 100%). We considered the study design to be different from an evaluation of the accuracy of diagnostic tests and therefore did not include this study in the review.*

54. O'Grady J et al. Evaluation of the Xpert MTB/RIF assay at a tertiary care referral hospital in a setting where tuberculosis and HIV infection are highly endemic. *Clinical Infectious Diseases*, 2012, 55:1171–1178. *This study evaluated Xpert MTB/RIF in patients able to produce sputum, irrespective of admission diagnosis, not presumed TB patients.*

55. Okumu A. Comparison of performance between Ziehl Neelsen (ZN) microscopy and the Xpert MTB/RIF assay in detection of M. tuberculosis in sputum at KEMRI/CDC TB lab. *Tropical Medicine and International Health* 2012, 17:23–24. *Abstract only.*

56. Perkins MD et al. Xpert MTB/RIF test for tuberculosis: author reply. Lancet 2011, 378:482–483. *Correspondence.*

57. Peter JG et al. 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:e39966 doi:10.1371/journal.pone.0039966). *This study evaluated Xpert MTB/RIF for the diagnosis of extrapulmonary TB.*

58. Peters D et al. Should Xpert MTB/RIF be rolled out in low-income countries? *International Journal of Tuberculosis and Lung Disease*, 2012, 16:702–703. *Correspondence.*

59. Rachow A et al. Increased and expedited case detection by Xpert MTB/RIF assay in childhood tuberculosis: a prospective cohort study. *Clinical Infectious Diseases*, 2012, 54:1388–1396. *This study evaluated Xpert MTB/RIF for the diagnosis of TB in children.*

60. Salvo F et al. Xpert MTB/RIF test for tuberculosis. *Lancet* 2011, 378:481–482; author reply, 482–483. *Editorial and comment.*

61. Scott LE et al. Dried culture spots for Xpert MTB/RIF external quality assessment: results of a phase 1 pilot study in South Africa. *Journal of Clinical Microbiology*, 2012, 49:4356–4360. *Technical paper.*

62. Swaminathan S, Rekha VV. Antigen detection as a point-of-care test for TB: the case of lipoarabinomannan. *Future Microbiology*, 2012, 7:559–564. *Not relevant to the systematic review.*

63. Tansey B. Cepheid unveils fast TB test to aid developing countries. *AIDS Reader,* 2009, 19:229. *Editorial and comment.*

64. Taylor N et al. Can a simple flotation method lower the limit of detection of Mycobacterium tuberculosis in extrapulmonary samples analyzed by the GeneXpert MTB/RIF assay? *Journal of Clinical Microbiology*, 2012, 50:2272–2276. *This study evaluated Xpert MTB/RIF for the diagnosis of extrapulmonary TB.*

65. Theron G, Peter J, Dheda K. Xpert MTB/RIF test for tuberculosis. *Lancet* 2011, 378:481; author reply, 482–483. *Editorial and comment.*

66. Theron G et al. Correlation of Mycobacterium tuberculosis specific and non-specific quantitative Th1 T-cell responses with bacillary load in a high burden setting. *PLoS One*, 2012, 7:e37436 (doi: 10.1371/journal.pone.0037436). *Study of treatment monitoring.*

67. Theron G et al. The use of an automated quantitative polymerase chain reaction (Xpert MTB/RIF) to predict the sputum smear status of tuberculosis patients. *Clinical Infectious Diseases*, 2012, 54:384–388. *Duplicate data.*

68. Theron G et al. Do adjunct tuberculosis tests, when combined with Xpert MTB/RIF, improve accuracy and the cost of diagnosis in a resource-poor setting? *European Respiratory Journal*, 2012, 40:161–168. *Duplicate data.*

69. Tortoli E et al. Clinical validation of Xpert MTB/RIF for the diagnosis of extrapulmonary tuberculosis. *European Respiratory Journal,* 2012, 40:442–447. *This study evaluated Xpert MTB/RIF for the diagnosis of extrapulmonary TB.*

70. Trebucq A et al. Xpert MTB/RIF for national tuberculosis programmes in low-income countries: when, where and how? *International Journal of Tuberculosis and Lung Disease,* 2012, 15:1567–1572. *Editorial and comment.*

71. Trebucq A, Harries AD, Rieder HL. Should Xpert MTB/RIF be rolled out in low-income countries? Reply. *International Journal of Tuberculosis and Lung Disease*, 2012, 16:703–704. *Correspondence.*

72. Vadwai V et al. Xpert MTB/RIF: a new pillar in diagnosis of extrapulmonary tuberculosis? *Journal of Clinical Microbiology*, 2011, 49:2540–2545. *This study evaluated Xpert MTB/RIF for the diagnosis of extrapulmonary TB.*

73. Vadwai V et al. Need to confirm isoniazid susceptibility in Xpert MTB/RIF rifampin susceptible cases. *Indian Journal of Medical Research*, 2012, 135:560–561. *Correspondence.*

74. Van Rie A et al. Xpert MTB/RIF(®) for point-of-care diagnosis of TB in high-HIV burden, resource-limited countries: hype or hope? *Expert Review of Molecular Diagnostics*, 2010, 10:937–946. *This review covered technical details of Xpert MTB/RIF and the test's potential value as a point-of-care test.*

75. Van Rie A. A single Xpert MTB/RIF test of sputum for diagnosis of tuberculosis and multidrug resistance shows high sensitivity and specificity and reduces diagnosis and treatment delays. *Evidence Based Medicine*, 2011, 16:174–175. *Case report.*

76. Van Zyl-Smit RN et al. Comparison of quantitative techniques including Xpert MTB/RIF to evaluate mycobacterial burden. *PLoS One*, 2011, 6:e28815 (doi: 10.1371/journal.pone.0028815). *Technical paper.*

77. Walters E et al. Rapid diagnosis of pediatric intrathoracic tuberculosis from stool samples using the Xpert MTB/RIF assay: a pilot study. *Pediatric Infectious Disease Journal,* 2012, 31:1316. *This study evaluated Xpert MTB/RIF for TB in children.*

78. Wang Y et al. Evaluation of GeneXpert MTB/RIF assay on AFB-smear positive respiratory samples. *Journal of Molecular Diagnostics*, 2010, 12:885–886. *Abstract only.*

79. Williamson DA et al. Clinical failures associated with *rpo*B mutations in phenotypically occult multidrug-resistant Mycobacterium tuberculosis. *International Journal of Tuberculosis and Lung Disease*, 2012, 16:216–220. *Case–control study.*

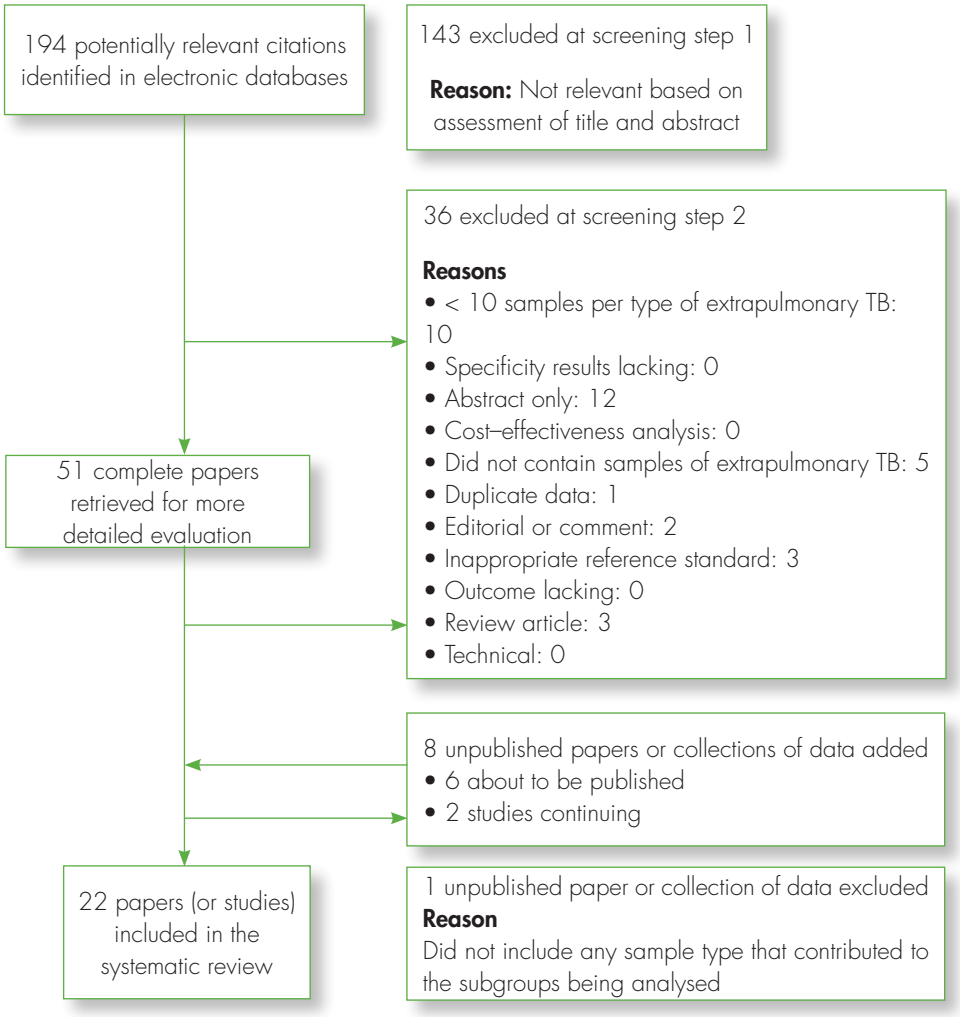
80. Wood R et al. Lipoarabinomannan in urine during tuberculosis treatment: association with host and pathogen factors and mycobacteriuria. *BMC Infectious Diseases*, 2012, 12:47 (doi:10.1186/1471-2334-12-47). *This study evaluated Xpert MTB/RIF for the diagnosis of extrapulmonary TB.*

81. Yoon C et al. Impact of Xpert MTB/RIF testing on tuberculosis management and outcomes in hospitalized patients in Uganda. *PLoS One*, 2012, 7:e48599 (doi:10.1371/journal.pone.0048599). *Duplicate data.*

*82.* Zar HJ et al. Rapid molecular diagnosis of pulmonary tuberculosis in children using nasopharyngeal specimens. *Clinical Infectious Diseases*, 2012, 55:1088–1095. *This study evaluated Xpert MTB/RIF for the diagnosis of TB in children.*

## Annex 5. Selection of studies evaluating the accuracy of Xpert MTB/RIF in diagnosing extrapulmonary TB and rifampicin resistance in adults and children

Figure 1. Flow diagram of studies included in the review



## Included studies

### *Published studies*

1. 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 MTB/RIF system. *Saudi Medical Journal*, 2012, 33:1100–1105.

2. Armand S et al. 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, 49:1772–1766.

3. Bates M et al. Assessment of the Xpert MTB/RIF assay for diagnosis of tuberculosis with gastric lavage aspirates in children in sub-Saharan Africa: a prospective descriptive study. *Lancet Infectious Diseases*, 2013, 13:36–42.

4. Causse M et al. Comparison of two molecular methods for rapid diagnosis of extrapulmonary tuberculosis. *Journal of Clinical Microbiology*, 2011,49:3065–3067.

5. Friedrich SO, von Groote-Bidlingmaier F, Diacon AH. Xpert MTB/RIF assay for diagnosis of pleural tuberculosis. *Journal of Clinical Microbiology*, 2011, 49:4341–4342.

6. Hanif SN et al. GeneXpert MTB/RIF for rapid detection of Mycobacterium tuberculosis in pulmonary and extra-pulmonary samples. *International Journal of Tuberculosis And Lung Disease,* 2011, 15:1274–1275.

7. Hillemann D et al. Rapid molecular detection of extrapulmonary tuberculosis by the automated GeneXpert MTB/RIF system. *Journal of Clinical Microbiology,* 2011, 49:1202–1205.

8. Ligthelm LJ et al. Xpert MTB/RIF for rapid diagnosis of tuberculous lymphadenitis from fine-needle-aspiration biopsy specimens. *Journal of Clinical Microbiology*, 2011, 49:3967–39670.

9. Malbruny B et al. Rapid and efficient detection of Mycobacterium tuberculosis in respiratory and non-respiratory samples. *International Journal of Tuberculosis and Lung Disease*, 2011, 15:553–555.

10. 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, 50:513–515.

11. Safianowska A et al. [Two selected commercially based nucleic acid amplification tests for the diagnosis of tuberculosis]. *Pneumonologia i Alergologia Polska*, 2012, 80:6–12.

12. Tortoli E et al. Clinical validation of Xpert MTB/RIF for the diagnosis of extrapulmonary tuberculosis. *European Respiratory Journal,* 2012,40:442–447.

13. Vadwai V et al. Xpert MTB/RIF: a new pillar in diagnosis of extrapulmonary tuberculosis? *Journal of Clinical Microbiology*, 2011, 49:2540–2545.

14. Walters E et al. Rapid diagnosis of pediatric intrathoracic tuberculosis from stool samples using the Xpert MTB/RIF assay: a pilot study. *Pediatric Infectious Disease Journal,* 2012, 31:1316.

15. 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, 49:4138–4141.

### *Unpublished studies*

The names of the authors of unpublished studies have been obscured to maintain confidentiality.

1. Caws M. Xpert MTB/RIF for the diagnosis of cerebrospinal tuberculosis. 2013 (submitted).

2. [redacted] Xpert MTB/RIF for the diagnosis of extrapulmonary tuberculosis (interim analysis).

3. [redacted] Comparison of rapid tools, including Xpert MTB/RIF assay, for the diagnosis of pleural tuberculosis. 2013 (submitted).

4. Oliveira EF, Trajman A, Bastos M. Accuracy of PCR for diagnosis of pleural tuberculosis. 2013 (submitted).

5. Patel V. Quantitative PCR (GeneXpert MTB/RIF) for the diagnosis of tuberculous meningitis in a high burden setting: a prospective study. 2013 (submitted).

6. [redacted] Xpert MTB/RIF for the diagnosis of extra-pulmonary tuberculosis (interim analysis).

7. [redacted] Xpert MTB/RIF for the diagnosis of extrapulmonary tuberculosis (interim analysis).

## Excluded studies with reasons for exclusion

1. Alvarez-Uria G et al. 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*, 2012:932862 (http://dx.doi.org/10.1155/2012/932862). *Inappropriate reference standard.*

2. Archontakis Z et al, Evaluation of GeneXpert MTB/RIF assay in respiratory and non-respiratory specimens for the rapid detection of Mycobacterium tuberculosis. *Clinical Microbiology and Infection*, 2011, 17 (Suppl. 4):S592. *Abstract only.*

3. Barnard M et al. Diagnostic performance of Genotype® MTBDRplus version 2 line probe assay is equivalent to the XpertMTB/RIF assay. *Journal of Clinical Microbiology*, 2012, 50:3712–3716. *Did not contain nonrespiratory samples.*

4. Barreales-Fonseca A et al. GeneXpert MTB/RIF system in pulmonary and extrapulmonary specimens: comparison with other nucleic acid technologies. *Clinical Microbiology and Infection*, 2012, 18:773. *Abstract only.*

5. Blakemore R et al. A multisite assessment of the quantitative capabilities of the Xpert MTB/RIF assay. *American Journal of Respiratory and Critical Care Medicine*, 2011, 184:1076–1084. *Did not contain nonrespiratory samples.*

6. Blakemore R et al. Evaluation of the analytical performance of the Xpert MTB/RIF assay. *Journal of Clinical Microbiology*, 2010, 48:2495–2501. *Did not contain non-respiratory samples*

7. Bowles E et al. Xpert MTB/RIF, a novel automated polymerase chain reaction-based tool for the diagnosis of tuberculosis. *International Journal of Tuberculosis and Lung Disease*, 2011, 15:988–989. *Fewer than 10 samples for each type of extrapulmonary TB.*

8. Ciftci IH, Aslan MH, Asik G. [Evaluation of Xpert MTB/RIF results for the detection of Mycobacterium tuberculosis in clinical samples]. *Mikrobiyoloji Bülteni*, 2011, 45:43–47. *Fewer than 10 samples for each type of extrapulmonary TB.*

9. Clemente MG et al. Evaluation of the Xpert MTB/RIF test for rapid detection of Mycobacterium tuberculosis. *American Journal of Respiratory and Critical Care Medicine*, 2011, 183:Meeting Abstracts. *Abstract only.*

10. De Lauzanne A et al. Rapid diagnosis of smear-negative tuberculous osteoarthritis by real-time polymerase chain reaction on bone tissue. *Pediatric Infectious Disease Journal,* 2011, 30:184. *Fewer than 10 samples for each type of extrapulmonary TB.*

11. Deforges L et al. Application of the rapid detection system for M. tuberculosis complex and rifampicin resistance Xpert MTB/RIF in decontaminated respiratory specimens, non-respiratory specimens and cultures. Special Issue: Abstracts of the 20th ECCMID, Vienna, Austria, 10–13 April 2010. *Clinical Microbiology and Infection*, 2010, 16(Suppl. S2):S53. *Abstract only.*

12. Dhasmana DJ et al. GeneXpert MTB/RIF assay improves the diagnostic yield of EBUS-TBNA in smear-negative intrathoracic tuberculous lymphadenopathy. *Thorax*, 2011, 66:A75. *Abstract only.*

13. Drobniewski F et al. Diagnosis of tuberculosis and drug resistance: what can new tools bring us? *International Journal of Tuberculosis and Lung Disease*, 2012, 16:860–870. *Review article.*

14. Ferrara G et al. Xpert MTB/RIF test for tuberculosis. *Lancet*, 2011, 378:482. *Editorial or comment.*

15. Gascoyne-Binzi DM et al. Comparison of on-demand PCR testing for Mycobacterium tuberculosis with culture and batched PCR. *Clinical Microbiology and Infection*, 2012, 18:543. *Fewer than 10 samples for each type of extrapulmonary TB.*

16. Gous N et al. Performance of the Roche LightCycler Real-Time PCR assay for diagnosing extrapulmonary tuberculosis. *Journal of Clinical Microbiology*, 2012, 50:2100–2103. *Fewer than 10 samples for each type of extrapulmonary TB.*

17. Guna MR et al. Detection of Mycobacterium tuberculosis complex by RT-PCR in non-respiratory clinical specimens in Health Department 9, Valencia (Spain). *Clinical Microbiology and Infection,* 2011, 17 (Suppl. 4):S591. *Abstract only.*

18. Ioannidis P et al. Cepheid GeneXpert MTB/RIF assay for Mycobacterium tuberculosis detection and rifampicin resistance identification in patients with substantial clinical indications of tuberculosis and smear-negative microscopy results. *Journal of Clinical Microbiology*, 2011, 49:3068–3070. *Fewer than 10 samples for each type of extrapulmonary TB.*

19. Lawn SD. Pre-screening with GeneXpert MTB/RIF may increase use of isoniazid preventive therapy in antiretroviral programmes. I*nternational Journal of Tuberculosis and Lung Disease*, 2011, 15:1272–1273. *Editorial or comment.*

20. Lawn SD et al. 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, 60:289–294. *Inappropriate reference standard.*

21. Lawn SD, Zumla AI. Diagnosis of extrapulmonary tuberculosis using the Xpert MTB/RIF assay. *Expert Review of Anti-Infective Therapy*, 2012, 10:631–635. *Review article.*

22. Marlowe EM, Novak-Weekley SM, Cumpio JV. Comparison of three molecular methods for the direct detection of Mycobacterium tuberculosis complex from respiratory & non-respiratory specimens. *Journal of Molecular Diagnostics*, 2010, 12:885. *Abstract only.*

23. Miller MB et al. Performance of Xpert MTB/RIF MTB/RIF RUO assay and IS6110 real-time PCR for Mycobacterium tuberculosis detection in clinical samples. *Journal of Clinical Microbiology,* 2011, 49:3458–3562. *Fewer than 10 samples for each type of extrapulmonary TB.*

24. Mokaddas E, Ahmed S, SaadEldeen H. Comparison of two nucleic acid amplification assays, the ProbeTec ET assay and Xpert MTB/RIF assay, for detection of Mycobacterium tuberculosis in pulmonary and extrapulmonary specimens. *Clinical Microbiology and Infection*, 2012, 18:547. A*bstract only.*

25. Mokaddas E, Haneef S, Saad Eldeen H. Evaluation of Gene Xpert MTB/RIF system for the direct detection of rifampicin-resistant Mycobacterium tuberculosis from pulmonary and extrapulmonary specimens in the National Tuberculosis Reference Laboratory in Kuwait. *Clinical Microbiology and Infection, 2011*, 17(Suppl. 4):S591. *Duplicate data.*

26. Moure R et al. Rapid detection of Mycobacterium tuberculosis complex and rifampicin resistance in smear-negative clinical samples by use of an integrated real-time PCR method. *Journal of Clinical Microbiology*, 2011, 49:1137–1139. *Fewer than 10 samples for each type of extrapulmonary TB.*

27. Naidoo S. Evaluation of GeneXpert MTB/RIF assay on pulmonary and extra-pulmonary samples in a high-throughput laboratory. Special Issue: Abstracts of the 20th ECCMID, Vienna, Austria, 10–13 April 2010. *Clinical Microbiology and Infection*, 2010, 16(Suppl. S2):S604–S605. *Abstract only.*

28. Nazli A, Cavusoglu C. Evaluation of GeneXpert MTB/RIF assay for the diagnosis of tuberculosis and rapid detection of rifampicin resistance in clinical specimens. *Clinical Microbiology and Infection,* 2011, 17(Suppl. 4):S592–S593. *Abstract only.*

29. Peter JG et al. 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:e39966 (doi:10.1371/journal.pone.0039966). *Inappropriate reference standard.*

30. Pinto P et al. Evaluation of a point-of-care molecular test for M. tuberculosis detection in an emergency setting. *Clinical Microbiology and Infection*, 2012, 18:542. *Abstract only.*

31. Rachow A et al. Increased and expedited case detection by Xpert MTB/RIF assay in childhood tuberculosis: a prospective cohort study. *Clinical Infectious Diseases*, 2012, 54:1388–1396. *Fewer than 10 samples for each type of extrapulmonary TB.*

32. Teo J et al. 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, 49:3659–3662. *Fewer than 10 samples for each type of extrapulmonary TB.*

33. Theron G. Cycle-threshold (CT) values of an automated TB-specific PCR platform (Xpert MTB/RIF) as a predictor of smear status and grade. *Tropical Medicine and International Health*, 2012, 17:24. *Did not contain nonrespiratory samples*.

34. Theron G et al. Do adjunct tuberculosis tests, when combined with Xpert MTB/RIF, improve accuracy and the cost of diagnosis in a resource-poor setting? *European Respiratory Journal*, 2012, 40:161–168. *Did not contain nonrespiratory samples*.

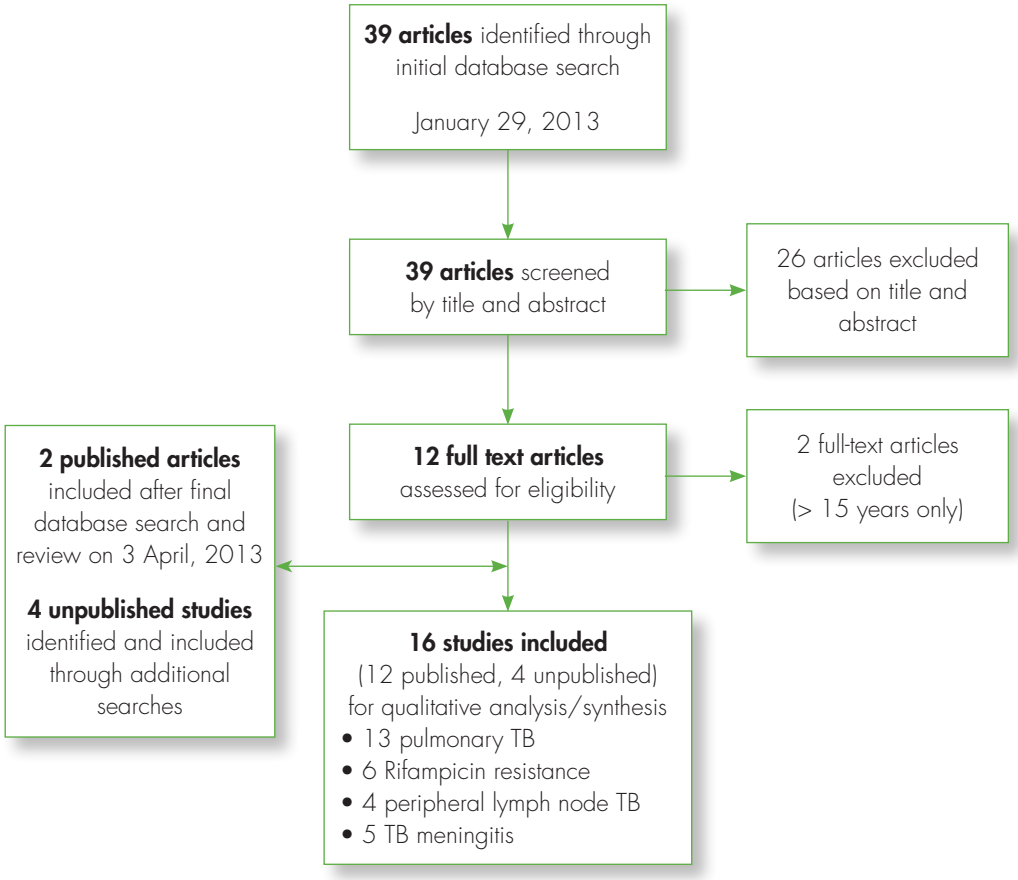
35. Theron GDV et al. Evaluation of the geneXpert MTB/RIF assay for the diagnosis of tuberculosis (TB) using sputum and extrapulmonary samples. *American Journal of Respiratory and Critical Care Medicine*, 2011, 183:Meeting Abstracts. *Abstract only.*

36. Yin, QQ, Shen AD.[Diagnostic value of Xpert MTB/RIF assay for tuberculosis]. *Zhonghua Jie He He Hu Xi Za Zhi*, 2012, 35:363-5. *Review article.*

37. Peter J. Xpert MTB/RIF for the diagnosis of pericardial tuberculosis. In: Denkinger C, ed. *No samples for subgroups included.*

# Annex 6. Selection of studies evaluating the accuracy of Xpert MTB/RIF in diagnosing pulmonary and extrapulmonary TB, and rifampicin resistance in children

Figure 1. Flow diagram of studies included in the review




## Included studies

### *Published studies*

1. Bates M et al. Assessment of the Xpert MTB/RIF assay for diagnosis of tuberculosis with gastric lavage aspirates in children in sub-Saharan Africa: a prospective descriptive study. *Lancet Infectious Diseases*, 2013, 13:36–42.

2. Causse M et al. Comparison of two molecular methods for rapid diagnosis of extrapulmonary tuberculosis. *Journal of Clinical Microbiology*, 2011, 49:3065–3067.

3. Ligthelm LJ et al. Xpert MTB/RIF for rapid diagnosis of tuberculous lymphadenitis from fine-needle-aspiration biopsy specimens. *Journal of Clinical Microbiology,* 2011, 49:3967–3970.

4. Malbruny B et al. Rapid and efficient detection of Mycobacterium tuberculosis in respiratory and non-respiratory samples. *International Journal of Tuberculosis and Lung Disease*, 2011, 15:553–555.

5. Nguyen TQ 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).

6. Nicol MP et al. Accuracy of the Xpert MTB/RIF test for the diagnosis of pulmonary tuberculosis in children admitted to hospital in Cape Town, South Africa: a descriptive study. *Lancet Infectious Diseases*, 2011, 11:819–824.

7. Rachow A et al. Increased and expedited case detection by Xpert MTB/RIF assay in childhood tuberculosis: a prospective cohort study. *Clinical Infectious Diseases,* 2012, 54:1388–1396.

8. Sekadde MP et al. Evaluation of the Xpert MTB/RIF test for the diagnosis of childhood pulmonary tuberculosis in Uganda: a cross-sectional diagnostic study. *BMC Infectious Diseases*, 2013;13:133 (doi: 10.1186/1471-2334-13-133).

9. Tortoli E et al. Clinical validation of Xpert MTB/RIF for the diagnosis of extrapulmonary tuberculosis. *European Respiratory Journal*, 2012, 40:442–447.

10. Vadwai V et al. Xpert MTB/RIF: a new pillar in diagnosis of extrapulmonary tuberculosis? *Journal of Clinical Microbiology*, 2011, 49:2540–2545.

11. Walters E et al. Rapid diagnosis of pediatric intrathoracic tuberculosis from stool samples using the Xpert MTB/RIF assay: a pilot study. *Pediatric Infectious Disease Journal,* 2012, 31:1316.

12. Zar HJ et al. Rapid molecular diagnosis of pulmonary tuberculosis in children using nasopharyngeal specimens. *Clinical Infectious Diseases*, 2012, 55:1088–1095.

### *Unpublished studies*

The names of the authors of unpublished studies have been obscured to maintain confidentiality.

1. [redacted], et al. Bacteremia and tuberculosis in severely malnourished Bangladeshi children presenting with acute pneumonia including the use of Xpert MTB/RIF assay (study continuing May 2013).

2. [redacted] CA. Xpert for the diagnosis of peripheral lymph node TB in children (study continuing May 2013).

3. [redacted] et al. GeneXpert MTB/RIF in the diagnosis of pulmonary TB among severely malnourished hospitalized Malawian children (study continuing May 2013).

4. [redacted] Diagnostic yield and treatment response in childhood intra-thoracic tuberculosis: effect of disease severity (study continuing May 2013).

## Excluded studies with reasons for exclusion

1. Armand S et al. 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, 49:1772–1776. *This was a case–control study that compared the Xpert MTB/RIF assay with in-house IS6110-based real-time PCR using TaqMan probes (IS6110-TaqMan assay) for TB detection.*

2. Boehme CC et al. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *Lancet*, 2011, 30;377:1495–1505. *Adults only.*

3. Boehme CC et al. Rapid molecular detection of tuberculosis and rifampin resistance. *New England Journal of Medicine*, 2010, 363:1005–15. *Adults only.*

4. Boehme CC et al. Rapid molecular detection of tuberculosis and rifampin resistance. *New England Journal of Medicine*, 2010, 9;363:1005–1015. *Adults only.*

5. Chang K et al. Rapid and effective diagnosis of tuberculosis and rifampicin resistance with Xpert MTB/RIF assay: a meta-analysis. *Journal of Infection*, 2012, 64:580–588. *Meta-analysis.*

6. Chegou NN et al. Tuberculosis assays: past, present and future. *Expert Review of Anti-Infective Therapy*, 2011, 9 :457–469. *Narrative review.*

7. Cuevas LE. The urgent need for new diagnostics for symptomatic tuberculosis in children. *Indian Journal of Pediatrics*, 2011, 78:449–455. *Narrative review.*

8. Dorjee K, Salvo F, Dierberg KL. Xpert® MTB/RIF diagnosed disseminated smear-negative MDR-TB in a sub-district hospital in India. *International Journal of Tuberculosis and Lung Disease*, 2012, 16:1560–1561. *Case report.*

9. Drobniewski F et al. Diagnosis of tuberculosis and drug resistance: what can new tools bring us? *International Journal of Tuberculosis and Lung Disease*, 2012, 16:860–870. *Review article.*

10. Friedrich SO et al. Suitability of Xpert MTB/RIF and genotype MTBDRplus for patient selection for a tuberculosis clinical trial. *Journal of Clinical Microbiology*, 2011, 49:2827–2831. *Adults only.*

11. Gotuzzo E. Xpert MTB/RIF for diagnosis of pulmonary tuberculosis. *Lancet Infectious Diseases*, 2011, 11:802–803. *Editorial and comment.*

12. Helb D et al. Rapid detection of Mycobacterium tuberculosis and rifampin resistance by use of on-demand, near-patient technology. *Journal of Clinical Microbiology*, 2010, 48:229–237. *Adults only.*

13. Hepple P, Ford N, McNerney R. Microscopy compared to culture for the diagnosis of tuberculosis in induced sputum samples: a systematic review. *International Journal of Tuberculosis and Lung Disease*, 2012, 16:579–588. *Review article.*

14. Hoek KG et al. Detecting drug-resistant tuberculosis: the importance of rapid testing. *Molecular Diagnosis & Therapy*, 2011, 15:189–194. *Narrative review.*

15. Lawn SD, Zumla AI. Diagnosis of extrapulmonary tuberculosis using the Xpert MTB/RIF assay. *Expert Review of Anti-Infective Therapy*, 2012, 10:631–635. *Review article.*

16. López Ávalos GG, Prado Montes De Oca E. Classic and new diagnostic approaches to childhood tuberculosis. *Journal of Tropical Medicine*, 2012:818219 (doi: 10.1155/2012/818219). *Review article.*

17. Nahid P, Menzies D. Update in tuberculosis and nontuberculous mycobacterial disease 2011. *American Journal of Respiratory and Critical Care Medicine*, 2012, 185:1266–1270. *Review article.*

18. Ntinginya EN et al. Performance of the Xpert® MTB/RIF assay in an active case-finding strategy: a pilot study from Tanzania. *International Journal of Tuberculosis and Lung Disease,* 2012, 16:1468–1470. *This study included both adults and children. The study used an active case-finding strategy involving previously known TB cases. The study design was considered to be different from a study of the accuracy of diagnostic tests, and therefore was not included in the review.*

19. Palomino JC. Current developments and future perspectives for TB diagnostics. Future Microbiology, 2012, 7:59–71. *Review article.*

20. Perez-Velez CM. Pediatric tuberculosis: new guidelines and recommendations. *Current Opinion in Pediatrics*, 2012, 24:319–328. *Review article.*

21. Raj A, Singh N. Tuberculous meningitis – an update on diagnostic approaches. International *Journal of Research in Pharmaceutical Sciences*, 2012, 3:312–318. *Review article.*

22. Singh V, Kaur S. Diagnosis of tuberculosis newer investigations. *Indian Journal of Practical Pediatrics*, 2012, 14:273–283. *Review article.*

23. Taylor N et al. Can a simple flotation method lower the limit of detection of Mycobacterium tuberculosis in extrapulmonary samples analyzed by the GeneXpert MTB/RIF assay? *Journal of Clinical Microbiology*, 2012, 50:2272–2276. *Used Xpert MTB/RIF for the diagnosis of extrapulmonary TB.*

24. Theron G et al. Evaluation of the Xpert MTB/RIF assay for the diagnosis of pulmonary tuberculosis in a high HIV prevalence setting. *American Journal of Respiratory and Critical Care Medicine*, 2011, 184:132–140. *Adults only.*

25. Theron G et al. Correlation of Mycobacterium tuberculosis specific and non-specific quantitative Th1 T-cell responses with bacillary load in a high burden setting. *PLoS One* 2012, 7:e37436 (doi: 10.1371/journal.pone.0037436). *Treatment monitoring.*

26. Turnbull ER et al. A model of tuberculosis screening for pregnant women in resource-limited settings using Xpert MTB/RIF. *Journal of Pregnancy,* 2012, 2012: 565049 (doi:10.1155/2012/565049). *Review article.*

27. Whittaker E, Zar HJ. Promising directions in the diagnosis of childhood tuberculosis. *Expert Review of Respiratory Medicine*, 2012, 6:385–395. *Review article.*

28. Zar HJ et al. Advances in childhood tuberculosis – contributions from the university of Cape Town. *South African Medical Journal*, 2012, 102:518–521. *Review article.*

29. Zeka AN, Tasbakan S, Cavusoglu C. Evaluation of the GeneXpert MTB/RIF assay for the rapid diagnosis of tuberculosis and detection of RIF-resistance in pulmonary and extrapulmonary specimens. *Journal of Clinical Microbiology,* 2011, 49:4138–4141. *Narrative review.*

## Annex 7. Literature review of affordability and cost effectiveness of Xpert MTB/RIF for the diagnosis of TB

1. Abimbola TO. Cost-effectiveness of tuberculosis diagnostic strategies to reduce early mortality among persons with advanced HIV infection initiating antiretroviral therapy. *Journal of Acquired Immune Deficiency Syndromes*, 2012, 60(1):e1-7 (doi: 10.1097/QAI.0b013e318246538f).

2. Andrews JR et al. The cost-effectiveness of routine tuberculosis screening with Xpert MTB/RIF prior to initiation of antiretroviral therapy in South Africa: a model-based analysis. AIDS, 2012, 26:987–995 (doi: 10.1097/QAD.0b013e3283522d47).

3. Menzies NA et al. Population health impact and cost-effectiveness of tuberculosis diagnosis with Xpert MTB/RIF: a dynamic simulation and economic evaluation. *PLoS Medicine*, 2012, 9:e1001347 (doi:10.1371/journal.pmed.1001347).

4. Meyer-Rath G et al. The impact and cost of scaling up GeneXpert MTB/RIF in South Africa. *PLoS One*, 2012, 7:e36966 (doi:10.1371/journal.pone.0036966).

5. Meyer-Rath G et al. *The incremental cost of introducing Xpert® MTB/RIF into the South African national tuberculosis programme*. Johannesburg, Health Economics and Epidemiology Research Office, 2011, http://www.tbonline.info/media/uploads/documents/hero-costofgenexpert.pdf

6. Pantoja A et al. Xpert MTB/RIF for diagnosis of tuberculosis and drug-resistant tuberculosis: a cost and affordability analysis. *European Respiratory Journal,* 2013, 42:708–720 (doi:10.1183/09031936.00147912).

7. Schnippel K et al. Scaling up Xpert MTB/RIF technology: the costs of laboratory- vs. clinic-based roll-out in South Africa. *Tropical Medicine and International Health,* 2012, 17: 1142–1151 (doi: 10.1111/j.1365-3156.2012.03028.x).

8. Schnippel K et al. Diagnosing Xpert MTB/RIF negative TB: impact and cost of alternative algorithms for South Africa. *South African Medical Journal*, 2013, 103:101–106.

9. Theron G et al. Do adjunct TB tests, when combined with Xpert MTB/RIF, improve accuracy and the cost of diagnosis in a resource-poor setting? *European Respiratory Journal,* 2012, 40:161–168 (doi: 10.1183/09031936.00145511).

10. Van Rie A et al. Point-of-care Xpert MTB/RIF for smear-negative tuberculosis suspects at a primary care clinic in South Africa. *International Journal of Tuberculosis and Lung Disease,* 2013, 17:368–372 (doi: 10.5588/ijtld.12.0392).

11. Vassall A et al. Rapid diagnosis of tuberculosis with the Xpert MTB/RIF assay in high burden countries: a cost-effectiveness analysis. *PLoS Medicine*, 2011, 8:e1001120 (doi:10.1371/journal.pmed.1001120).

12. Winetsky DE et al. Screening and rapid molecular diagnosis of tuberculosis in prisons in Russia and Eastern Europe: a cost-effectiveness analysis. *PLoS Medicine*, 2012, 9: e1001348 (doi:10.1371/journal.pmed.1001348).

World Health Organization

**Global TB Programme**
World Health Organization
Avenue Appia 20, CH-1211
Geneva-27, Switzerland

Information Resource Centre HTM/GTB:
Email: tbdocs@who.int
Website: www.who.int/tb