

Automated real-time nucleic acid
amplification technology for rapid
and simultaneous detection of tuberculosis
and rifampicin resistance:

Xpert MTB/RIF assay for the diagnosis
of pulmonary and extrapulmonary TB
in adults and children

NEW DIAGNOSTIC TESTS
POLICY UPDATE
TUBERCULOSIS
TB
PULMONARY TB
RIFAMPICIN
DIAGNOSIS
RESISTANCE
DRUG-RESISTANCE
TB/HIV
PERFORMANCE
RAPID TB TEST
ACCURACY
RECOMMENDATIONS
MYCOBACTERIUM
MOLECULAR DIAGNOSTICS



World Health
Organization

Automated real-time nucleic acid amplification
technology for rapid and simultaneous detection
of tuberculosis and rifampicin resistance:
Xpert MTB/RIF assay for the diagnosis
of pulmonary and extrapulmonary TB
in adults and children

Policy update



World Health
Organization

WHO Library Cataloguing-in-Publication Data

Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF assay for the diagnosis of pulmonary and extra-pulmonary TB in adults and children. Policy update.

Revision of 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.

1.Tuberculosis, Multidrug-resistant - diagnosis. 2.Tuberculosis - diagnosis. 3.Rifampin - pharmacology. 4.Mycobacterium tuberculosis - isolation and purification. 5.HIV infections - diagnosis. 6.Sensitivity and specificity. 7.Guideline. I. World Health Organization.

ISBN: 978 92 4 150633 5

(NLM classification: WF 310)

© 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

Printed in France

WHO/HTM/TB/2013.16

Contents

EXECUTIVE SUMMARY	XI
WHO'S POLICY RECOMMENDATIONS	XV
1. BACKGROUND	1
2. METHODS	3
2.1 EVIDENCE SYNTHESIS	3
2.2 EXPERT GROUP MEETING	7
2.3 EXTERNAL REVIEW	7
2.4 PREPARING THE POLICY UPDATE	7
3. SCOPE	8
3.1 DATE OF REVIEW: 2017	8
4. EVIDENCE BASE FOR POLICY FORMULATION	9
4.1 USING XPERT MTB/RIF TO DIAGNOSE PULMONARY TB AND RIFAMPICIN RESISTANCE IN ADULTS	9
4.1.1 USING XPERT MTB/RIF AS A REPLACEMENT TEST FOR SMEAR MICROSCOPY	10
4.1.2 USING XPERT MTB/RIF AS AN ADD-ON TEST FOLLOWING MICROSCOPY	12
4.1.3 USING XPERT MTB/RIF TO DETECT SMEAR-POSITIVE CULTURE-POSITIVE TB	12
4.1.4 USING XPERT MTB/RIF TO DETECT SMEAR-NEGATIVE CULTURE-POSITIVE TB	12
4.1.5 USING XPERT MTB/RIF TO DETECT PULMONARY TB IN HIV-NEGATIVE AND HIV- POSITIVE INDIVIDUALS	12
4.1.6 USING XPERT MTB/RIF TO DETECT RIFAMPICIN RESISTANCE	14
4.1.7 EFFECT OF THE VERSION OF THE XPERT MTB/RIF ASSAY	15
4.1.8 ACCURACY OF THE XPERT MTB/RIF G4 CARTRIDGE	15
4.1.9 ACCURACY OF THE REFERENCE STANDARDS	16
4.2 USING XPERT MTB/RIF TO DIAGNOSE EXTRAPULMONARY TB AND RIFAMPICIN RESISTANCE IN ADULTS AND CHILDREN	17
4.2.1 DETECTING LYMPH NODE TB IN SAMPLES FROM BIOPSY OR FINE-NEEDLE ASPIRATION	18
4.2.2 DETECTING PLEURAL TB IN PLEURAL FLUID	19
4.2.3 DETECTING TB IN SAMPLES OF CSF	20
4.2.4 DETECTING TB IN GASTRIC FLUID	22
4.2.5 DETECTING TB IN TISSUE SAMPLES	22
4.2.6 DETECTING RIFAMPICIN RESISTANCE	23

4.3 USING XPERT MTB/RIF TO DIAGNOSE TB AND RIFAMPICIN RESISTANCE IN CHILDREN	23
4.3.1 USING XPERT MTB/RIF TO DIAGNOSE PULMONARY TB IN CHILDREN	24
4.3.2 XPERT MTB/RIF COMPARED WITH SMEAR MICROSCOPY	26
4.3.3 PERFORMANCE OF XPERT MTB/RIF IN SMEAR-POSITIVE AND SMEAR-NEGATIVE CHILDREN	27
4.3.4 XPERT MTB/RIF IN CHILDREN AGED 0–4 YEARS AND 5–15 YEARS	29
4.3.5 XPERT MTB/RIF IN HIV-POSITIVE AND HIV-NEGATIVE CHILDREN	30
4.3.6 USING XPERT MTB/RIF TO DETECT PERIPHERAL LYMPH NODE TB IN CHILDREN	32
4.3.7 USING XPERT MTB/RIF TO DETECT TB MENINGITIS IN CHILDREN	32
4.3.8 USING XPERT MTB/RIF TO DETECT RIFAMPICIN RESISTANCE IN CHILDREN	32
4.4 AFFORDABILITY AND COST EFFECTIVENESS OF USING XPERT MTB/RIF TO DIAGNOSE TB	33
5. WHO'S POLICY RECOMMENDATIONS	38
5.1 USING XPERT MTB/RIF TO DIAGNOSE PULMONARY TB AND RIFAMPICIN RESISTANCE IN ADULTS AND CHILDREN	38
5.2 USING XPERT MTB/RIF TO DIAGNOSE EXTRAPULMONARY TB AND RIFAMPICIN RESISTANCE IN ADULTS AND CHILDREN	39
6. IMPLEMENTATION CONSIDERATIONS	39
6.1 RESEARCH NEEDS	42
6.2 PLANS FOR SUPPORTING SCALE-UP OF THE IMPLEMENTATION OF XPERT MTB/RIF	42
7. GRADE TABLES	44
ANNEXES	74
ANNEX 1. MEMBERS OF THE EXPERT GROUP	74
ANNEX 2. WHO STAFF MEMBERS	76
ANNEX 3. MEMBERS OF THE STRATEGIC AND TECHNICAL ADVISORY GROUP FOR TUBERCULOSIS (STAG-TB)	76
ANNEX 4. DECLARATIONS OF INTERESTS	78

Tables

TABLE 1.	META-ANALYSIS OF THE SENSITIVITY AND SPECIFICITY OF XPERT MTB/RIF IN DIAGNOSING EXTRAPULMONARY TB AND RIFAMPICIN RESISTANCE IN ADULTS AND CHILDREN COMPARED AGAINST CULTURE AS A REFERENCE STANDARD AS WELL AS AGAINST A COMPOSITE REFERENCE STANDARD, BY TYPE OF EXTRAPULMONARY SPECIMEN	xiii
TABLE 2.	POOLED SENSITIVITY AND SPECIFICITY OF THE XPERT MTB/RIF ASSAY FOR DETECTING PULMONARY TB AND RIFAMPICIN RESISTANCE	12
TABLE 3.	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	27
TABLE 4.	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	29
TABLE 5.	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	31
TABLE 6.	METAREGRESSION MODEL FOR XPERT MTB/RIF USING SAMPLES OF EXPECTORATED OR INDUCED SPUTUM FROM CHILDREN, CONTROLLING FOR SMEAR STATUS AND HIV STATUS	31
TABLE 7.	OVERVIEW OF STUDIES COMPARING THE COSTS OF USING XPERT MTB/RIF AND FOLLOW-ON TESTS WITH CURRENT DIAGNOSTIC ALGORITHMS FOR DIAGNOSING TB AND MDR-TB	34
TABLE 8.	GRADE EVIDENCE PROFILE: ACCURACY OF XPERT MTB/RIF IN DIAGNOSING PULMONARY TB IN ADULTS	44
TABLE 9.	GRADE EVIDENCE PROFILE: ACCURACY OF XPERT MTB/RIF IN DIAGNOSING PULMONARY TB IN SPUTUM SMEAR-POSITIVE ADULTS	45
TABLE 10.	GRADE EVIDENCE PROFILE: ACCURACY OF XPERT MTB/RIF IN DIAGNOSING PULMONARY TB IN SPUTUM SMEAR-NEGATIVE ADULTS	46
TABLE 11.	GRADE EVIDENCE PROFILE: ACCURACY OF XPERT MTB/RIF IN DIAGNOSING PULMONARY TB IN ADULTS LIVING WITH HIV	47
TABLE 12.	GRADE EVIDENCE PROFILE: ACCURACY OF XPERT MTB/RIF IN DIAGNOSING PULMONARY TB IN ADULTS WITHOUT HIV INFECTION	48
TABLE 13.	GRADE EVIDENCE PROFILE: THE INCREMENTAL YIELD OF XPERT MTB/RIF COMPARED WITH MICROSCOPY IN PATIENTS WITH CULTURE-CONFIRMED TB	49
TABLE 14.	GRADE EVIDENCE PROFILE: ACCURACY OF XPERT MTB/RIF IN DIAGNOSING PULMONARY TB IN ADULTS AS AN ADD-ON TEST FOLLOWING NEGATIVE SPUTUM-SMEAR MICROSCOPY	50
TABLE 15.	SENSITIVITY OF XPERT MTB/RIF IN SMEAR-NEGATIVE CULTURE-CONFIRMED PULMONARY TB IN INDIVIDUALS, BY HIV STATUS	51

TABLE 16.	GRADE EVIDENCE PROFILE: ADDITIONAL YIELD OF XPERT MTB/RIF OVER MICROSCOPY IN SMEAR-NEGATIVE TB	51
TABLE 17.	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	52
TABLE 18.	ACCURACY OF XPERT MTB/RIF IN DETECTING TB IN LYMPH NODE FLUID AND TISSUE (A. EVIDENCE PROFILE, B. SUMMARY OF FINDINGS)	53
TABLE 19.	ACCURACY OF XPERT MTB/RIF IN DETECTING TB IN PLEURAL FLUID (A. EVIDENCE PROFILE, B. SUMMARY OF FINDINGS)	55
TABLE 20.	ACCURACY OF XPERT MTB/RIF IN DETECTING TB IN CEREBROSPINAL FLUID (A. EVIDENCE PROFILE, B. SUMMARY OF FINDINGS)	57
TABLE 21.	GRADE EVIDENCE PROFILE: ACCURACY OF XPERT MTB/RIF IN DETECTING TB IN GASTRIC FLUID	59
TABLE 22.	GRADE EVIDENCE PROFILE: ACCURACY OF XPERT MTB/RIF IN DETECTING TB IN TISSUE SAMPLES	60
TABLE 23.	GRADE EVIDENCE PROFILE: ACCURACY OF XPERT MTB/RIF IN DETECTING RIFAMPICIN RESISTANCE IN NONRESPIRATORY SPECIMENS	61
TABLE 24.	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)	62
TABLE 25.	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)	63
TABLE 26.	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)	64
TABLE 27.	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)	70
TABLE 28.	GRADE EVIDENCE PROFILE: ACCURACY OF XPERT MTB/RIF IN DETECTING RIFAMPICIN RESISTANCE IN RESPIRATORY SPECIMENS FROM CHILDREN	71
TABLE 29.	GRADE EVIDENCE PROFILE: ACCURACY OF XPERT MTB/RIF IN DETECTING PERIPHERAL LYMPH NODE TB IN CHILDREN	72
TABLE 30.	GRADE EVIDENCE PROFILE AND SUMMARY OF FINDINGS: ACCURACY OF XPERT MTB/RIF IN DETECTING TB MENINGITIS IN CHILDREN	73

Figures

FIGURE 1.	STEPS IN USING THE XPERT MTB/RIF ASSAY	2
FIGURE 2.	SELECTION OF STUDIES EVALUATING THE ACCURACY OF XPERT MTB/RIF IN DIAGNOSING PULMONARY TB AND RIFAMPICIN RESISTANCE IN ADULTS: FLOW DIAGRAM OF STUDIES IDENTIFIED BY THE INITIAL LITERATURE SEARCHES	9
FIGURE 3.	SELECTION OF STUDIES EVALUATING THE ACCURACY OF XPERT MTB/RIF IN DIAGNOSING PULMONARY TB AND RIFAMPICIN RESISTANCE IN ADULTS: FLOW DIAGRAM OF STUDIES IDENTIFIED BY THE UPDATED LITERATURE SEARCH	10
FIGURE 4.	FOREST PLOT OF THE SENSITIVITY AND SPECIFICITY OF XPERT MTB/RIF FOR DETECTING PULMONARY TB IN 27 STUDIES (36 STUDY CENTRES)	11
FIGURE 5.	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)	13
FIGURE 6.	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)	14
FIGURE 7.	SELECTION OF STUDIES EVALUATING THE ACCURACY OF XPERT MTB/RIF IN DIAGNOSING EXTRAPULMONARY TB AND DETECTING RIFAMPICIN RESISTANCE IN ADULTS AND CHILDREN: FLOW DIAGRAM OF STUDIES INCLUDED IN THE REVIEW	17
FIGURE 8.	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	18
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 A COMPOSITE REFERENCE STANDARD	19
FIGURE 10.	FOREST PLOT OF THE SENSITIVITY AND SPECIFICITY OF XPERT MTB/RIF IN DETECTING TB USING PLEURAL FLUID COMPARED WITH CULTURE AS A REFERENCE STANDARD	20
FIGURE 11.	FOREST PLOT OF THE SENSITIVITY AND SPECIFICITY OF XPERT MTB/RIF IN DETECTING TB IN PLEURAL FLUID COMPARED WITH A COMPOSITE REFERENCE STANDARD	20
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	21
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	21
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	22
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	23

FIGURE 16.	SELECTION OF STUDIES EVALUATING THE ACCURACY OF XPERT MTB/RIF IN DIAGNOSING TB AND RIFAMPICIN RESISTANCE IN CHILDREN: FLOW DIAGRAM OF STUDIES INCLUDED IN THE REVIEW	24
FIGURE 17.	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	25
FIGURE 18.	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	26
FIGURE 19.	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	28
FIGURE 20.	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	30
FIGURE 21.	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	32

Abbreviations

AFB	acid-fast bacilli
CI	confidence interval
CrI	credible interval
CRS	composite reference standard
CSF	cerebrospinal fluid
DOI	Declaration of Interests
DST	drug-susceptibility testing
FINN	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
PCR	polymerase chain reaction
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analyses
<i>rpoB</i>	gene encoding for the β -subunit of the DNA-dependent RNA polymerase of <i>Mycobacterium tuberculosis</i>
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
STAG-TB	Strategic and Technical Advisory Group for Tuberculosis
TB	tuberculosis
WHO	World Health Organization
XDR-TB	extensively drug-resistant tuberculosis

Acknowledgements

This document was prepared by a WHO Steering Group comprising Christopher Gilpin, Karin Weyer, Wayne van Gemert and Fuad Mirzayev (all from WHO's Global TB Programme) on the basis of consensus agreed at an Expert Group Meeting convened by WHO in Geneva during 20–21 May 2013.

The findings and recommendations from the meeting were presented to WHO's *Strategic and Technical Advisory Group for Tuberculosis (STAG-TB) in June 2013* (Annex 3). STAG-TB agreed with the recommendations made by the Expert Group on using Xpert MTB/RIF to diagnose TB and rifampicin resistance in pulmonary and extrapulmonary TB in adults and children, and advised WHO to produce a policy update.

This document was finalized following consideration of all comments and suggestions from the participants of the Expert Group and the members of STAG-TB.

WHO gratefully acknowledges the contributions of the Chairperson of the Expert Group (Holger Schünemann), the members of the Expert Group (Annex 1) and STAG-TB, as well as the WHO staff members who developed this policy update (Annex 2). Karen Steingart (systematic reviewer for pulmonary TB), Claudia Denkinger (systematic reviewer for extrapulmonary TB), Anne Deijten and Anna Mandalakas (systematic reviewers for paediatric TB) and Andrea Pantoja (reviewer for affordability and cost effectiveness of the test) are thanked for preparing the systematic reviews and presenting their findings to the members of the Expert Group.

Funding from the United States Agency for International Development is gratefully acknowledged through USAID-WHO Consolidated Grant No. GHA-G-00-09-00003/US 2012 0392.

Declarations of Interests

The members of the Expert Group, technical resource consultants and members of STAG-TB completed Declarations of Interests (DOIs). These were reviewed by the WHO Steering Group prior to the meeting of the Expert Group and prior to preparing the current policy update. The review of each DOI assessed whether an interest had been declared and, if so, whether it was insignificant or potentially significant. If the interest was assessed as being significant or potentially significant, the declaration was referred to WHO's Legal Department, and their advice on the meeting's procedures was followed. A summary of DOI statements is provided in Annex 4. 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 in 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 the STAG-TB or in drafting WHO's policy update.

Executive summary

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 accuracy of 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 Xpert MTB/RIF were issued by WHO early in 2011,¹ supported by an operational how-to document² and a checklist for implementation at the country level.³

WHO's current policies and guidance recommend 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). 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. Given the amount of additional data on Xpert MTB/RIF that have emerged since 2010, an update of WHO's policies and guidance was warranted. WHO's Global TB Programme therefore commissioned three systematic reviews to update and revise the guidance; these reviews examined the utility of Xpert MTB/RIF in 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. 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 a detailed meeting report is available at:

http://www.who.int/tb/laboratory/policy_statements/en/

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).

Using Xpert MTB/RIF to diagnose pulmonary TB and rifampicin resistance in adults

Twenty seven unique studies involving 9 558 participants were included in the review. The reference standards for detecting pulmonary TB were solid culture or liquid culture. The reference standard for detecting rifampicin resistance was phenotypic culture-based drug-susceptibility testing (DST).

When used as an initial diagnostic test replacing smear microscopy, Xpert MTB/RIF achieved an overall pooled sensitivity of 88% (95% credible interval [CrI], 84–92%)⁴ and a 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, 1 952 participants); for smear-negative culture-positive TB, the pooled sensitivity was 68% (95% CrI, 61–74%) (23 studies, 7 151 participants).

For people living with HIV, the pooled sensitivity of Xpert MTB/RIF was 79% (95% CrI, 70–86%) (7 studies, 1 789 participants); for people without HIV infection, the pooled sensitivity was 86% (95% CrI, 76–92%) (7 studies, 1 470 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, including true negatives and false positives).

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

Fifteen published studies and 7 unpublished studies, involving 5 922 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]) were included in the meta-analysis with a comparison against culture and against a composite reference standard (CRS). In the different studies, the CRS included some combination of a nucleic acid amplification test (NAAT) other than Xpert MTB/RIF, histology, smear, culture, biochemical testing, presenting signs, or a response to treatment with anti-TB therapy (Table 1).

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

Table 1. Meta-analysis of the sensitivity and specificity of Xpert MTB/RIF in diagnosing extrapulmonary TB and rifampicin resistance in adults and children compared against culture as a reference standard as well as against a composite reference standard, by type of extrapulmonary specimen

Specimen type	Comparison (No. of studies, No. of samples)	Median (%) pooled sensitivity (pooled 95% CrI)	Median (%) pooled specificity (pooled 95% CrI)
Lymph node tissue and aspirate	Xpert MTB/RIF compared against culture (14 studies, 849 samples)	84.9 (72–92)	92.5 (80–97)
	Xpert MTB/RIF compared against a composite reference standard (5 studies, 1 unpublished)	83.7 (74–90)	99.2 (88–100)
Cerebrospinal fluid	Xpert MTB/RIF compared against culture (16 studies, 709 samples)	79.5 (62–90)	98.6 (96–100)
	Xpert MTB/RIF compared against a composite reference standard (6 studies, 512 samples)	55.5 (51–81)	98.8 (95–100)
Pleural fluid	Xpert MTB/RIF compared against culture (17 studies, 1385 samples)	43.7 (25–65)	98.1 (95–99)
	Xpert MTB/RIF compared against a composite reference standard (7 studies, 698 samples)	17 (8–34)	99.9 (94–100)
Gastric lavage and aspirate	Xpert MTB/RIF compared against culture (12 studies, 1258 samples)	83.8 (66–93)	98.1 (92–100)
Other tissue samples	Xpert MTB/RIF compared against culture (12 studies, 699 samples)	81.2 (68–90)	98.1 (87–100)

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

The data for additional sample types (such as, ascitic fluid, pericardial fluid, urine, blood and stool) were limited and therefore not considered in the analysis.

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.

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 presumed to have TB was 66% in 10 studies where expectorated sputum or induced sputum was used (95% CrI, 52–77%); the pooled sensitivity was 66% in 7 studies where samples from gastric lavage or aspiration 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 against clinical TB used as the reference standard was very low at 4% for samples of expectorated or induced sputum (8 studies), and 15% for samples from gastric lavage or aspiration (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%).

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.

WHO's policy recommendations

Box 1. Using Xpert MTB/RIF to diagnose pulmonary TB and rifampicin resistance in adults and children

These recommendations should be read in conjunction with the remarks in section 5.1.

- 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 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, very low-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).
- Xpert MTB/RIF may be used rather than conventional microscopy and culture as the initial diagnostic test in all children suspected of having TB (conditional recommendation acknowledging resource implications, very low-quality evidence).
- Xpert MTB/RIF may be used as a follow-on test to microscopy in adults suspected of having TB but not at risk of MDR-TB or HIV-associated TB, especially when further testing of smear-negative specimens is necessary (conditional recommendation acknowledging resource implications, high-quality evidence).

Box 2. Using Xpert MTB/RIF to diagnose extrapulmonary TB and rifampicin resistance in adults and children

These recommendations should be read in conjunction with the remarks in section 5.2.

- Xpert MTB/RIF should be used in preference to conventional microscopy and culture as the initial diagnostic test for CSF specimens from patients suspected of having TB meningitis (strong recommendation given the urgency for rapid diagnosis, very low-quality evidence).
- Xpert MTB/RIF may be used as a replacement test for usual practice (including conventional microscopy, culture or histopathology) for testing specific nonrespiratory specimens (lymph nodes and other tissues) from patients suspected of having extrapulmonary TB (conditional recommendation, very low-quality evidence).

POLICY UPDATE

1. 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 on using the test. Policy recommendations on using the Xpert MTB/RIF assay were issued by WHO early in 2011⁵, supported by an operational how-to document⁶ and a checklist for implementation at the country level⁷.

WHO's current policies and guidance recommend 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). 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, the recommendation generalizes from data on adults to include the use of Xpert MTB/RIF in children.

Extrapulmonary TB accounts for about 25% of all cases of TB and an even higher

percentage of cases in children and in people who are immunocompromised. Diagnosing extrapulmonary TB is often challenging, requiring the clinician to obtain specimens for microscopy, culture and histopathology from the suspected sites of involvement. However, the availability of these tests is limited and the need for alternative tests to use to diagnose TB in nonrespiratory samples is great. In 2011, the global burden of TB in children was estimated to be 500 000 cases, representing approximately 6% of all cases of TB. However, in all likelihood this burden has been underestimated due to the difficulties associated with 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.

Since 2010, at least 85 peer-reviewed research papers have been published about using Xpert MTB/RIF to diagnose pulmonary, extrapulmonary and paediatric TB, and studies continue to be performed. Given the amount of additional data on Xpert MTB/RIF that have emerged since 2010, an update of WHO's policies and 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

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).

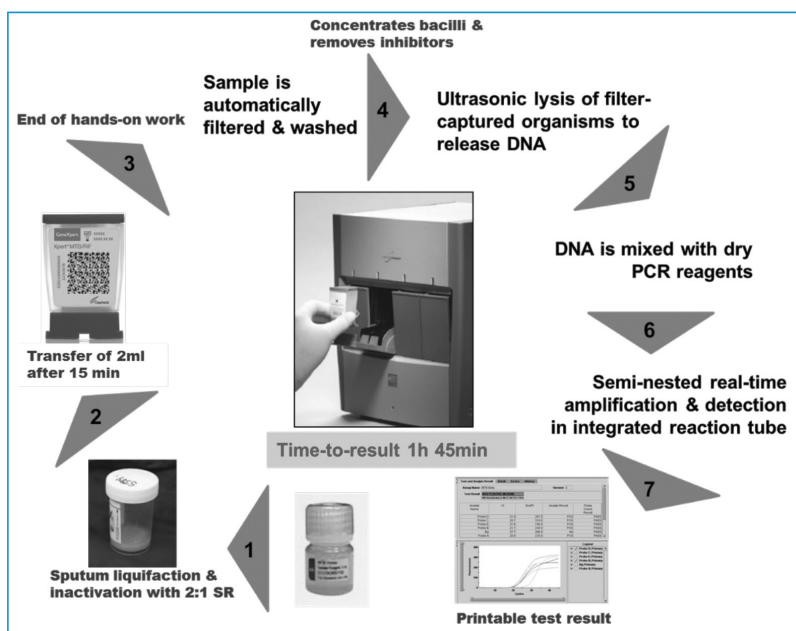
for diagnosing TB and rifampicin resistance in pulmonary, extrapulmonary and paediatric TB. Published studies on the affordability and cost effectiveness of Xpert MTB/RIF were also reviewed. WHO convened an Expert Group to review the evidence at Les Pensières, Veyrier-du-Lac, France, during 20–21 May 2013.

Xpert MTB/RIF is an automated polymerase chain reaction (PCR) test (that is, a molecular test) utilizing the GeneXpert platform (Cepheid, Sunnyvale, CA, United States). Xpert MTB/RIF is a single test that can detect both *Mycobacterium tuberculosis* complex and rifampicin resistance within 2 hours after starting the assay, with minimal hands-on technical time (Figure 1). Unlike conventional nucleic acid amplification tests (NAATs), in Xpert MTB/RIF sample processing, PCR amplification and detection are integrated into a single self-enclosed test unit, which is the Xpert MTB/RIF cartridge. Following sample loading, all steps in the assay are automated and

contained within the cartridge. In addition, the assay's sample reagent, used to liquefy sputum, is tuberculocidal (that is, it has the ability to kill TB bacteria), which largely eliminates concerns about biosafety during the test procedure. These features allow the technology to be taken out of a central laboratory or reference laboratory and to be 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 modules⁸.

The test procedure may be used directly on clinical specimens, either fresh sputum samples or sputum pellets (also called sputum sediment), which are obtained 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 15 minutes. After incubation, 2 ml of the treated sample are transferred to the cartridge, and the run is initiated.

Figure 1. Steps in using the Xpert MTB/RIF assay^a



^a Figure used with permission from the Foundation for Innovative New Diagnostics (FIND).

8 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).

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 *rpoB* gene of TB. Five different coloured beacons are used, each covering a separate nucleic acid sequence within the amplified *rpoB* 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; TB detected, rifampicin resistance indeterminate; and an invalid result.

2. Methods

2.1 Evidence synthesis

In May 2013, a guideline development group (referred to as the Expert Group in this document) was convened by WHO's Global TB Programme to assess the data on Xpert MTB/RIF with a view to updating WHO's 2011 policy recommendations on its use. WHO commissioned three systematic reviews on the use of Xpert MTB/RIF to diagnose pulmonary, extrapulmonary and paediatric TB and to detect rifampicin resistance, as well as a review of the affordability and cost effectiveness of Xpert MTB/RIF.

In accordance with WHO's standards for assessing evidence when formulating policy recommendations, the Grading of Recommendations Assessment, Development and Evaluation (GRADE)

system⁹ was used for the evidence synthesis process to provide a systematic, structured framework for evaluating both the accuracy of the test and the test's impact on patients and public health. The evaluations used the GRADE system to determine the quality of the evidence and provide information on the strength of the recommendations using a priori questions (that is, 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 3.

9 Schünemann HJ et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ*, 2008, 336:1106-1110.

Box 3. 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 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.

Three systematic reviews were conducted according to the standards outlined by the Cochrane Collaboration in the Cochrane handbook.¹⁰ A comprehensive search of the following databases was performed: Cochrane Infectious Diseases Group Specialized Register, MEDLINE, Embase, Thomson Reuters (formerly ISI) Web of Knowledge, Medion, LILACS, BIOSIS and Scopus. Searches of the metaRegister of Controlled Trials and the search portal of WHO's International Clinical Trials Registry were also performed to identify any continuing trials. Searches included all studies without restrictions on language.

A comprehensive search of the literature for studies on the cost effectiveness and affordability of Xpert MTB/RIF included PubMed, the Health Economics Evaluation Database, the United Kingdom's National Health Service Economic Evaluation Database, the Cost-effectiveness Analysis Registry and the Research Papers in Economics database.

Where feasible, meta-analysis was used to summarize the results of independent studies, and these results have been displayed as forest plots. Where meta-analysis was not feasible due to heterogeneity, the evidence has been presented in a narrative synthesis.

¹⁰ Higgins JPT, Green S, eds. Cochrane handbook of systematic reviews for interventions, version 5.1.0. Cochrane Collaboration, 2011 (available at <http://www.cochrane-handbook.org>).

Each reviewer prepared GRADE evidence profiles for each PICO question. GRADE evidence profiles were prepared to assess the diagnostic accuracy of Xpert MTB/RIF for each systematic review. In systematic reviews that assess the accuracy of diagnostic tests, the choice of an optimal reference standard is critical, since the reference standard is used to determine the presence or absence of the target condition. The reference standards used for the different systematic reviews are described below.

- **Using Xpert MTB/RIF to detect pulmonary TB and rifampicin resistance:** the reference standards used were conventional culture and drug-susceptibility testing (DST). Culture using either solid media or commercial liquid media, as recommended by WHO, was considered to be an acceptable reference standard. The reference methods for DST for rifampicin resistance were those recommended by WHO, and these included molecular line probe assays.
- **Using Xpert MTB/RIF to detect extrapulmonary TB in adults and children:** the reference standard used was either conventional culture (as described above) or a composite reference standard (CRS) defined by the authors of the individual studies that were included in the systematic review; 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.
- **Using Xpert MTB/RIF to detect TB and rifampicin resistance in children:** the reference standard used was either conventional culture (as described above) or a clinical TB reference standard, recognizing the limitations of mycobacterial culture in children. Children were categorized as positive using the clinical TB standard if they had started anti-TB therapy as a result of a clinical diagnosis of

TB. This broad clinical reference standard was chosen in order to accommodate the heterogeneous methods and clinical definitions used in the studies. Children assigned to the group clinical not TB (that is, negative according to the reference standard clinical TB) 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.

Using the GRADE framework, results for sensitivity and specificity were used as proxy measures for outcomes seen as important to patients; these outcomes were based on the relative importance or impact of false-positive and false-negative results. Poor sensitivity would result in false-negative results so that patients with TB or MDR-TB would be missed, and this would have negative consequences in terms of morbidity, mortality and transmission of disease. Poor specificity would result in false-positive results so that patients without TB or MDR-TB would be prescribed unnecessary treatment, and this would have negative consequences that might include serious adverse events related to treatment with 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 to have 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 alternative diagnosis;

- false positives – the likelihood of anxiety and morbidity caused by additional testing, 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.
- critical outcomes – diagnostic accuracy as reflected by true-positive, true-negative, false-positive and false-negative results; time to diagnosis;
- important outcome – cost.

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. The detailed report was prepared by the Steering Group; the report went through several iterations before being finally signed off by members of the Expert Group.

For each outcome, the quality of evidence according to GRADE was initially regarded as high since all studies were cross-sectional or cohort studies, prospectively enrolling patients suspected of having TB or MDR-TB. The quality of the evidence and the limitations of the studies were assessed using six GRADE criteria: (1) study design, (2) risk of bias, (3) directness, (4) inconsistency, (5) imprecision, and (6) publication or reporting bias.

Each review used the Quality Assessment of Diagnostic Accuracy Studies (QUADAS -2) tool to appraise the studies.¹¹ This tool consists of four domains: patient selection, index test, reference standard, and flow and timing.

2.2 Expert Group meeting

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 the Expert Group to identify any other important 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:

2.3 External review

The findings and recommendations from the Expert Group's meeting were presented in June 2013 to WHO's Strategic and Technical Advisory Group for TB (STAG-TB), members of which served as the external review group. STAG-TB agreed with the Expert Group's recommendations, and advised WHO to develop and disseminate an updated policy on using Xpert MTB/RIF to diagnose TB and rifampicin resistance in pulmonary and extrapulmonary TB in adults and children. STAG-TB also recommended that WHO should continue its global coordination and implementation plan aimed at scaling up the use of the Xpert MTB/RIF assay.¹²

2.4 Preparing the policy update

The draft policy update, which was based on the consensus recommendations made by the Expert Group, was subsequently prepared by the Steering Group and circulated to the Expert Group and members of STAG-TB following an iterative process similar to that described above.

11 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.

12 Strategic and Technical Advisory Group for Tuberculosis (STAG-TB): report of the 13th meeting. Geneva, World Health Organization, 2013 (WHO/HTM/TB/2013.09) available at http://www.who.int/tb/advisory_bodies/STAG_report2013.pdf

3. Scope

This document replaces WHO's 2011 policy statement on Xpert MTB/RIF¹³, and provides a pragmatic summary of the updated evidence and recommendations on using Xpert MTB/RIF to diagnose pulmonary and extrapulmonary TB, and to detect rifampicin resistance in adults and children. It should be read in conjunction with the detailed findings from the 2013 report of the Expert Group's meeting¹⁴ and the framework for implementing TB diagnostics. These documents are available at <http://www.who.int/tb/dots/laboratory/policy/en>; they provide guidance on implementing the diagnostic tools and methods approved by WHO within the context of a country's infrastructure, resources, epidemiology of TB and MDR-TB, and on TB policy reform. This policy update will be supported by the second edition of WHO's Xpert MTB/RIF implementation manual (a publication derived from the guidance on policy).¹⁵

None of the existing tools for diagnosing TB are mutually exclusive, and they can be implemented in various combinations in screening and diagnostic algorithms, which can be tailored to be highly specific to each country's settings and resources. Therefore, input from experts working in TB laboratories is needed to define the most cost-effective and efficient algorithms for individual countries; these algorithms should be guided by WHO's standards and procedures, and

implemented within a framework of integrated activities aimed at strengthening laboratories.

This policy guidance should be used to support the implementation of Xpert MTB/RIF technology to diagnose TB and detect rifampicin resistance within TB and TB–HIV programmes. The policies are intended to be used by managers and laboratory directors working in these programmes in coordination with external laboratory consultants, donor agencies, technical advisers, laboratory technicians, procurement officers for laboratory equipment, service providers in the private sector, relevant government sectors, and implementation partners that are involved in country-level strengthening of TB laboratories. Individuals responsible for programme planning, budgeting, mobilizing resources and implementing training activities for TB and TB–HIV diagnostic services may also benefit from this document.

3.1 Date of review: 2017

Additional data from implementation sites will be reviewed annually, and the guidance will continue to be refined based on more extensive field evaluations of the new technology that will be conducted after implementation; these evaluations will include country-specific cost-effectiveness and cost-benefit analyses.

13 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).

14 Using the Xpert MTB/RIF assay to detect pulmonary and extrapulmonary tuberculosis and rifampicin resistance in adults and children: Expert Group meeting report. Geneva, World Health Organization, 2013 (available at http://www.who.int/tb/laboratory/policy_statements/en/)

15 Xpert MTB/RIF implementation manual. Technical and operational 'how-to': practical considerations. 2nd ed. Geneva, World Health Organization, 2014 (available at http://www.who.int/tb/laboratory/policy_statements/en/)

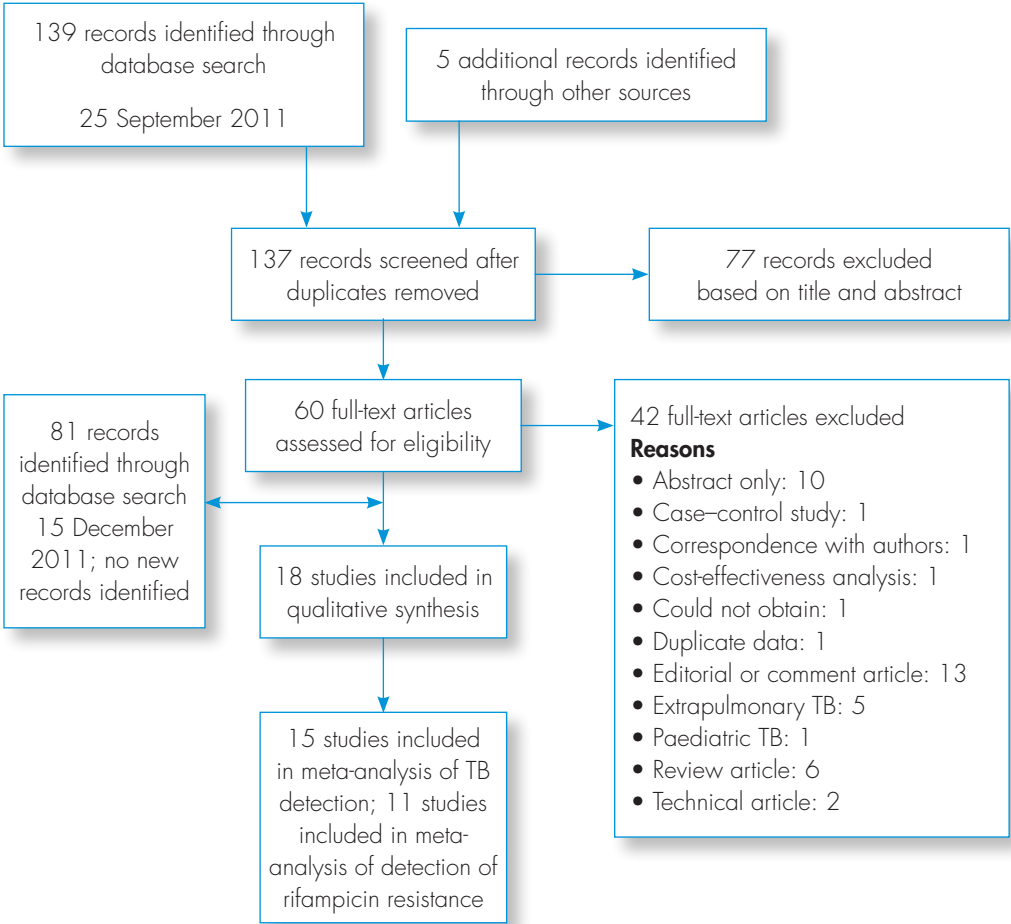
4. Evidence base for policy formulation

4.1 Using Xpert MTB/RIF to diagnose pulmonary TB and rifampicin resistance in adults

A Cochrane review published 31 January 2013¹⁶ evaluated 18 studies (Figure 2). An additional literature search was performed on

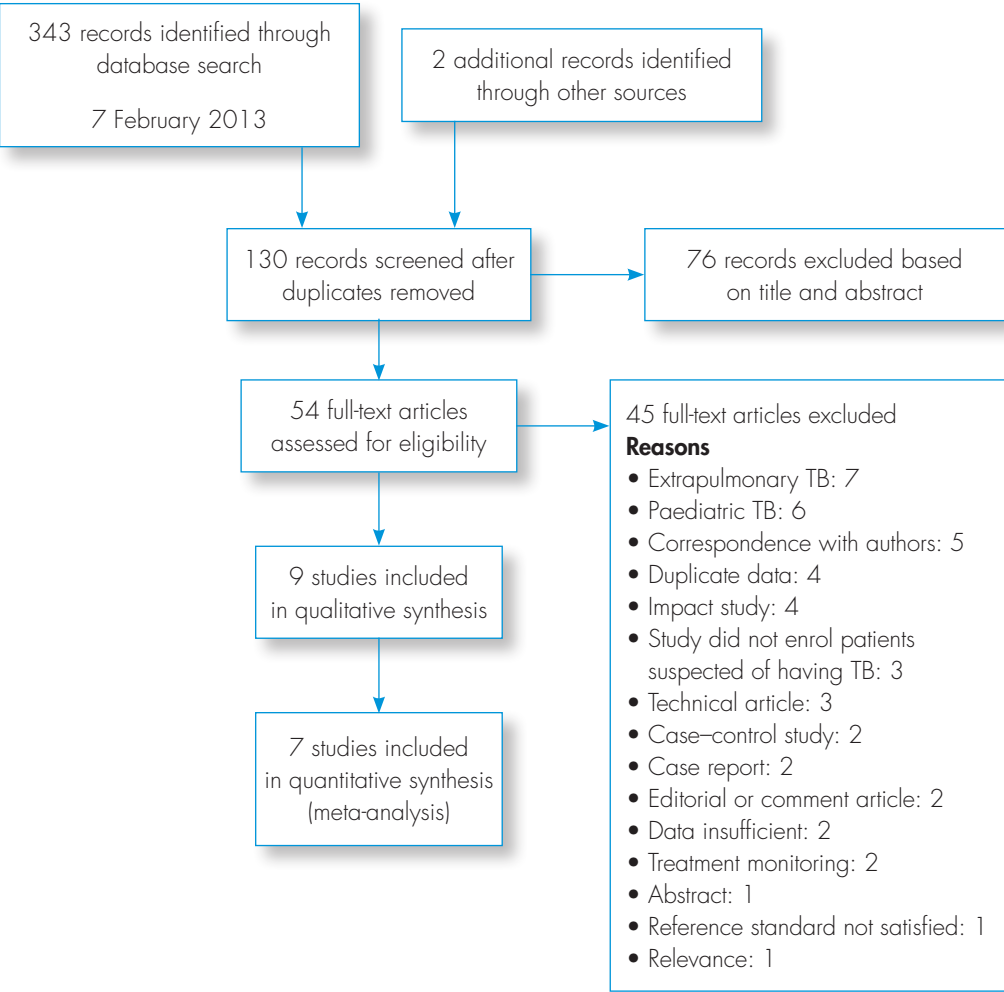
7 February 2013, which identified 9 additional studies (Figure 3). From the combined literature searches, 27 relevant studies in 36 study centres (26 published studies and 1 unpublished study) were identified and evaluated.

Figure 2. Selection of studies evaluating the accuracy of Xpert MTB/RIF in diagnosing pulmonary TB and rifampicin resistance in adults: flow diagram of studies identified by the initial literature searches



16 Steingart KR et al. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. Cochrane Database of Systematic Reviews, 2013, (1):CD009593 (doi:10.1002/14651858.CD009593.pub2).

Figure 3. Selection of studies evaluating the accuracy of Xpert MTB/RIF in diagnosing pulmonary TB and rifampicin resistance in adults: flow diagram of studies identified by the updated literature search

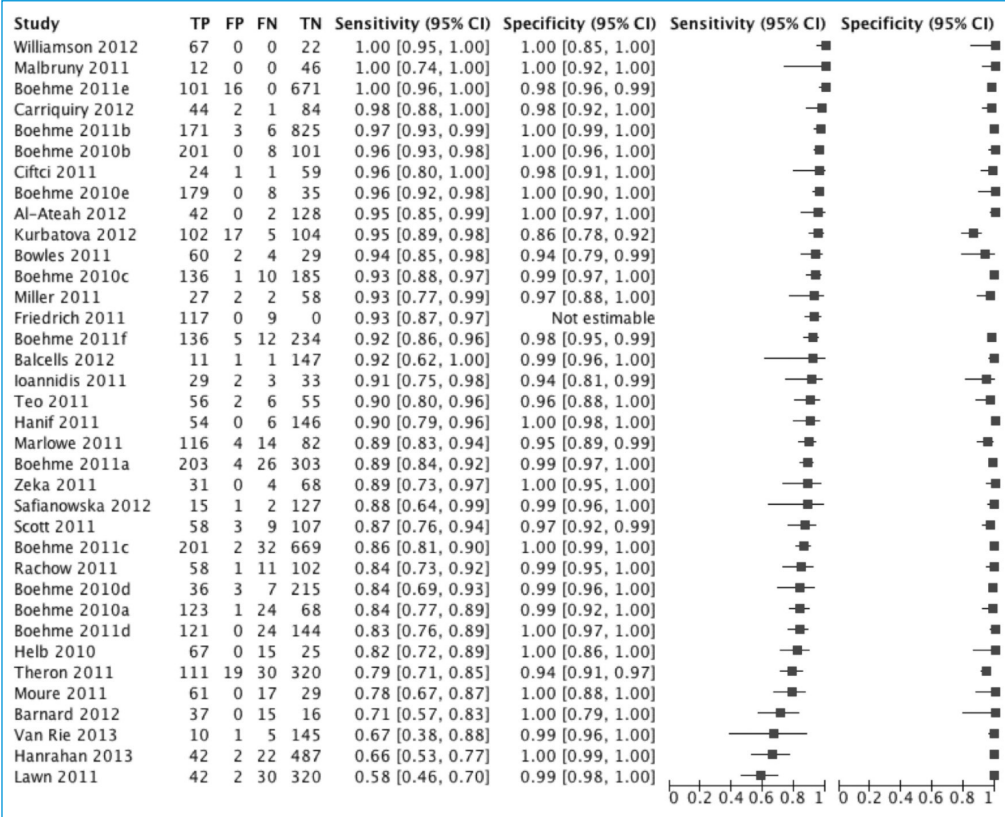


To assess the accuracy of Xpert MTB/RIF in detecting TB, data were reviewed from 27 studies that included 9558 participants; 22 of the studies, involving 9008 participants, were included in the meta-analysis. Five studies that enrolled primarily smear-positive or smear-

negative patients were excluded. The pooled sensitivity of Xpert MTB/RIF for detecting TB was 88% (95% credible interval [CrI], 84–92%); the pooled specificity for detecting TB was 99% (95% CrI, 98–99%) (Figure 4).¹⁷

¹⁷ The credible interval (CrI) is the Bayesian equivalent of the confidence interval (CI).

Figure 4. 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).

4.1.1 Using Xpert MTB/RIF as a replacement test for smear microscopy

Twenty-one studies (8880 participants) provided data that compared the sensitivity of Xpert MTB/RIF with 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%) (Table 2).

Table 2. 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).

4.1.2 Using Xpert MTB/RIF as an add-on test following microscopy

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%). In other words, 68% of smear-negative culture-confirmed cases of TB were detected using Xpert MTB/RIF following smear microscopy, which increased case-detection by 68% (95% CrI, 61–74%) in this group (Table 2).

4.1.3 Using Xpert MTB/RIF to detect smear-positive culture-positive TB

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

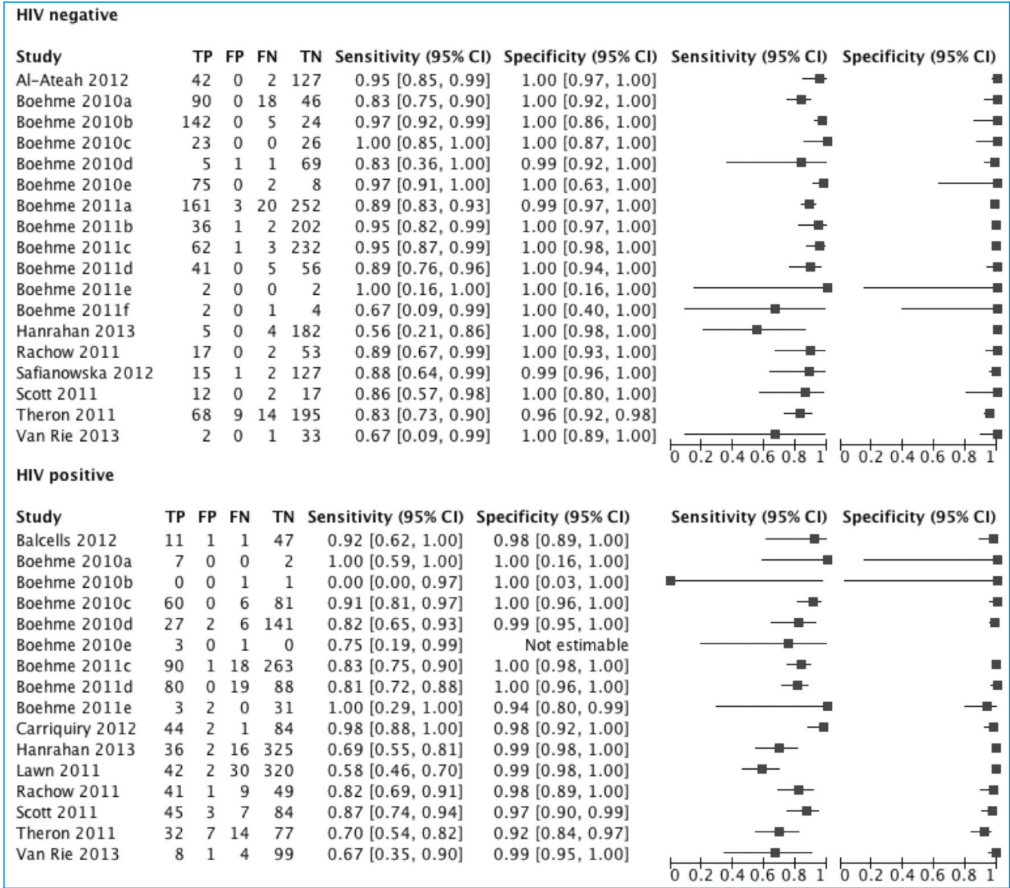
4.1.4 Using Xpert MTB/RIF to detect smear-negative culture-positive TB

Twenty four studies (33 study centres, 7247 participants) reported data on smear-negative cases. 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 allowed for direct comparison between subgroups that were smear-positive and those that were smear-negative. The pooled estimate of sensitivity for smear-negative culture-positive TB was 68% (95% CrI, 61–74%).

4.1.5 Using Xpert MTB/RIF to detect pulmonary TB in HIV-negative and HIV-positive individuals

Nine studies (18 study centres, 2555 participants) reported data on HIV-negative individuals, and 10 studies (16 study centres, 2378 participants) reported on HIV-positive individuals (Figure 5). There was variability in sensitivity in both the HIV-negative subgroup (range, 56–100%) and 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: from 96% to 100% in the HIV-negative subgroup, and from 92% to 100% in the HIV-positive subgroup.

Figure 5. 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).

The meta-analysis included 7 studies that provided data on both HIV-negative individuals (1470 participants) and HIV-positive individuals (1789 participants). The pooled sensitivity for the HIV-negative subgroup was 86% (95% CrI, 76–92%); for the HIV-positive subgroup it was 79% (95% CrI, 70–86%) .

The corresponding pooled specificities were similar: for the HIV-negative subgroup the specificity was 99% (95% CrI, 98–100%); for 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, suggesting that some of the differences between the HIV-positive subgroup and the HIV-negative subgroup could be attributed to differences in smear status.

Five studies reported data from which it was possible to assess the accuracy of Xpert MTB/RIF in HIV-positive individuals with culture-positive smear-negative TB. The sensitivity of Xpert MTB/RIF in HIV-positive individuals with smear-negative culture-positive TB ranged from 43% to 93%; in HIV-positive individuals with smear-positive culture-positive TB it ranged from 91% to 100%. Data were sufficient to perform a univariate meta-

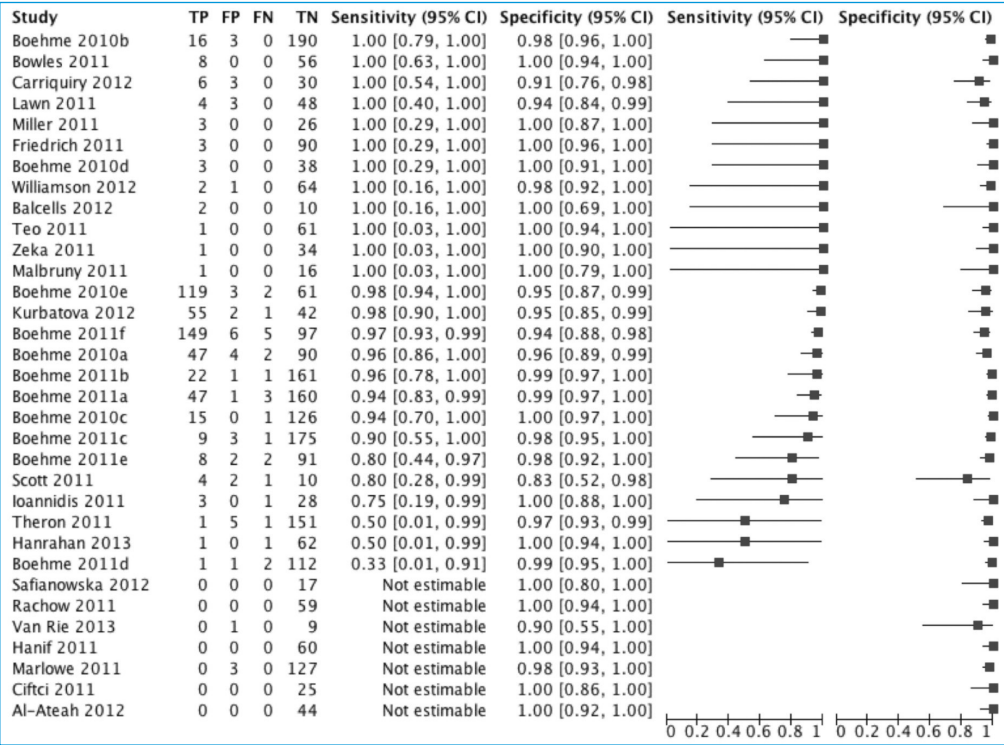
analysis to assess the sensitivity of Xpert MTB/RIF. Among people living with HIV, the pooled sensitivity of Xpert MTB/RIF for smear-negative culture-positive TB was 61% (95% CrI, 42–79%) compared with 97% (95% CrI, 91–99%) for smear-positive culture-positive TB; this was a statistically significant result (data not shown). Hence, among people who are coinfectd 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.

4.1.6 Using Xpert MTB/RIF to detect rifampicin resistance

Of the 27 studies, 24 studies (33 study centres, 2969 participants) provided data on detecting

rifampicin resistance, and included 555 rifampicin-resistant specimens. Figure 6 shows the forest plots of sensitivity and specificity for this analysis. Although there was variability in the estimates of sensitivity (range, 33–100%), in general the poorer estimates of sensitivity were related to study centres that had only a low number of rifampicin-resistant specimens. There was less variability in specificity (range, 83–100%). The pooled sensitivity by univariate analysis was 95% (95% CrI, 90–97%); the pooled specificity was 98% (95% CrI, 97–99%). The pooled sensitivity and specificity were the same when bivariate analysis was used for the subset of studies that provided data on both sensitivity and specificity (17 studies, 2624 participants).

Figure 6. 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).

4.1.7 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 configuration of the Xpert MTB/RIF system 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 Xpert MTB/RIF G2 and G3 cartridges). Another modification was implemented in late 2011 (using the Xpert MTB/RIF G4 cartridge), which changed the molecular beacon sequence of Probe B to improve detection of rifampicin resistance when there were fluctuations in the annealing temperature. Changes to fluidics and software virtually eliminated the signal-loss detection error (known as the 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. By 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 the estimates of either sensitivity or specificity for the Xpert MTB/RIF G1 cartridge when compared with later versions of the cartridge.

4.1.8 Accuracy of the Xpert MTB/RIF G4 cartridge

Two studies used the Xpert MTB/RIF G4 cartridge and provided data suitable to determine specificity. 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 6).

FIND evaluated the diagnostic accuracy of the G4 cartridge¹⁸ in a study that tested 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 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 culture using the BACTEC MGIT (mycobacterial growth indicator tube) 960 Mycobacterial Detection System (Becton Dickinson, Franklin Lakes, NJ, United States), with *M. tuberculosis* species confirmed using Capilia (Tauns Laboratories, Shizuoka, Japan), GenoType

18 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>).

MTBDRplus (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 the BACTEC MGIT 960 and, in a few cases, using only the GenoType MTBDRplus 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-resistant TB 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 them as resistant), the *rpoB* 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 them as rifampicin sensitive, sequencing of the *rpoB* region was performed; discordant results resolved in eight of these cases in favour of Xpert MTB/RIF.

4.1.9 Accuracy of the reference standards

Culture is regarded as the best reference standard for active TB, and was the reference standard used for TB in the systematic review on pulmonary TB. Phenotypic culture-based DST methods, using WHO's recommended critical concentrations, were the reference standards for rifampicin resistance.¹⁹

Three recent studies have raised concerns about using phenotypic DST to detect rifampicin

resistance, in particular the automated BACTEC MGIT 960 system. Van Deun and colleagues reported that the BACTEC 460 and the BACTEC MGIT 960 missed certain strains associated with low-level rifampicin resistance.²⁰ 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 *rpoB* gene but whose results from the BACTEC MGIT 960 indicated that the isolates were rifampicin susceptible. In this study, 2/49 (4.1%) patients whose isolates did not have apparent *rpoB* gene mutations experienced treatment failure compared with 3/3 (100%) patients whose isolates did have *rpoB* gene mutations and had been deemed rifampicin-susceptible using phenotypic methods.²¹

In a study involving retreatment patients, Van Deun and colleagues found that disputed *rpoB* 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.²² 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).

Therefore, using only phenotypic DST as a reference to determine the specificity of a molecular method of DST may underestimate the specificity of the molecular method of DST. In light of these findings, it is unclear whether and to what extent Xpert MTB/RIF might outperform phenotypic methods of DST in assessing rifampicin resistance.

WHO continues to collect and evaluate emerging data on this issue, and will formally review the

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

²⁰ Van Deun A et al. Mycobacterium tuberculosis strains with highly discordant rifampin susceptibility test results. *Journal of Clinical Microbiology*, 2009, 47:3501–3506.

²¹ 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.

²² Van Deun A et al. Rifampicin drug resistance tests for tuberculosis: challenging the gold standard. *Journal of Clinical Microbiology*, 2013, 51:2633–2640.

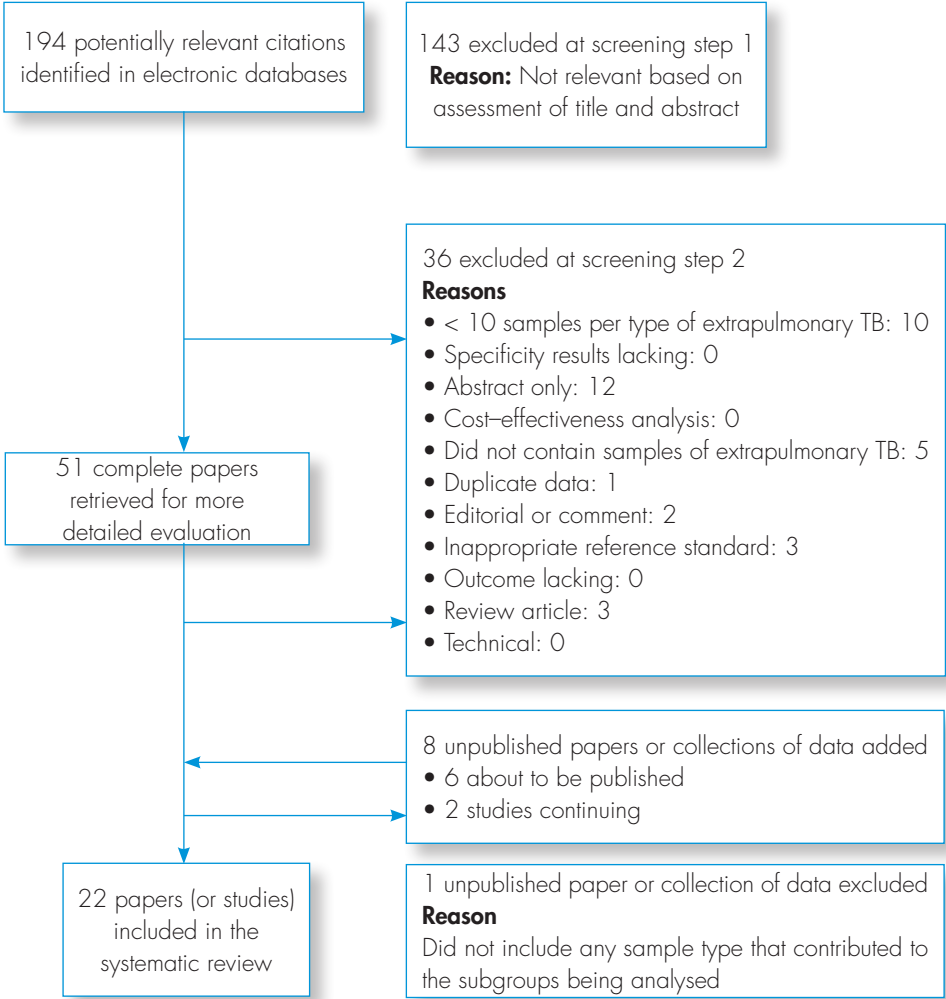
accuracy of phenotypic resistance standards for DST once sufficient data become available.

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

A literature search was performed for published and unpublished reports of studies made available

from 1 January 2007 until 14 December 2012. From the literature searches, 22 relevant studies that included 5922 specimens were identified. Figure 7 shows the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) diagram with the flow of the studies.

Figure 7. Selection of studies evaluating the accuracy of Xpert MTB/RIF in diagnosing extrapulmonary TB and detecting rifampicin resistance in adults and children: flow diagram of studies included in the review



Thirteen studies (59%) were conducted in low-income and middle-income countries. All studies were performed in tertiary care centres or reference laboratories. Two studies (one by Bates and one by Walters) included only children;²³ nine studies included no children. In the remaining 11 studies the percentages of children in the study population ranged from 2% to 34%. Three published studies and four unpublished studies included only one type of sample (for example, only pleural fluid). The remainder of the studies included varying percentages of different types of samples. 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 frozen, archived samples ; 15 used fresh samples; and 1 study used both fresh and frozen samples.

The studies reviewed were diverse with respect to both the different types of samples tested and their relative percentages in each study. The heterogeneity in the performance characteristics of Xpert MTB/RIF, primarily in sensitivity, across the different types of samples was substantial. Therefore, combining these studies to obtain an overall estimate of the accuracy of Xpert MTB/

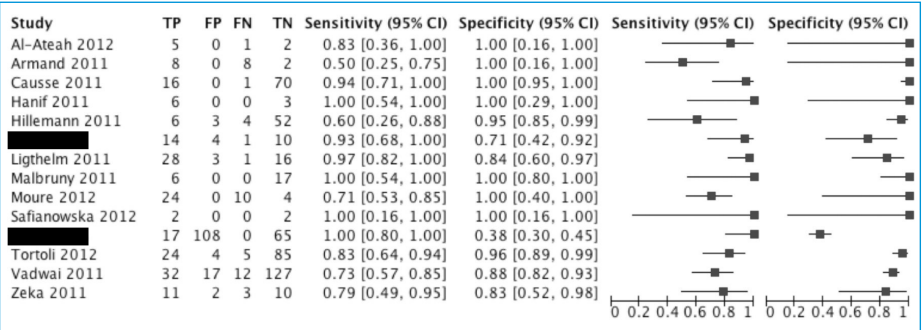
RIF in diagnosing extrapulmonary TB would not be meaningful.

An analysis of predefined subgroups of sample types (that is, pleural fluid, lymph node aspirate or tissue, CSF, gastric fluid, and tissue other than lymph node) was undertaken to account for the heterogeneity among the 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.

4.2.1 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) compared against culture as a reference standard (Figure 8). 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) the estimates for sensitivity ranged from 50% to 100%. The pooled sensitivity across studies was 84.9% (95% CI, 72.1–92.4%); the pooled specificity was 92.5% (95% CI, 80.3–97.4%).

Figure 8. 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.

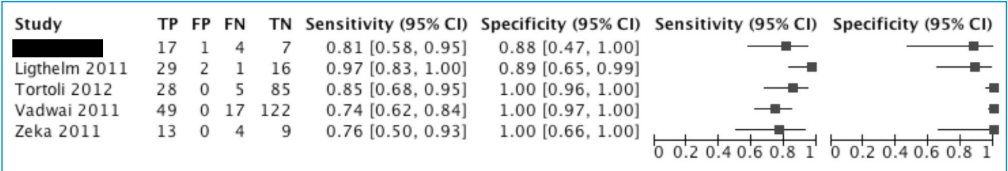
^a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

²³ Additional information about the studies can be found in the annexes to Using the Xpert MTB/RIF assay to detect pulmonary and extrapulmonary tuberculosis and rifampicin resistance in adults and children: Expert Group meeting report. Geneva, World Health Organization, 2013 (available at: http://www.who.int/tb/laboratory/policy_statements/en/)

Five studies (one unpublished) assessed the accuracy of Xpert MTB/RIF on samples from lymph nodes compared against an author-defined CRS (Figure 9). 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%), and the pooled specificity was 99.2% (95% CI, 88.4–100%).

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 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).

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 included 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. Given the limited amount of data for the group that had a prevalence of HIV greater than 10%, a summary estimate was not determined.

4.2.2 Detecting pleural TB in pleural fluid

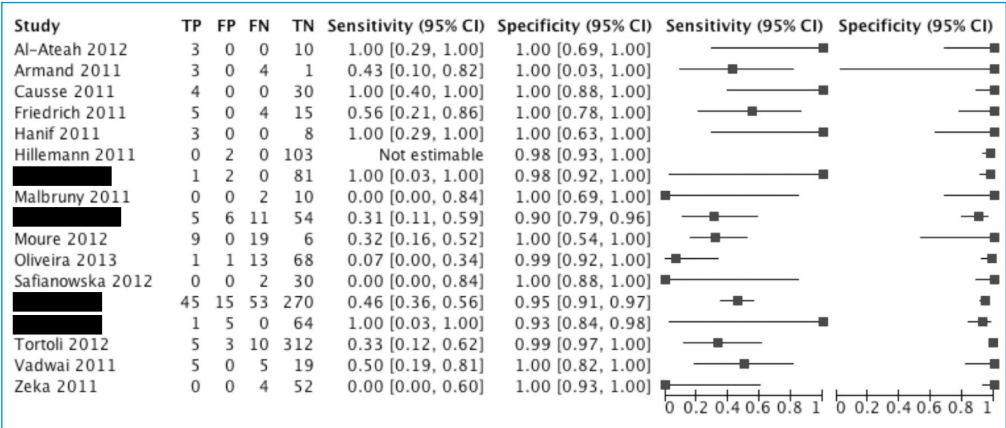
Seventeen studies (1385 samples, 217 culture positive) provided data that could be used to estimate the sensitivity and specificity of Xpert MTB/RIF in testing pleural fluid. Results from the assessment of the accuracy of Xpert MTB/RIF using samples from pleural biopsy were integrated into the assessment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

lymph node.

The sensitivity of Xpert MTB/RIF in testing pleural fluid varied from 0% to 100% among the studies. The outliers at the lower end of the range and the upper end were studies with few culture-confirmed cases of TB. One study was excluded from the meta-analysis because the sensitivity could not be estimated; a second study was also excluded because it included fewer than 10 specimens of pleural fluid. The pooled sensitivity was low at 43.7%, with wide confidence intervals (95% CI, 24.8–64.7%); the pooled specificity was high at 98.1% (95% CI, 95.3–99.2%) (Figure 10).

Seven studies (4 published and 3 unpublished) with 698 samples (188 culture positive) evaluated Xpert MTB/RIF in 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

Figure 10. 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

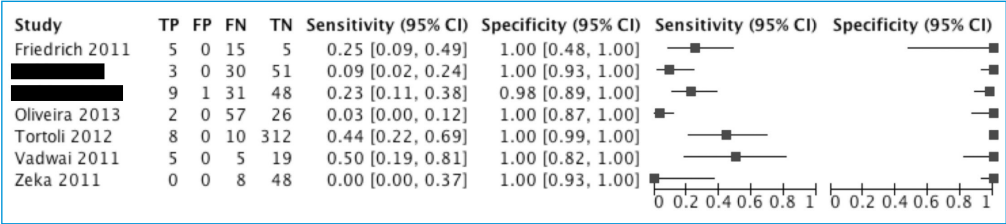


TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval.
a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test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

reference standard, the CRS subgroup yielded an even lower pooled sensitivity (17.0%; 95% CI, 7.5–34.2%), albeit with a high specificity (99.9%; 95% CI, 93.7–100.0%) (Figure 11).

Pleural fluid is not regarded as a suitable specimen for the microbiological diagnosis of pleural TB. Pleural biopsy is the preferred sample type for diagnosing pleural TB.

Figure 11. Forest plot of the sensitivity and specificity of Xpert MTB/RIF in detecting TB in pleural 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).

4.2.3 Detecting TB in samples of CSF

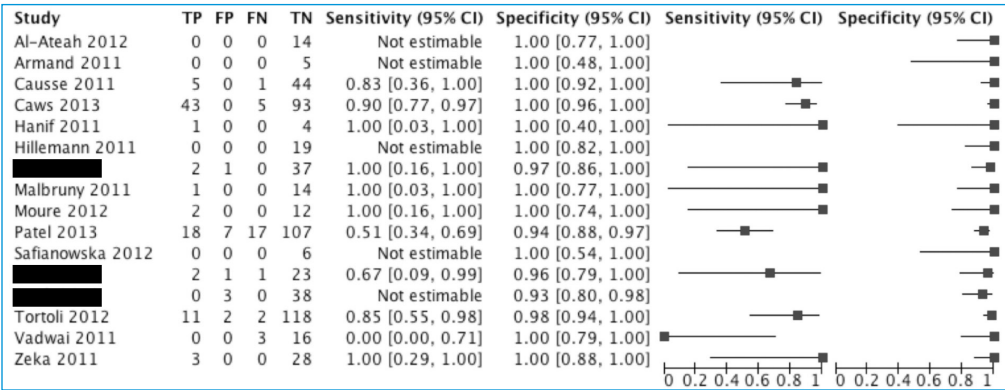
In total, 709 CSF samples in 16 studies were tested with Xpert MTB/RIF, and the results were compared against culture as a reference standard (13 studies had more than 10 samples, and 10 of these provided information on both sensitivity and specificity). Only 117 culture-confirmed cases of TB were found. Estimates of sensitivity varied

widely and ranged from 51% to 100%; one study with 19 samples (3 false negatives) was an outlier at 0%. The pooled sensitivity across studies was 79.5% (95% CI, 62.0–90.2%) and the 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 (Figure 12).

Ten of the sixteen studies comparing Xpert MTB/RIF with culture as a 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 Xpert MTB/RIF (82%; 95% CI, 71–93% for concentrated samples versus 56%; 95% CI 36-77% for unconcentrated samples), although the confidence intervals overlapped. The use of a concentration step did not affect the specificity.

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

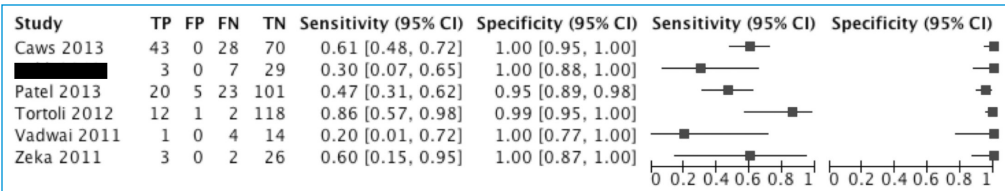


TP, true positive; FP, false positive; FN, false negative; TN, true negative; CI, confidence interval.
a The figure shows the estimated sensitivity and specificity of each study (blue square) and its 95% CI (black horizontal line). The names of unpublished studies have been obscured. Values for test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

Only 6 studies (3 unpublished) assessed the accuracy of using Xpert MTB/RIF to test 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 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).

4.2.4 Detecting TB in gastric fluid

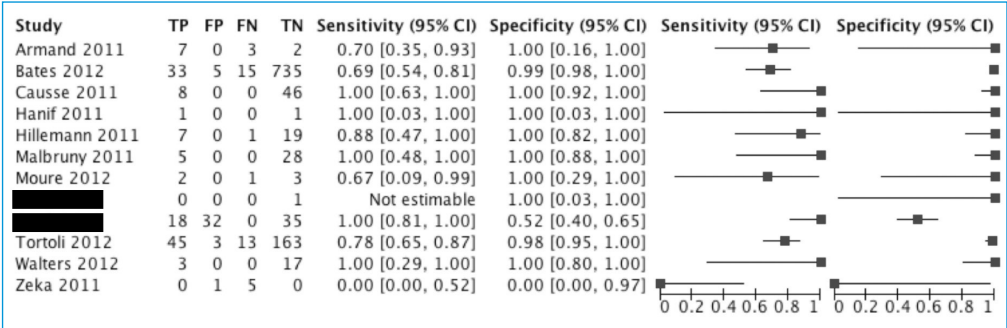
Twelve studies (1258 samples) examined the performance of Xpert MTB/RIF in samples of

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 gastric-fluid samples. The estimates of sensitivity varied from 69% to 100%; specificity

varied from 98% to 100%, with one unpublished study an outlier reporting 52% specificity. The pooled sensitivity across studies was 83.8% (95% CI, 65.9–93.2%), and the pooled specificity was 98.1% (95% CI, 92.3–99.5%) (Figure 14).

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).

All studies used a concentration step during specimen processing and prior to using the Xpert MTB/RIF assay; therefore, this was not considered to be a source of heterogeneity within the data. The condition of the gastric-fluid specimen (fresh versus frozen) did not appear to affect the performance of Xpert MTB/RIF. The five studies testing fresh specimens achieved a pooled sensitivity of 83% (95% CI, 71–93%). A further three studies used frozen specimens and had a pooled sensitivity of 79% (95% CI, 57–100%). The condition of the specimen (fresh or frozen) did not affect the specificity.

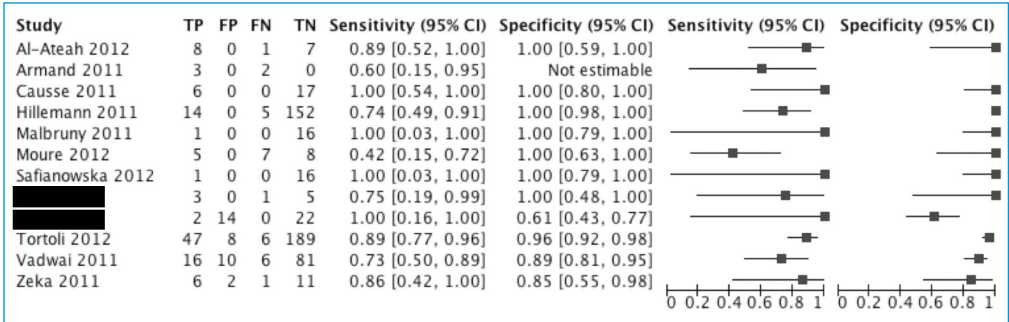
4.2.5 Detecting TB in tissue samples

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%. The pooled estimate of sensitivity was calculated as 81.2% (95% CI, 67.7–89.9%). The pooled specificity was 98.1% (95% CI, 87.0–99.8%) (Figure 15).

The condition of the specimen (fresh versus frozen) did not appear to affect the performance of Xpert MTB/RIF. The five studies testing fresh specimens achieved a pooled sensitivity of 79% (95% CI, 64–94%). A further three studies used frozen specimens and had a pooled sensitivity of 76% (95% CI, 58–94%). The condition of the specimen (fresh or frozen) did not affect the specificity.

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).

4.2.6 Detecting rifampicin resistance

Data on detecting rifampicin resistance were 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 identified as susceptible by phenotypic DST. Five of these six samples underwent sequencing of the *rpoB* gene, and four were found to have a

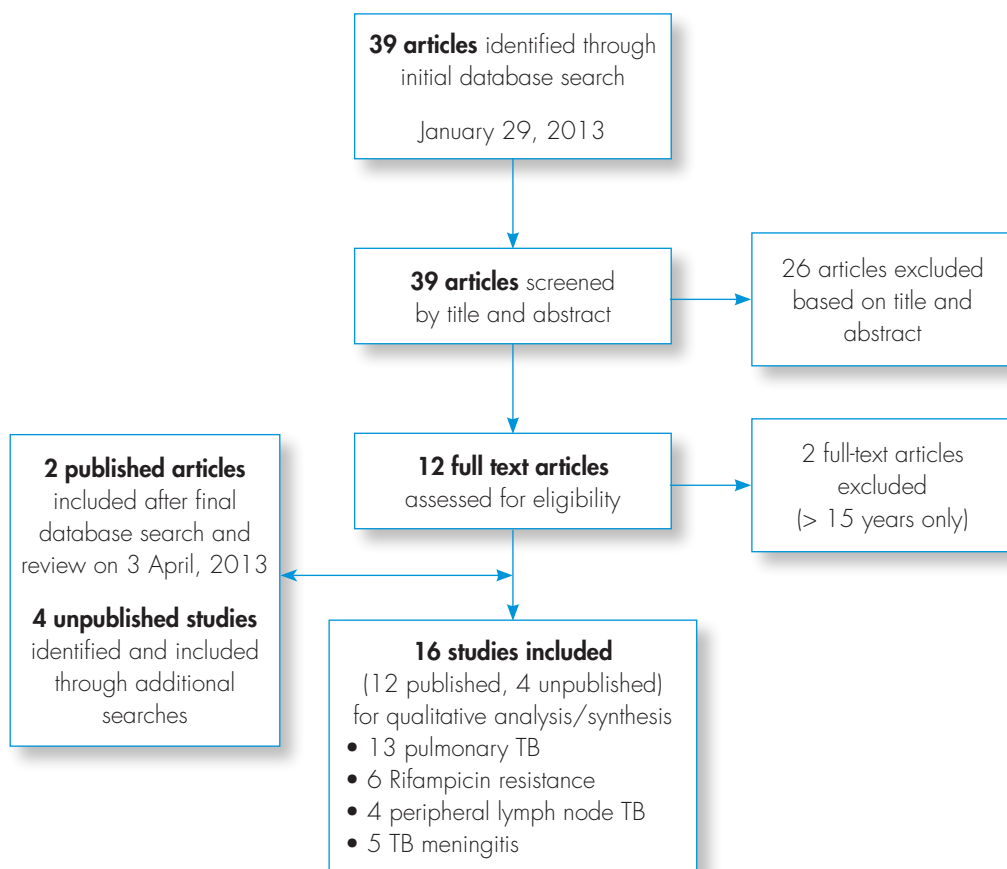
mutation in the same region of the *rpoB* gene at codon 533. Hence, Xpert MTB/RIF detected four additional rifampicin-resistant strains that would have been missed by phenotypic DST alone.

See section 4.1.9 for more information about the accuracy of the reference standards used to detect rifampicin resistance.

4.3 Using Xpert MTB/RIF to diagnose TB and rifampicin resistance in children

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 12 articles that underwent a full-text review, 10 were included 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. In total, 16 studies were included (Figure 16).

Figure 16. Selection of studies evaluating the accuracy of Xpert MTB/RIF in diagnosing TB and rifampicin resistance in children: flow diagram of studies included in the review



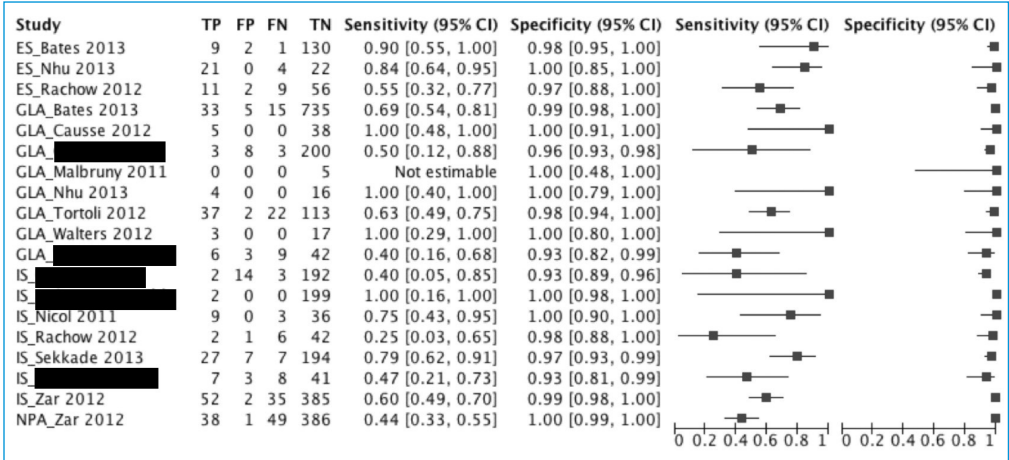
4.3.1 Using Xpert MTB/RIF to diagnose pulmonary TB in children

The utility of Xpert MTB/RIF in diagnosing paediatric 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: 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).

For samples of expectorated sputum, sensitivity varied from 55% to 90%; for induced sputum sensitivity varied from 40% to 100%; and for gastric lavage or aspirate it varied from 40% to 100%. 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% (Figure 17).

Figure 17. 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).

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

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

4.3.1.1 Differences in the reference standards used

In all of 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 based on a clinical diagnosis of TB. This broad clinical reference standard was chosen in order to accommodate the heterogeneous study methods and clinical definitions used 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 did not worsen after at least 1 month of follow-up after enrolment.

4.3.1.2 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 credible intervals indicated there was a high level of heterogeneity among the studies. The specificity values for Xpert MTB/RIF compared against culture as the reference standard were all at least 98%, with narrow credible intervals.

4.3.1.3 Accuracy of Xpert MTB/RIF compared with clinical TB as the reference standard

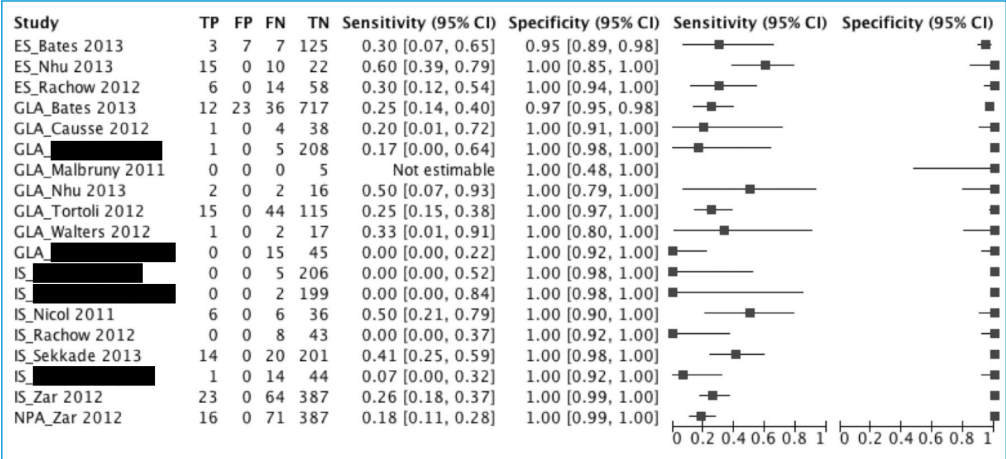
The sensitivity of Xpert MTB/RIF in culture-negative samples from paediatric patients compared against clinical TB as the reference standard was 4% for samples of expectorated or induced sputum and 15% for gastric lavage or aspirate, with all confidence intervals being wide and

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.

4.3.2 Xpert MTB/RIF compared with smear microscopy

The diagnostic accuracy of smear microscopy was calculated and compared against culture as the reference standard. Sensitivities varied from 30% to 60% for samples of expectorated sputum, 0% to 50% for induced sputum, 0% to 50% for gastric lavage or aspirate, and the sensitivity was 18% in the one cohort providing samples of nasopharyngeal aspirate. The 95% confidence intervals were wide and overlapping. The specificity was 95% or higher for all studies and specimen types (Figure 18).

Figure 18. 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, 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).

For expectorated and induced sputum the pooled sensitivity was 29% (95% CrI, 16–42%); for gastric lavage or aspirate it was 22% (95% CrI, 12–35%); the pooled credible intervals were wide for both expectorated or induced sputum and gastric lavage or aspirate, indicating a high

level of heterogeneity. The pooled specificity of Xpert MTB/RIF for samples of gastric lavage or aspirate was 99% (95% CrI, 97–100%); for expectorated and induced sputum it was 100% (95% CrI, 99–100%) (Table 3).

Table 3. 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.

4.3.3 Performance of Xpert MTB/RIF in smear-positive and smear-negative children

Seven studies reporting data from 1083 children were included in the analysis of using Xpert MTB/RIF to test samples of expectorated or induced sputum.

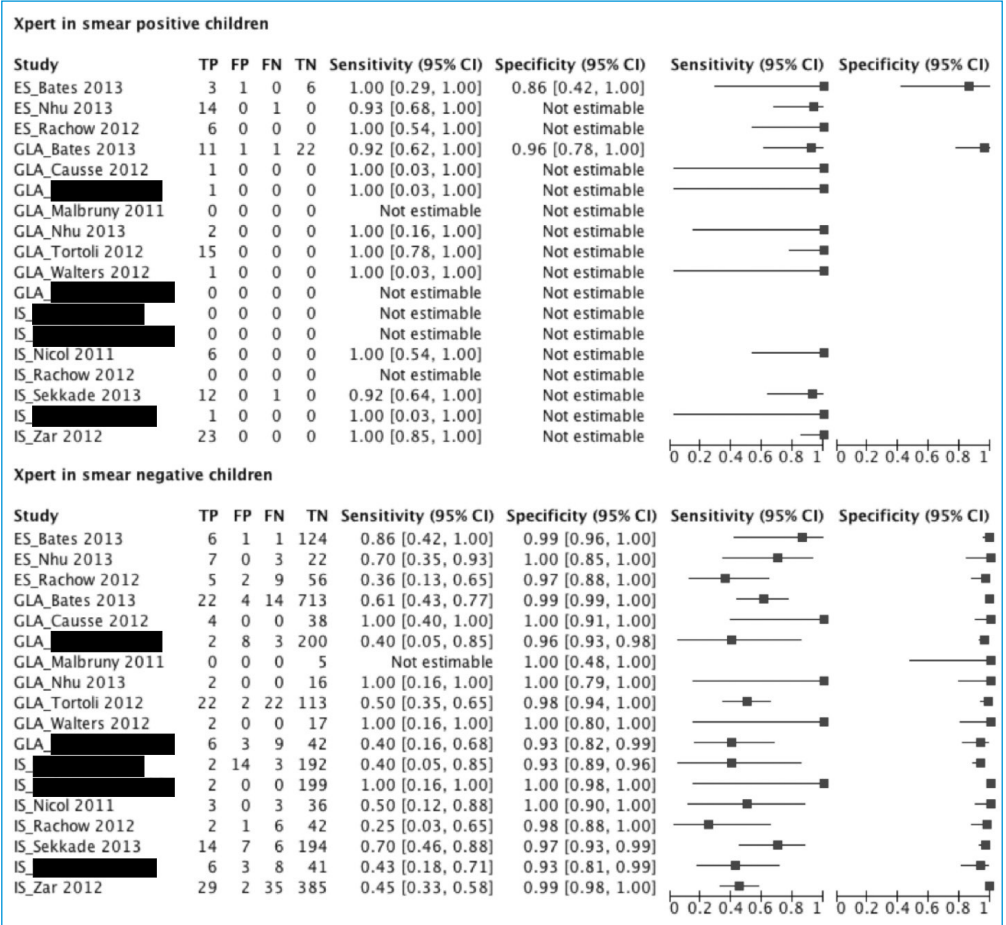
Six studies with data from 1259 children were included in the analysis of using Xpert MTB/RIF to test samples of gastric lavage or aspirate. All studies that were included reported results for both smear-positive and smear-negative children (Figure 19).

The sensitivity of Xpert MTB/RIF on samples of expectorated or induced sputum from children with smear-negative results ranged from 25% to 86%. In contrast, the sensitivity of Xpert MTB/RIF in testing samples of either expectorated or induced sputum from smear-positive children ranged from 92% to 100%.

The pooled estimate of sensitivity for smear-positive children was 96% (95% CrI, 90–99%); for smear-negative children it was 55% (95% CrI, 41–69%).

The findings were similar when Xpert MTB/RIF was used to test samples of gastric lavage or aspirate: for smear-positive children the overall sensitivity was 95% (95% CrI, 83–99%), and for smear-negative children it was 62% (95% CrI, 44–80%). The credible intervals were wide (indicating variability), but did not overlap (Table 4).

Figure 19. 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).

Table 4. 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; additional information about the studies can be found in the annexes to Using the Xpert MTB/RIF assay to detect pulmonary and extrapulmonary tuberculosis and rifampicin resistance in adults and children: Expert Group meeting report. Geneva, World Health Organization, 2013
(available at: http://www.who.int/tb/laboratory/policy_statements/en/)

^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.

4.3.4 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 to test samples of expectorated or

induced sputum. The estimated accuracy of Xpert MTB/RIF in testing samples of gastric lavage or aspirate 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% for Xpert MTB/RIF used to test samples of expectorated or induced sputum (95% CrI, 36–74%) and gastric lavage or aspirate (95% CrI, 38–75%). 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 credible intervals.

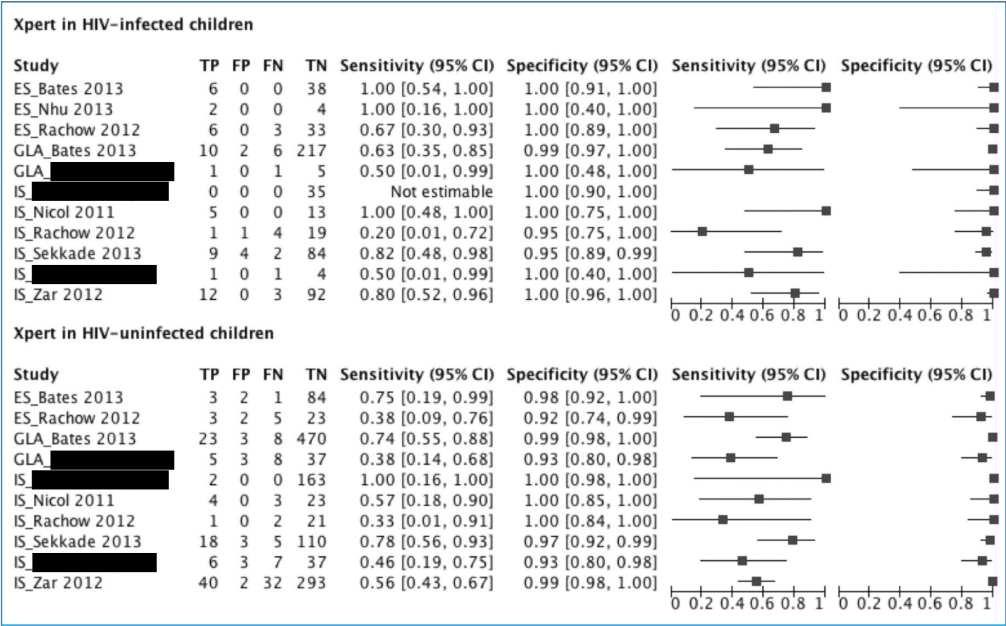
4.3.5 Xpert MTB/RIF in HIV-positive and HIV-negative children

Seven studies reporting data from 1074 children were included in the analysis of the accuracy of Xpert MTB/RIF in testing samples of expectorated

or 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 20).

Figure 20. 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 5).

Table 5. 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; additional information about the studies can be found in the annexes to Using the Xpert MTB/RIF assay to detect pulmonary and extrapulmonary tuberculosis and rifampicin resistance in adults and children: Expert Group meeting report. Geneva, World Health Organization, 2013 (available at: http://www.who.int/tb/laboratory/policy_statements/en/)

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

A metaregression model that simultaneously controlled for smear status and HIV status using Xpert MTB/RIF to test 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 6).

Table 6. 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 perform 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 samples from FNA or lymph node biopsy to diagnose peripheral lymph node TB was evaluated in five studies. Two studies were excluded from the meta-analysis because their samples included fewer than five participants. Therefore, the analysis included 3 studies with data from 172 children (Figure 21). The pooled sensitivity of Xpert MTB/RIF compared against culture as the reference standard was 86% (95% CrI, 65–96%); the pooled specificity was 81% (95% CrI, 54–93%). Credible intervals were wide for both sensitivity and specificity, indicating heterogeneity.

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

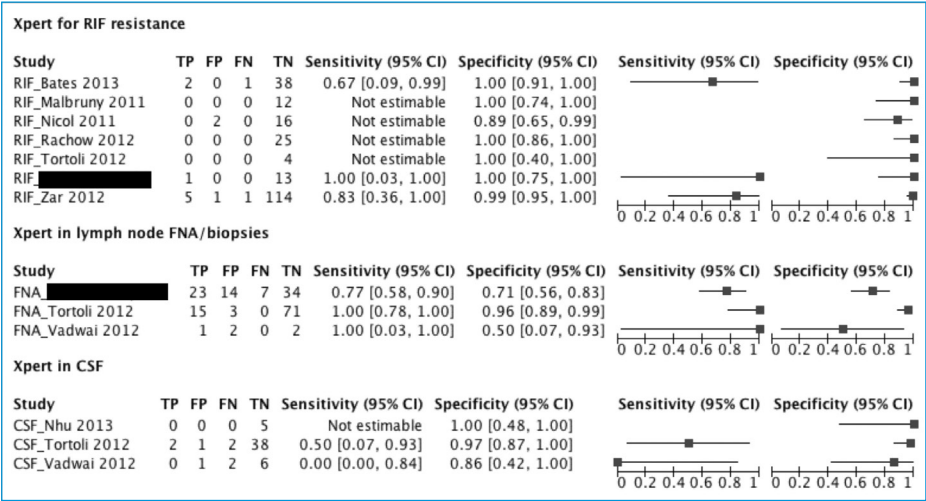
The use of CSF to diagnose TB meningitis was evaluated in five studies that included 61 children.

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 five children. Hence, there were insufficient data to calculate sensitivity from the two remaining studies in which 2/6 (33%) culture-positive children had positive results from Xpert MTB/RIF (Figure 21). 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 21). 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 21. 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).

4.4 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 current diagnostic algorithms. 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 was done (Table 7).

On a global level, the review of the evidence showed that using Xpert MTB/RIF to diagnose TB and MDR-TB was cost effective for all individuals suspected of having TB including those suspected to be coinfected with HIV when compared with current practices. Using Xpert MTB/RIF to diagnose MDR-TB would cost US\$ 0.09 billion per year globally, which is less than the cost of conventional diagnostics used globally and in all countries that have a high burden of TB.

Diagnosing TB in HIV-positive people using Xpert MTB/RIF would cost about US\$ 0.10 billion per

year, and would cost less than the conventional diagnostics used globally in almost all of the countries with a high burden of TB. Testing everyone who has signs and symptoms of TB would cost almost US\$ 0.47 billion per year globally, which is much more than the cost of conventional diagnostics. Nevertheless, in European countries, Brazil and South Africa, the cost would represent less than 10% of current funding for TB programmes.

Introducing Xpert MTB/RIF to diagnose MDR-TB and to diagnose TB in HIV-positive people is warranted in many countries. Using it to test everyone who has signs and symptoms of TB is affordable in several middle-income countries, but its financial viability in low-income countries would require large increases in TB funding or further reductions in the price of the test, or both.

At the country level, the majority of costing and cost-effectiveness studies were from South Africa, so there remains a need for further evidence to be collected from other countries and epidemiological settings.

Table 7. Overview of studies comparing the costs of using Xpert MTB/RIF and follow-on tests with current diagnostic algorithms for diagnosing TB and MDR-TB

Reference ^a	Primary research question	Setting	Population	Baseline diagnostic algorithm	Main comparison	Time horizon	TB transmission considered	Conclusion
Abimbola et al 2012	Cost-effectiveness analysis of using Xpert MTB/RIF to reduce early mortality in individuals initiating ART	Sub-Saharan Africa	HIV-positive individuals initiating ART	Smear microscopy, CXR (for sputum smear-negative individuals), culture	Xpert MTB/RIF as first-line test	6 months	Yes	Xpert MTB/RIF is cost effective as a first-line test
Andrews et al 2012	Cost-effectiveness analysis of using Xpert MTB/RIF to screen for TB in HIV-positive individuals initiating ART	South Africa	HIV-positive individuals initiating ART	Smear microscopy, culture	Xpert MTB/RIF as first-line test	Lifetime	Yes	Xpert MTB/RIF is cost effective as a first-line test
Menzies et al 2012	Cost-effectiveness analysis of using Xpert MTB/RIF to diagnose TB	Botswana, Lesotho, Namibia, South Africa and Swaziland	Individuals suspected to have TB	Smear microscopy, culture for sputum smear-negative individuals	Xpert MTB/RIF as first-line test	10 years and 20 years	Yes	Xpert MTB/RIF is cost effective as a first-line test
Vassall et al 2011	Cost-effectiveness analysis of using Xpert MTB/RIF	India, South Africa and Uganda	Individuals suspected to have TB	Sputum smear-negative individuals with 2 negative smears plus positive clinical diagnosis and/or CXR and culture-based DST for retreatment cases	(1) Xpert MTB/RIF in addition to smear microscopy; (2) Xpert MTB/RIF replacing smear microscopy	Lifetime	No	Xpert MTB/RIF is cost effective in both comparisons

Winetsky et al 2012	Cost-effectiveness analysis of using Xpert MTB/RIF	Prisons in the former Soviet Union	Individuals suspected to have TB (prisoners in a setting with a high burden of MDR-TB)	Symptom screening, self-referral, mass miniature radiography	Xpert MTB/RIF as first-line test	10 years	Yes	Xpert MTB/RIF is cost effective as a first-line test
Meyer-Rath et al 2012	Cost of scaling up use of Xpert MTB/RIF nationally	South Africa	All individuals suspected to have TB	Smear microscopy, CXR for sputum smear-negative individuals, culture	Xpert MTB/RIF as first-line test	Time taken to obtain a diagnosis (duration of diagnostic pathway)	No	Scaling up Xpert MTB/RIF adds 35% more to budget in 2011 (US\$ 293 million versus US\$ 218 million); scale-up increases substantially the number of TB cases and MDR-TB cases diagnosed, and patients started on treatment
Meyer-Rath et al 2011 (policy brief)	Cost of scaling up use of Xpert MTB/RIF nationally	South Africa	All individuals suspected to have TB	Smear microscopy, CXR for sputum smear-negative individuals, culture	Xpert MTB/RIF as first-line test	Time taken to obtain a diagnosis (duration of diagnostic pathway)	No	Annual additional budget required is around US\$ 34–79 million for accelerated scale-up

Pantoja et al 2012	Cost analysis of using Xpert MTB/RIF for target populations	Global analysis, and 36 countries with a high burden of TB and a high burden of MDR-TB	All individuals suspected to have TBT, HIV-positive individuals suspected to have TB, individuals suspected to have MDR-TB	Smeat microscopy, CXR for sputum smear-negative individuals, culture, DST	Xpert MTB/RIF as first-line test for each target population	1 year	No	Using Xpert MTB/RIF to diagnose MDR-TB and TB in HIV-positive individuals costs less than current conventional diagnosis; using Xpert MTB/RIF for all individuals suspected to have TB is much more costly but is affordable in middle-income countries
Schnippel et al 2012	Placement of Xpert MTB/RIF: cost at laboratory level and at clinic level	South Africa	All individuals suspected to have TB	Not applicable	Xpert MTB/RIF at laboratory level and at clinic level	1 year	No	Cost of Xpert MTB/RIF at point-of-treatment (clinic) level is 5.1% more expensive than placement at subdistrict laboratories (US\$ 107 million versus US\$ 71 million)
Schnippel et al 2013	Cost of a second Xpert MTB/RIF test to individuals who were negative by first Xpert MTB/RIF	South Africa	HIV-positive individuals suspected to have TB	Culture for individuals negative by Xpert MTB/RIF	Second Xpert MTB/RIF test for individuals negative by first Xpert MTB/RIF of diagnostic pathway)	Time taken to obtain a diagnosis	No	Cost per TB patient on treatment of offering second Xpert MTB/RIF test to those who initially tested negative is 1.2% less than offering alternative culture

Theron et al 2012	Accuracy and cost of tests used in combination with Xpert MTB/RIF	South Africa	All individuals suspected to have TB	Smear microscopy, CXR, interferon γ release assays (adjunct tests)	Xpert MTB/RIF combined with adjunct tests	Not available	No	Using Xpert MTB/RIF for individuals who are sputum smear-negative had the lowest cost of any strategy, and the highest accuracy
Van Rie et al 2013	Cost, yield and turnaround time of using Xpert MTB/RIF to diagnose sputum smear-negative individuals	South Africa	All individuals suspected to have TB: sputum smear-negatives	Smear microscopy culture	Xpert MTB/RIF used to test sputum smear-negative individuals suspected to have TB	Not available	No	Cost per case diagnosed is similar in both strategies; Xpert MTB/RIF is cost-saving for patients

ART, antiretroviral therapy; MDR-TB, multidrug-resistant tuberculosis; CXR, chest X-ray; DST, drug-susceptibility testing.
 a Additional information about the studies can be found in the annexes to Using the Xpert MTB/RIF assay to detect pulmonary and extrapulmonary tuberculosis and rifampicin resistance in adults and children: Expert Group meeting report. Geneva, World Health Organization, 2013 (available at: http://www.who.int/tb/laboratory/policy_statements/en/).

5. WHO's policy recommendations

The GRADE process confirmed that there is a solid evidence base to support the widespread use of Xpert MTB/RIF to detect TB and rifampicin resistance. WHO therefore recommends the use of Xpert MTB/RIF as described below.

5.1 Using Xpert MTB/RIF to diagnose pulmonary TB and rifampicin resistance in adults and children

- 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 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, very low-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).
- Xpert MTB/RIF may be used rather than conventional microscopy and culture as the initial diagnostic test in all **children** suspected of having TB (conditional recommendation acknowledging resource implications, very low-quality evidence).
- Xpert MTB/RIF may be used as a follow-on test to microscopy in **adults** suspected of having TB but who are not at risk of MDR-TB or HIV-associated TB, especially when further testing of smear-negative specimens is necessary (conditional recommendation acknowledging resource implications, high-quality evidence).

Remarks

These recommendations apply to the use of Xpert MTB/RIF for specimens of processed and unprocessed sputum.

These recommendations also apply to specimens of gastric lavage and aspirate from adults and children; the recommendation for adults is based on the generalization of data from children.

These recommendations support the use of a single sputum specimen for diagnostic testing, acknowledging that processing multiple specimens increases the sensitivity of Xpert MTB/RIF but has resource implications.

Children suspected of having pulmonary TB but who have had a single negative result by Xpert MTB/RIF should undergo further diagnostic testing, and a child for whom there is a high clinical suspicion of TB should be treated even if an Xpert MTB/RIF result is negative or if the test is not available.

Conventional microscopy and culture remain essential for monitoring therapy and for performing DST for anti-TB agents other than rifampicin (including for isoniazid and second-line anti-TB agents).

Expanding the scope of the use of Xpert MTB/RIF and its placement in diagnostic algorithms will have significant implications for operational implementation, and its use should be phased in within the context of national strategic plans for TB.

Emerging data have shown that Xpert MTB/RIF detects some rifampicin-resistant strains that are identified as susceptible by phenotypic DST. Sequencing these discordant results usually resolves in favour of Xpert MTB/RIF, and patients missed by phenotypic DST have poor treatment outcomes on first-line treatment.

5.2 Using Xpert MTB/RIF to diagnose extrapulmonary TB and rifampicin resistance in adults and children

- Xpert MTB/RIF should be used in preference to conventional microscopy and culture as the initial diagnostic test for **CSF specimens** from patients suspected of having TB meningitis (strong recommendation given the urgency of rapid diagnosis, very low-quality evidence).
- Xpert MTB/RIF may be used as a replacement test for usual practice (including conventional microscopy, culture or histopathology) for testing specific nonrespiratory specimens (**lymph nodes and other tissues**) from patients suspected of having extrapulmonary TB (conditional recommendation, very low-quality evidence).

Remarks

Individuals suspected of having extrapulmonary TB but who have had a single negative result from Xpert MTB/RIF should undergo further diagnostic testing, and those for whom there is a high clinical suspicion for TB (especially children) should be treated even if an Xpert MTB/RIF result is negative or if the test is not available.

For CSF specimens, Xpert MTB/RIF should be preferentially used instead of culture if the sample

volume is low or if additional specimens cannot be obtained in order to make a quick diagnosis. If sufficient volume of material is available, concentration methods should be used to increase the yield.

Pleural fluid is a suboptimal sample for the bacterial confirmation of pleural TB regardless of the method used. A pleural biopsy is the preferred sample. The sensitivity of Xpert MTB/RIF in testing samples of pleural fluid is very low. Nevertheless, any individual with a positive result from pleural fluid tested by Xpert MTB/RIF should be treated for pleural TB; those with a negative result from Xpert MTB/RIF should have other tests.

Conventional microscopy and culture are essential for monitoring therapy and for performing DST for anti-TB agents other than rifampicin (including for isoniazid and second-line anti-TB agents).

Emerging data have shown that Xpert MTB/RIF detects some rifampicin-resistant strains that are found to be susceptible by phenotypic DST. Sequencing these discordant results usually resolves in favour of Xpert MTB/RIF, and patients missed by phenotypic DST have poor treatment outcomes on first-line treatment.

These recommendations do not apply to samples of stool, urine or blood, given the lack of data on the utility of Xpert MTB/RIF for these specimens.

6. Implementation considerations

A range of implementation issues were identified that need to be addressed to ensure that the use of Xpert MTB/RIF will be optimal.

- Although Xpert MTB/RIF is suitable for use at all levels of the health system, and testing with Xpert MTB/RIF does not require additional laboratory equipment, the sophisticated nature of the device requires

care in handling – that is, a stable and uninterrupted electrical supply is necessary to avoid interrupting the procedure and losing results; security against theft is necessary as is adequate storage space for the cartridges; staff must be assigned to perform the testing; and biosafety procedures must be put in place that are similar to those used for microscopy.

- Scaling up the implementation of Xpert MTB/RIF does not eliminate the need for capacity for conventional TB microscopy, culture and DST. Microscopy or culture, or both, remain essential for monitoring treatment since molecular tests based on DNA detection have not been shown to be suitable. Therefore, **Xpert MTB/RIF should not be used to monitor treatment**. In addition, conventional culture and DST are required to detect resistance to anti-TB agents other than rifampicin.
- The decision to scale-up the implementation of Xpert MTB/RIF should be made by ministries of health **within the context of national plans for managing TB, MDR-TB and HIV-associated TB**; plans for scale-up should consider the country's specific epidemiology, the screening strategies used, how to ensure timely access to quality-assured first-line and second-line anti-TB agents, and whether care-delivery mechanisms are appropriate.
- Countries already using the molecular line probe assay to rapidly diagnose rifampicin resistance may introduce Xpert MTB/RIF at lower levels of the laboratory service, given that the line probe assay is suitable for high-throughput testing at the central or regional laboratory level (for additional information, see http://www.who.int/tb/laboratory/policy_statements/en/).
- The negative predictive value of Xpert MTB/RIF²⁴ used to test for rifampicin resistance is more than 99% both in settings with a low prevalence of rifampicin resistance and in those with a high prevalence of rifampicin resistance – that is, **a negative result accurately excludes the possibility of rifampicin resistance and no further testing is needed** to confirm the negative results.
- The positive predictive value²⁵ of Xpert MTB/RIF used to test for susceptibility to rifampicin exceeds 90% in settings or patient groups where the underlying prevalence of rifampicin resistance is greater than 15%. In settings or patient groups where rifampicin resistance is rare, the positive predictive value is adversely affected. The positive predictive values ranges from 71% to 84% in settings where the prevalence of rifampicin resistance is between 5% and 10%, and it diminishes to less than 70% when the prevalence of underlying rifampicin resistance falls below 5%. **The positive predictive value can be greatly improved by engaging in a careful risk assessment for individual patients and by using targeted testing.**
- It is important to differentiate between new cases of TB and previously treated cases; previously treated cases have a much higher likelihood of MDR-TB. **Even in groups in which the prevalence of MDR-TB is low, testing previously treated TB cases with Xpert MTB/RIF will result in a high positive predictive value for detecting rifampicin resistance.** Testing new cases not at risk for MDR-TB in groups in which the prevalence of MDR-TB is low will result in a lower positive predictive value for rifampicin susceptibility.
- It is uncommon to detect rifampicin resistance in groups in which there is **a low prevalence of MDR-TB**. In patients in a low-prevalence group who initially test positive by Xpert MTB/RIF, **a second Xpert MTB/RIF test** may be used to control for preanalytical and postanalytical errors, and to improve the clinicians' confidence in deciding on an appropriate regimen.
 - In patients whose Xpert MTB/RIF result repeatedly detects rifampicin resistance,

24 The negative predictive value for rifampicin resistance is the proportion of cases diagnosed as rifampicin-susceptible that are truly susceptible.

25 The positive predictive value for rifampicin resistance is the proportion of cases diagnosed as rifampicin-resistant that are truly resistant.

a WHO-recommended treatment regimen for MDR-TB (including isoniazid) should be initiated and optimized following confirmatory testing for rifampicin resistance and additional DST to test for resistance to isoniazid and second-line anti-TB agents.

– Patients with discordant results for rifampicin resistance reported by Xpert MTB/RIF should be commenced on a WHO-recommended regimen of first-line agents.

- Based on the considerations above, confirmatory testing for rifampicin resistance in groups with a low prevalence of MDR-TB shown by Xpert MTB/RIF using other genotypic testing or conventional DST techniques may be unreliable and can produce discordant results which require resolution by DNA sequencing (section 4.1.9).
- Because Xpert MTB/RIF detects resistance only to rifampicin, countries with documented or suspected cases of **extensively drug-resistant TB** (XDR-TB) should establish or expand their **capacity for conventional culture and DST for second-line agents** to ensure quality assured testing of second-line agents, following WHO's policy guidance.
- Scale-up of Xpert MTB/RIF should be implemented within the context of national plans for laboratory strengthening, considering that the GeneXpert system may also provide a technology platform for other diagnostic services, and reduce the costs involved in providing integrated laboratory services.
- Xpert MTB/RIF cartridges and the specimen reagent should be stored at 2–28 °C, following the manufacturer's recommendations. The cartridges are bulky and require substantial storage space.
- The maximum capacity of a single, 4-module GeneXpert instrument is 16–20 specimens per day. Therefore, busier sites will either need several 4-module instruments or larger instruments (16 modules or more),

and these additional requirements will have associated implications for cost and storage.

- The manufacturer's recommended **ambient operating temperature** for the GeneXpert instrument is limited to a maximum of 30 °C. In settings where the ambient temperature regularly exceeds 30 °C, the room where the assay is being done may need to be air conditioned.
- The Xpert MTB/RIF cartridges have a shelf-life of 12 months, posing a challenge in relatively inaccessible areas that have complex customs-clearance procedures. Therefore, inventory will need to be managed carefully by considering the usage, shelf-life and lead time for the delivery of orders.
- The GeneXpert **modules require annual calibration**, which must be performed by an authorized service provider or carried out by exchanging the modules. A remote calibration option is available from the manufacturer. This option does not require that modules be exchanged, and it can be done using a specially designed calibration kit. The calibration kit contains special cartridges that can be run on each module (without a sample) when calibration is due. During this run (which lasts approximately 20 minutes), the instrument will be automatically calibrated. However, if this calibration fails, it will be necessary to exchange the module. The calibration kit has the same shelf-life as the usual cartridges; in 2013 the cartridges and kit could be ordered together at the preferential pricing of US\$ 450.00 for a kit sufficient to calibrate up to 4 modules.
- **The initial capital** cost for the GeneXpert unit (US\$ 17 000 per 4-module desktop unit or US\$ 17 500 per 4-module laptop unit) is significantly higher than the cost of microscopy (around US\$ 1 500 per microscope), but it is much lower than

for conventional culture and DST (up to US\$ 1.4 million for a new laboratory or up to US\$ 300 000 for an established laboratory), given the need for extensive biosafety equipment and the infrastructure required for conventional testing.

- A detailed commercial sales contract and **customer support plan** should be negotiated with the supplier, guaranteeing an ample and continual supply of cartridges, customs clearance, maintenance and calibration, and appropriate repair and replacement.
- **Mechanisms for rapidly reporting** Xpert MTB/RIF results to clinicians, and offering timely access to appropriate treatment, must be established to provide patients with the benefits of an early diagnosis.
- Health-care staff must be trained how to properly select persons to test with the Xpert MTB/RIF assay in different epidemiological settings; they must also be trained to interpret the results of rifampicin-resistance testing for persons with and without risk factors for drug-resistant TB, understand when to initiate a WHO-recommended MDR-TB regimen based on the results from Xpert MTB/RIF, and how to use additional testing (either a second Xpert MTB/RIF assay or another test) to ensure early initiation of the appropriate WHO-recommended first-line regimen or MDR-TB treatment regimen.

6.1 Research needs

A series of operational research questions have been identified related to the introduction and scale-up of Xpert MTB/RIF, and its impact on the diagnosis of TB, MDR-TB and the management of patients.

No additional studies, at country level, on the diagnostic accuracy of Xpert MTB/RIF are needed.

Further operational research should focus on the following priorities:

- evaluation of diagnostic algorithms in different epidemiological and geographical settings, and patient populations;
- country-specific cost-effectiveness and cost-benefit analyses of Xpert MTB/RIF in different programmatic settings;
- prospective evaluations of using Xpert MTB/RIF for nonrespiratory specimens, including blood, urine and stool;
- evaluation of the impact of Xpert MTB/RIF in reducing the diagnostic delay and the time until appropriate treatment is initiated;
- evaluation of the impact of Xpert MTB/RIF on case-detection, treatment access and treatment outcomes in hard-to-reach populations.

6.2 Plans for supporting scale-up of the implementation of Xpert MTB/RIF

WHO will continue to evaluate the usefulness and impact of Xpert MTB/RIF, with active collaboration from the Global Laboratory Initiative of the Stop TB Partnership²⁶, the WHO and Global Laboratory Initiative's TB Supranational Reference Laboratory Network, and several organizations that have implemented the Xpert MTB/RIF system (including technical agencies, research organizations, nongovernmental organizations and donors). Implementing organizations are brought together by WHO during an annual meeting to share their experiences, identify operational constraints and solutions, and contribute to the global collection of data by WHO on the scale-up of Xpert MTB/RIF in different settings.

In addition, WHO's Global TB Programme maintains a website²⁷ dedicated to monitoring the roll-out of Xpert MTB/RIF in order to inform and facilitate coordination among implementing organizations including countries, technical

26 For additional information, see <http://www.stoptb.org/wg/gli/>.

27 For additional information, see <http://www.who.int/tb/laboratory/mtbrifrollout>.

agencies, nongovernmental agencies and other partners.

A second edition of WHO's implementation document for Xpert MTB/RIF is being developed by WHO's Global TB Programme based on this policy update; it will be disseminated widely through e-mail and on the web sites of stakeholders. The Xpert MTB/RIF implementation manual outlines the requirements for systematic scale-up of Xpert MTB/RIF in varying epidemiological and resource settings. The document suggests ways to developing diagnostic algorithms and manage patients; it also discusses the operational and logistical aspects that need to be addressed during implementation of Xpert MTB/RIF. This policy update will also be incorporated into a comprehensive training package on Xpert MTB/RIF (being developed by WHO and the Global Laboratory Initiative) that will be made available to country and technical partners.

WHO and the Global Drug Facility have initiated

a unified forecasting initiative for Xpert MTB/RIF that collects forecasts of orders from major programmes and projects; these forecasts are submitted quarterly to the manufacturer. Reliable forecasting is needed to ensure effective scale-up of Xpert MTB/RIF by aiding the manufacturer in planning to meet demand.

Given the public-health importance and the ease of use of the Xpert MTB/RIF technology, expanded uptake at country level is anticipated and, therefore, the impact on TB case-detection will be monitored by WHO's Global TB Programme. In addition, improvements in diagnosing drug-resistant TB using this technology will need to be accompanied by scale-up of access to appropriate treatment, and it will be essential that scale-up is closely coordinated with care-delivery mechanisms at the country level. Therefore, implementation of the new policy guidance will be harmonized with the Global TB Programme's efforts to scale-up treatment of MDR-TB.

Box 4. Online resources

Strategic and Technical Advisory Group for Tuberculosis (STAG-TB): report of the 13th meeting. Geneva, World Health Organization, 2013 (WHO/HTM/TB/2013.09) (http://www.who.int/entity/tb/advisory_bodies/STAG_report2013.pdf).

Using the Xpert MTB/RIF assay to detect pulmonary and extrapulmonary tuberculosis and rifampicin resistance in adults and children: Expert Group meeting report. Geneva, World Health Organization, 2013 (available at http://www.who.int/tb/laboratory/policy_statements/en/).

Policy framework for implementing new tuberculosis diagnostics. Geneva, World Health Organization, 2010 (available at: http://www.who.int/tb/laboratory/whopolicy_framework_mar2011.pdf).

Xpert MTB/RIF implementation manual. Technical and operational 'how-to': practical considerations, 2nd ed. Geneva, World Health Organization, 2014 (available at http://www.who.int/tb/laboratory/policy_statements/en/).

Xpert MTB/RIF training package. Geneva, Global Laboratory Initiative, 2014 (available at <http://www.stoptb.org/wg/gli/>).

7. GRADE tables

Table 8. 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					Number of results/1000 individuals tested (95% CrI) ^a			
		Limitations	Indirectness	Inconsistency	Imprecision	Publication bias	Quality of evidence	Prevalence 25/1000 ^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)
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)
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)
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)

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 9. 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	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	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 10. 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 ^b	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 ^b	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 ^b	10 (10-20)	10 (10-19)	9 (9-18)
True negatives (individuals without TB)	Cross-sectional	None ^a	None ^b	None	None	Undetected ^b	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 11. 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	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	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	20 (10-39)	19 (10-38)	18 (9-36)
True negatives (individuals without TB)	Cross-sectional	None ^a	None ^b	None	None	Undetected ^e	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 12. 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 (19-23)	Prevalence 50/1000 (38-46)	Prevalence 100/1000 (76-92)
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 13. 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 14. 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 15. 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 16. 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 17. 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

PlCO 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	84 (45-9)	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	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	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	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 (discordant 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.

Table 18. 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 19. 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
False negatives (individuals without TB)	Majority cross-sectional	Serious (–1) ^b	None ^c	Serious (–1) ^d	Serious (–1) ^e	Undetected ^f
False positives (individuals incorrectly classified as having TB)	Majority cross-sectional	Serious (–1) ^b	None ^c	Serious (–1) ^d	Serious (–1) ^e	Undetected ^f
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

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

PO 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
False negatives (individuals without TB)	Majority cross-sectional	Serious (–1) ^b	None ^c	Serious (–1) ^d	Serious (–1) ^e	Undetected ^f
False positives (individuals incorrectly classified as having TB)	Majority cross-sectional	Serious (–1) ^b	None ^c	Serious (–1) ^d	Serious (–1) ^e	Undetected ^f
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

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 QUADAS2 tool was used to assess the risk of bias.

c 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.

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 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.

f 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.

g 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 21. 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 (1 258 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 22. 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 23. 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
False negatives (individuals without TB)	Majority cross-sectional	None ^c	Serious (–1) ^d	Serious (–1) ^e	Serious (–1) ^f	Undetected ^g
False positives (individuals incorrectly classified as having TB)	Majority cross-sectional	None ^c	Serious (–1) ^d	Serious (–1) ^e	Serious (–1) ^f	Undetected ^g
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

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.

o 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.

f 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.

Table 24. GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting TB in children compared with culture as a reference standard (A. Expecto-rated 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. Expecto-rated 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
False negatives (individuals incorrectly classified as not having TB)	Cross-sectional	None ^b	Serious (–1) ^c	Serious (–1) ^d	None ^e	Undetected ^f
False positives (individuals incorrectly classified as having TB)	Cross-sectional	None ^b	Serious (–1) ^c	Serious (–1) ^d	None ^e	Undetected ^f
True negatives (individuals without TB)	Cross-sectional	None ^b	Serious (–1) ^c	Serious (–1) ^d	None ^e	Undetected ^f
Time to diagnosis ^g		Only descriptive data available				Not rated

CrI, credible interval.

^a Bachrow 2012¹ 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 only TB treatment (7/71037), however, in one study (LaCourse, unpublished) 99/300 children were excluded because they were lost to follow up. Another study (Bates 2013)² enrolled a number of children who had already initiated anti-TB treatment (7/71037) 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. Heterogeneity: quality evidence may be lowered if there are important differences in the populations studied and the tests used, and if these aspects of the studies are compared with the results of the studies in 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)³ included a broader group of children, approximately 20% of children suspected of having TB had been identified through contact tracing. Two unpublished studies (Christi and LaCourse) included children who were severely malnourished, and TB was one differential referral diagnosis. (Children in the Christi 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 53% (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 75% with 95% CrI, 57–88%, 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 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 a minimal concern 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 Lowenstein-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
False negatives (individuals incorrectly classified as not having TB)	Cross-sectional	None ^a	Serious (–1) ^b	Serious (–1) ^c	None (serious) ^d	Undetected ^e
False positives (individuals incorrectly classified as having TB)	Cross-sectional	None ^a	Serious (–1) ^b	Serious (–1) ^c	None ^d	Undetected ^e
True negatives (individuals without TB)	Cross-sectional	None ^a	Serious (–1) ^b	Serious (–1) ^c	None ^d	Undetected ^e

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		Quality of evidence		Number of results/1000 individuals tested		Quality of evidence	
	Prevalence 10/1000 ^b	Prevalence 50/1000 ^b (95% CrI) ^a			Prevalence 10/1000 ^b	Prevalence 50/1000 ^b (95% CrI) ^a		
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 25. GRADE evidence profile: accuracy of Xpert MTB/RIF in detecting TB in children compared with a clinical reference standard (A. Expecterated 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. Expecterated 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. Expecterated 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
False negatives (individuals incorrectly classified as not having TB)	Cross-sectional	Serious (–1) ^b	Serious (–1) ^c	Serious (–1) ^d	None ^e	Undetected ^f
False positives (Individuals incorrectly classified as having TB)	Cross-sectional	Serious (–1) ^b	Serious (–1) ^c	Serious (–1) ^d	None ^e	Undetected ^f
True negatives (Individuals without TB)	Cross-sectional	Serious (–1) ^b	Serious (–1) ^c	Serious (–1) ^d	None ^e	Undetected ^f

CrI, credible interval.

a. Rachow 2012 reported testing one cohort using expecterated sputum and one using induced sputum; this was counted as two studies.
b. The QUADAS2 tool was used to assess the risk of bias. Six of seven studies included in the metaanalysis enrolled individuals consecutively [seven studies used specimens of expecterated 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 taCourse) 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 expecterated 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 expecterated 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 expecterated or induced sputum was within the +/-10% margin; however, for gastric lavage and aspirate it was wide. There were no concerns for expecterated 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 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		Quality of evidence		Number of results/1000 individuals tested		Quality of evidence	
	Prevalence 10/1000 ^a	Prevalence 50/1000 ^b			Prevalence 10/1000 ^b	Prevalence 50/1000 ^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 26. 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. Expectedated 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	Expectedated or induced sputum: none ^d Gastric lavage or aspirate: serious (–1) ^d	Undetected ^e Expectedated 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	Expectedated or induced sputum: none ^d Gastric lavage or aspirate: serious (–1) ^d	Undetected ^e Expectedated 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 expectedated 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 expectedated 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.

^bAll 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.

^cNo study directly addressed the question of inconsistency by performing prioritising 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.

^dThe 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.

^eTwo 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		Quality of evidence	Prevalence 100/1000 ^b	Number of results/1000 individuals tested		Prevalence 100/1000 ^b	Quality of evidence
	Prevalence 10/1000 ^a	Prevalence 50/1000 ^a (95% CrI) ^a			Prevalence 10/1000 ^a	Prevalence 50/1000 ^a (95% CrI) ^a		
True positives (individuals with TB)	6 (4-7)	28 (21-35)	Low ⊕⊕○○	55 (41-69)	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)	Low ⊕⊕○○	45 (31-59)	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)	Moderate ⊕⊕⊕○	18 (9-36)	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)	Moderate ⊕⊕⊕○	882 (864-891)	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 27. 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) ^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)	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 28. 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 MTBDRplus (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					Number of results/1000 individuals tested (95% CrI)	
		Limitations	Indirectness	Inconsistency	Imprecision	Publication bias	Quality of evidence	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 ^c	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 ^c	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 ^c	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 29. 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	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	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	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	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 30. 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.

Annexes

Annex 1. Members of the Expert Group

Sevim Ahmedov (Observer)

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: lbarrera2000@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 (Observer)

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 (Observer)

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 (Observer)

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 (systematic reviewer)

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

Anne Detjen (systematic reviewer)

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 (Observer)

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 (systematic reviewer)

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 (systematic reviewer)

Consultant
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 (systematic reviewer)

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

Annex 2. 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 3. Members of the Strategic and Technical Advisory Group for Tuberculosis (STAG-TB)

Dr Draurio Barreira

Head
National TB Control Program
Ministry of Health
Brasilia-DF
Brazil

Dr Amy Bloom

Senior Technical Adviser
US Agency for International Development (USAID)
Washington, DC
United States

Prof. Gavin Churchyard

Chief Executive Officer
The Aurum Institute NPC
Parktown, Johannesburg
South Africa

Dr Daniela M Cirillo

Head
Emerging Bacterial Pathogens Unit
San Raffaele del Monte Tabor
Foundation
San Raffaele Scientific Institute
Milan
Italy

Professor Elizabeth Corbett

Tropical Epidemiology
London School of Hygiene & Tropical
Medicine and
Malawi Liverpool Wellcome
Trust Clinical Research Programme
Blantyre
Malawi

Dr Charles L Daley

Chief
Division of Mycobacterial and
Respiratory Infections
National Jewish Medical and
Research Center
Denver, CO
United States

Professor Stephen Graham

International Child Health
University of Melbourne
Department of Paediatrics
Royal Children's Hospital
Parkville, Melbourne
Australia

Dr Akramul Islam

Associate Director
Health, Nutrition and Population
Program
Bangladesh Rural Advancement Committee
Centre
Dhaka
Bangladesh

Dr Michael Kimerling

Senior Program Officer, TB
Global Health Program
Bill and Melinda Gates Foundation
Seattle, WA
United States

Dr Jaime Lagahid

National Center for Disease,
Prevention and Control
Department of Health
Manila
The Philippines

Dr Chakaya J Muhwa

(STAG-TB Chair)
Chief Research Officer
Centre for Respiratory Diseases
Research
Kenya Medical Research Institute
Nairobi
Kenya

Dr A Prakash

Joint Secretary
Disease Control
Ministry of Health and Family
Welfare
New Delhi
India

Dr Ejaz Qadeer

Manager
National TB Control Programme
Federal Ministry of Health
Islamabad
Pakistan

Dr Joseph Sitienei

National TB Programme Manager
Division of Leprosy, Tuberculosis and
Lung Disease
Ministry of Health
Nairobi
Kenya

Dr Alena Skrahina

Scientific Director
Republican Research and Practical
Centre for Pulmonology and
Tuberculosis
Minsk
Belarus

Dr Soumya Swaminathan

Director
National Institute for Research in Tuberculosis
Indian Council for Medical Research
Chennai
India

Dr Francis Varaine

Coordinator of MSF Working Group on
Tuberculosis
Médecins Sans Frontières
Paris
France

Dr Dalene von Delft

Medical Doctor
TB PROOF
Somerset West
Cape Town
South Africa

Dr Maarten van Cleeff

Program Director TB Care
KNCV Tuberculosis Foundation
The Hague
The Netherlands

Annex 4. 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/RIF test.

Claudia Denking: 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.



**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

ISBN 978 92 4 150633 5



9 789241 506335