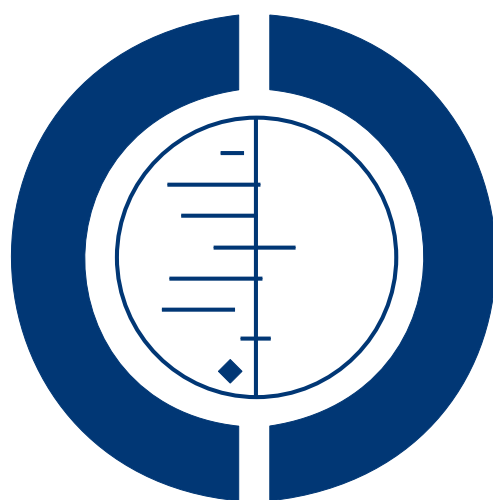


Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults (Review)

Steingart KR, Sohn H, Schiller I, Kloda LA, Boehme CC, Pai M, Dendukuri N



**THE COCHRANE
COLLABORATION®**

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2013, Issue 1

<http://www.thecochranelibrary.com>



Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults (Review)
Copyright © 2013 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
BACKGROUND	3
Figure 1.	5
OBJECTIVES	7
METHODS	7
RESULTS	11
Figure 2.	12
Figure 3.	13
Figure 4.	14
Figure 5.	16
Figure 6.	17
Figure 7.	18
Figure 8.	19
Figure 9.	20
Figure 10.	21
Figure 11.	22
Figure 12.	24
Figure 13.	25
DISCUSSION	30
AUTHORS' CONCLUSIONS	33
ACKNOWLEDGEMENTS	33
REFERENCES	34
CHARACTERISTICS OF STUDIES	40
DATA	98
Test 1. TB detection, all studies.	98
Test 2. Add on.	100
Test 3. Smear positive.	101
Test 4. Smear negative.	102
Test 5. HIV positive.	103
Test 6. HIV negative.	104
Test 7. TB detection, condition of specimen.	105
Test 8. TB detection, specimen preparation.	106
Test 9. TB prevalence.	107
Test 10. Income status.	108
Test 11. Rifampicin resistance.	109
Test 12. RIF resistance prevalence.	110
ADDITIONAL TABLES	110
APPENDICES	114
Figure 14.	125
Figure 15.	126
CONTRIBUTIONS OF AUTHORS	130
DECLARATIONS OF INTEREST	130
SOURCES OF SUPPORT	130
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	130

[Diagnostic Test Accuracy Review]

Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults

Karen R Steingart¹, Hojoon Sohn², Ian Schiller³, Lorie A Kloda⁴, Catharina C Boehme⁵, Madhukar Pai², Nandini Dendukuri²

¹Department of Health Services, University of Washington, School of Public Health, Seattle, Washington, USA. ²Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada. ³Department of Clinical Epidemiology, McGill University Health Centre, Montreal, Canada. ⁴Library, McGill University, Montreal, Canada. ⁵Foundation for Innovative New Diagnostics (FIND), Geneva, Switzerland

Contact address: Karen R Steingart, Department of Health Services, University of Washington, School of Public Health, Seattle, Washington, 98195-7230, USA. karenst@uw.edu.

Editorial group: Cochrane Infectious Diseases Group.

Publication status and date: New, published in Issue 1, 2013.

Review content assessed as up-to-date: 7 September 2012.

Citation: Steingart KR, Sohn H, Schiller I, Kloda LA, Boehme CC, Pai M, Dendukuri N. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database of Systematic Reviews* 2013, Issue 1. Art. No.: CD009593. DOI: 10.1002/14651858.CD009593.pub2.

Copyright © 2013 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Accurate and rapid detection of tuberculosis (TB) and drug resistance are critical for improving patient care and decreasing the spread of TB. Xpert® MTB/RIF assay (Xpert) is a rapid, automated test that can detect both TB and rifampicin resistance, within two hours after starting the test, with minimal hands-on technical time, but is more expensive than conventional sputum microscopy.

Objectives

To assess the diagnostic accuracy of Xpert for pulmonary TB (TB detection), both where Xpert was used as an initial test replacing microscopy, and where Xpert was used as an add-on test following a negative smear microscopy result.

To assess the diagnostic accuracy of Xpert for rifampicin resistance detection where Xpert was used as the initial test, replacing conventional culture-based drug susceptibility testing.

The population of interest was adults suspected of having pulmonary TB or multidrug-resistant TB (MDR-TB), with or without HIV infection.

Search methods

We performed a comprehensive search of the following databases: Cochrane Infectious Diseases Group Specialized Register; MEDLINE; EMBASE; ISI Web of Knowledge; MEDION; LILACS; BIOSIS; and SCOPUS. We also searched the metaRegister of Controlled Trials (mRCT) and the search portal of the WHO International Clinical Trials Registry Platform to identify ongoing trials. We performed searches on 25 September 2011 and we repeated them on 15 December 2011, without language restriction.

Selection criteria

We included randomized controlled trials, cross-sectional, and cohort studies that used respiratory specimens to compare Xpert with culture for detecting TB and Xpert with conventional phenotypic drug susceptibility testing for detecting rifampicin resistance.

Data collection and analysis

For each study, two review authors independently extracted a set of data using a standardized data extraction form. When possible, we extracted data for subgroups by smear and HIV status. We assessed the quality of studies using the QUADAS-2 tool. We carried out meta-analyses to estimate the pooled sensitivity and specificity of Xpert separately for TB detection and rifampicin resistance detection using a bivariate random-effects model. We estimated the median pooled sensitivity and specificity and their 95% credible intervals (CrI).

Main results

We identified 18 unique studies as eligible for this review, including two multicentre international studies, one with five and the other with six distinct study centres. The majority of studies (55.6%) were performed in low-income and middle-income countries. In 17 of the 18 studies, Xpert was performed by trained technicians in reference laboratories.

When used as an initial test replacing smear microscopy (15 studies, 7517 participants), Xpert achieved a pooled sensitivity of 88% (95% CrI 83% to 92%) and pooled specificity of 98% (95% CrI 97% to 99%). As an add-on test following a negative smear microscopy result (14 studies, 5719 participants), Xpert yielded a pooled sensitivity of 67% (95% CrI 58% to 74%) and pooled specificity of 98% (95% CrI 97% to 99%). In clinical subgroups, we found the following accuracy estimates: the pooled sensitivity was 98% (95% CrI 97% to 99%) for smear-positive, culture-positive TB and 68% (95% CrI 59% to 75%) for smear-negative, culture-positive TB (15 studies); the pooled sensitivity was 80% (95% CrI 67% to 88%) in people living with HIV and 89% (95% CrI 81% to 94%) in people without HIV infection (four studies). For rifampicin resistance detection (11 studies, 2340 participants), Xpert achieved a pooled sensitivity of 94% (95% CrI 87% to 97%) and pooled specificity of 98% (95% CrI 97% to 99%). In a separate analysis, Xpert could distinguish between TB and nontuberculous mycobacteria (NTM) in clinical samples with high accuracy: among 139 specimens with NTM, Xpert was positive in only one specimen that grew NTM.

In a hypothetical cohort of 1000 individuals suspected of having rifampicin resistance (a proxy for MDR-TB), where the prevalence of rifampicin resistance is 30%, we estimated that on average Xpert would wrongly identify 14 patients as being rifampicin resistant. In comparison, where the prevalence of rifampicin resistance is only 2%, we estimated that the number of individuals wrongly identified as rifampicin resistant would increase to 20, an increase of 43%.

Authors' conclusions

This review shows that Xpert used as an initial diagnostic test for TB detection and rifampicin resistance detection in patients suspected of having TB, MDR-TB, or HIV-associated TB is sensitive and specific. Xpert may also be valuable as an add-on test following microscopy for patients who have previously been found to be smear-negative. An Xpert result that is positive for rifampicin resistance should be carefully interpreted and take into consideration the risk of MDR-TB in a given patient and the expected prevalence of MDR-TB in a given setting.

Studies in this review mainly assessed sensitivity and specificity of the test when used in reference laboratories in research investigations. Most studies were performed in high TB burden countries. Ongoing use of Xpert in high TB burden countries will contribute to the evidence base on the diagnostic accuracy and clinical impact of Xpert in routine programmatic and peripheral health care settings, including settings where the test is performed at the point of care.

BACKGROUND

Tuberculosis (TB) is one of the world's most important infectious causes of morbidity and mortality among adults. When TB is detected and effectively treated, the disease is largely curable. However, in 2010, 8.8 million people developed TB disease (active TB) for the first time (WHO Global Report 2011). Of the total new TB cases, approximately 13% occurred among people living with HIV. Among people without HIV infection, 1.1 million people (14%) died of TB and among people with HIV infection, 350,000 people (31%) died of TB (WHO Global Report 2011).

Drug-resistant TB, including multidrug-resistant TB (MDR-TB, defined as resistance to at least isoniazid and rifampicin, the two most important first-line anti-TB drugs) and extensively drug-resistant TB (XDR-TB, defined as MDR-TB plus resistance to any fluoroquinolone (such as ofloxacin or moxifloxacin) and to at least one of three injectable second-line drugs (amikacin, capreomycin, or kanamycin)) has emerged as a serious threat to global health (Zumla 2012). In 2010, there were an estimated 650,000 cases of MDR-TB (WHO Global Report 2011). Recently, the World Health Organization (WHO) reported the highest rates (greater than 65% in people who had previously received TB treatment) of MDR-TB ever recorded in several areas of the former Soviet Union (Zignol 2012). Worldwide, a substantial percentage (~35%) of patients with drug-susceptible TB remain undiagnosed and a staggering percentage (~85%) of patients with MDR-TB remain undiagnosed (WHO Global Report 2011). Of the people diagnosed with TB, less than 3% are tested to determine the pattern of drug resistance (Chaisson 2012). In addition to drug resistance, another major challenge is the accurate detection of smear-negative disease which disproportionately occurs in HIV-positive people with TB (Harries 2004).

Accurate and rapid detection of TB, including smear-negative TB and drug resistant-TB, are critical for improving patient outcomes (increased cure and decreased mortality, additional drug resistance, treatment failure, and relapse) and decreasing TB transmission. Mycobacterial culture is generally considered the best available reference standard for TB diagnosis and is the first step in detecting drug resistance. However, this is a relatively complex and slow procedure. Solid culture typically takes four to eight weeks for results and liquid culture, though more rapid than solid culture, takes days for results. Liquid culture is, however, more prone to contamination (WHO Policy Framework 2010). In addition, culture requires specialized laboratories and highly skilled staff. In 2010, WHO endorsed a novel, rapid, automated, cartridge-based nucleic acid amplification test (NAAT), the Xpert® MTB/RIF assay (Cepheid, Sunnyvale, USA) (hereafter referred to as Xpert), that can simultaneously detect TB and rifampicin resistance (WHO Policy Xpert 2011), WHO recommends that Xpert be used as the initial diagnostic test in individuals suspected of MDR-TB or HIV-associated TB. If Xpert detects rifampicin resistance in patients considered at risk of MDR-TB, an appropriate MDR-TB regimen should be started while additional sputum specimens

are obtained for culture and drug susceptibility testing. Subsequent testing will confirm the presence of rifampicin resistance and enable testing for drug resistance to isoniazid and other first-line drugs and second-line drugs. Ideally, Xpert should be used at the district or subdistrict health facility level (WHO Policy Xpert 2011).

Target condition being diagnosed

Tuberculosis

TB is caused by the bacterium *Mycobacterium tuberculosis* and is spread from person to person through the air. TB most commonly affects the lungs (pulmonary TB), but may affect any organ or tissue, such as the brain or bones, outside of the lungs (extrapulmonary TB). Signs and symptoms of pulmonary TB include cough for at least two weeks, fever, chills, night sweats, weight loss, haemoptysis (coughing up blood), and fatigue. Signs and symptoms of extrapulmonary TB depend on the site of disease. TB treatment regimens must contain multiple drugs to which the organisms are sensitive to be effective. The treatment of MDR-TB is complex, usually requiring two years or more of therapy and drugs that are less potent and more toxic than the drugs used to treat drug-susceptible TB. International guidelines for TB treatment are issued by WHO and regularly updated. Current WHO guidelines on TB treatment are based on evidence assessed according to the GRADE (Grading of Recommendations Assessment, Development and Evaluation) approach for developing health care recommendations (Guyatt 2011; WHO Guidelines Drug-resistant TB 2011; WHO Guidelines TB Treatment 2010).

Rifampicin resistance

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

Index test(s)

Xpert is an automated polymerase chain reaction (PCR) test (molecular test) utilizing the GeneXpert® platform (Blakemore 2010; Cepheid 2009; Helb 2010). Xpert is a single test that can both detect *M. tuberculosis* complex and rifampicin resistance within two hours after starting the test, with minimal hands-

on technical time. Unlike conventional nucleic acid amplification tests (NAATs), Xpert is unique because sample processing and PCR amplification and detection are integrated into a single self-enclosed test unit, the GeneXpert cartridge. Following sample loading, all steps in the assay are completely automated and self-contained. In addition, the assay's sample reagent, used to liquefy sputum, has potent tuberculocidal (the ability to kill TB bacteria) properties and so largely eliminates biosafety concerns during the test procedure (Banada 2010). These features allow the technology to be taken out of a reference laboratory and used nearer to the patient (Small 2011). Xpert requires an uninterrupted and stable electrical power supply, temperature control, and yearly calibration of the cartridge modules (WHO Rapid Implementation 2011).

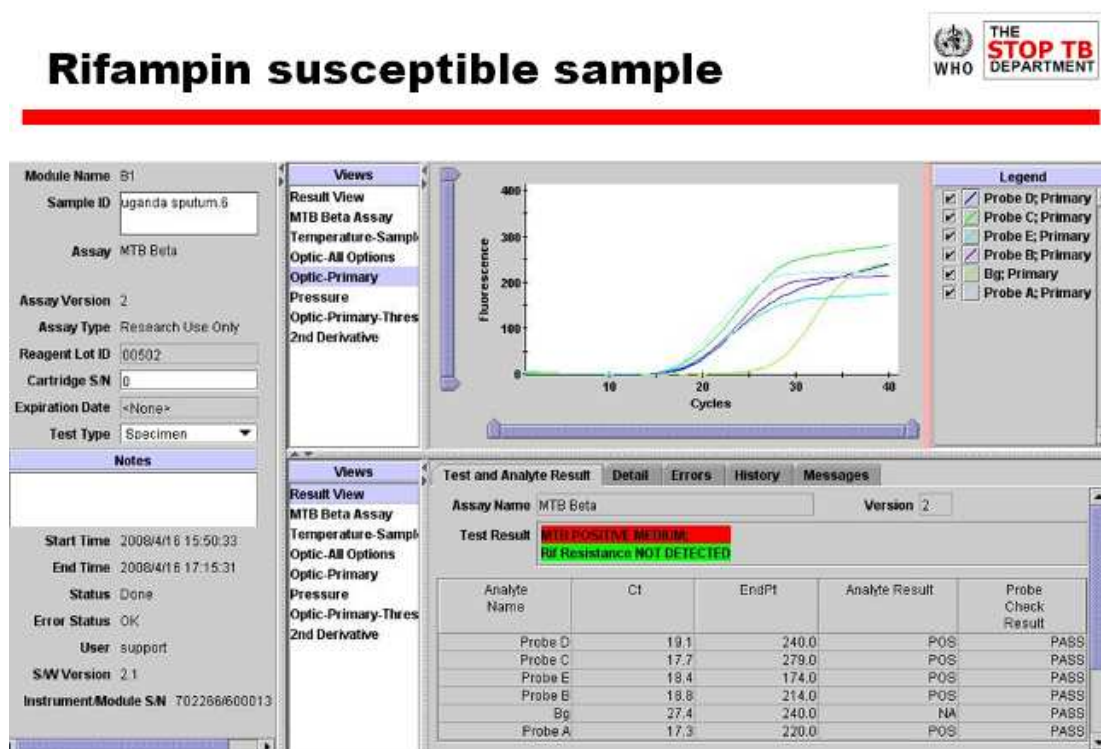
The test procedure may be used directly on clinical specimens, either raw sputum samples or sputum pellets (also called sputum sediment), samples created after decontaminating and concentrating the sputum (Blakemore 2010). In both cases, the test material is combined with the assay sample reagent, mixed by hand or vortex, and incubated at room temperature for 15 minutes. After the incubation step, 2 mL of the treated sample are transferred to the cartridge and the run initiated (Helb 2010).

Xpert's limit of detection, "the lowest number of colony forming units per sample that can be reproducibly distinguished from negative samples with 95% confidence" (Cepheid 2009), is 5 genome copies of purified DNA per reaction or 131 colony forming units

per mL in *M. tuberculosis* spiked sputum (Helb 2010). In comparison, to see TB bacilli by microscopic examination requires at least 10,000 bacilli per mL of sputum (Toman 2004). Xpert detects both live and dead bacteria (Miotto 2012).

Xpert uses molecular beacon technology to detect rifampicin resistance. Molecular beacons are nucleic acid probes that recognize and report the presence or absence of the normal, rifampicin-susceptible, 'wild type' sequence of the *rpoB* gene of TB. 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. Failure of the beacon to bind or delayed binding to the matching sequence indicates potential rifampicin resistance. The number and timing of detection (when the fluorescent signal rises above a pre-determined baseline cycle threshold) of positive beacons as well as results of sample processing controls allows the test to distinguish among the following results: 'No TB'; 'TB detected, rifampicin resistance detected'; 'TB detected, no rifampicin resistance detected'; and an 'invalid result' (Figure 1). As mentioned, a single Xpert run will provide both detection of TB and detection of rifampicin resistance. One cannot deselect testing for rifampicin resistance and only run the assay for TB detection, although it is possible for the laboratory to omit results for rifampicin resistance when reporting to the healthcare provider.

Figure 1. Readout of Xpert MTB/RIF assay for a TB positive, rifampicin-susceptible specimen. Courtesy: Karin Weyer, The WHO STOP TB Department.



1

Since Xpert was released, there have been four generations of the test involving different software and cartridge combinations. Studies using the first three generations (G1, G2, and G3 cartridges) are included in this systematic review. A newer cartridge, G4, to be used with software version 4.0 or higher, has been released and is now included in all Xpert kits (personal communication; Ellen Jo Baron, Cepheid, 26 June 2012). Studies using G4 and newer versions of the assay will be included in updates of this review.

Clinical pathway

Patients suspected of having pulmonary TB or MDR-TB.

Prior test(s)

For TB detection, Xpert could be used as an initial test or as an add-on test following a negative smear microscopy result.

Role of index test(s)

Consistent with WHO recommendations (WHO Policy Xpert 2011), we were interested in the following purposes for testing:

I. Xpert for TB detection

- A. Xpert used as an initial test replacing smear microscopy in a population unselected by smear status
- B. Xpert used as an add-on test following a negative smear microscopy result

II. Xpert for rifampicin resistance detection

- A. Xpert used as an initial test for rifampicin resistance replacing conventional phenotypic drug susceptibility testing as the initial test.

Xpert does not eliminate the need for subsequent culture and phenotypic drug susceptibility testing, which are required to monitor treatment progress and to detect resistance to drugs other than rifampicin.

Alternative test(s)

In this section, we describe selected alternative tests for detection of TB and rifampicin resistance. For a comprehensive review of these tests, we refer the reader to several excellent resources (Drobniewski 2012; Nahid 2012; UNITAID 2012).

Smear microscopy, the direct examination of sputum smears with Ziehl-Neelsen staining for acid-fast bacilli (TB bacteria), is the most commonly used test for TB detection in resource-limited settings ([International Standards 2009](#)). Advantages of smear microscopy include its simplicity, low cost, speed, and high specificity in high TB burden areas. In addition, smear microscopy identifies the most infectious TB patients. Smear microscopy can be performed in basic laboratories. Drawbacks of smear microscopy include the need for specialized training and its relatively low sensitivity, 50% to 60% on average. Although, the sensitivity of microscopy can be improved by approximately 10% with fluorescence ([Steingart 2006](#)), a large number of TB cases still go undiagnosed. Smear microscopy contributes little to the diagnosis of paediatric TB and does not, by definition, identify smear-negative TB which may account for 24% to 61% of all pulmonary cases in people living with HIV ([Getahun 2007](#); [Perkins 2007](#); [Steingart 2006a](#)). Microscopy cannot distinguish between drug-susceptible TB and drug-resistant TB.

Nucleic acid amplification tests (NAATs) are molecular systems that can detect small quantities of genetic material (DNA or RNA) from microorganisms, such as TB. A variety of molecular amplification methods are available, of which PCR is the most common. NAATs are available as commercial kits and in-house tests (based on a protocol developed in a non-commercial laboratory) and are used routinely in high-income countries for TB detection. In-house PCR is widely used in developing countries because these tests are less expensive than commercial kits. However, in-house PCR is known to produce highly inconsistent results ([Flores 2005](#)). The use of NAATs has recently been recommended as standard practice in the United States ([CDC 2009](#)). The main advantage of NAATs is that they can provide results several weeks earlier than culture ([CDC 2009](#)). Drawbacks are that these tests are often too expensive and complex for routine use by TB programmes in resource-limited settings. In addition, although the specificity of NAATs is high, some NAATs have shown variable and low sensitivity, especially in sputum smear-negative patients ([Flores 2005](#); [Greco 2006](#); [Ling 2008a](#)).

Loop-mediated isothermal amplification (LAMP), a molecular method for TB detection, amplifies DNA under isothermal conditions. The large quantity of DNA generated makes it possible to detect amplification and interpret the assay result by visual inspection for fluorescence or turbidity ([Boehme 2007](#)). Advantages of LAMP are that the method is rapid (results can be obtained within several hours) and highly specific for *M. tuberculosis* complex. In addition, LAMP can be used in areas with limited resources such as district level laboratories because it is relatively inexpensive, does not require the use of a thermal cycler, and is simple to perform ([UNITAID 2012](#)). Drawbacks include the difficulty in interpreting test results via the colorimetric change. A decision and potential policy guidance from the WHO Expert Group that reviewed the available evidence on LAMP is forthcoming (personal communication; Wayne Van Gemert, 20 June 2012).

Alternative molecular methods for drug susceptibility testing include the commercial line probe assays, INNO-LiPA Rif.TB (Innogenetics, Ghent, Belgium) and GenoType® MTBDRplus assay (Hain LifeScience GmbH, Nehren, Germany). The INNO-LiPA Rif.TB assay targets common mutations in the *rpoB* gene associated with rifampicin resistance while the GenoType® MTBDRplus assay also targets the common mutations in *katG* and *inhA* genes associated with isoniazid resistance in addition to the mutations in the *rpoB* gene ([UNITAID 2012](#)). Advantages of line probe assays are that they can provide a result for detection of TB and drug resistance in 1 to 2 days. Line probe assays have both high sensitivity (greater than 97%) and high specificity (greater than 99%) for the detection of rifampicin resistance alone, or in combination with isoniazid (sensitivity greater than 90%; specificity greater than 99%), on TB isolates and smear-positive sputum specimens ([Ling 2008](#)). Drawbacks are that line probe assays are expensive and must be used in reference laboratories ([Nahid 2012](#)). These tests have been endorsed by WHO ([WHO Policy Line Probe Assays 2008](#)).

Several alternative culture-based methods for TB detection and drug susceptibility testing are available, including the microscopic observation drug susceptibility (MODS) assay which is available in both a non-commercial and commercial version (Hardy Diagnostics, USA), non-commercial colorimetric redox indicator methods and the non-commercial nitrate reductase assay. MODS is used for both detection of TB and detection of resistance to rifampicin and isoniazid. The basic principle underlying MODS is that *M. tuberculosis* complex bacteria grow as a corded mass while most nontuberculous mycobacteria (NTM) do not. The morphology of the bacteria can be observed with an inverted microscope ([Moore 2006](#)). Colorimetric redox indicator methods use the principle of a change in the colour of a dye to indicate the growth of viable mycobacteria in the medium ([Martin 2007](#)). The nitrate reductase assay is based on the ability of *M. tuberculosis* to reduce nitrate to nitrite. A dye is used to indicate bacterial growth ([Martin 2008](#)). All three methods have high accuracy for detection of rifampicin resistance, are inexpensive, and are relatively rapid (results are generally available in less than 10 days). Drawbacks of these tests include biosafety requirements, the need for specialized training, and the absence of standard methods for colorimetric methods and the nitrate reductase assay. Culture-based methods also involve challenges with preparing, diluting and storing drug solutions, and a need for a consistent power supply to maintain incubator temperatures ([UNITAID 2012](#)). WHO has endorsed MODS and the nitrate reductase assay for direct drug susceptibility testing of sputum specimens and all three methods for indirect drug susceptibility testing of TB isolates grown in conventional culture ([WHO Policy Noncommercial Culture 2011](#)). WHO considers the use of these tests as an interim solution while TB programmes build capacity for molecular tests or automated liquid culture and drug susceptibility testing ([WHO Policy Noncommercial Culture 2011](#)).

Rationale

Xpert, if accurate, would provide obvious benefits for patients (earlier diagnosis and the opportunity to begin earlier, appropriate treatment) and for public health (opportunities to interrupt TB transmission), especially in developing countries. To our knowledge, at the time of writing this review, one systematic review and meta-analysis on the diagnostic accuracy of Xpert has been published (Chang 2012). However, the authors used statistical methods for meta-analysis other than the currently recommended bivariate random-effects models (Macaskill 2010).

In September 2010, the WHO Stop TB Department convened an Expert Group meeting to review the available evidence on Xpert for the purpose of formulating recommendations to guide the use of the test (WHO Policy Xpert 2011). Since that time, the evidence has rapidly accumulated and additional studies on Xpert have been published. We conducted a systematic review to synthesize this body of evidence using currently recommended methods.

OBJECTIVES

Primary objectives

Since Xpert can detect both TB and rifampicin resistance, we had two review questions with the following primary objectives:

I. Xpert for TB detection

To determine summary estimates of the diagnostic accuracy of Xpert for the diagnosis of pulmonary TB in adults

II. Xpert for rifampicin resistance detection

To determine summary estimates of the diagnostic accuracy of Xpert for detection of rifampicin resistance in adults

Setting of testing

We were interested in how Xpert performed in patients who were evaluated in peripheral laboratories or health facilities. Laboratory services for TB may be described at three levels: peripheral level (typically district or subdistrict laboratories or more decentralized peripheral microscopy centres); intermediate level (typically regional laboratories); and central level (typically national or reference laboratories) (WHO Policy Framework 2010). Diagnostic tests often perform well when initially evaluated in reference laboratories; however tests may not perform as well when they are run in settings of intended use. As mentioned above, Xpert is intended for use in laboratories or health facilities at the district or subdistrict level. Xpert is currently not intended for use at peripheral microscopy centres.

METHODS

Criteria for considering studies for this review

Types of studies

We included primary studies that assessed the diagnostic accuracy of Xpert for pulmonary TB and/or rifampicin resistance. Diagnostic accuracy studies are typically cross-sectional in design. However, we also searched for randomized controlled trials (RCTs) and cohort studies. We only included studies that reported data comparing Xpert to an acceptable reference standard from which we could extract true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values. Xpert could be assessed alone or together with other tests.

We excluded studies with a case-control design because these types of studies are prone to bias, in particular, studies enrolling patients with severe disease and healthy participants without disease. We also excluded studies reported only in abstracts.

Participants

Studies had to include adult or predominantly adult patients suspected of having pulmonary TB or MDR-TB, with or without HIV infection. We considered adults to be 15 years of age or older. We included studies that assessed the diagnostic accuracy of Xpert using sputum and other respiratory specimens, such as fluid obtained from bronchial alveolar lavage and tracheal aspiration. We included studies from all settings and all countries. Data on age of study participants were available for the majority of included studies. We considered it highly likely that studies that did not report age data involved all or predominantly adults for the following reasons: the vast majority of specimens evaluated with Xpert were sputum specimens and children have difficulty producing sputum; we excluded data on specimens obtained by gastric aspiration, as this specimen collection method is used mostly for investigating TB in children; we excluded studies that specifically evaluated the use of Xpert in children; and we performed a sensitivity analysis by dropping studies that did not report age data to check whether the accuracy results changed (Sensitivity analyses).

Index tests

Xpert was the index test under evaluation.

Target conditions

The target conditions were active pulmonary TB and rifampicin resistance.

Reference standards

For TB, acceptable reference standards utilized solid media: Löwenstein-Jensen, Middlebrook 7H10 or 7H11, or Ogawa media; or a commercial liquid culture system: such as BACTEC™

460TB System or BACTEC™ MGIT™ (mycobacterial growth indicator tube) 960 Mycobacterial Detection System, BD, USA; BacT/ALERT® System, bioMérieux, France; or both solid media and a commercial liquid culture system; or VersaTREK® Mycobacteria Detection & Susceptibility, Thermo Fisher Scientific, USA.

For rifampicin resistance, the reference standards were conventional phenotypic drug susceptibility testing methods as recommended by WHO ([WHO Policy DST 2008](#)). Acceptable methods used solid media, such as Löwenstein-Jensen, Middlebrook 7H10 or 7H11, or Ogawa media and/or a commercial liquid culture system, such as BACTEC™ 460TB System or BACTEC™ MGIT™ (mycobacterial growth indicator tube) 960 Mycobacterial Detection System, BD, USA; BacT/ALERT® System, bioMérieux, France; or VersaTREK® Mycobacteria Detection & Susceptibility, Thermo Fisher Scientific, USA.

Search methods for identification of studies

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

Electronic searches

Vittoria Lutje, (VL) the Information Specialist from the Cochrane Infectious Diseases Group, searched the following databases on 25 September 2011 using the strategy described in [Appendix 1](#): Cochrane Infectious Diseases Group Specialized Register; MEDLINE; EMBASE; ISI Web of Knowledge; MEDION; LILACS; BIOSIS; and SCOPUS. She also searched the metaRegister of Controlled Trials (mRCT) and the search portal of the WHO International Clinical Trials Registry Platform, to identify ongoing trials. She repeated searches on 15 December 2011. We limited all searches to 2007 onward because the development of Xpert was completed in 2009 and the first paper describing its clinical use was published electronically in 2009 ([Helb 2010](#)). VL performed the searches without language restriction.

Searching other resources

To identify additional published, unpublished, and ongoing studies, we performed the following tasks:

- reviewed reference lists of included articles and review articles identified through the above methods;
- contacted Cepheid, the test manufacturer;
- handsearched WHO reports on Xpert;
- contacted researchers at the Foundation for Innovative New Diagnostics (FIND), members of the Stop TB Partnership's New Diagnostics Working Group, and other experts in the field of TB diagnostics.

While preparing this review, the Xpert MTB/RIF Mapping Tool became available (<http://xrmt.treattb.org/projects-listing2>). This tool has been developed by the International Union Against Tuberculosis and Lung Disease, through the USAID-funded TREAT TB initiative, and supported by WHO. The tool allows researchers and policy makers to globally map ongoing research activities related to Xpert. We may use this tool to identify studies for updates of this review.

Data collection and analysis

Selection of studies

Two review authors (KRS and HS) independently scrutinized titles and abstracts identified by electronic literature searching to identify potentially eligible studies (screen 1). We selected any citation identified by either review author during screen 1 for full-text review. Next, we retrieved the full paper of each potentially eligible article identified in the search. Two review authors (KRS and HS) independently assessed articles for inclusion using predefined inclusion and exclusion criteria (screen 2). In screen 2, we resolved any discrepancies by discussion between the review authors, or if necessary, by decision of a third review author (MP). We maintained a list of excluded studies and their reasons for exclusion.

We named studies according to the surname of the first author and year of publication. For multicentre studies, the study-naming scheme uniquely identifies multiple study centres from within each study ([Boehme 2010a](#); [Boehme 2010b](#)), each of which reported data separately for a distinct population at a given study site. Hence, the number of study centres exceeds the number of studies. (Please note that when cited in the text, “Boehme [year]a” refers to an overall paper, consisting of all centres).

Data extraction and management

We extracted data on the following characteristics:

- author; publication year; study design; case country of residence; country income status classified by the World Bank List of Economies ([World Bank 2011](#)); clinical setting; laboratory setting;
- population, age, gender, HIV status, smear status, and follow-up;
- reference standard;
- parameter value in the Xpert system associated with determining rifampicin resistance;
- specimen collection (expectorated sputum, induced sputum)
- condition of the specimen (fresh or frozen);
- preparation of the specimen (processed or unprocessed);
- QUADAS-2 items ([Whiting 2011](#));

- data for two-by-two tables for Xpert, including results reported as indeterminate (results reported as invalid, error, or no result);
- time to diagnosis (time from specimen collection until there is an available TB result in laboratory or clinic);
- time to treatment initiation (time from specimen collection until time patient starts treatment).

Whenever possible, we extracted TP, FP, FN, and TN values based on one Xpert result for one specimen provided by one patient. However, in some of the studies, the number of specimens (and Xpert results) exceeded the number of patients, suggesting that a single patient may have provided multiple specimens. We therefore compared pooled sensitivity and specificity for TB detection in all studies with pooled sensitivity and specificity in the subset of studies that provided one Xpert result based on one specimen provided by one patient (see [Sensitivity analyses](#)).

When data were available, we extracted FP and TN values for participants without TB disease by their smear and HIV status. Concerning the definition of smear positivity, as the vast majority of included studies performed Xpert in reference laboratories, we assumed these studies adhered to the revised definition of a new sputum smear-positive pulmonary TB case based on the presence of at least one acid-fast bacillus in at least one sputum sample in countries with a well functioning external quality assurance system ([WHO Policy Smear-positive TB Case 2007](#)).

We developed a standardized data extraction form and piloted the form with four studies (22%). Based upon the pilot, we finalized the form. Two review authors (KRS and HS) independently extracted data from each study using the final form. We contacted authors of studies for missing data and clarifications. We entered all data into Excel and [SPSS 2006](#). The final data extraction form is included in [Appendix 2](#).

Assessment of methodological quality

We appraised the quality of included studies with the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool ([Whiting 2011](#)). QUADAS-2 consists of four domains: patient selection, index test, reference standard, and flow and timing. We assessed all domains for the potential for risk of bias, and, in addition, we assessed the first three domains for concerns regarding applicability. We used questions, called signalling questions, for each domain to form judgments about the risk of bias. As recommended, we first developed guidance on how to appraise each signalling question and interpret this information. Then, one review author (KRS) piloted the tool with four (22%) of the included studies. Based on experience gained from the pilot, we finalized the tool. Two review authors independently assessed the methodological quality of the included studies with the finalized tool. We presented results in the text, in graphs, and in a table. We did not generate a summary “quality score” because of problems associated with such numeric scores ([Juni 1999](#); [Whiting 2005](#)).

The domains of the QUADAS-2 tool and their interpretation are provided in [Appendix 3](#).

Statistical analysis and data synthesis

We performed descriptive analyses for the results of the included studies using [SPSS 2006](#) and present key study characteristics in [Characteristics of included studies](#). We used data reported in the two-by-two tables to calculate sensitivity and specificity estimates and 95% confidence intervals (CI) for individual studies and to generate forest plots using [Review Manager 5](#). We chose to use data that were not subject to discrepant analyses (ie unresolved data), since resolved data after discrepant analyses are a potential for risk of bias ([Hadgu 2005](#)).

We carried out meta-analyses to estimate the pooled sensitivity and specificity of Xpert separately for TB detection (I. A. and I. B) and rifampicin resistance detection (II. A.). We determined pooled estimates using an adaptation of the bivariate random-effects model ([Reitsma 2005](#)) to allow for a hierarchical structure for the two multicentre studies ([Boehme 2010a](#); [Boehme 2011a](#)). The bivariate random-effects approach allowed us to calculate the pooled estimates of sensitivity and specificity while dealing with potential sources of variation caused by (1) imprecision of sensitivity and specificity estimates within individual studies; (2) correlation between sensitivity and specificity across studies; and (3) variation in sensitivity and specificity between studies.

We estimated all models using a Bayesian approach with non-subjective prior distributions and implemented using WinBUGS (Version 1.4.3) ([Lunn 2000](#)). Under the Bayesian approach, all unknown parameters must be provided a prior distribution that defines the range of possible values of the parameter and the likelihood of each of those values based on information external to the data. In order to let the observed data determine the final results, we chose to use low-information prior distributions over the pooled sensitivity and specificity parameters and their between-study standard deviation parameters. The model we used is summarized in the Statistical Appendix together with the WinBUGS program used to implement it ([Appendix 4](#)). Information from the prior distribution is combined with the likelihood of the observed data in accordance with Bayes Theorem to obtain a posterior distribution for each unknown parameter.

Using a sample from the posterior distribution we can obtain various descriptive statistics of interest. We estimated the median pooled sensitivity and specificity and their 95% credible intervals (CrI). The median or the 50% quantile is the value below which 50% of the posterior sample lies. We chose to report the median because the posterior distributions of some parameters may be skewed and the median would be considered a better point estimate of the unknown parameter than the mean in such cases. The 95% CrI is the Bayesian equivalent of the classical (frequentist) 95% CI. (We have indicated 95% CI for individual study estimates and 95% CrI for pooled study estimates as appropriate). The 95%

CrI may be interpreted as an interval that has a 95% probability of capturing the true value of the unknown parameter given the observed data and the prior information. We also extracted estimates of the 'predicted' sensitivity and specificity in a future study together with their 95% CrIs. The predicted value gives an idea of heterogeneity at the study level. We can compare the predicted intervals to the pooled intervals to get an idea of the heterogeneity. With a large number of studies, the pooled interval may be narrow. However, if there is considerable variability in sensitivity and specificity estimates between studies, this variability will be reflected in a wide predicted interval despite the large number of studies. We generated the plots using R (version 2.15.1) (R 2008).

I. B. Xpert used as an add-on test following a negative smear microscopy result

To determine the value of Xpert as an add-on test, we estimated its sensitivity and specificity among smear-negative patients. We did this by evaluating two types of studies: 1) studies that enrolled participants selected to be predominantly smear negative by prior microscopic examination and 2) studies that enrolled unselected participants who were evaluated by both Xpert and microscopy concurrently. In the second group, we used only the data for those individuals who had a negative microscopy result. As described above, we determined pooled estimates using an adaptation of the bivariate random-effects model (Reitsma 2005).

Approach to indeterminate index test results

We excluded indeterminate test results from the analyses for determination of sensitivity and specificity for both TB detection and rifampicin resistance detection. We used a hierarchical model for a single proportion to estimate the proportion of indeterminate index test results.

Investigations of heterogeneity

I. A. TB detection

We first investigated heterogeneity through visual examination of forest plots of sensitivity and specificity. We then explored the possible influence of clinical factors by analyses of the subgroups described in the protocol. We expected the majority of studies to report TP, FP, FN, and TN values by smear status and several studies to report values by HIV status. Therefore, we fit the meta-analysis model separately within subgroups defined by smear and HIV status to examine the effects of these covariates on the pooled sensitivity and specificity. To study the impact of the remaining covariates of interest, we extended the meta-analysis model to a meta-regression model for the sake of statistical efficiency. We did this by expressing the pooled logit(sensitivity) and logit(specificity)

as linear functions of a dichotomous covariate to examine the effects of individual covariates on the pooled sensitivity and specificity. The covariates considered were all dichotomous variables and related to the condition of the specimens (fresh versus frozen), preparation of the specimens (unprocessed versus processed), TB prevalence (low $\leq 30\%$ versus high $> 30\%$), and country income status (low-/middle-income versus high-income).

II. A. Detection of rifampicin resistance

As mentioned, there have been four versions of Xpert, G1, G2, G3, and G4 (the current version). At the time we performed this review, there were no published studies of G4. Studies included in this review used Xpert versions G1, G2, and G3. The different Xpert versions involve software and cartridge processing adjustments with consequent changes in the way the presence or absence of rifampicin resistance is determined in the Xpert system. One of several factors influencing the determination of rifampicin resistance in the earlier Xpert versions (those included in this review) was a parameter value, also called delta cycle threshold cutoff adjusted to provide the optimum tradeoff between sensitivity for detecting 'rifampicin resistant' samples and specificity for detecting 'rifampicin susceptible' samples. A parameter value of 3.5 was incorporated in the algorithm for Xpert version G1, and a parameter value of 5, in the algorithm for Xpert versions G2 and G3. We investigated the influence of these two parameter values (3.5 and 5) on sensitivity and specificity by including a dichotomous covariate in the regression model.

We also explored the influence of rifampicin resistance prevalence on the pooled sensitivity and specificity estimates by including a dichotomous covariate, high prevalence $> 15\%$ and low prevalence $\leq 15\%$, in the regression model.

Sensitivity analyses

We performed sensitivity analyses by limiting inclusion in the meta-analysis to: 1) studies that provided data by age that explicitly met the age criterion for participants; 2) studies where consecutive patients were selected; 3) studies where a single specimen yielded a single Xpert result for a given patient; and 4) studies that explicitly represented the use of Xpert for the diagnosis of patients suspected of having TB.

Assessment of reporting bias

We chose not to carry out formal assessment of publication bias using methods such as funnel plots or regression tests because such techniques have not been found to be helpful for diagnostic test accuracy studies (Macaskill 2010). However, Xpert is produced by only one manufacturer and, being a new test for which there has been considerable attention and scrutiny, we believe reporting bias was minimal.

Other analyses

NTM

NTM, such as *M. avium* complex and *M. intracellulare*, comprise a multi-species group of human pathogens that are ubiquitous in water and soil. NTM can cause severe pulmonary and other diseases that share clinical signs with TB but are treated differently from TB, including drug-resistant TB. People living with HIV with severe immunosuppression are particularly vulnerable to infections caused by NTM (Gopinath 2010). NTMs were not mentioned in the majority of studies, and, when reported, were usually excluded from specificity determinations. We, therefore, summarized separately data for NTM by determining the percent of false-positive Xpert results in samples that grew NTMs (see Other analyses: NTM).

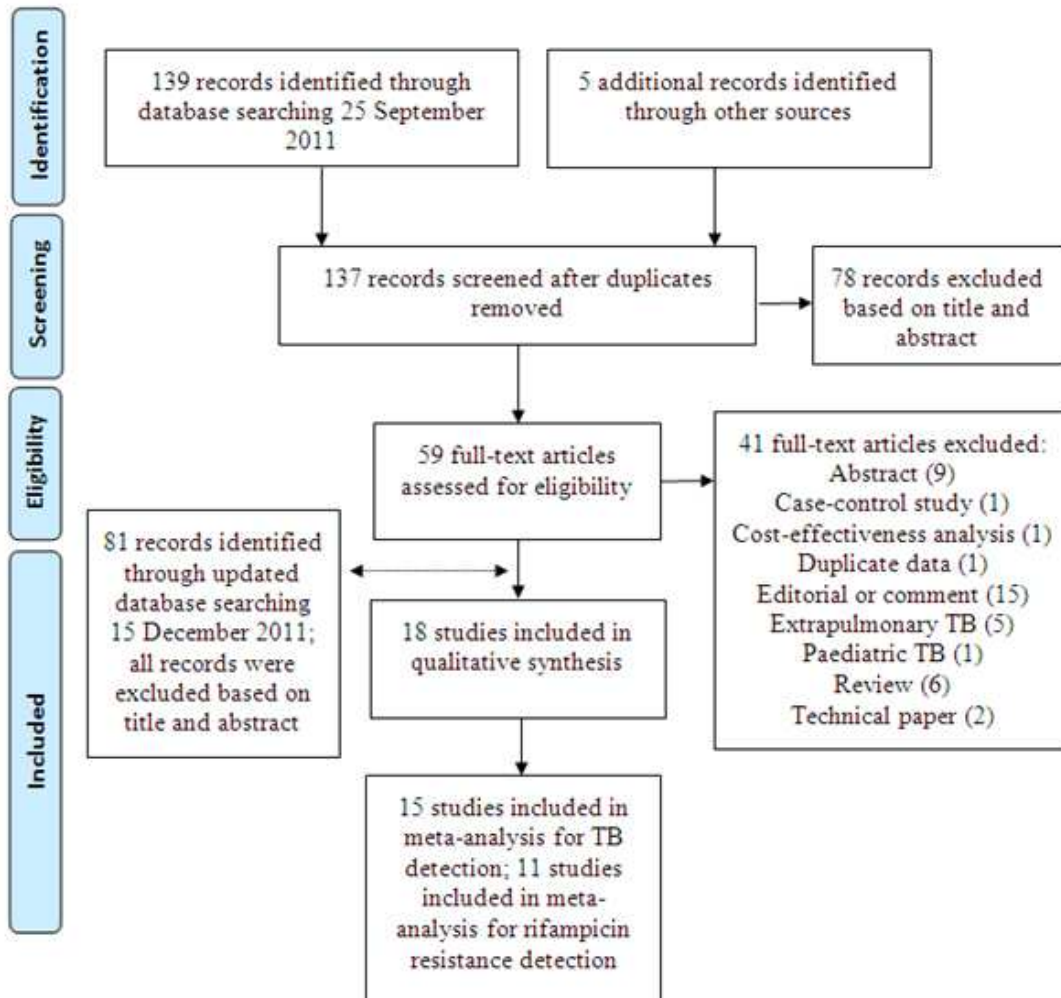
RESULTS

Results of the search

We performed the initial electronic search on 25 September 2011 and we identified 139 titles. We added five titles through reference

checking and correspondence with experts in the field. After we removed duplicates, 137 titles remained of which we excluded 78 titles based on a review of title and abstract. We retrieved full-text articles for 59 citations, of which we excluded 41 for the following reasons: abstract (nine studies); case control design (one study); cost-effectiveness analysis (one study); duplicate title (one study); editorials or comments (15 studies); evaluation of Xpert for extrapulmonary TB (five studies) or paediatric TB (one study); narrative reviews (six studies); and technical aspects (two studies) (Characteristics of excluded studies describes selected papers). We performed an updated search on 15 December 2011 that yielded 81 titles, all of which had been identified during the previous search or were ineligible based on title or abstract. Thus, we included 18 relevant studies in this review, (Figure 2). Of the total 18 studies, two were international multicentre studies (Boehme 2010a; Boehme 2011a) carried out at five and six study centres, respectively. The two studies by Boehme involved different patients. For Boehme 2010a and Boehme 2011a descriptive results and methodological quality were presented at the study centre level. Meta-analysis results were presented at the study level. One other study, conducted at three sites, presented accuracy data for the three sites combined and was considered as a single study and a single study centre (Marlowe 2011). Hence there were 27 study centres in the review.

Figure 2. Flow diagram of studies in the review.



Methodological quality of included studies

Figure 3 shows the overall quality of the 27 study centres included in the review. In the patient selection domain, we considered 20 study centres (74%) to be at low risk of bias because participants were enrolled consecutively. We considered the majority of the remaining study centres to be at high risk of bias, mainly because the sampling method was by convenience. With regard to applicability (patient characteristics and setting), we judged six study centres (one study) to be of low concern because the study centres in this study evaluated Xpert in decentralized laboratories associated with health clinics and provincial hospitals (Boehme 2011a). The remaining study centres in the review ran Xpert in reference laboratories, thus their results may not be generalizable to other

settings that lack specialized laboratory infrastructure and highly trained personnel. We also noted that no study centre used Xpert at the point of care in a health facility outside a laboratory. In the index test domain, we considered all study centres to be at low concern for both risk of bias and applicability. In the reference standard domain, we judged 23 study centres (85%) to be at low risk of bias for TB and 24 study centres (89%) to be at low risk of bias for rifampicin resistance. In the reference standard domain, we judged Hanif 2011 to be at high risk of bias for TB as the target condition (objective I) because blinding was not reported, but low risk of bias for rifampicin resistance as the target condition (objective II) since this study used BACTEC 460, a system that provides an automated test result. We judged applicability to be

of low concern for all studies in the reference standard domain. In the flow and timing domain, we considered 23 study centres (85%) to be of low concern for risk of bias because all patients were accounted for in the analysis and information about indeterminate results was provided. The quality assessment results for the individual study centres can be found in [Figure 4](#). Note that we had nearly complete information for all study centres.

Figure 3. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across the 27 included study centres (18 studies). The reference standard domain pertains to TB as the target condition. See text for the reference standard pertaining to rifampicin resistance.

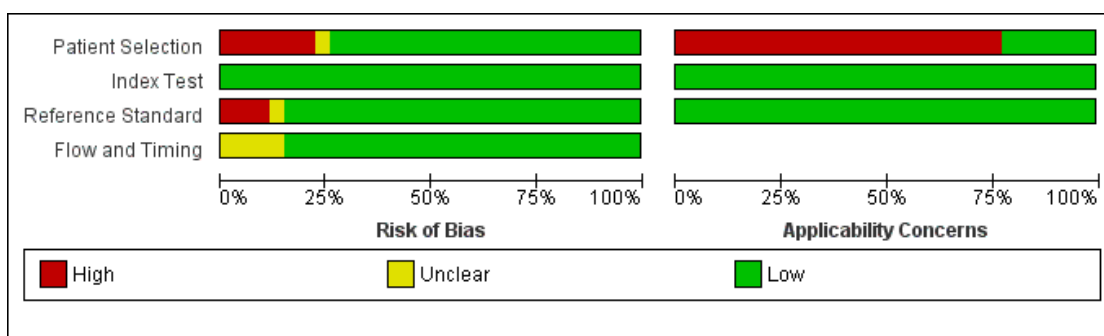


Figure 4. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study centre.

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Boehme 2010a	+	+	+	+	-	+	+
Boehme 2010b	+	+	+	+	-	+	+
Boehme 2010c	+	+	+	+	-	+	+
Boehme 2010d	+	+	+	+	-	+	+
Boehme 2010e	+	+	+	+	-	+	+
Boehme 2011a	+	+	+	+	+	+	+
Boehme 2011b	+	+	+	+	+	+	+
Boehme 2011c	+	+	+	+	+	+	+
Boehme 2011d	+	+	+	+	+	+	+
Boehme 2011e	+	+	+	+	+	+	+
Boehme 2011f	+	+	+	+	+	+	+
Bowles 2011	-	+	+	?	-	+	+
Ciftci 2011	?	+	+	?	-	+	+
Friedrich 2011	-	+	+	+	-	+	+
Hanif 2011	+	+	-	+	-	+	+
Helb 2010	+	+	+	?	-	+	+
Ioannidis 2011	-	+	+	+	-	+	+
Lawn 2011	+	+	+	+	-	+	+
Malbruny 2011	+	+	+	+	-	+	+
Marlowe 2011	-	+	+	+	-	+	+
Miller 2011	-	+	+	+	-	+	+
Moure 2011	-	+	?	+	-	+	+
Rachow 2011	+	+	+	?	-	+	+
Scott 2011	+	+	+	+	-	+	+
Teo 2011	+	+	-	+	-	+	+
Theron 2011	+	+	+	+	-	+	+
Zeka 2011	+	+	-	+	-	+	+

- High
 ? Unclear
 + Low

Findings

Eighteen studies including 7816 participants (median number in the studies 117, Interquartile range (IQR) 93, 214), evaluated Xpert for TB detection. The prevalence of TB in the 18 studies ranged from 18.3% (Lawn 2011) to 100% (Friedrich 2011), median 37.4% (IQR 29.4, 59.8). Of the total 18 studies, 11 studies including 2340 participants provided data for 2 x 2 tables for rifampicin resistance detection. Of the remaining studies, four studies reported that all specimens were found to be drug sensitive (no rifampicin resistant specimens) (Ciftci 2011; Hanif 2011; Marlowe 2011; Rachow 2011); two studies provided data jointly for pulmonary and extrapulmonary specimens (Miller 2011; Moure 2011); and one study did not report information on rifampicin resistance (Helb 2010). The prevalence of rifampicin resistance in the 11 studies ranged from 0.8% (Teo 2011) to 29.4% (Scott 2011), median 7.3% (IQR 3.0, 19.6). All studies used a cross-sectional study design relevant to determining the diagnostic accuracy of Xpert. The majority of studies used expectorated (coughed-up) sputum not induced sputum.

[Characteristics of included studies](#) presents key characteristics for the 27 study centres. Nineteen study centres (70.4%) (10 studies (55.6%)) were located in low-income or middle-income countries. In the countries represented by the 27 study centres, TB incidence rates per 100,000 population ranged from 4.1 (USA) to 981 (South Africa). The percent MDR-TB among new TB cases ranged from 0.9% (Greece and South Africa, Cape Town) to 22.3% (Azerbaijan) and among retreatment cases, ranged from 0% (Tanzania) to 55.8% (Azerbaijan) (Wright 2009; WHO Drug

Resistance 2008; WHO M/XDR-TB 2010; Zignol 2012).

I. A. TB detection, Xpert used as an initial test replacing smear microscopy

Forest plots of Xpert sensitivity and specificity for TB detection are presented for the total 18 studies (27 study centres) in [Figure 5](#). Sensitivity estimates varied from 58% to 100% and specificity estimates, from 94% to 100%. We included fifteen of the total 18 studies, including 7517 participants, in this meta-analysis. We excluded one study because it preferentially enrolled smear-positive patients (Friedrich 2011) and two studies because they preferentially enrolled smear-negative patients (Ioannidis 2011; Moure 2011). The pooled median sensitivity and specificity were 88% (95% CrI 83% to 92%) and 98% (95% CrI 97% to 99%), respectively, ([Table 1](#)). The predicted sensitivity and specificity for Xpert for TB detection were 88% (95% CrI 66% to 97%) and 98% (95% CrI 92% to 100%), respectively, the wider 95% CrIs around the predicted values suggesting some variability between studies particularly in sensitivity. [Figure 6](#) presents the pooled and predicted sensitivity and specificity estimates together with the credible and prediction regions for Xpert for TB detection. The summary point appears close to the upper left-hand corner of the plot, suggesting high accuracy of Xpert for TB detection. The 95% credible region around the summary (pooled) value of sensitivity and specificity, the region that contains likely combinations of the pooled sensitivity and specificity, is relatively narrow. The 95% prediction region is wider, displaying more uncertainty as to where the likely values of sensitivity and specificity might occur for individual studies.

Figure 5. Forest plots of Xpert sensitivity and specificity for TB detection, Xpert used as an initial test replacing smear microscopy. The individual studies are ordered by decreasing sensitivity. TP = True Positive; FP = False Positive; FN = False Negative; TN = True Negative. Between brackets are the 95% CI of sensitivity and specificity. The figure shows the estimated sensitivity and specificity of the study (blue square) and its 95% CI (black horizontal line). Xpert specificity could not be estimated in one study.

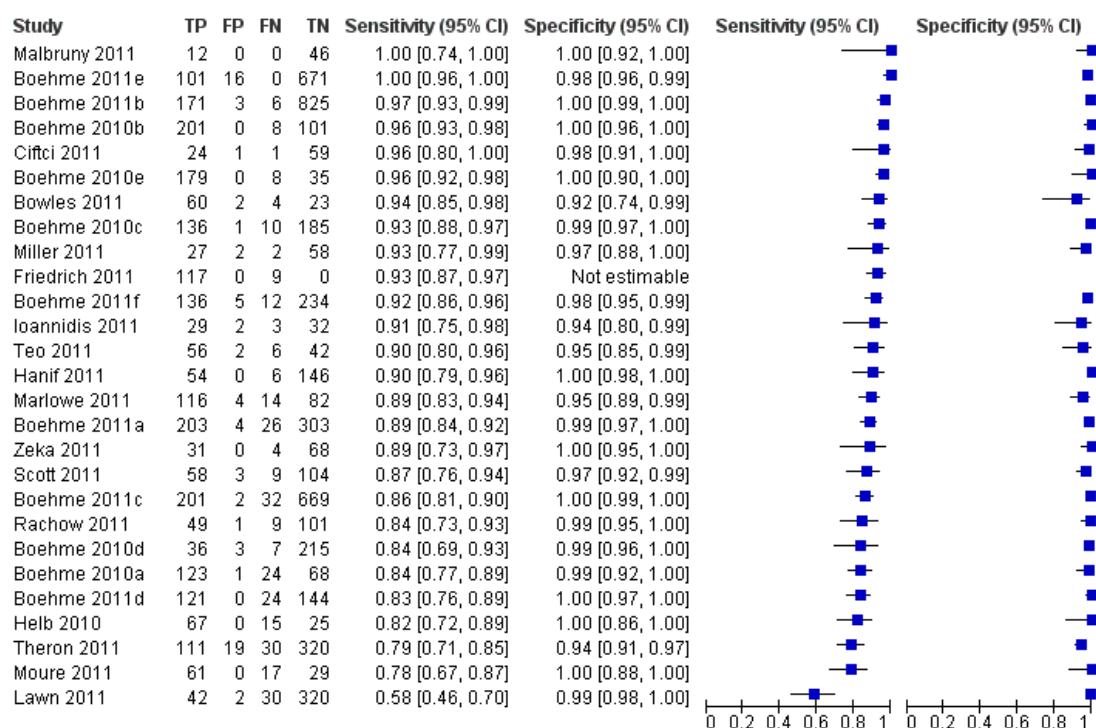
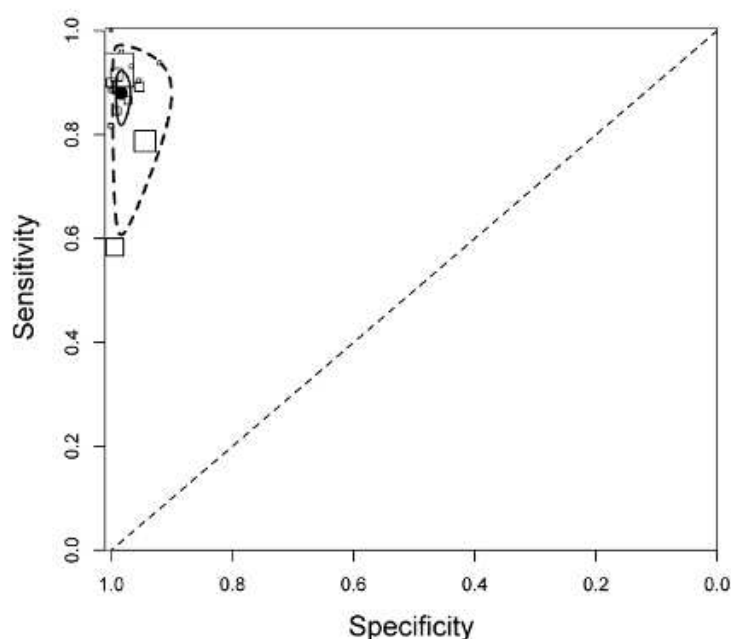


Figure 6. Summary plots of Xpert sensitivity and specificity for TB detection, Xpert used as an initial test replacing smear microscopy. Each individual study is represented by an empty square. The size of the square is proportional to the sample size of the study such that larger studies are represented by larger squares. The filled circle is the pooled median estimate for sensitivity and specificity. The solid curves represent the 95% credible region around the summary estimate; the dashed curves represent the 95% prediction region.



I. B. TB detection, Xpert used as an add-on test following a negative smear microscopy result

We presented forest plots of Xpert sensitivity and specificity for TB detection for studies reporting data for smear-negative patients for 16 studies (25 study centres) in [Figure 7](#). Two studies initially performed microscopy and, for those patients found to be smear-negative, subsequently ran the Xpert test ([Ioannidis 2011](#); [Moure 2011](#)). Both studies were laboratory-based assessments performed in high-income countries. [Ioannidis 2011](#) ($n = 52$) reported a higher sensitivity (83%, 95% CI 59% to 86%) and lower specificity (94%, 95% CI 80% to 99%) than [Moure 2011](#) ($n = 107$) (sensitivity of 78%, 95% CI 67% to 87%; specificity of 100%, 95% CI 88% to 100%). As noted above, we considered a diagnostic strategy that enrolled unselected participants who were evaluated by both Xpert and microscopy concurrently to be a proxy for

a strategy in which Xpert was used as an add-on test following a negative smear microscopy result. We included fourteen studies, including 5719 participants, in this meta-analysis. We excluded [Ioannidis 2011](#), [Moure 2011](#), and two additional studies from this analysis: [Ciftci 2011](#) because data by smear status were not provided and [Friedrich 2011](#) because this study preferentially selected smear-positive patients. In the meta-analysis, the pooled sensitivity was 67% (95% CrI 58% to 74%) and the pooled specificity was 98% (95% CrI 97% to 99%) ([Table 1](#)). [Figure 8](#) presents the pooled and predicted sensitivity and specificity estimates together with the credible and prediction regions for this analysis. The summary point is relatively far from the upper left-hand corner of the plot, suggesting low accuracy of Xpert when used as an add-on test following microscopy. The 95% credible region around the summary value of sensitivity and specificity is relatively wide.

Figure 7. Forest plots of Xpert for TB detection, Xpert used as an add-on test following a negative smear microscopy result. TP = True Positive; FP = False Positive; FN = False Negative; TN = True Negative. Between brackets the 95% CI of sensitivity and specificity. The figure shows the estimated sensitivity and specificity of the study (blue square) and its 95% CI (black horizontal line).

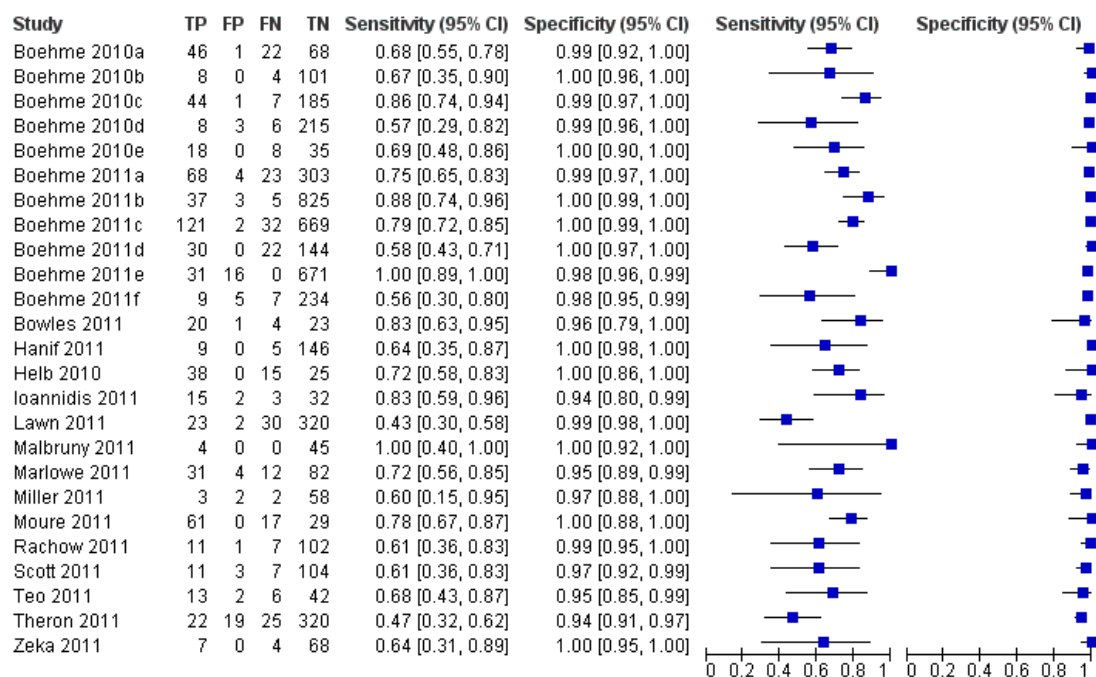
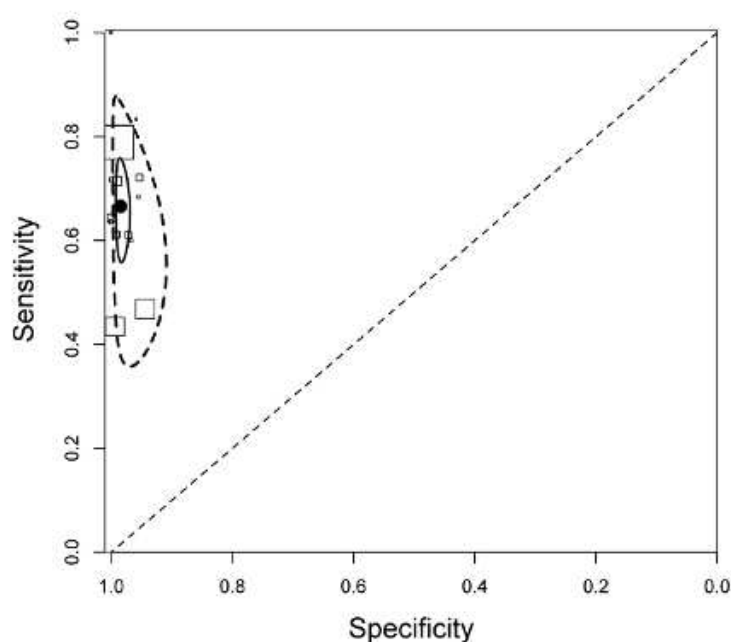


Figure 8. Summary plots of Xpert sensitivity and specificity for TB detection, Xpert used as an add-on test following a negative smear microscopy result. Each individual study is represented by an empty square. The size of the square is proportional to the sample size of the study such that larger studies are represented by larger squares. The filled circle is the pooled median estimate for sensitivity and specificity. The solid curve represents the 95% credible region around the summary estimate; the dashed curves represent the 95% prediction region.



Indeterminate results

Of the total 18 studies, four studies reported zero indeterminate results; 10 studies reported the number of indeterminate test results; and four studies did not provide information about indeterminate results. Data were available only at the study level for multicentre studies. Of 13,277 tests performed, the pooled proportion of indeterminate test results was very low (1.1%, 95% CrI 0.04% to 2.0%).

Investigations of heterogeneity, TB detection

Clinical subgroups

It is possible that the accuracy of Xpert in distinct subgroups of patients could differ causing heterogeneity in Xpert performance. We, therefore, determined sensitivity and specificity estimates for patients grouped by smear or HIV status.

TB detection in smear-positive and smear-negative patients

Figure 9 displays the forest plots for studies reporting data for smear-positive patients. There was little heterogeneity in sensitivity estimates, range 95% to 100%. For determination of Xpert specificity, only five of the 15 studies reported the cross-tabulation between smear and culture and, as expected, very few (≤ 1) smear-positive, culture negative participants were recorded (Bowles 2011; Hanif 2011; Ioannidis 2011; Malbruny 2011; Teo 2011). We included 1735 participants in the meta-analysis of Xpert for smear-

positive, culture-positive TB. The pooled sensitivity estimate was very high at 98% (95% CrI 97% to 99%) (Table 2). We could not estimate Xpert pooled specificity in the studies in this subgroup as the participants were all considered true TB positive.

Figure 9. Forest plot of Xpert sensitivity for TB detection in smear-positive subgroup. The squares represent the sensitivity and specificity of one study, the black line its CI. TP = true positive; FP = false positive; FN = false negative; TN = true negative. Xpert specificity could not be estimated in these studies.

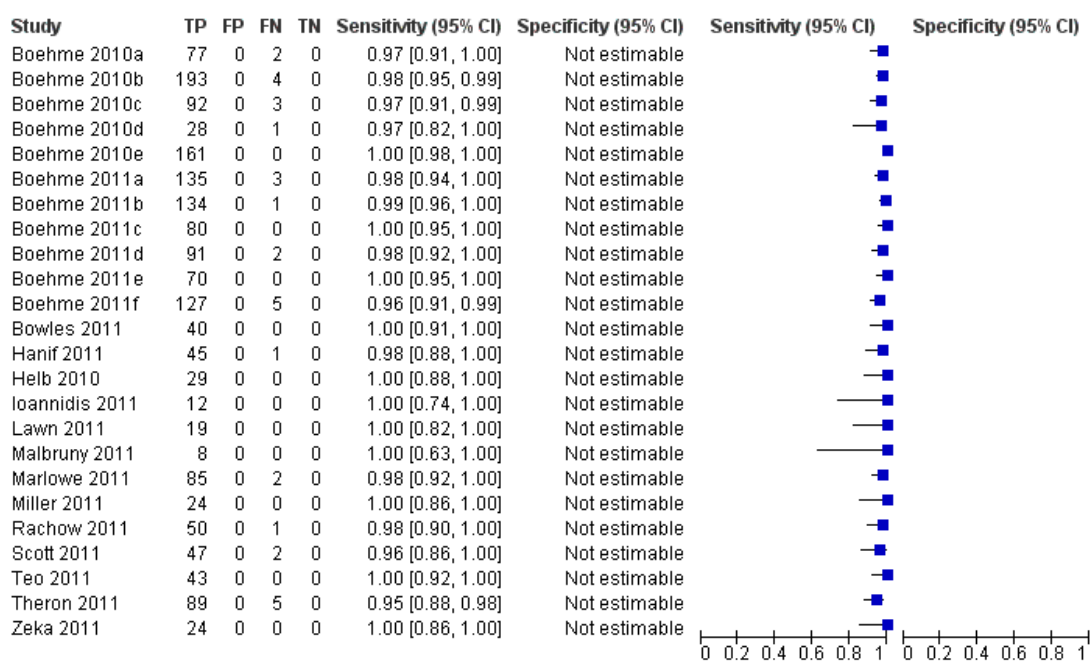


Figure 7 displays the forest plots for studies reporting data for smear-negative patients. There was considerable variability in sensitivity estimates, 43% to 100%. Specificity estimates showed far less variation, ranging from 94% to 100%. The study by Lawn 2011, which showed the lowest sensitivity, evaluated the use of Xpert to screen HIV-infected patients with advanced immunodeficiency enrolling in antiretroviral therapy services regardless of symptoms. Although the majority of patients in the study had TB symptoms, the inclusion of asymptomatic patients may explain the lower sensitivity in this study. For the meta-analysis, 15 studies (24 study centres) provided a within-study comparison of smear-positive and smear-negative participants. We excluded three studies from this analysis: one study that pre-selected smear-positive patents (Friedrich 2011), one study that enrolled only smear-negative patients (Moure 2011), and one study that did not provide data by smear status (Ciftci 2011). We included 5771 participants

in the meta-analysis of Xpert for smear-negative, culture-positive TB. The pooled sensitivity estimate for smear-negative TB was 68% (95% CrI 59% to 75%), considerably lower than the pooled sensitivity estimate for smear-positive TB of 98% (95% CrI 97% to 99%) (Table 2). The 95% CrI for the difference in Xpert sensitivity in smear-positive and smear-negative subgroups lies entirely above 0 (Table 2).

TB detection among people living with HIV

We identified five studies (11 study centres) including 1557 HIV-positive participants and four studies (13 study centres) including 1981 HIV-negative participants (Figure 10). In the HIV-positive subgroup, there was variability in sensitivity estimates, ranging from 70% to 100%. A possible explanation for this heterogeneity is the presence of smear-negative specimens in this subgroup (see TB detection among HIV infected, smear-negative patients be-

low). In the HIV-negative subgroup, there was less heterogeneity in sensitivity estimates, ranging from 83% to 100%. Specificity varied less than sensitivity in both subgroups, ranging from 92% to 100% in studies including people living with HIV and from 96% to 100% in studies including people without HIV infection. In the HIV-positive subgroup, the pooled sensitivity was 76% (95% CrI 63% to 85%), lower than the estimate in the HIV-negative subgroup (89%, 95% CrI 81% to 94%). We estimated the probability that the sensitivity in the HIV-positive subgroup was greater than that in the HIV-negative subgroup was 0.94 (Table

2). When we excluded the study by Lawn 2011 and limited the meta-analysis to the four studies that provided data for both HIV-infected and uninfected participants (within-study comparisons), the pooled sensitivity in the HIV-positive subgroup (1163 participants) increased to 80% (95% CrI 67% to 88%) (Table 2). Using within-study comparisons, the pooled specificity was slightly lower in the HIV-positive subgroup (97%, 95% CrI 93% to 99%) than the HIV-negative subgroup (99%, 95% CrI 96% to 99%) (Table 2).

Figure 10. Forest plots of Xpert sensitivity and specificity for TB detection in HIV-positive and HIV-negative subgroups. The squares represent the sensitivity and specificity of one study and the black line represent its CI. TP = true positive; FP = false positive; FN = false negative; TN = true negative.

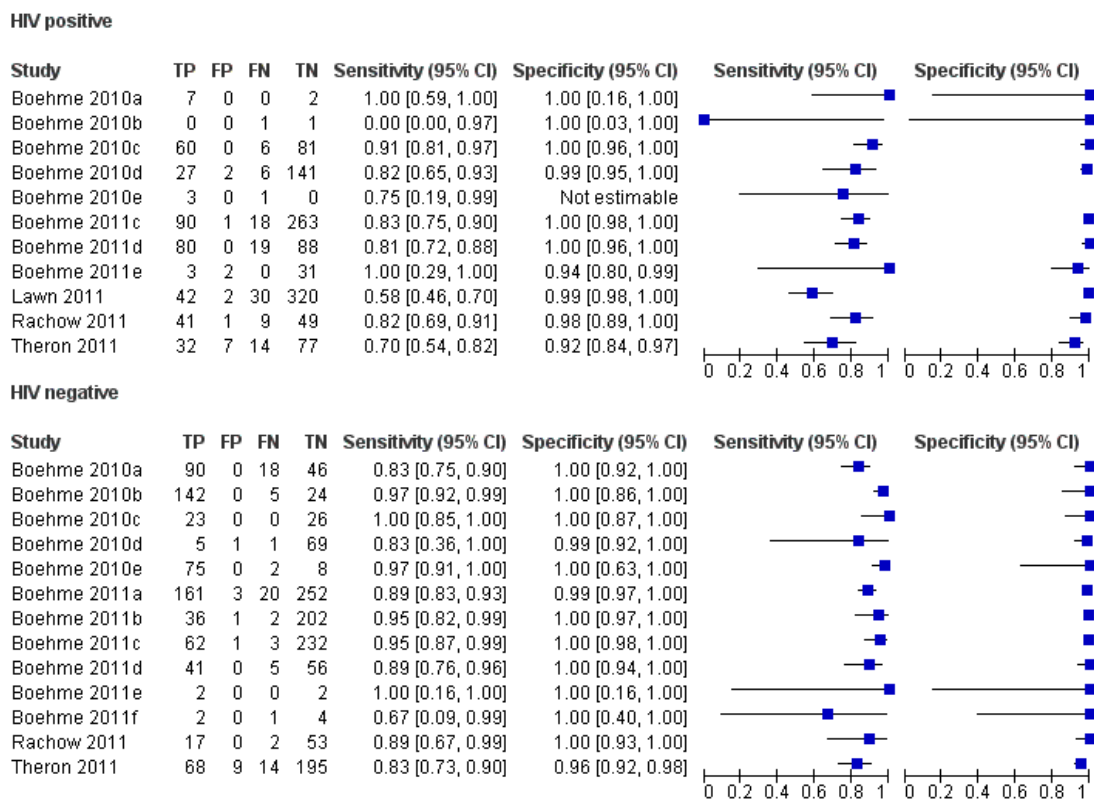
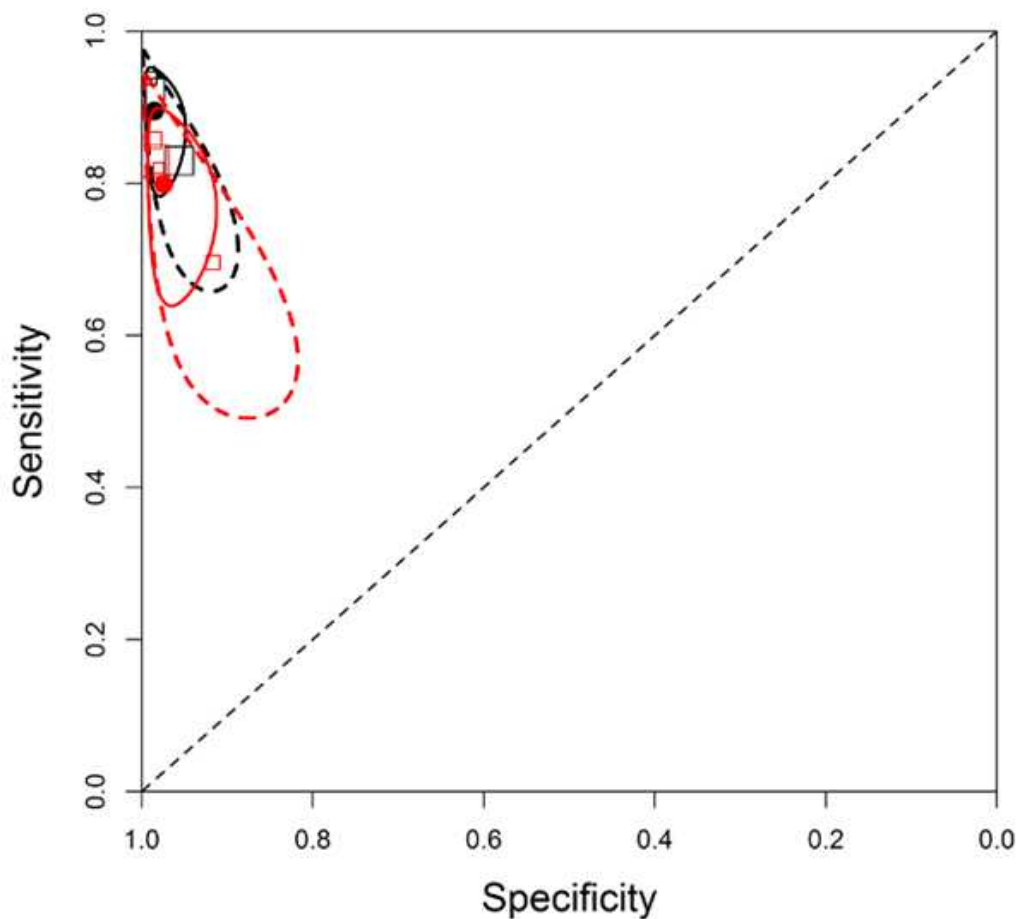


Figure 11 presents the pooled and predicted estimates of sensitivity and specificity together with the credible and prediction region for the HIV-positive and HIV-negative subgroups (four within-study comparisons). The 95% credible region around the pooled

estimate of sensitivity and specificity is relatively wide. The 95% prediction region is even wider, displaying more uncertainty as to where the likely values of sensitivity and specificity may occur for individual studies. Based on the pooled value and the credible re-

gion, it appeared that Xpert sensitivity had a greater probability of being lower in the HIV-positive than the HIV-negative subgroup. However, there was considerable heterogeneity between studies and the prediction region for the HIV-positive subgroup completely encompassed that for the HIV-negative subgroup. Therefore, there is too much uncertainty in the pooled estimates in the two subgroups to clearly distinguish between them.

Figure 11. Summary plots of Xpert sensitivity and specificity for TB detection in HIV-positive (red colour) and HIV-negative subgroups (black colour). Each individual study is represented by an empty square. The size of the square is proportional to the sample size of the study such that larger studies are represented by larger squares. The filled circles are the median estimates for sensitivity and specificity. The solid curves represent the 95% credible region around the summary estimates; the dashed curves represent the 95% prediction region.



TB detection among HIV-positive smear-negative patients

Two studies reported data from which to assess the accuracy of Xpert in HIV-positive patients with culture-positive, smear-negative TB. In [Lawn 2011](#), among people living with HIV, Xpert sensitivity for smear-negative, culture-positive TB was 43% (95% CI 30% to 58%; 23/53) compared with 100% (95% CI 82% to 100%; 19/19) for smear-positive, culture-positive TB. In [Theron 2011](#), among people living with HIV, Xpert sensitivity for smear-negative, culture-positive TB was 47% (95% CI 29% to 67%; 11/23) compared with 91% (95% CI 72% to 99%; 21/23) for smear-positive, culture-positive TB. We did not perform a meta-analysis because of the small number of studies.

Effect of condition of the specimen

We included twelve studies in this meta-analysis. Again, we excluded three studies that made an effort to select smear-positive patients ([Friedrich 2011](#)) or smear-negative patients ([Ioannidis 2011](#); [Moure 2011](#)). In addition, we excluded three studies that reported results for fresh and frozen specimens jointly ([Bowles 2011](#); [Malbruny 2011](#); [Marlowe 2011](#)). The pooled sensitivity was 88% (95% CrI 80% to 93%) in studies using fresh specimens (six studies), slightly higher than the pooled sensitivity of 85% (95% CrI 77% to 91%) in studies using frozen specimens (six studies) ([Table 2](#)). The probability that the pooled sensitivity in fresh specimens exceeds that in frozen specimens was estimated at 0.73 ([Table 2](#)). The pooled specificity was 99% (95% CrI 98% to 100%) for fresh specimens, higher than the pooled specificity of 97% (95% CrI 95% to 99%) for frozen specimens. The probability that the pooled specificity in fresh specimens exceeds that in frozen specimens was estimated at 0.98 ([Table 2](#)). Thus, in comparison with frozen specimens, fresh specimens had a significantly higher specificity but not sensitivity.

Effect of specimen preparation

We included fifteen studies in this meta-analysis (again excluding the three studies that made an effort to select smear-positive patients ([Friedrich 2011](#)) or smear-negative patients ([Ioannidis 2011](#); [Moure 2011](#))). Five studies used samples prepared from unprocessed (untreated) specimens and ten studies used samples prepared from processed specimens, usually pellets processed by N-acetyl-cysteine and sodium hydroxide (NALC-NaOH) centrifugation. The pooled sensitivity estimate was 92% (95% CrI 87% to 96%) in studies using unprocessed specimens, higher than the pooled sensitivity estimate of 85% (95% CrI 79% to 90%) in studies using processed specimens ([Table 2](#)). The probability that the pooled sensitivity in unprocessed specimens exceeds that in processed specimens was estimated at 0.98. The 95% CrI for the difference in Xpert sensitivity in the unprocessed and processed

subgroups lies entirely above 0. Thus, unprocessed specimens had a significantly higher sensitivity than processed specimens. The pooled specificity was 99% (95% CrI 97% to 99%) for unprocessed specimens, similar to the pooled specificity of 98% (95% CrI 96% to 99%) for processed specimens ([Table 2](#)).

Effect of TB prevalence

We included fifteen studies in this meta-analysis (again excluding the three studies that made an effort to select smear-positive patients ([Friedrich 2011](#)) or smear-negative patients ([Ioannidis 2011](#); [Moure 2011](#))). We considered nine studies to have high TB prevalence (> 30%) and six studies to have low TB prevalence (\leq 30%). The pooled sensitivity was 89% (95% CrI 84% to 93%) for high TB settings, higher than the pooled sensitivity of 86% (95% CrI 77% to 92%) for low TB settings ([Table 2](#)). The probability that the pooled median sensitivity in high TB settings exceeds that in low TB settings was estimated at 0.80 ([Table 2](#)). The pooled specificity estimates were similar, 98% (95% CrI 96% to 99%) for high TB settings and 99% (95% CrI 97% to 99%) for low TB settings.

Effect of country income status

We included fifteen studies in this meta-analysis (again excluding the three studies that made an effort to select smear-positive patients ([Friedrich 2011](#)) or smear-negative patients ([Ioannidis 2011](#); [Moure 2011](#))). The pooled sensitivity was 85% (95% CrI 79% to 90%) for low/middle-income countries (nine studies), lower than the pooled sensitivity of 92% (95% CrI 86% to 96%) for high-income countries (six studies) ([Table 2](#)). The probability that the pooled sensitivity in high income countries exceeds that for low/middle-income countries was estimated at 0.96; the specificities were similar for the two income levels ([Table 2](#)).

II. A. Rifampicin resistance detection, Xpert used as an initial test replacing conventional drug susceptibility testing

As noted above, of the total 18 studies, 11 studies (20 study centres) including 2340 participants provided data for Xpert for rifampicin resistance detection. Of the remaining studies, four studies reported that all specimens were found to be drug sensitive (no rifampicin resistant specimens) ([Ciftci 2011](#); [Hanif 2011](#); [Marlowe 2011](#); [Rachow 2011](#)) and one study did not report information on rifampicin resistance ([Helb 2010](#)). Two studies provided data jointly for pulmonary and extrapulmonary specimens: [Miller 2011](#) reported that Xpert detected four rifampicin-resistant specimens of which three were positive by the reference standard and [Moure 2011](#) reported that Xpert detected 86% (6/7) of the rifampicin-resistant specimens. The 11 studies in this analysis included 485 rifampicin-resistant specimens, median four specimens, (IQR 2, 7).

Two studies, [Boehme 2010a](#) and [Boehme 2011a](#) accounted for the vast majority (94%) of the rifampicin-resistant specimens. [Figure 12](#) shows the forest plots of sensitivity (presence of rifampicin-resistant TB) and specificity (presence of rifampicin-susceptible TB) for this analysis. The individual study centres in the plots are ordered by decreasing sensitivity and decreasing number of true positive results. Although, there was considerable heterogeneity in sensitivity estimates (ranging from 33% to 100%), in general there was less variability among study centres with a larger number of rifampicin-resistant specimens. Specificity showed less variability

than sensitivity, ranging from 83% to 100%. The pooled sensitivity and specificity were 94% (95% CrI 87% to 97%) and 98% (95% CrI 97% to 99%), respectively ([Table 1](#)). [Figure 13](#) presents the pooled and predicted estimates of sensitivity and specificity together with the credible and prediction region for Xpert for rifampicin resistance detection. The summary point appears close to the upper left-hand corner of the plot, suggesting high accuracy of Xpert for rifampicin resistance detection. The 95% credible region around the summary (pooled) value of sensitivity and specificity is relatively narrow.

Figure 12. Forest plots of Xpert sensitivity and specificity for detection of rifampicin resistance, Xpert used as an initial test replacing conventional drug susceptibility testing as the initial test. The individual studies are ordered by decreasing sensitivity and decreasing number of true positives. The squares represent the sensitivity and specificity of one study, the black line its CI. TP = true positive; FP = false positive; FN = false negative; TN = true negative.

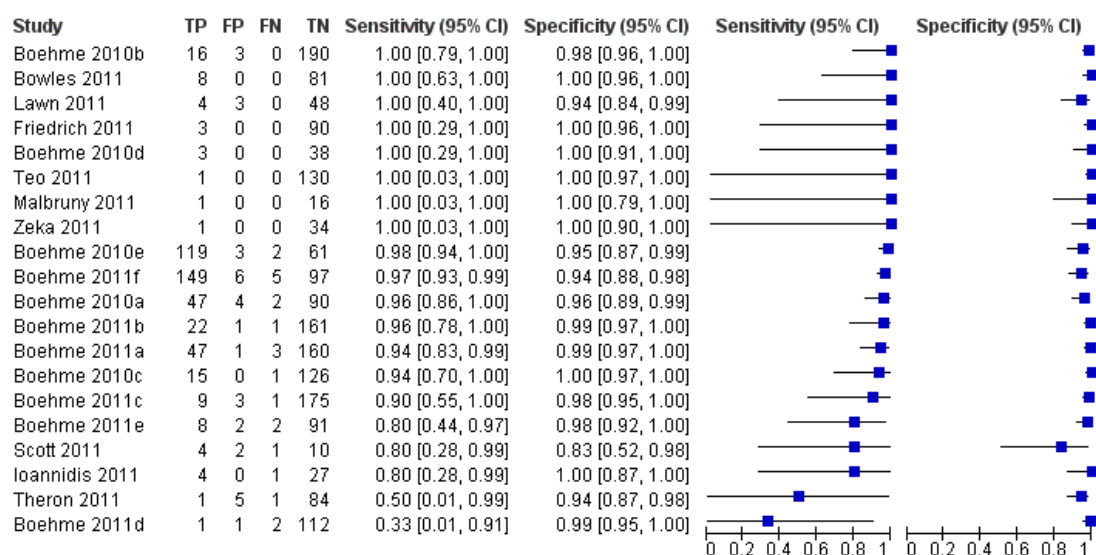
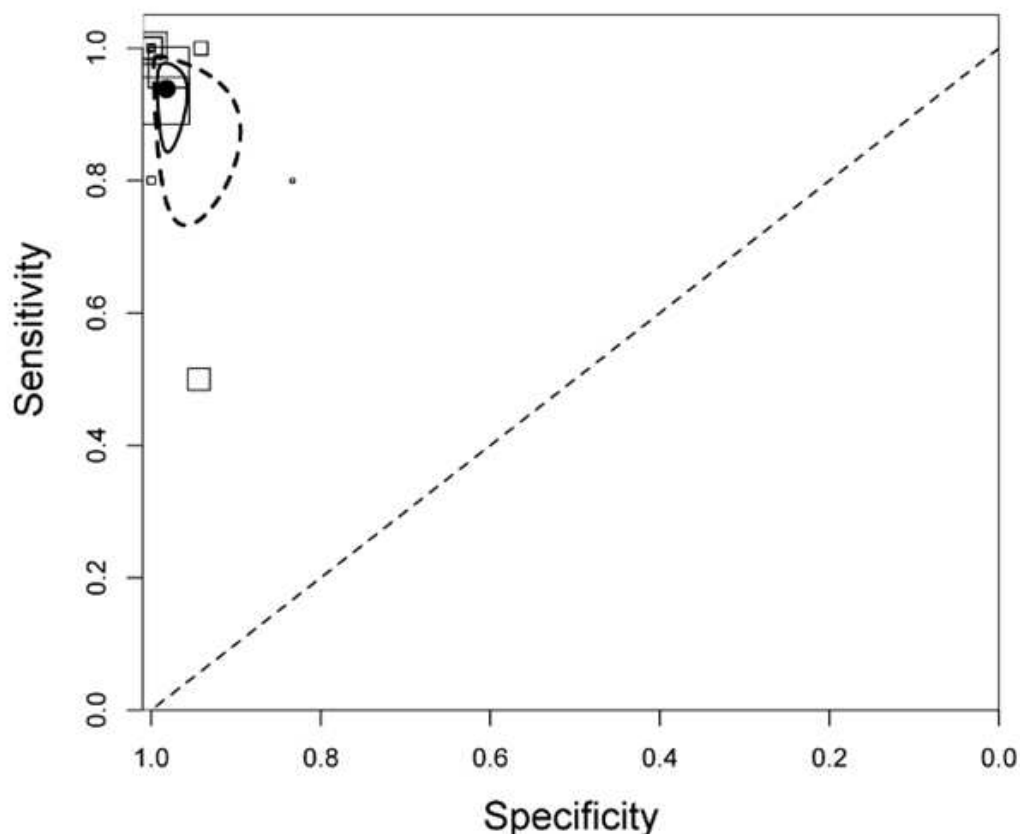


Figure 13. Summary plots of Xpert sensitivity and specificity for detection of rifampicin resistance, Xpert used as an initial test replacing conventional drug susceptibility testing as the initial test. Each individual study is represented by an empty square. The size of the square is proportional to the sample size of the study such that larger studies are represented by larger squares. The filled circle is the pooled median estimate for sensitivity and specificity. The solid curves represent the 95% credible region around the summary estimate; the dashed curves represent the 95% prediction region.



Investigations of heterogeneity, rifampicin resistance detection

Effect of delta cycle threshold cutoff

A major source of heterogeneity in systematic reviews of diagnostic test accuracy is the difference in values (thresholds) used to define a positive test between studies. We explored the effect of the parameter value (delta cycle threshold cutoff) in the Xpert system on the determination of rifampicin resistance. We included eleven studies in this analysis: four studies used a parameter value of 3.5 and seven studies a value of 5. The pooled sensitivity was 96%

(95% CrI 81% to 100%) for studies using 3.5 and 94% (95% CrI 86% to 97%) for studies using 5 (Table 3). The pooled specificities were high for both parameter values, 100% (95% CrI 98% to 100%) for studies using 3.5 and 98% (95% CrI 96% to 99%) for studies using 5. Thus, there was considerable overlap between the estimates of Xpert performance for the two parameter values and no apparent difference between them.

Effect of rifampicin resistance prevalence

The pooled sensitivities for both levels of rifampicin resistance prevalence were similar: 94% (95% CrI 85% to 98%) for high rifampicin resistance prevalence settings (> 15%, four studies) and

93% (95% CrI 80% to 99%) for low rifampicin resistance prevalence settings ($\leq 15\%$, seven studies) (Table 3). The pooled specificities were also similar: 98% (95% CrI 95% to 99%) for high rifampicin resistance settings and 98% (95% CrI 96% to 99%) for low rifampicin resistance settings.

Sensitivity analyses

We performed sensitivity analyses by limiting inclusion in the meta-analysis to: 1) studies that provided data by age that met the criterion for adults; 2) studies where consecutive patients were selected; 3) studies where a single specimen yielded a single Xpert result for a given patient; and 4) studies that clearly represented the use of the test for diagnosis of patients with suspected TB. We found no differences in the findings in these subgroups compared with the overall results. Concerning the diagnostic setting, 17 of the included studies evaluated Xpert in reference laboratories. One multicentre study, evaluated Xpert in district or subdistrict health

facilities (Boehme 2011a). The results from this study may not be generalizable to other basic laboratories because this study was performed under research conditions. Nonetheless, in the meta-analysis, the exclusion of results from this single study did not make a difference in the findings.

Other analyses

NTM

Eight studies provided data on a variety of NTM that grew from the specimens tested to look for evidence of cross-reactivity: six NTM (Bowles 2011); one NTM (Ioannidis 2011); 41 NTM (Marlowe 2011); 20 NTM (Moure 2011); 45 NTM (Rachow 2011); five NTM (Scott 2011); 13 NTM (Teo 2011); and eight NTM (Theron 2011). Among these studies comprising 139 NTM, Xpert was positive in only one (0.7%) specimen that grew NTMs (Rachow 2011).

Summary of findings

Review question: What is the diagnostic accuracy of Xpert MTB/RIF assay for detection of pulmonary TB and detection of rifampicin resistance?

Patients/population: Adults suspected of having pulmonary TB or MDR-TB (for TB detection); confirmed TB cases (for rifampicin resistance detection)

Purpose: TB detection: Xpert MTB/RIF assay used as an initial test replacing microscopy and used as an add-on test following a negative smear microscopy result. Rifampicin resistance detection: Xpert MTB/RIF assay as an initial test replacing conventional phenotypic drug susceptibility testing

Setting: Basic laboratories and primary health facilities (peripheral health services level)

Index test: Xpert MTB/RIF assay

Importance: Compared with culture and conventional drug susceptibility testing, Xpert MTB/RIF assay could have considerable advantages for scaling up programmatic management of TB by offering rapid diagnosis nearer to the point of care, standardized testing, potential for high throughput, and fewer requirements for laboratory biosafety

Reference standards: TB: solid or liquid culture; rifampicin resistance: phenotypic drug susceptibility testing

Studies: Cross-sectional or cohort

Diagnostic accuracy for TB detection

Type of analysis	Effect (95% credible interval)	No. of participants (studies)	What do these results mean given 5% prevalence of TB among individuals suspected of having pulmonary TB?	What do these results mean given 15% prevalence of TB among individuals suspected of having pulmonary TB?	What do these results mean given 30% prevalence of TB among individuals suspected of having pulmonary TB?
TB detection, Xpert used as an initial test replacing microscopy	Pooled median sensitivity 88% (83, 92) and pooled median specificity 98% (97, 98)	7517 (15)	With a prevalence of 5%, 50/1000 individuals will have pulmonary TB. Of these, 44 (TP) will be identified; 6 (FN) will be missed. Of the 950 individuals without TB, 931 (TN) will not be treated; 19 (FP) may be unnecessarily treated	With a prevalence of 15%, 150/1000 individuals will have pulmonary TB. Of these, 132 (TP) will be identified; 18 (FN) will be missed. Of the 850 individuals without TB, 833 (TN) will not be treated; 17 (FP) may be unnecessarily treated	With a prevalence of 30%, 300/1000 individuals will have pulmonary TB. Of these, 264 (TP) will be identified; 36 (FN) will be missed. Of the 700 individuals without TB, 686 (TN) will not be treated; 14 (FP) may be unnecessarily treated
TB detection, Xpert used as an add-on test following a negative smear microscopy result	Pooled median sensitivity 67% (58, 74) and pooled median specificity 98% (97, 98)	5719 (14)	With a prevalence of 5%, 50/1000 individuals will have pulmonary TB. Of these, 34 (TP) will be identified; 16 (FN) will be missed. Of the 950 individuals without TB, 931 (TN) will not be treated; 19 (FP) may be	With a prevalence of 15%, 150/1000 individuals will have pulmonary TB. Of these, 101 (TP) will be identified; 49 (FN) will be missed. Of the 850 individuals without TB, 833 (TN) will not be treated; 17 (FP) will not be treated; 17 (FP)	With a prevalence of 30%, 300/1000 individuals will have pulmonary TB. Of these, 201 (TP) will be identified; 99 (FN) will be missed. Of the 700 individuals without TB, 686 (TN) will not be treated; 14 (FP)

		unnecessarily treated	may be unnecessarily treated	may be unnecessarily treated	
Smear-positive, culture-positive subgroup, within-study comparison	Pooled median sensitivity 98% (97, 99); specificity of Xpert could not be estimated in these studies	1735 (15)	With a prevalence of 5%, 50/1000 individuals will have pulmonary TB. Of these, 49 (TP) will be identified; 1 (FN) will be missed	With a prevalence of 15%, 150/1000 individuals will have pulmonary TB. Of these, 147 (TP) will be identified; 3 (FN) will be missed	With a prevalence of 30%, 300/1000 individuals will have pulmonary TB. Of these, 294 (TP) will be identified; 6 (FN) will be missed
Smear-negative, culture-positive subgroup	Pooled median sensitivity 68% (59, 75) and pooled median specificity 98% (97, 99)	5771 (15)	With a prevalence of 5%, 50/1000 individuals will have pulmonary TB. Of these, 34 (TP) will be identified; 16 (FN) will be missed. Of the 950 individuals without TB, 931 (TN) will not be treated; 19 (FP) may be unnecessarily treated	With a prevalence of 15%, 150/1000 individuals will have pulmonary TB. Of these, 102 (TP) will be identified; 48 (FN) will be missed. Of the 850 individuals without TB, 833 (TN) will not be treated; 17 (FP) may be unnecessarily treated	With a prevalence of 30%, 300/1000 individuals will have pulmonary TB. Of these, 204 (TP) will be identified; 96 (FN) will be missed. Of the 700 individuals without TB, 686 (TN) will not be treated; 14 (FP) may be unnecessarily treated
HIV-positive subgroup	Pooled median sensitivity 80% (67, 88) and pooled median specificity 97% (93, 99)	1163 (4)	With a prevalence of 5%, 50/1000 individuals will have pulmonary TB. Of these, 40 (TP) will be identified; 10 (FN) will be missed. Of the 950 individuals without TB, 922 (TN) will not be treated; 28 (FP) may be unnecessarily treated	With a prevalence of 15%, 150/1000 individuals will have pulmonary TB. Of these, 120 (TP) will be identified; 30 (FN) will be missed. Of the 850 individuals without TB, 825 (TN) will not be treated; 25 (FP) may be unnecessarily treated	With a prevalence of 30%, 300/1000 individuals will have pulmonary TB. Of these, 240 (TP) will be identified; 60 (FN) will be missed. Of the 700 individuals without TB, 679 (TN) will not be treated; 21 (FP) may be unnecessarily treated
HIV-negative subgroup	Pooled median sensitivity 89% (81, 94) and pooled median specificity 99% (96, 99)	1981 (4)	With a prevalence of 5%, 50/1000 individuals will have pulmonary TB. Of these, 45 (TP) will be identified; 5 (FN) will be missed. Of the 950 individuals without TB, 941 (TN) will not be treated; 9 (FP) may be unnecessarily treated	With a prevalence of 15%, 150/1000 individuals will have pulmonary TB. Of these, 134 (TP) will be identified; 16 (FN) will be missed. Of the 850 individuals without TB, 842 (TN) will not be treated; 8 (FP) may be unnecessarily treated	With a prevalence of 30%, 300/1000 individuals will have pulmonary TB. Of these, 267 (TP) will be identified; 33 (FN) will be missed. Of the 700 patients individuals without TB, 693 (TN) will not be treated; 7 (FP) may be unnecessarily treated
Diagnostic accuracy for detection of rifampicin resistance					

Type of analysis	Effect (95% credible interval)	No. of participants (studies)	What do these results mean given 2% prevalence of rifampicin resistance among individuals with TB?	What do these results mean given 15% prevalence of rifampicin resistance among individuals with TB?	What do these results mean given 30% prevalence of rifampicin resistance among individuals with TB?
Rifampicin resistance detection, Xpert used as an initial test replacing conventional drug susceptibility testing	Pooled median sensitivity 94% (87, 97) and pooled median specificity 98% (97, 99)	2340 (11)	With a prevalence of 2%, 20/1000 will have rifampicin resistance. Of these, 19 (TP) will be identified; 1 (FN) will be missed. Of the 980 patients with TB sensitive to rifampicin, 960 (TN) will not be treated; 20 (FP) may be unnecessarily treated	With a prevalence of 15%, 150/1000 will have rifampicin resistance. Of these, 141 (TP) will be identified; 9 (FN) will be missed. Of the 850 patients with TB sensitive to rifampicin, 833 (TN) will not be treated; 17 (FP) may be unnecessarily treated	With a prevalence of 30%, 300/1000 will have rifampicin resistance. Of these, 282 (TP) will be identified; 18 (FN) will be missed. Of the 700 patients with TB sensitive to rifampicin, 686 (TN) will not be treated; 14 (FP) may be unnecessarily treated

TP, true positive; FN, false negative; TN, true negative; FP, false positive

DISCUSSION

Early diagnosis of TB is critical for reducing TB transmission and incidence, and rapid molecular diagnostics offer great promise in this area. In 2010, the WHO endorsed the Xpert assay, a novel, rapid diagnostic system using an automated PCR test that is simple enough to be run in basic laboratories and clinics, outside of a reference laboratory setting. Until the development of Xpert, smear microscopy had been the sole method used for TB diagnosis in most laboratories in developing countries, where over 95% of TB-related deaths occur. Xpert can, generally within two hours, simultaneously detect TB and resistance to rifampicin, considered to be a good proxy for MDR-TB (WHO Policy Xpert 2011). Following endorsement, the WHO unveiled a strategic plan for Xpert implementation (WHO Rapid Implementation 2011). The plan recommended that: “1. Xpert MTB/RIF should be used as the initial diagnostic test in individuals suspected of having MDR-TB or HIV-associated TB (Strong recommendation) and 2. Xpert MTB/RIF may be considered as a follow-on test to microscopy in settings where MDR-TB or HIV is of lesser concern, especially in further testing of smear-negative specimens. (Conditional recommendation acknowledging major resource implications)” (WHO Rapid Implementation 2011).

The findings in this systematic review lend support to the WHO recommendations on the use of Xpert as an initial diagnostic test for TB detection and rifampicin resistance detection in patients suspected of having MDR-TB or HIV-associated TB. When used as an add-on test following a negative smear microscopy result, Xpert yielded a sensitivity of 67%. Although a sensitivity of 67% may be considered low, it is precisely in smear-negative patients that improvements in diagnostic tests are needed and where any improved diagnostic test will have the biggest impact. Therefore, Xpert may also be valuable as an add-on test following smear microscopy. When used as an initial test replacing conventional drug susceptibility testing, we found high sensitivity and specificity of Xpert for rifampicin resistance detection. Xpert, however, does not eliminate the need for subsequent culture and phenotypic drug susceptibility testing, which are required to monitor treatment progress and to detect resistance to drugs other than rifampicin. We found that, in comparison with processed specimens, unprocessed specimens had a significantly higher sensitivity. This may be an effect of the difference in buffer:sample ratio (3:1 for processed and 2:1 for unprocessed sputum as per package insert) and the resulting lower input volume for processed compared to unprocessed sputum.

We wish to underscore that the findings in this review should be interpreted with caution as they are based on previous versions of Xpert. There were no studies of the current version of Xpert, G4, included in this review. It is possible that the performance of Xpert G4 will be different. Also, for Xpert diagnosis to be effective, ideally, the reporting of Xpert results should occur rapidly so treatment can be started.

Xpert has now begun to be rolled out in 21 countries via UNITAID, with a price drop from \$16.86 to \$9.98 (US) per cartridge, a price that will remain in effect until 2022 (The Gates Foundation 2012; UNITAID 2012a) (UNITAID is a global health initiative working to increase access for tests and medicines for HIV/AIDS, TB, and malaria). Since Xpert was endorsed by WHO, a large number of studies have been performed and country-level policy makers are making decisions about adoption and scale-up.

This systematic review represents the most comprehensive review on the diagnostic accuracy of Xpert and provides evidence that may help countries make decisions about scaling up Xpert for programmatic management of TB and drug-resistant TB. Although, the information in this review will help to inform, other factors such as level of deployment in the health system, cost, and operational considerations, including the ability to maintain an uninterrupted and stable electrical power supply, temperature control, and maintenance of the cartridge modules, will also influence those decisions, as discussed in a recent commentary (Trebuć 2011).

This review focused on diagnostic accuracy. Sensitivity and specificity, however, depend on the performance of a test in a particular situation, defined by the population, the setting, and prior testing. In a different population or setting or with a different testing strategy, the sensitivity and specificity are likely to change (Bossuyt 2008). Our review question concerned the performance of Xpert in peripheral laboratories and health facilities. We identified only one study that used Xpert in this diagnostic setting and no studies that performed Xpert in primary health care clinics at the point of care. The remaining studies ran Xpert in reference laboratories. Hence, the findings from this review may not be applicable in resource-constrained areas that lack high-quality laboratory infrastructure and highly trained personnel. Also, since the majority of studies included in the analysis were from low- and middle-income countries where advanced, smear-positive TB is common, the findings may not directly apply to high-income countries where most TB cases are smear-negative, have minimal disease, and induced sputum samples are often necessary for diagnosis. With regard to income status, we found that the pooled sensitivity in high-income countries was higher than that for low-/middle-income countries. The presence of specialized laboratory services and highly trained staff may be one reason contributing to the higher sensitivity of Xpert in high-income versus low- and middle-income countries.

We acknowledge that patient outcomes are clearly important to patients, decision makers, and the wider TB community. Outcomes in addition to diagnostic accuracy, however, could not be systematically addressed in this review as they would have required a different methodology. Nonetheless, we looked for and summarized two ‘time to event’ outcomes (time to result and time to treatment) when data were provided by the included studies (eight studies, Table 4). Xpert results for TB detection were usually re-

ported within two hours or on the same day, compared with liquid culture results which were reported in around 16 to 20 days. Two studies reported on time to detection of rifampicin resistance and both studies found that, compared with conventional methods, Xpert greatly decreased the time to diagnosis (Boehme 2011a, Lawn 2011). However, early detection of rifampicin resistance may not lead to improved patient outcomes if the result is not linked to appropriate treatment, services, and supervision (WHO Xpert Checklists 2011). One study did provide information about time to treatment initiation; for smear-negative, culture-positive TB, the median delay in beginning treatment was 56 days (IQR, 39, 81) before Xpert was introduced, compared with five days (IQR, 2, 8) after Xpert was introduced (Boehme 2011a). Although we did not systematically review patient outcomes for the current review, data regarding delays in switching from the standard regimen for drug-susceptible TB to an appropriate regimen for MDR-TB would be also be useful because of the potential harms to patients being treated with the wrong drug regimen.

In addition to diagnostic accuracy, cost and cost-effectiveness of Xpert are major concerns for national TB programmes and policy makers. However, our systematic review was not intended to address this outcome. Several cost-effectiveness studies have been published recently (Abimbola 2012; Andrews 2012; Dowdy 2011; Meyer-Rath 2012; Schnippel 2012; Vassall 2011) and these will need to be systematically reviewed.

Summary of main results

The main results are summarized in the Summary of Results table (Summary of findings).

- When used as an initial test replacing smear microscopy, Xpert achieved modest sensitivity (88%) and high specificity (98%) for TB detection.
- When used as an add-on test following smear microscopy, Xpert yielded a sensitivity of 67%.
 - Xpert sensitivity for smear-positive, culture-positive TB was very high and consistent (98%); Xpert sensitivity for smear-negative, culture-positive TB was lower and more variable (68%).
 - Xpert detected 80% of pulmonary TB cases in people living with HIV and 89% of pulmonary TB cases in people without HIV infection.
 - When used as an initial test replacing conventional drug susceptibility testing, Xpert detected 94% of rifampicin-resistant TB with high specificity (98%).
 - The proportion of indeterminate Xpert results was very low (1.1%).

Application of the meta-analysis to a hypothetical cohort

The Summary of findings summarizes the findings of the review by applying the results to a hypothetical cohort of 1000 individuals

suspected of having pulmonary TB or MDR-TB. We present three different scenarios: for Xpert used as an initial test for TB detection or as an add-on test following microscopy, the prevalence of TB in the setting or patient subgroup varies from 5% to 15% to 30%; for Xpert for rifampicin resistance detection, the prevalence of rifampicin resistance in the setting varies from 2% to 15% to 30%. The consequence of a false positive result is that a patient may be unnecessarily treated for MDR-TB, with the possibility of developing severe adverse reactions. The consequence of a false negative test is that the patient may not receive treatment, resulting in more severe disease, death, and spread of TB.

I. A. TB detection, Xpert used as an initial test replacing smear microscopy

TB prevalence of 5%: if the point estimates for Xpert are applied to a hypothetical cohort of 1000 patients suspected of having TB, where 5% actually do have TB, then Xpert would be expected to miss six cases and falsely diagnose 19 cases.

TB prevalence of 15%: if the point estimates for Xpert are applied to a hypothetical cohort of 1000 patients suspected of having TB, where 15% actually do have TB, then Xpert would be expected to miss 18 cases and falsely diagnose 17 cases.

TB prevalence of 30%: if the point estimates for Xpert are applied to a hypothetical cohort of 1000 patients suspected of having TB, where 30% actually do have TB, then Xpert would be expected to miss 36 cases and falsely diagnose 14 cases.

I. B. TB detection, Xpert used as an add-on test following a negative smear microscopy result

TB prevalence of 5%: if the point estimates for Xpert are applied to a hypothetical cohort of 1000 patients suspected of having TB, where 50 patients actually do have TB, then Xpert would be expected to miss 16 cases and falsely diagnose 19 cases.

TB prevalence of 15%: if the point estimates for Xpert are applied to a hypothetical cohort of 1000 patients suspected of having TB, where 150 patients actually do have TB, then Xpert would be expected to miss 49 cases and falsely diagnose 17 cases.

TB prevalence of 30%: if the point estimates for Xpert are applied to a hypothetical cohort of 1000 patients suspected of having TB, where 300 patients actually do have TB, then Xpert would be expected to miss 99 cases and falsely diagnose 14 cases.

II. A. Rifampicin resistance detection, Xpert used as an initial test replacing conventional drug susceptibility testing

Rifampicin resistance prevalence of 2%: if the point estimates for Xpert are applied to a hypothetical cohort of 1000 patients suspected of having rifampicin resistant-TB, where 20 patients actually have rifampicin resistance, then Xpert would be expected to miss one case and falsely diagnose 20 cases.

Rifampicin resistance prevalence of 15%: if the point estimates for Xpert are applied to a hypothetical cohort of 1000 patients

suspected of having rifampicin resistant-TB, where 150 patients actually have rifampicin resistance, then Xpert would be expected to miss nine cases and falsely diagnose 17 cases.

Rifampicin resistance prevalence of 30%: if the point estimates for Xpert are applied to a hypothetical cohort of 1000 patients suspected of having rifampicin resistant-TB, where 300 patients actually have rifampicin resistance, then Xpert would be expected to miss 18 cases and falsely diagnose 14 cases.

Strengths and weaknesses of the review

The findings of this review are based on comprehensive searching, strict inclusion criteria, and standardized data extraction. The strength of this review is that it allows an assessment to be made of the diagnostic accuracy of Xpert for detection of TB when Xpert is used as a replacement test for smear microscopy or as an add-on test following smear microscopy. In addition, the review allows for a determination of the accuracy of Xpert for detection of rifampicin resistance when Xpert is used as an initial test replacing conventional drug susceptibility testing.

Completeness of evidence

This is a reasonably complete data set that involved comprehensive searching and correspondence with experts in the field and the test manufacturer to identify additional studies, as well as repeated correspondence with study authors to obtain additional data and information that was missing in the papers. The search strategy included studies published in all languages. However, we acknowledge that we may have missed some studies despite the comprehensive search. We are aware of several ongoing studies and papers published after our search period and will include these data in an update of the review. For the clinical subgroups, one study that was conducted in a high HIV prevalence setting only reported summary values and derivation of data for the 2 x 2 tables by HIV status was not possible. Several studies that investigated Xpert for both pulmonary and extrapulmonary TB could not be included in all analyses because we were unable to extract data pertaining to pulmonary TB alone. Lastly, the evidence in this review is mostly derived from high TB incidence countries and should be carefully extrapolated to low incidence settings.

Accuracy of the reference standards used

Culture is regarded as the best available reference standard for TB and was the reference standard used for TB in this review. Culture-based phenotypic drug susceptibility testing methods using WHO recommended critical concentrations were the reference standards for rifampicin resistance (WHO Policy DST 2008). Concerning the latter, two recent studies have raised concerns about phenotypic susceptibility testing methods for rifampicin using the recommended critical concentrations. Van Deun 2009 reported that

certain conventional drug susceptibility methods missed low-level rifampicin resistance. Furthermore, using Xpert and gene sequencing, Williamson 2012 identified four patients (three with clinical information available) whose TB isolates contained mutations to the *rpoB* gene but appeared to be rifampicin susceptible using phenotypic methods. 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 were deemed rifampicin susceptible with phenotypic methods. In light of these findings, it is unclear whether and to what extent Xpert might out-perform the conventional reference standards for drug susceptibility testing.

Quality and quality of reporting of the included studies

The majority (56%) of studies used consecutive selection of participants and > 75% of the studies interpreted Xpert results without knowledge of the results of the reference standard. Xpert results are generated automatically, without requiring subjective interpretation. Only one study performed Xpert in health facilities and provincial hospitals, settings that matched the review question. In general, studies were fairly well reported, though we corresponded with almost all authors for additional data and missing information. We encourage authors of future studies to follow the recommendations in the STARD statement to improve the quality of reporting (Bossuyt 2003).

Completeness and relevance of the review

We noted that the vast majority of studies were conducted in central reference laboratories whereas, for our review question, we were interested in how Xpert performed in peripheral laboratories and health facilities. Therefore, the findings in this review may not be applicable to peripheral settings. Studies using the first three generations of Xpert (G1, G2, and G3 cartridges) are included in this systematic review. A newer cartridge, G4, to be used with software version 4.0 or higher, has been released and is now included in all Xpert kits. Therefore, the findings in this review should be viewed with caution as they may not pertain to the newest Xpert version. An updated Cochrane review will cover studies using Xpert G4 and any subsequent versions. This systematic review did not address the use of Xpert in children. A systematic review on the diagnostic accuracy of Xpert for extrapulmonary TB is underway.

Applicability of findings to the review question

We found high specificity (98%) of Xpert for both detection of TB and rifampicin resistance. When used as an initial test for TB detection replacing smear microscopy, Xpert achieved a sensitivity

of 88%. When used as an add-on test following a negative smear microscopy result, Xpert yielded a lower sensitivity of 67%. In clinical subgroups, we found very high sensitivity (98%) of Xpert for smear-positive, culture-positive TB and lower sensitivity of 68% for smear-negative, culture-positive TB. We found modest sensitivity of Xpert (80%) in people living with HIV and higher sensitivity (89%) in people without HIV infection. An Xpert result that is positive for rifampicin resistance should be carefully interpreted and take into consideration the risk of MDR-TB in a given patient and the expected prevalence of MDR-TB. For example, in a hypothetical cohort of 1000 individuals suspected of having rifampicin resistance (a proxy for MDR-TB), where the prevalence of rifampicin resistance is 30%, we found that, on average, Xpert would wrongly identify 14 patients as being rifampicin resistant. In most of the countries represented in this review, among patients who had previously never been treated for TB, the prevalence of MDR-TB was only around 1% to 2%. In a hypothetical cohort, where the prevalence of rifampicin resistance is 2%, the number of individuals wrongly identified as rifampicin resistant would increase from 14 (when rifampicin resistance prevalence is 30%) to 20, an increase of 43%. These individuals may be unnecessarily treated with second-line anti-TB drugs and experience serious adverse events. Thus, in a setting with very low prevalence of MDR-TB, an Xpert result indicating rifampicin resistance should prompt confirmation by a more definitive test. Studies in this review assessed sensitivity and specificity of the test when used in laboratories in research studies; the accuracy of Xpert may be lower in routine practice settings.

AUTHORS' CONCLUSIONS

Implications for practice

The high sensitivity in smear-positive TB and modest sensitivity in smear-negative TB, along with the high specificity of Xpert mean that Xpert may be used as the initial diagnostic test for TB detection in individuals suspected of having TB, MDR-TB, or HIV-associated TB. Xpert may also be valuable as an add-on test following a negative smear microscopy result in patients suspected of having TB. The high sensitivity and high specificity of Xpert for

rifampicin resistance detection mean that Xpert may be used as an initial diagnostic test for rifampicin resistance detection. An Xpert result that is positive for rifampicin resistance should be carefully interpreted and take into consideration the risk of MDR-TB in a given patient and the expected prevalence of MDR-TB in a given setting. Policy makers will also need to take into account other factors relating to cost and operational concerns such as feasibility of the use of Xpert in peripheral laboratories and health centres.

Implications for research

Future studies should assess the diagnostic accuracy of Xpert in peripheral laboratories and clinical settings, such as primary health facilities with microscopy laboratories, TB screening centres, and antiretroviral clinics, especially settings where the test is performed at the point of care. Studies on Xpert in children are emerging and will need to be separately reviewed (Nicol 2011; Rachow 2012; Zar 2012). A systematic review on the diagnostic accuracy of extrapulmonary TB is underway to summarize the emerging studies (Causse 2011; Friedrich 2011; Hanif 2011; Hillemann 2011; Ioannidis 2011; Ligthelm 2011; Miller 2011; Tortoli 2012; Vadwai 2011). Future systematic reviews should also summarize the growing body of evidence on patient outcomes (clinical impact), cost, and cost effectiveness.

ACKNOWLEDGEMENTS

We wish to thank all authors of the included studies for answering our questions and providing additional data. We are grateful to Vittoria Lutje, Liverpool School of Tropical Medicine, for help with the search strategy. We also wish to thank Edward Desmond, California State Department of Health, for his comments on alternative TB tests, and Ellen Jo Baron, Cepheid, for her comments on the index test. In addition, we wish to thank Matteo Zignol, WHO Stop TB Department, for providing the prevalence rates for MDR-TB. We also wish to thank Selcan Alptekin, Rothamsted Research, for her help with translation.

The editorial base of the Cochrane Infectious Diseases Group is funded by UKaid from the UK Government for the benefit of developing countries.

REFERENCES

References to studies included in this review

- Boehme 2010a** *{published data only}*
Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, et al. Rapid molecular detection of tuberculosis and rifampin resistance. *New England Journal of Medicine* 2010;**363**:1005–15.
- Boehme 2010b** *{published data only}*
Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, et al. Rapid molecular detection of tuberculosis and rifampin resistance. *New England Journal of Medicine* 2010;**363**:1005–15.
- Boehme 2010c** *{published data only}*
Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, et al. Rapid molecular detection of tuberculosis and rifampin resistance. *New England Journal of Medicine* 2010;**363**:1005–15.
- Boehme 2010d** *{published data only}*
Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, et al. Rapid molecular detection of tuberculosis and rifampin resistance. *New England Journal of Medicine* 2010;**363**:1005–15.
- Boehme 2010e** *{published data only}*
Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, et al. Rapid molecular detection of tuberculosis and rifampin resistance. *New England Journal of Medicine* 2010;**363**:1005–15.
- Boehme 2011a** *{published data only}*
Boehme CC, Nicol MP, Nabeta P, Michael JS, Gotuzzo E, Tahirli R, et al. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *Lancet* 2011;**377**(9776): 1495–505.
- Boehme 2011b** *{published data only}*
Boehme CC, Nicol MP, Nabeta P, Michael JS, Gotuzzo E, Tahirli R, et al. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *Lancet* 2011;**377**(9776): 1495–505.
- Boehme 2011c** *{published data only}*
Boehme CC, Nicol MP, Nabeta P, Michael JS, Gotuzzo E, Tahirli R, et al. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *Lancet* 2011;**377**(9776): 1495–505.
- Boehme 2011d** *{published data only}*
Boehme CC, Nicol MP, Nabeta P, Michael JS, Gotuzzo E, Tahirli R, et al. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *Lancet* 2011;**377**(9776): 1495–505.
- Boehme 2011e** *{published data only}*
Boehme CC, Nicol MP, Nabeta P, Michael JS, Gotuzzo E, Tahirli R, et al. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *Lancet* 2011;**377**(9776): 1495–505.
- Boehme 2011f** *{published data only}*
Boehme CC, Nicol MP, Nabeta P, Michael JS, Gotuzzo E, Tahirli R, et al. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *Lancet* 2011;**377**(9776): 1495–505.
- Bowles 2011** *{published data only}*
Bowles EC, Freyee B, van Ingen J, Mulder B, Boeree M J, van Soolingen D. Xpert MTB/RIF®, a novel automated polymerase chain reaction-based tool for the diagnosis of tuberculosis. *International Journal of Tuberculosis and Lung Disease* 2011;**15**:988–9.
- Ciftci 2011** *{published data only}*
Ciftci IH, Aslan MH, Asik G. [Evaluation of Xpert MTB/RIF results for the detection of Mycobacterium tuberculosis in clinical samples]. *Mikrobiyoloji bulteni* 2011;**45**(1):43–7.
- Friedrich 2011** *{published data only}*
Friedrich SO, Venter A, Kayigire XA, Dawson R, Donald PR, Diacon AH. Suitability of Xpert MTB/RIF and genotype MTBDRplus for patient selection for a tuberculosis clinical trial. *Journal of Clinical Microbiology* 2011;**49**(8):2827–31.
- Hanif 2011** *{published data only}*
Hanif SN, Eldeen HS, Ahmad S, Mokaddas E. GeneXpert MTB/RIF for rapid detection of Mycobacterium tuberculosis in pulmonary and extra-pulmonary samples. *International Journal of Tuberculosis and Lung Disease* 2011;**15**(9):1274–5.
- Helb 2010** *{published data only}*
Helb D, Jones M, Story E, Boehme C, Wallace E, Ho K, et al. Rapid detection of Mycobacterium tuberculosis and rifampin resistance by use of on-demand, near-patient technology. *Journal of Clinical Microbiology* 2010;**48**(1): 229–37.
- Ioannidis 2011** *{published data only}*
Ioannidis P, Papaventsis D, Karabela S, Nikolaou S, Panagi M, Raftopoulou E, et al. Cepheid GeneXpert MTB/RIF assay for Mycobacterium tuberculosis detection and rifampin resistance identification in patients with substantial clinical indications of tuberculosis and smear-negative microscopy results. *Journal of Clinical Microbiology* 2011;**49**(8): 3068–70.
- Lawn 2011** *{published data only}*
Lawn SD, Brooks SV, Kranzer K, Nicol MP, Whitelaw A, Vogt M, et al. Screening for HIV-associated tuberculosis and rifampicin resistance before antiretroviral therapy using

the Xpert MTB/RIF assay: a prospective study. *PLoS Medicine* 2011;**8**(7):e1001067.

Malbruny 2011 {published data only}

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

Marlowe 2011 {published data only}

Marlowe EM, Novak-Weekley SM, Cumpio J, Sharp SE, Momeny MA, Babst A, et al. Evaluation of the Cepheid Xpert MTB/RIF assay for direct detection of *Mycobacterium tuberculosis* complex in respiratory specimens. *Journal of Clinical Microbiology* 2011;**49**(4):1621–3.

Miller 2011 {published data only}

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

Moure 2011 {published data only}

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

Rachow 2011 {published data only}

Rachow A, Zumla A, Heinrich N, Rojas-Ponce G, Mtafya B, Reither K, et al. Rapid and accurate detection of *Mycobacterium tuberculosis* in sputum samples by Cepheid Xpert MTB/RIF assay--a clinical validation study. *PLoS One* 2011;**6**(6):e20458.

Scott 2011 {published data only}

Scott LE, McCarthy K, Gous N, Nduna M, Van Rie A, Sanne I, et al. Comparison of Xpert MTB/RIF with other nucleic acid technologies for diagnosing pulmonary tuberculosis in a high HIV prevalence setting: a prospective study. *PLoS Medicine* 2011;**8**(7):e1001061.

Teo 2011 {published data only}

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

Theron 2011 {published data only}

Theron G, Peter J, van Zyl-Smit R, Mishra H, Streicher E, Murray S, et al. Evaluation of the Xpert MTB/RIF assay for the diagnosis of pulmonary tuberculosis in a high HIV prevalence setting. *American Journal of Respiratory Critical Care Medicine* 2011;**184**(1):132–40.

Zeka 2011 {published data only}

Zeka AN, Tasbakan S, Cavusoglu C. Evaluation of the GeneXpert MTB/RIF assay for the rapid diagnosis of tuberculosis and detection of RIF-resistance in pulmonary

and extrapulmonary specimens. *Journal of Clinical Microbiology* 2011;**49**(12):4138–41.

References to studies excluded from this review

Armand 2011 {published data only}

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

Blakemore 2011 {published data only}

Blakemore R, Nabeta P, Davidow A L, Vadwai V, Tahiri R, Munsamy V, et al. A multi-site assessment of the quantitative capabilities of the Xpert(R) MTB/RIF assay. *American Journal of Respiratory Critical Care Medicine* 2011;**184**(9):1076–84.

Causse 2011 {published data only}

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

Friedrich 2011a {published data only}

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

Hillemann 2011 {published data only}

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

Lawn 2011b {published data only}

Lawn SD, Nicol MP. Xpert® MTB/RIF assay: development, evaluation and implementation of a new rapid molecular diagnostic for tuberculosis and rifampicin resistance. *Future Microbiology* 2011;**6**(9):1067–82.

Ligthelm 2011 {published data only}

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

Nicol 2011 {published data only}

Nicol MP, Workman L, Isaacs W, Munro J, Black F, Eley B, et al. Accuracy of the Xpert MTB/RIF test for the diagnosis of pulmonary tuberculosis in children admitted to hospital in Cape Town, South Africa: a descriptive study. *Lancet Infectious Diseases* 2011;**11**(11):819–24.

Vadwai 2011 {published data only}

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

Van Rie 2010 {published data only}

Van Rie A, Page-Shipp L, Scott L, Sanne I, Stevens W. Xpert (®) MTB/RIF for point-of-care diagnosis of TB in high-HIV burden, resource-limited countries: hype or hope?. *Expert Review of Molecular Diagnostics* 2010;**10**(7):937–46.

References to ongoing studies**Dheda 2012** {unpublished data only}

Multicentre randomised control trial of point-of-treatment (Clinic-based) Xpert MTB/RIF assay. Ongoing study 7 July 2011.

Luetkemeyer 2012 {unpublished data only}

Evaluation of Xpert MTB/RIF assay for the rapid identification of TB and TB rifampin resistance in HIV-infected and HIV-uninfected pulmonary tuberculosis suspects. Ongoing study 24 April 2012.

Peter 2012 {unpublished data only}

A randomised control trial of sputum induction, and new and emerging technologies in a high HIV prevalence primary care setting. Ongoing study August 2009.

Additional references**Abimbola 2012**

Abimbola TO, Marston BJ, Date AA, Blandford JM, Sangrujee N, Wiktor SZ. Cost-effectiveness of tuberculosis diagnostic strategies to reduce early mortality among persons with advanced HIV infection initiating antiretroviral therapy. *Journal of Acquired Immune Deficiency Syndromes* 2012;**60**(1):e1–7.

Andrews 2012

Andrews JR, Lawn SD, Rusu C, Wood R, Noubary F, Bender MA, et al. The cost-effectiveness of routine tuberculosis screening with Xpert MTB/RIF prior to initiation of antiretroviral therapy: a model-based analysis. *AIDS* 2012;**26**(8):987–95.

Banada 2010

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

Blakemore 2010

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

Boehme 2007

Boehme CC, Nabeta P, Henostroza G, Raqib R, Rahim Z, Gerhardt M, et al. Operational feasibility of using loop-mediated isothermal amplification for diagnosis of pulmonary tuberculosis in microscopy centers of developing countries. *Journal of Clinical Microbiology* 2007;**45**(6):1936–40.

Bossuyt 2003

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. The STARD statement for reporting

studies of diagnostic accuracy: explanation and elaboration. *Clinical Chemistry* 2003;**49**(1):7–18.

Bossuyt 2008

Bossuyt PM, Loefflang MM. Chapter 6: Developing Criteria for Including Studies. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.4 [updated September 2008]*. The Cochrane Collaboration, 2008.

Buzaianu 2008

Buzaianu M, Kadane JB. Adjusting for Verification Bias in Diagnostic Test Evaluation: A Bayesian Approach. *Statistics in Medicine* 2008;**27**:2453–2473.

CDC 2009

Centers for Disease Control and Prevention (CDC). Updated guidelines for the use of nucleic acid amplification tests in the diagnosis of tuberculosis. *Morbidity and Mortality Weekly Report* 2009;**58**(1):7–10.

Cepheid 2009

Cepheid. Brochure: Xpert®MTB/RIF. Two-hour detection of MTB and resistance to rifampicin. http://www.cepheid.com/media/files/eu/brochures/XpertMTB_Broch_R9_EU.pdf. Sunnysvale, Accessed 17 June 2012.

Chaisson 2012

Chaisson RE, Nuermberger EL. Confronting multidrug-resistant tuberculosis. *New England Journal of Medicine* 2012;**366**(23):2223–4.

Chang 2012

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

Chu 2009

Chu H, Chen S, Louis TA. Random Effects Models in a Meta-Analysis of the Accuracy of Two Diagnostic Tests Without a Gold Standard. *J Am Stat Association* 2009;**104**:512–23.

Dowdy 2011

Dowdy DW, Cattamanchi A, Steingart KR, Pai M. Is scale-up worth it? Challenges in economic analysis of diagnostic tests for tuberculosis. *PLoS Medicine* 2011; Vol. 8, issue 7: e1001063.

Drobniewski 2012

Drobniewski F, Nikolayevskyy V, Balabanova Y, Bang D, Papaventsis D. Diagnosis of tuberculosis and drug resistance: what can new tools bring us?. *International Journal of Tuberculosis and Lung Disease* 2012;**16**(7):860–70.

Flores 2005

Flores LL, Pai M, Colford JM Jr, Riley LW. In-house nucleic acid amplification tests for the detection of Mycobacterium tuberculosis in sputum specimens: meta-analysis and meta-regression. *BMC Microbiology* 2005;**5**:55.

Getahun 2007

Getahun H, Harrington M, O'Brien R, Nunn P. Diagnosis of smear-negative pulmonary tuberculosis in people with HIV infection or AIDS in resource-constrained settings:

- informing urgent policy changes. *Lancet* 2007;**369**(9578): 2042–9.
- Gopinath 2010**
Gopinath K, Singh S. Non-tuberculous mycobacteria in TB-endemic countries: Are we neglecting the danger?. *PLoS Neglected Tropical Diseases* 2010;**4**(4):e615.
- Greco 2006**
Greco S, Girardi E, Navarra A, Saltini C. Current evidence on diagnostic accuracy of commercially based nucleic acid amplification tests for the diagnosis of pulmonary tuberculosis. *Thorax* 2006;**61**(9):783–90.
- Guyatt 2011**
Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology* 2011;**64**(4):383–94.
- Hadgu 2005**
Hadgu A, Dendukuri N, Hilden J. Evaluation of nucleic acid amplification tests in the absence of a perfect gold-standard test: a review of the statistical and epidemiologic issues. *Epidemiology* 2005;**16**(5):604–12.
- Harries 2004**
Harries A. How does the diagnosis of tuberculosis in persons infected with HIV differ from diagnosis in persons not infected with HIV?. In: Frieden T editor(s). *Toman's tuberculosis: case detection, treatment, and monitoring - questions and answers*. WHO/HTM/TB/2004.334. Geneva: World Health Organization, 2004:80–3.
- Hartmann 1967**
Hartmann G, Honikel KO, Knusel F, Nuesch J. The specific inhibition of the DNA-directed RNA synthesis by rifamycin. *Biochimica et Biophysica Acta* 1967;**145**(3): 843–4.
- International Standards 2009**
Tuberculosis coalition for technical assistance. *International Standards for Tuberculosis Care (ISTC)*. Second Edition. The Hague: Tuberculosis coalition for technical assistance, 2009.
- Juni 1999**
Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *Journal of the American Medical Association* 1999;**282**(11):1054–60.
- Ling 2008**
Ling DI, Zwerling AA, Pai M. GenoType MTBDR assays for the diagnosis of multidrug-resistant tuberculosis: a meta-analysis. *European Respiratory Journal* 2008;**32**(5):1165–74.
- Ling 2008a**
Ling DI, Flores LL, Riley LW, Pai M. Commercial nucleic-acid amplification tests for diagnosis of pulmonary tuberculosis in respiratory specimens: meta-analysis and meta-regression. *PLoS One* 2008;**3**(2):e1536.
- Lunn 2000**
Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* 2000;**10**:325–37.
- Macaskill 2010**
Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Y Takwoingi. Analysing and Presenting Results. In: Deeks JJ, Bossuyt PM, C Gatsonis editor(s). *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. Version 0.9.0*. The Cochrane Collaboration, 2010.
- Martin 2007**
Martin A, Portaels F, Palomino JC. Colorimetric redox-indicator methods for the rapid detection of multidrug resistance in *Mycobacterium tuberculosis*: a systematic review and meta-analysis. *Journal of Antimicrobial Chemotherapy* 2007;**59**(2):175–83.
- Martin 2008**
Martin A, Panaiotov S, Portaels F, Hoffner S, Palomino JC, Angeby K. The nitrate reductase assay for the rapid detection of isoniazid and rifampicin resistance in *Mycobacterium tuberculosis*: a systematic review and meta-analysis. *Journal of Antimicrobial Chemotherapy* 2008;**62**(1):56–64.
- Meyer-Rath 2012**
Meyer-Rath G, Schnippel K, Long L, MacLeod W, Sanne I, Stevens W, et al. The impact and cost of scaling up GeneXpert MTB/RIF in South Africa. *PLoS One* 2012; Vol. 7, issue 5:e36966.
- Miotto 2012**
Miotto P, Bigoni S, Migliori GB, Matteelli A, Cirillo DM. Early tuberculosis treatment monitoring by Xpert® MTB/RIF. *European Respiratory Journal* 2012;**39**(5):1269–71.
- Moore 2006**
Moore DA, Evans CA, Gilman RH, Caviedes L, Coronel J, Vivar A, et al. Microscopic-observation drug-susceptibility assay for the diagnosis of TB. *New England Journal of Medicine* 2006;**355**(15):1539–50.
- Nahid 2012**
Nahid P, Kim PS, Evans CA, Alland D, Barer M, Diefenbach J, et al. Clinical research and development of tuberculosis diagnostics: moving from silos to synergy. *Journal of Infectious Disease* 2012;**205**(Suppl 2):S159–68.
- Perkins 2007**
Perkins MD, Cunningham J. Facing the crisis: improving the diagnosis of tuberculosis in the HIV era. *Journal of Infectious Diseases* 2007;**196**(Suppl 1):S15–27.
- R 2008**
R Development Core Team. R: A language and environment for statistical computing. Version 2.15.1. Vienna: R Foundation for Statistical Computing, 2008.
- Rachow 2012**
Rachow A, Clowes P, Saathoff E, Mtafya B, Michael E, Ntinginya EN, et al. Increased and expedited case detection by Xpert MTB/RIF assay in childhood tuberculosis: a prospective cohort study. *Clinical Infectious Diseases* 2012; **54**(10):1388–96.
- Reitsma 2005**
Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in

- diagnostic reviews. *Journal of Clinical Epidemiology* 2005; **58**(10):982–90.
- Review Manager 5**
The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). 5.0. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2008.
- Schnippel 2012**
Schnippel K, Meyer-Rath G, Long L, MacLeod W, Sanne I, et al. Scaling up Xpert MTB/RIF technology: the costs of laboratory- vs. clinic-based roll-out in South Africa. *Tropical Medicine and International Health* 2012;**17**(9): 1142–51.
- Small 2011**
Small PM, Pai M. Tuberculosis diagnosis - time for a game change. *New England Journal of Medicine* 2011;**363**(11): 1070–1.
- SPSS 2006**
SPSS Inc. SPSS for Windows. 14.0.1.366. Chicago: SPSS Inc., 2006.
- Steingart 2006**
Steingart KR, Henry M, Ng V, Hopewell PC, Ramsay A, Cunningham J, et al. Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. *Lancet Infectious Diseases* 2006;**6**(9):570–81.
- Steingart 2006a**
Steingart KR, Ng V, Henry M, Hopewell PC, Ramsay A, Cunningham J, et al. Sputum processing methods to improve the sensitivity of smear microscopy for tuberculosis: a systematic review. *Lancet Infectious Diseases* 2006;**6**(10): 664–74.
- Telenti 1993**
Telenti A, Imboden P, Marchesi F, Lowrie D, Cole S, Colston MJ, et al. Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. *Lancet* 1993;**341** (8846):647–50.
- The Gates Foundation 2012**
The Gates Foundation. Public-Private Partnership Announces Immediate 40 Percent Cost Reduction for Rapid TB Test. <http://www.gatesfoundation.org/press-releases/Pages/public-private-partnership-40-percent-reduction-TB-test.aspx> 6 August 2012.
- Toman 2004**
Toman K. How many bacilli are present in a sputum specimen found positive by smear microscopy?. In: Frieden T editor(s). *Toman's tuberculosis: case detection, treatment, and monitoring - questions and answers*. WHO/HTM/TB/2004.334. Geneva: World Health Organization, 2004: 11–13.
- Tortoli 2012**
Tortoli E, Russo C, Piersimoni C, Mazzola E, Dal Monte P, Pascarella M, et al. Clinical validation of Xpert MTB/RIF for the diagnosis of extrapulmonary tuberculosis. *European Respiratory Journal* 2012;**40**(2):442–7.
- Trébucq 2011**
Trébucq A, Enarson DA, Chiang CY, Van Deun A, Harries AD, Boillot F, et al. Xpert® MTB/RIF for national tuberculosis programmes in low-income countries: when, where and how?. *International Journal of Tuberculosis and Lung Disease* 2011;**15**(12):1567–72.
- UNITAID 2012**
Boyle D, Pai M. *UNITAID: Tuberculosis diagnostic technology landscape*. Geneva: WHO, 2012.
- UNITAID 2012a**
UNITAID. UNITAID approves US\$30 million for innovative project to roll out ground-breaking tuberculosis test at reduced cost. <http://www.unitaid.eu/resources/news/releases/943-unitaid-approves-us-30-million-for-innovative-project-to-roll-out-ground-breaking-tuberculosis-test-at-reduced-cost> 13 June 2012.
- Van Deun 2009**
Van Deun A, Barrera L, Bastian I, Fattorini L, Hoffmann H, Kam KM, et al. *Mycobacterium tuberculosis* strains with highly discordant rifampin susceptibility test results. *Journal of Clinical Microbiology* 2009;**47**(11):3501–6.
- Vassall 2011**
Vassall A, van Kampen S, Sohn H, Michael JS, John KR, den Boon S, et al. Rapid diagnosis of tuberculosis with the Xpert MTB/RIF assay in high burden countries: a cost-effectiveness analysis. *PLoS Medicine* 2011; Vol. 8, issue 11:e1001120.
- Whiting 2005**
Whiting P, Harbord R, Kleijnen J. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Medical Research Methodology*. 2005;**5**:19.
- Whiting 2011**
Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529–36.
- WHO Drug Resistance 2008**
World Health Organization. *Anti-tuberculosis drug resistance in the world: fourth global report*. WHO/HTM/TB/2008.394. Geneva: World Health Organization, 2008.
- WHO Global Report 2011**
World Health Organization. *Global tuberculosis control: WHO report 2011*. WHO/HTM/TB/2011.16. Geneva: World Health Organization, 2011.
- WHO Guidelines Drug-resistant TB 2011**
World Health Organization. *Guidelines for the programmatic management of drug-resistant tuberculosis- 2011 update*. WHO/HTM/TB/2011.6. Geneva: World Health Organization, 2011.
- WHO Guidelines TB Treatment 2010**
World Health Organization. *Treatment of tuberculosis: guidelines WHO/HTM/TB/2009.420*. 4th Edition. Geneva: World Health Organization, 2010.

WHO M/XDR-TB 2010

World Health Organization. *Multidrug and extensively drug-resistant TB (M/XDR-TB): 2010 global report on surveillance and response*. WHO/HTM/TB/2010.3. Geneva: World Health Organization, 2010.

WHO Policy DST 2008

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

WHO Policy Framework 2010

World Health Organization. *Policy framework for implementing new tuberculosis diagnostics*. Geneva: World Health Organization, 2011.

WHO Policy Line Probe Assays 2008

World Health Organization. *Molecular line probe assays for rapid screening of patients at risk of multidrug-resistant tuberculosis (MDR-TB)*. Geneva: World Health Organization, 2008.

WHO Policy Noncommercial Culture 2011

World Health Organization. *Noncommercial culture and drug-susceptibility testing methods for screening patients at risk for multidrug-resistant tuberculosis*. Geneva: World Health Organization, 2011.

WHO Policy Smear-positive TB Case 2007

World Health Organization. Definition of a new sputum smear-positive TB case. http://www.who.int/tb/laboratory/policy_sputum_smearpositive_tb_case/en/index.html 2007.

WHO Policy Xpert 2011

World Health Organization. *Policy statement: automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF system*. WHO/HTM/TB/2011.4. Geneva: World Health Organization, 2011.

WHO Rapid Implementation 2011

World Health Organization. *Rapid implementation of the Xpert MTB/RIF diagnostic test. Technical and operational How-to. Practical considerations*. WHO/HTM/TB/2011.2. Geneva: World Health Organization, 2011.

WHO Xpert Checklists 2011

World Health Organization. *Checklist of prerequisites to country implementation of Xpert MTB/RIF and key action points at country level*. WHO/HTM/TB/2011.12. Geneva: World Health Organization, 2011.

Williamson 2012

Williamson DA, Roberts SA, Bower JE, Vaughan R, Newton S, Lowe O, et al. Clinical failures associated with rpoB mutations in phenotypically occult multidrug-resistant *Mycobacterium tuberculosis*. *International Journal of Tuberculosis and Lung Disease* 2012;**16**:216–20.

World Bank 2011

World Bank. *World Bank List of Economies* <http://siteresources.worldbank.org/DATASTATISTICS/Resources/CLASS.XL>. Washington (District of Columbia): World Bank, 2011.

Wright 2009

Wright A, Zignol M, Van Deun A, Falzon D, Gerdes SR, Feldman K, et al. Epidemiology of antituberculosis drug resistance 2002-07: an updated analysis of the Global Project on Anti-Tuberculosis Drug Resistance Surveillance. *Lancet* 2009;**373**(9678):1861–73.

Zar 2012

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

Zignol 2012

Zignol M, van Gemert W, Falzon D, Sismanidis C, Glaziou P, Floyd K, et al. Surveillance of anti-tuberculosis drug resistance in the world: an updated analysis, 2007-2010. *Bulletin of the World Health Organization* 2012;**90**(2): 111–119D.

Zumla 2012

Zumla A, Abubakar I, Raviglione M, Hoelscher M, Ditiu L, McHugh TD, et al. Drug-resistant tuberculosis—current dilemmas, unanswered questions, challenges, and priority needs. *Journal of Infectious Diseases* 2012;**205**(Suppl 2): S228–40.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

Boehme 2010a

Study characteristics			
Patient sampling	Prospective study with consecutive enrolment of participants; site in a multicentre study		
Patient characteristics and setting	<p>Presenting signs and symptoms: Persistent productive cough for ≥ 2 weeks</p> <p>Age: median 37 years; range 20 to 69 years</p> <p>Sex, female: 0%</p> <p>HIV infection: 4.7%</p> <p>History of TB: 54.6%</p> <p>Sample size: 216</p> <p>Clinical setting: Special treatment facility for prisoners, high MDR-TB setting</p> <p>Laboratory: Reference laboratory</p> <p>Country: Azerbaijan</p> <p>World Bank Income Classification: Middle-/low-income</p> <p>TB incidence rate: 110 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 22.3% and among retreatment cases = 55.8% (Source: survey in Baku, 2007)</p> <p>TB prevalence in study centre: 68.1%</p>		
Index tests	<p>Index: Xpert MTB/RIF assay</p> <p>Specimen condition: Fresh</p> <p>Specimen preparation: Unprocessed</p> <p>Parameter value for rifampicin resistance: 3.5</p>		
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB</p> <p>Reference standard for pulmonary TB: Löwenstein-Jensen culture and MGIT 960</p> <p>Target condition: Rifampicin resistance</p> <p>Reference standard for rifampicin resistance: Proportion method on Löwenstein-Jensen media</p>		
Flow and timing			
Comparative			
Notes	<p>Women were not included, but otherwise considered representative spectrum</p> <p>Data for one specimen per patient were provided by the study author</p>		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Boehme 2010a (Continued)

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

Boehme 2010a (Continued)

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

Boehme 2010b

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of participants; site in a multicentre study
Patient characteristics and setting	<p>Presenting signs and symptoms: Persistent productive cough for ≥ 2 weeks</p> <p>Age: median 31 years; range 18 to 79 years</p> <p>Sex, female: 43.3%</p> <p>HIV infection: 1.7%</p> <p>History of TB: 23.7%</p> <p>Sample size: 310</p> <p>Clinical setting: Primary health care DOTS (directly observed treatment, short-course) centres in shanty towns</p> <p>Laboratory: Reference laboratory</p> <p>Country: Peru</p> <p>World Bank Income Classification: Middle-/low-income</p> <p>TB incidence rate: 106 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 5.3% and among retreatment cases = 23.6% (Source: Nationwide survey, 2006)</p> <p>TB prevalence in study centre: 67.4%</p>
Index tests	<p>Index: Xpert MTB/RIF assay</p> <p>Specimen condition: Fresh</p> <p>Specimen preparation: Unprocessed</p> <p>Parameter value for rifampicin resistance: 3.5</p>
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB</p> <p>Reference standard for pulmonary TB: Löwenstein-Jensen culture and MGIT 960</p> <p>Target condition: Rifampicin resistance</p> <p>Reference standard for rifampicin resistance: Proportion method on Löwenstein-Jensen media</p>
Flow and timing	
Comparative	
Notes	Data for one specimen per patient were provided by the study author
Methodological quality	

Boehme 2010b (Continued)

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

Boehme 2010b (Continued)

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

Boehme 2010c

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of participants; site in a multicentre study
Patient characteristics and setting	<p>Presenting signs and symptoms: Persistent productive cough for ≥ 2 weeks</p> <p>Age: median 36 years; range 18 to 80 years</p> <p>Sex, female: 34.1%</p> <p>HIV infection: 76.1%</p> <p>History of TB: 43.0%</p> <p>Sample size: 332</p> <p>Clinical setting: Clinic, high HIV setting</p> <p>Laboratory: Reference laboratory</p> <p>Country: South Africa, Cape Town</p> <p>World Bank Income Classification: Middle-/low-income</p> <p>TB incidence rate: 981 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 0.9% and among retreatment cases = 4.0% (Source: Survey in Western Cape Province, 2002)</p> <p>TB prevalence in study centre: 44.0%</p>
Index tests	<p>Index: Xpert MTB/RIF assay</p> <p>Specimen condition: Fresh</p> <p>Specimen preparation: Unprocessed</p> <p>Parameter value for rifampicin resistance: 3.5</p>
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB</p> <p>Reference standard for pulmonary TB: Löwenstein-Jensen culture and MGIT 960</p> <p>Target condition: Rifampicin resistance</p> <p>Reference standard for rifampicin resistance: MGIT 960</p>
Flow and timing	
Comparative	
Notes	Data for one specimen per patient were provided by the study author

Boehme 2010c (Continued)

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

Boehme 2010c (Continued)

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

Boehme 2010d

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of participants; site in a multicentre study
Patient characteristics and setting	<p>Presenting signs and symptoms: Persistent productive cough for ≥ 2 weeks</p> <p>Age: median 32 years; range 18 to 68 years</p> <p>Sex, female: 59.4%</p> <p>HIV infection: 71.4%</p> <p>History of TB: 45.1%</p> <p>Sample size: 261</p> <p>Clinical setting: TB clinics, high HIV setting</p> <p>Laboratory: Reference laboratory</p> <p>Country: South Africa, Durban</p> <p>World Bank Income Classification: Middle-/low-income</p> <p>TB incidence rate: 981 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 1.7% and among retreatment cases = 7.7% (Source: Survey in Kwazulu-Natal Province, 2002)</p> <p>TB prevalence in study centre: 16.5%</p>
Index tests	<p>Index: Xpert MTB/RIF assay</p> <p>Specimen condition: Fresh</p> <p>Specimen preparation: Unprocessed</p> <p>Parameter value for rifampicin resistance: 3.5</p>
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB</p> <p>Reference standard for pulmonary TB: Middlebrook 7H11 culture and MGIT 960</p> <p>Target condition: Rifampicin resistance</p> <p>Reference standard for rifampicin resistance: Proportion method on Löwenstein-Jensen media</p>
Flow and timing	
Comparative	
Notes	Data for one specimen per patient were provided by the study author

Boehme 2010d (Continued)

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

Boehme 2010d (Continued)

Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		

Boehme 2010e

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of participants; site in a multicentre study
Patient characteristics and setting	<p>Presenting signs and symptoms: Persistent productive cough for ≥ 2 weeks</p> <p>Age: median 30 years; range 17 to 88 years</p> <p>Sex, female: 39.1%</p> <p>HIV infection: 4.4%</p> <p>History of TB: 75.2%</p> <p>Sample size: 222</p> <p>Clinical setting: Tertiary hospital, high MDR-TB setting</p> <p>Laboratory: Reference laboratory</p> <p>Country: India</p> <p>World Bank Income Classification: Middle-/low-income</p> <p>TB incidence rate: 185 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 2.1% and among retreatment cases = 11.9% (Source: Survey in Andhra Pradesh, 2009)</p> <p>TB prevalence in study centre: 84.2%</p>
Index tests	<p>Index: Xpert MTB/RIF assay</p> <p>Specimen condition: Fresh</p> <p>Specimen preparation: Unprocessed</p> <p>Parameter value for rifampicin resistance: 3.5</p>
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB</p> <p>Reference standard for pulmonary TB: Löwenstein-Jensen culture and MGIT 960</p> <p>Target condition: Rifampicin resistance</p> <p>Reference standard for rifampicin resistance: MGIT 960</p>
Flow and timing	
Comparative	
Notes	Data for one specimen per patient were provided by the study author

Boehme 2010e (Continued)

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

Boehme 2010e (Continued)

Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		

Boehme 2011a

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of participants; site in a multicentre study
Patient characteristics and setting	<p>Presenting signs and symptoms: Cough lasting at least 2 weeks</p> <p>Age: median 36 years; interquartile range 30 to 44 years</p> <p>Sex, female: < 1%</p> <p>HIV infection: < 1%</p> <p>History of TB: Not stated</p> <p>Sample size: 536 for detection of MTB; 211 for detection of rifampicin resistance</p> <p>Clinical setting: MDR-TB screening facility</p> <p>Laboratory: Microscopy area of MDR-TB screening facility</p> <p>Country: Azerbaijan</p> <p>World Bank Income Classification: Middle-/low-income</p> <p>TB incidence rate: 110 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 22.3% for new cases and among retreatment cases = 55.8% (Source: survey in Baku, 2007)</p> <p>TB prevalence in study centre: 42.7%</p>
Index tests	<p>Index: Xpert MTB/RIF assay</p> <p>Specimen condition: Fresh</p> <p>Specimen preparation: Unprocessed</p> <p>Parameter value for rifampicin resistance: 5</p>
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB</p> <p>Reference standard for pulmonary TB: MGIT 960</p> <p>Target condition: Rifampicin resistance</p> <p>Reference standard for rifampicin resistance: MGIT 960</p>
Flow and timing	
Comparative	

Boehme 2011a (Continued)

Notes	Follow-up reported for all sites combined: 24/153 patients with culture-negative, clinically diagnosed TB had positive results on MTB/RIF testing. 20/24 patients had follow-up, and all 20 improved on TB treatment		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results for TB detection interpreted without knowledge of the results of the index test?	Yes		
Were the reference standard results for rifampicin resistance detection interpreted without knowledge of the results of the index test?	Yes		

Boehme 2011a (Continued)

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

Boehme 2011b

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of participants; site in a multicentre study
Patient characteristics and setting	<p>Presenting signs and symptoms: Cough lasting at least 2 weeks Age: median 37 years; interquartile range 26 to 53 years Sex, female: 49% HIV infection: < 1% History of TB: Not stated Sample size: 1005 for detection of TB; 185 for detection of rifampicin resistance Clinical setting: Two health centres and one district hospital Laboratory: Microscopy area of health centres and district hospital Country: Peru World Bank Income Classification: Middle-/low-income TB incidence rate: 106 per 100,000 MDR-TB prevalence: Percent MDR-TB among new TB cases = 5.3% and among retreatment cases = 23.6% (Source: Nationwide survey, 2006) TB prevalence in study centre: 17.6%</p>
Index tests	<p>Index: Xpert MTB/RIF assay Specimen condition: Fresh Specimen preparation: Unprocessed Parameter value for rifampicin resistance: 5</p>
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB Reference standard for pulmonary TB: MGIT 960 Target condition: Rifampicin resistance Reference standard for rifampicin resistance: MGIT 960</p>
Flow and timing	

Boehme 2011b (Continued)

Comparative			
Notes	Follow-up was reported for all sites combined, see Boehme 2011a		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results for TB detection interpreted without knowledge of the results of the index test?	Yes		
Were the reference standard results for rifampicin resistance detection interpreted without knowledge of the results of the index test?	Yes		

Boehme 2011b (Continued)

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

Boehme 2011c

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of participants; site in a multicentre study
Patient characteristics and setting	<p>Presenting signs and symptoms: Cough lasting at least 2 weeks</p> <p>Age: median 36 years; interquartile range 29 to 46 years</p> <p>Sex, female: 49%</p> <p>HIV infection: 38%</p> <p>History of TB: Not stated</p> <p>Sample size: 904 for detection of TB; 188 for detection of rifampicin resistance</p> <p>Clinical setting: One health centre and one provincial hospital</p> <p>Laboratory: Microscopy area of health centre and provincial hospital</p> <p>Country: South Africa, Cape Town</p> <p>World Bank Income Classification: Middle-/low-income</p> <p>TB incidence rate: 981 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 0.9% and among retreatment cases = 4.0% (Source: Survey in Western Cape Province, 2002)</p> <p>TB prevalence in study centre: 25.8%</p>
Index tests	<p>Index: Xpert MTB/RIF assay</p> <p>Specimen condition: Fresh</p> <p>Specimen preparation: Unprocessed</p> <p>Parameter value for rifampicin resistance: 5</p>
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB</p> <p>Reference standard for pulmonary TB: MGIT 960</p> <p>Target condition: Rifampicin resistance</p> <p>Reference standard for rifampicin resistance: MGIT 960 and MTBDRplus</p>
Flow and timing	

Boehme 2011c (Continued)

Comparative			
Notes	Follow-up was reported for all sites combined, see Boehme 2011a MTBDRplus was done on culture isolates for smear-negative sputum		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results for TB detection interpreted without knowledge of the results of the index test?	Yes		
Were the reference standard results for rifampicin resistance detection interpreted without knowledge of the results of the	Yes		

Boehme 2011c (Continued)

index test?				
				Low
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Did all patients receive the same reference standard?	No			
Were all patients included in the analysis?	Yes			

Boehme 2011d

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of participants; site in a multicentre study
Patient characteristics and setting	<p>Presenting signs and symptoms: Cough lasting at least 2 weeks</p> <p>Age: median 32 years; interquartile range 26 to 38 years</p> <p>Sex, female: < 46%</p> <p>HIV infection: < 68%</p> <p>History of TB: Not stated</p> <p>Sample size: 289 for detection of TB; 116 for detection of rifampicin resistance</p> <p>Clinical setting: Emergency unit of referral hospital</p> <p>Laboratory: Microscopy area of referral hospital</p> <p>Country: Uganda</p> <p>World Bank Income Classification: Middle-/low-income</p> <p>TB incidence rate: 226 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 1.1% and among retreatment cases = 11.7% (Source: Survey in Kampala, 2008)</p> <p>TB prevalence in study centre: 50.2%</p>
Index tests	<p>Index: Xpert MTB/RIF assay</p> <p>Specimen condition: Fresh</p> <p>Specimen preparation: Unprocessed</p> <p>Parameter value for rifampicin resistance: 5</p>
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB</p> <p>Reference standard for pulmonary TB: Löwenstein-Jensen culture and MGIT 960</p> <p>Target condition: Rifampicin resistance</p> <p>Reference standard for rifampicin resistance: Proportion method on Löwenstein-Jensen media and line probe assay</p>

Boehme 2011d (Continued)

Flow and timing			
Comparative			
Notes	Follow-up was reported for all sites combined, see Boehme 2011a Line-probe assay and, for 10% of culture positive patients (every tenth patient), Löwenstein-Jensen proportion was performed on MGIT isolates (except when only positive on Löwenstein-Jensen)		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results for TB detection interpreted without knowledge of the results of the index test?	Yes		

Boehme 2011d (Continued)

Were the reference standard results for rifampicin resistance detection interpreted without knowledge of the results of the index test?	Yes			
				Low
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Did all patients receive the same reference standard?	No			
Were all patients included in the analysis?	Yes			

Boehme 2011e

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of participants; site in a multicentre study
Patient characteristics and setting	<p>Presenting signs and symptoms: Cough lasting at least 2 weeks Age: median 45 years; interquartile range 32 to 58 years Sex, female: 30% HIV infection: 4% History of TB: Not stated Sample size: 788 for detection of TB; 103 for detection of rifampicin resistance Clinical setting: Health centre Laboratory: Microscopy area of health centre Country: India World Bank Income Classification: Middle-/low-income TB incidence rate: 185 per 100,000 MDR-TB prevalence: Percent MDR-TB among new TB cases = 2.1% and among retreatment cases = 11.9% (Source: Survey in Andhra Pradesh, 2009) TB prevalence in study centre: 12.8%</p>
Index tests	<p>Xpert MTB/RIF assay Specimen condition: Fresh Specimen preparation: Unprocessed Parameter value for rifampicin resistance: 5</p>

Boehme 2011e (Continued)

Target condition and reference standard(s)	Target condition: Pulmonary TB Reference standard for pulmonary TB: Löwenstein-Jensen culture Target condition: Rifampicin resistance Reference standard for rifampicin resistance: Proportion method on Löwenstein-Jensen media		
Flow and timing			
Comparative			
Notes	Follow-up was reported for all sites combined, see Boehme 2011a		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results for TB detection interpreted without knowledge of	Yes		

Boehme 2011e (Continued)

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

Boehme 2011f

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of participants; site in a multicentre study
Patient characteristics and setting	<p>Presenting signs and symptoms: Cough lasting at least 2 weeks Age: median 47 years; interquartile range 34 to 58 years Sex, female: 36% HIV infection: < 1% History of TB: Not stated Sample size: 387 for detection of TB; 257 for detection of rifampicin resistance Clinical setting: MDR-TB screening facility Laboratory: Microscopy area of MDR-TB screening facility Country: Philippines World Bank Income Classification: Middle-/low-income TB incidence rate: 275 per 100,000 MDR-TB prevalence: Percent MDR-TB among new TB cases = 4.0% and among retreatment cases = 20.9% (Source: Nationwide survey, 2004) TB prevalence in study centre: 38.2%</p>
Index tests	<p>Index: Xpert MTB/RIF assay Specimen condition: Fresh Specimen preparation: Unprocessed Parameter value for rifampicin resistance: 5</p>

Boehme 2011f (Continued)

Target condition and reference standard(s)	Target condition: Pulmonary TB Reference standard for pulmonary TB: Ogawa culture and MGIT 960 Target condition: Rifampicin resistance Reference standard for rifampicin resistance: Proportion method on Löwenstein-Jensen media		
Flow and timing			
Comparative			
Notes	Follow-up was reported for all sites combined, see Boehme 2011a		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			Low
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results for TB detection interpreted without knowledge of	Yes		

Boehme 2011f (Continued)

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

Bowles 2011

Study characteristics	
Patient sampling	Prospective and retrospective study with enrolment of participants by convenience
Patient characteristics and setting	<p>Presenting signs and symptoms: Not reported</p> <p>Age: Not stated</p> <p>Sex, female: Not stated</p> <p>HIV infection: Not stated</p> <p>History of TB: Not stated</p> <p>Sample size: 89</p> <p>Clinical setting: Laboratory-based evaluation of respiratory specimens (predominantly sputum specimens) from a TB reference clinic</p> <p>Laboratory: Reference laboratory</p> <p>Country: Netherlands</p> <p>World Bank Income Classification: High-income</p> <p>TB incidence rate: 7.3 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 1.3% and among retreatment cases = 3.4% (Source: Nationwide surveillance, 2010)</p> <p>TB prevalence in study: 71.9%</p>
Index tests	<p>Index: Xpert MTB/RIF assay</p> <p>Specimen Condition: 26 fresh and 63 frozen (previously stored) samples</p> <p>Specimen Preparation: Unprocessed</p>

Bowles 2011 (Continued)

	Parameter value for rifampicin resistance: 3.5		
Target condition and reference standard(s)	Target condition: Pulmonary TB Reference standard for pulmonary TB: MGIT 960 Target condition: Rifampicin resistance Reference standard for rifampicin resistance: MGIT 960		
Flow and timing			
Comparative			
Notes	Sample included 2 extrapulmonary specimens (1 pleural fluid and 1 gastric aspirate) One patient whose sample was smear and culture-negative was culture-positive on a sample 11 days later		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Bowles 2011 (Continued)

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

Ciftci 2011

Study characteristics

Patient sampling	Prospective study; the sampling method was unclear
Patient characteristics and setting	<p>Presenting signs and symptoms: Symptoms suggestive of TB</p> <p>Age: Not stated</p> <p>Sex, female: Not stated</p> <p>HIV infection: Not stated</p> <p>History of TB: Not stated</p> <p>Sample size: 85</p> <p>Clinical setting: Laboratory-based evaluation of respiratory specimens (predominantly sputum) at a university hospital</p> <p>Laboratory: Reference laboratory</p> <p>Country: Turkey</p> <p>World Bank Income Classification: Middle/low income</p> <p>TB incidence rate: 28 per 100,000</p> <p>MDR-TB prevalence: Data not available</p> <p>TB prevalence in study: 29.4%</p>

Ciftci 2011 (Continued)

Index tests	Index: Xpert MTB/RIF assay Specimen Condition: Fresh Specimen Preparation: Unprocessed Parameter value for rifampicin resistance: 3.5		
Target condition and reference standard(s)	Target condition: Pulmonary TB Reference standard for pulmonary TB: BACTEC 460 Target condition: Rifampicin resistance Reference standard for rifampicin resistance: BACTEC 460		
Flow and timing			
Comparative			
Notes	Paper was written in Turkish: sample included 10 extrapulmonary specimens (5 pleural fluid and 5 urine samples); no patients were found to have rifampicin resistance		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
			Low
DOMAIN 3: Reference Standard			

Ciftci 2011 (Continued)

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

Friedrich 2011

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of participants
Patient characteristics and setting	<p>Presenting signs and symptoms: Patients recently diagnosed with smear-positive first time TB, untreated</p> <p>Age: Eligible aged 18 to 65 years</p> <p>Sex, female: Not stated</p> <p>HIV infection: Not stated</p> <p>History of TB: Not stated</p> <p>Sample size: 126</p> <p>Clinical setting: Two medical centres</p> <p>Laboratory: Reference laboratory</p> <p>Country: South Africa, Cape Town</p> <p>World Bank Income Classification: Middle-/low-income</p>

Friedrich 2011 (Continued)

	<p>TB incidence rate: 981 per 100,000 MDR-TB prevalence: Percent MDR-TB among new TB cases = 0.9% and among retreatment cases = 4.0% (Source: Survey in Western Cape Province, 2002) TB prevalence in study: 100.0%</p>		
Index tests	<p>Index: Xpert MTB/RIF assay Specimen Condition: Fresh Specimen Preparation: Processed Parameter value for rifampicin resistance: 5</p>		
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB Reference standard for pulmonary TB: MGIT 960 Target condition: Rifampicin resistance Reference standard for rifampicin resistance: MGIT 960</p>		
Flow and timing			
Comparative			
Notes	<p>The aim of this study was to assess NAATs for selecting patients for clinical trials of anti-TB medication. Patients with severe comorbidities were excluded. This study was used only for determination of sensitivity because all enrolled patients were predetermined to have TB disease</p>		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Unclear		
Did the study avoid inappropriate exclusions?	No		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		

Friedrich 2011 (Continued)

					Low
DOMAIN 3: Reference Standard					
Is the reference standards likely to correctly classify the target condition?	Yes				
Were the reference standard results for TB detection interpreted without knowledge of the results of the index test?	Yes				
Were the reference standard results for rifampicin resistance detection interpreted without knowledge of the results of the index test?	Yes				
					Low
DOMAIN 4: Flow and Timing					
Was there an appropriate interval between index test and reference standard?	Yes				
Did all patients receive the same reference standard?	Yes				
Were all patients included in the analysis?	Yes				

Hanif 2011

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of participants
Patient characteristics and setting	<p>Presenting signs and symptoms: Suspected TB based on presence of cough and radiographic findings</p> <p>Age: range 20 to 57 years old</p> <p>Sex, female: Not stated</p> <p>HIV infection: Not stated</p> <p>History of TB: Not stated</p> <p>Sample size: 206</p> <p>Clinical setting: Laboratory-based evaluation of respiratory specimens (predominantly sputum) at</p>

Hanif 2011 (Continued)

	<p>a university hospital Laboratory: Reference laboratory Country: Kuwait World Bank Income Classification: High-income TB incidence rate: 41 per 100,000 MDR-TB prevalence: Percent MDR-TB among new TB cases = 1.1% and among retreatment cases = 0% (Source: Nationwide surveillance, 2010) TB prevalence in study: 29.1%</p>		
Index tests	<p>Index: Xpert MTB/RIF assay Specimen Condition: Fresh Specimen Preparation: Unprocessed Parameter value for rifampicin resistance: 3.5</p>		
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB Reference standard for pulmonary TB: Löwenstein-Jensen culture and MGIT 960 Target condition: Rifampicin resistance Reference standard for rifampicin resistance: BACTEC 460</p>		
Flow and timing			
Comparative			
Notes	No patients were found to have rifampicin resistance		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Hanif 2011 (Continued)

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

Helb 2010

Study characteristics	
Patient sampling	Retrospective study with consecutive enrolment of participants
Patient characteristics and setting	<p>Presenting signs and symptoms: Cough lasting at least 2 weeks</p> <p>Age: median 34 years; range 18 to 76 years</p> <p>Sex, female: 30.8%</p> <p>HIV infection: 0.9%</p> <p>History of TB: 1.9%</p>

	<p>Sample size: 107 Clinical setting: TB hospital Laboratory: Reference laboratory Country: Vietnam World Bank Income Classification: Middle/low income TB incidence rate: 180 per 100,000 MDR-TB prevalence: Percent MDR-TB among new TB cases = 2.7% and among retreatment cases = 19.3% (Source: Nationwide survey, 2006) TB prevalence in study: 76.6%</p>		
Index tests	<p>Index: Xpert MTB/RIF assay Specimen Condition: Frozen Specimen Preparation: Processed Parameter value for rifampicin resistance: 3.5</p>		
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB Reference standard: Löwenstein-Jensen culture and MGIT 960</p>		
Flow and timing			
Comparative			
Notes	Rifampicin resistance data were not reported		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Helb 2010 (Continued)

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

Ioannidis 2011

Study characteristics	
Patient sampling	Prospective and retrospective study with enrolment of participants by convenience
Patient characteristics and setting	<p>Presenting signs and symptoms: High suspicion of TB in patients found to be predominantly smear negative by microscopy examination</p> <p>Age: Not stated</p> <p>Sex, female: Not stated</p> <p>HIV infection: Not stated</p>

Ioannidis 2011 (Continued)

	<p>History of TB: Not stated Sample size: 66 Clinical setting: Laboratory-based evaluation in routine hospital setting Laboratory: Reference laboratory Country: Greece World Bank Income Classification: High-income TB incidence rate: 4.6 per 100,000 MDR-TB prevalence: Percent MDR-TB among new TB cases = 0.9% and among retreatment cases = 6.7% (Source: Nationwide surveillance, 2010) TB prevalence in study: 48.0%</p>		
Index tests	<p>Index: Xpert MTB/RIF assay Condition: Fresh Preparation: Processed Parameter value for rifampicin resistance: 5</p>		
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB Reference standard for pulmonary TB: Löwenstein-Jensen culture and MGIT 960 Target condition: Rifampicin resistance Reference standard for rifampicin resistance: Proportion method on Löwenstein-Jensen media</p>		
Flow and timing			
Comparative			
Notes	Specimens were predominantly smear-negative		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	No		
Did the study avoid inappropriate exclusions?	No		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference stan-	Yes		

Ioannidis 2011 (Continued)

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

Lawn 2011

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of participants
Patient characteristics and setting	Presenting signs and symptoms: HIV-infected patients with advanced immunodeficiency; the majority of patients had one or more of the following TB symptoms: current cough, fever, night sweats, or weight loss

	<p>Age: median 34 years; interquartile range 28 to 41 years Sex, female: 65.4% HIV infection: 100% History of TB: 26.5% Sample size: 394 Clinical setting: HIV anti-retroviral clinic Laboratory: Reference laboratory Country: South Africa, Cape Town World Bank Income Classification: Middle-/low-income TB incidence rate: 981 per 100,000 MDR-TB prevalence: Percent MDR-TB among new TB cases = 0.9% and among retreatment cases = 4.0% (Source: Survey in Western Cape Province, 2002) TB prevalence in study: 18.3%</p>		
Index tests	<p>Index: Xpert MTB/RIF assay Specimen Condition: Fresh Specimen Preparation: Processed Parameter value for rifampicin resistance: 5</p>		
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB Reference standard for pulmonary TB: MGIT 960 Target condition: Rifampicin resistance Reference standard for rifampicin resistance: MGIT 960</p>		
Flow and timing			
Comparative			
Notes	<p>This study evaluated the use of Xpert to screen HIV-infected patients with advanced immunodeficiency enrolling in antiretroviral therapy services regardless of symptoms, although the majority of patients in the study had TB symptoms. Of three patients with apparent false-positive Xpert results, on follow-up, two patients had overt pulmonary and systemic symptoms suggestive of TB and improved on anti-TB treatment. The third patient was lost to follow-up Median CD4 cell count, 171 cells/ml; interquartile range 102-236</p>		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		

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

Malbruny 2011

Study characteristics			
Patient sampling	Prospective and retrospective study with consecutive enrolment of participants		
Patient characteristics and setting	<p>Presenting signs and symptoms: Clinical symptoms suggestive of TB Age: median 52 years Sex, female: 40.2% HIV infection: Not stated History of TB: Not stated Sample size: 58 Clinical setting: Laboratory-based evaluation of respiratory specimens at a university hospital Laboratory: Reference laboratory Country: France World Bank Income Classification: High-income TB incidence rate: 9.3 per 100,000 MDR-TB prevalence: Percent MDR-TB among new TB cases = 1.0% and among retreatment cases = 13.2% (Source: Nationwide surveillance, 2009) TB prevalence in study: 20.7%</p>		
Index tests	<p>Index: Xpert MTB/RIF assay Specimen Condition: Fresh and frozen Specimen Preparation: Processed Parameter value for rifampicin resistance: 5</p>		
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB Reference standard for pulmonary TB: Solid culture, type unspecified, and MGIT 960 Target condition: Rifampicin resistance Reference standard for rifampicin resistance: MGIT 960</p>		
Flow and timing			
Comparative			
Notes	31/58 (53.4%) of samples were bronchial aspirates One rifampicin resistant isolate was identified		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		

Malbruny 2011 (Continued)

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

Marlowe 2011

Study characteristics			
Patient sampling	Prospective and retrospective study with selection of specimens by convenience at two sites and consecutive selection of smear-positive specimens at one site		
Patient characteristics and setting	<p>Presenting signs and symptoms: Not reported</p> <p>Age: Not stated</p> <p>Sex, female: Not stated</p> <p>HIV infection: Not stated</p> <p>History of TB: Not stated</p> <p>Sample size: 216</p> <p>Clinical setting: Laboratory-based evaluation of respiratory samples</p> <p>Laboratory: Three different reference laboratories</p> <p>Country: USA</p> <p>World Bank Income Classification: High income</p> <p>TB incidence rate: 4.1 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 1.1% and among retreatment cases = 4.4% (Source: Nationwide surveillance, 2010)</p> <p>TB prevalence in study: 60.2%</p>		
Index tests	<p>Index: Xpert MTB/RIF assay</p> <p>Condition: Fresh and frozen</p> <p>Preparation: Processed</p> <p>Parameter value for rifampicin resistance: 5</p>		
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB</p> <p>Reference standard for pulmonary TB: Löwenstein-Jensen culture, Middlebrook 7H11 culture, and MGIT 960</p> <p>Target condition: Rifampicin resistance</p> <p>Reference standard for rifampicin resistance: phenotypic drug susceptibility testing with agar-based solid media and MGIT 960</p>		
Flow and timing			
Comparative			
Notes	<p>Unit of analysis was specimen</p> <p>Different reference standards were used at each of the three sites</p>		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Marlowe 2011 (Continued)

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

Marlowe 2011 (Continued)

Were all patients included in the analysis?	Yes		

Miller 2011

Study characteristics			
Patient sampling	Retrospective study; with enrolment of participants by convenience		
Patient characteristics and setting	<p>Presenting signs and symptoms: Not reported</p> <p>Age: Data provided for patients with pulmonary and extrapulmonary combined; 95% of patients were 15 years and older</p> <p>Sex, female: Not stated</p> <p>HIV infection: Not stated</p> <p>History of TB: Not stated</p> <p>Sample size: 89 pulmonary specimens (study also included 23 extrapulmonary specimens)</p> <p>Clinical setting: Laboratory-based evaluation of clinical specimens at a university hospital</p> <p>Laboratory: Reference laboratory</p> <p>Country: USA</p> <p>World Bank Income Classification: High-income</p> <p>TB incidence rate: 4.1 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 1.1% and among retreatment cases = 4.4% (Source: Nationwide surveillance, 2010)</p> <p>TB prevalence in study: 32.6%</p>		
Index tests	<p>Index: Xpert MTB/RIF assay</p> <p>Condition: Frozen</p> <p>Preparation: Processed</p> <p>Parameter value for rifampicin resistance: 5</p>		
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB</p> <p>Reference standard for pulmonary TB: Löwenstein-Jensen culture and MGIT 960</p> <p>Target condition: Rifampicin resistance</p> <p>Reference standard for rifampicin resistance: MGIT 960</p>		
Flow and timing			
Comparative			
Notes	Of specimens tested, four were positive by Xpert for rifampicin resistance; three were positive by the reference standard		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns

Miller 2011 (Continued)

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

Miller 2011 (Continued)

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

Moire 2011

Study characteristics	
Patient sampling	Retrospective study with enrolment of participants by convenience
Patient characteristics and setting	<p>Presenting signs and symptoms: Patients found to be smear negative by microscopy examination</p> <p>Age: All patients were 15 years of age or older</p> <p>Sex, female: Not stated</p> <p>HIV infection: Not stated</p> <p>History of TB: Not stated</p> <p>Sample size: 107</p> <p>Clinical setting: Laboratory-based evaluation of clinical specimens at a university hospital</p> <p>Laboratory: Reference laboratory</p> <p>Country: Spain</p> <p>World Bank Income Classification: High income</p> <p>TB incidence rate: 16 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 0.2% and among retreatment cases = 1.5% (Source: Survey in Galicia region, 2005)</p> <p>TB prevalence in study: 72.9%</p>
Index tests	<p>Index: Xpert MTB/RIF assay</p> <p>Specimen Condition: Frozen</p> <p>Specimen Preparation: Processed</p> <p>Parameter value for rifampicin resistance: 3.5</p>
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB</p> <p>Reference standard for pulmonary TB: Lowenstein-Jensen culture and MGIT 960</p> <p>Target condition: Rifampicin resistance</p> <p>Reference standard for rifampicin resistance: Bactec 460</p>
Flow and timing	
Comparative	
Notes	<p>Sample set Include 1 pulmonary biopsy specimen</p> <p>Of 85 pulmonary and extrapulmonary specimens tested, 6 were positive by Xpert for rifampicin resistance: 7 specimens were positive by the reference standard</p>

Moire 2011 (Continued)

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

Moire 2011 (Continued)

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

Rachow 2011

Study characteristics	
Patient sampling	Retrospective study with consecutive enrolment of participants
Patient characteristics and setting	<p>Presenting signs and symptoms: Suspected pulmonary TB based on clinical and radiographic findings</p> <p>Age: mean 39 years (SD = 13.8)</p> <p>Sex, female: 51.7%</p> <p>HIV infection: 58.9%</p> <p>History of TB: Not stated</p> <p>Sample size: 172</p> <p>Clinical setting: Referral hospital, high HIV setting</p> <p>Laboratory: Reference laboratory</p> <p>Country: United Republic of Tanzania</p> <p>World Bank Income Classification: Middle-/low-income</p> <p>TB incidence rate: 177 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 1.1% and among retreatment cases = 0% (Source: Nationwide survey, 2007)</p> <p>TB prevalence in study: 40.1%</p>
Index tests	<p>Index: Xpert MTB/RIF assay</p> <p>Specimen Condition: Frozen</p> <p>Specimen Preparation: Processed</p> <p>Parameter value for rifampicin resistance: 3.5</p>
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB</p> <p>Reference standard for pulmonary TB: Löwenstein-Jensen culture and MGIT 960</p> <p>Target condition: Rifampicin resistance</p> <p>Reference standard for rifampicin resistance: MGIT 960</p>
Flow and timing	
Comparative	

Rachow 2011 (Continued)

Notes	Patients were followed for a period of 56 days. Among 77 patients classified as smear negative, culture negative 'clinical TB', Xpert was positive in seven (9.1%) patients No patients were found to have rifampicin resistance		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results for TB detection interpreted without knowledge of the results of the index test?	Yes		
Were the reference standard results for rifampicin resistance detection interpreted without knowledge of the results of the index test?	Yes		

Rachow 2011 (Continued)

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

Scott 2011

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of participants
Patient characteristics and setting	<p>Presenting signs and symptoms: Suspected TB, presenting with cough, fever, night sweats, and/or weight loss</p> <p>Age: mean 32 years; range 19 to 75 years</p> <p>Sex, female: 41.1%</p> <p>HIV infection: 69.0%</p> <p>History of TB: Not stated</p> <p>Sample size: 177</p> <p>Clinical setting: Primary health care clinic</p> <p>Laboratory: Reference laboratory</p> <p>Country: South Africa, Johannesburg</p> <p>World Bank Income Classification: Middle-/low-income</p> <p>TB incidence rate: 981 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 1.8% and among retreatment cases = 6.7% (Source: Nationwide survey, 2002)</p> <p>TB prevalence in study: 37.9%</p>
Index tests	<p>Index: Xpert MTB/RIF assay</p> <p>Specimen Condition: Frozen</p> <p>Specimen Preparation: Processed</p> <p>Parameter value for rifampicin resistance: 3.5</p>
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB</p> <p>Reference standard for pulmonary TB: MGIT 960</p> <p>Target condition: Rifampicin resistance</p> <p>Reference standard for rifampicin resistance: MGIT 960</p>

Scott 2011 (Continued)

Flow and timing			
Comparative			
Notes	One follow-up visit was performed approximately 60 days after enrolment Xpert was performed on frozen specimens while MGIT culture and smear microscopy were performed on fresh specimens		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results for TB detection interpreted without knowledge of the results of the index test?	Yes		

Scott 2011 (Continued)

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

Teo 2011

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of participants
Patient characteristics and setting	<p>Presenting signs and symptoms: Patients suspected of having TB based on symptoms and radiographic findings</p> <p>Age: Not stated</p> <p>Sex, female: Not stated</p> <p>HIV infection: Not stated</p> <p>History of TB: Not stated</p> <p>Sample size: 106</p> <p>Clinical setting: University hospital</p> <p>Laboratory: Reference laboratory</p> <p>Country: Singapore</p> <p>World Bank Income Classification: High-income</p> <p>TB incidence rate: 35 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 2.3% and among retreatment cases = 6.4% (Source: Nationwide surveillance, 2010)</p> <p>TB prevalence in study: 58.5%</p>
Index tests	<p>Index: Xpert MTB/RIF assay</p> <p>Specimen Condition: Fresh</p> <p>Specimen Preparation: Processed</p> <p>Parameter value for rifampicin resistance: 3.5</p>

Teo 2011 (Continued)

Target condition and reference standard(s)	Target condition: Pulmonary TB Reference standard for pulmonary TB: Löwenstein-Jensen culture and MGIT 960 Target condition: Rifampicin resistance Reference standard for rifampicin resistance: Gene sequencing		
Flow and timing			
Comparative			
Notes	Respiratory specimens (predominantly sputum) submitted for routine testing; only one rifampicin-resistant isolate was identified		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
			Low
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Teo 2011 (Continued)

Were the reference standard results for TB detection interpreted without knowledge of the results of the index test?	No			
Were the reference standard results for rifampicin resistance detection interpreted without knowledge of the results of the index test?	No			
Low				
DOMAIN 4: Flow and Timing				
Was there an appropriate interval between index test and reference standard?	Yes			
Did all patients receive the same reference standard?	Yes			
Were all patients included in the analysis?	Yes			

Theron 2011

Study characteristics

Patient sampling	Retrospective study with consecutive enrolment of participants
Patient characteristics and setting	<p>Presenting signs and symptoms: Suspected TB based on compatible signs and symptoms</p> <p>Age: median 36 years; range 18 to 83 years</p> <p>Sex, female: 32.3%</p> <p>HIV infection: 31.3%</p> <p>History of TB: 34.3%</p> <p>Sample size: 480</p> <p>Clinical setting: Two primary care clinics in a high HIV prevalence area</p> <p>Laboratory: Reference laboratory</p> <p>Country: South Africa, Cape Town</p> <p>World Bank Income Classification: Middle-/low-income</p> <p>TB incidence rate: 981 per 100,000</p> <p>MDR-TB prevalence: Percent MDR-TB among new TB cases = 0.9% and among retreatment cases = 4.0% (Source: Survey in Western Cape Province, 2002)</p> <p>TB prevalence in study: 29.4%</p>

Theron 2011 (Continued)

Index tests	Index: Xpert MTB/RIF assay Specimen Condition: Frozen Specimen Preparation: Processed Parameter value for rifampicin resistance: 3.5		
Target condition and reference standard(s)	Target condition: Pulmonary TB Reference standard for pulmonary TB: MGIT 960 Target condition: Rifampicin resistance Reference standard for rifampicin resistance: MGIT 960		
Flow and timing			
Comparative			
Notes	Short-term follow-up cultures were obtained; 16 of 19 Xpert-positive culture-negative patients were considered likely to be TB cases based on follow-up cultures, gene sequencing, and the presence of characteristic radiographic features using a standardised scoring system		
Methodological quality			
Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
			High
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
			Low
DOMAIN 3: Reference Standard			

Theron 2011 (Continued)

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

Zeka 2011

Study characteristics	
Patient sampling	Prospective study with consecutive enrolment of patients
Patient characteristics and setting	<p>Presenting signs and symptoms: Clinical findings of possible TB</p> <p>Age: median 48 years; range 25 to 70 years</p> <p>Sex, female: 42.4%</p> <p>HIV infection: Not stated</p> <p>History of TB: Not stated</p> <p>Sample size: 103</p> <p>Clinical setting: Laboratory-based evaluation of routine sputum specimens at a university hospital</p> <p>Laboratory: Reference laboratory</p> <p>Country: Turkey</p> <p>World Bank Income Classification: Middle-/low-income</p> <p>TB incidence rate: 28 per 100,000</p>

Zeka 2011 (Continued)

	<p>MDR-TB prevalence: Data not available TB prevalence in study: 34.0%</p>
Index tests	<p>Index: Xpert MTB/RIF assay Specimen Condition: Frozen Specimen Preparation: Processed Parameter value for rifampicin resistance: 3.5</p>
Target condition and reference standard(s)	<p>Target condition: Pulmonary TB Reference standard for pulmonary TB: Löwenstein-Jensen culture and MB/MBacT liquid medium Target condition: Rifampicin resistance Reference standard for rifampicin resistance: Proportional method on 7H10 media</p>
Flow and timing	
Comparative	
Notes	Only one rifampicin resistant isolate was identified. Data for sputum specimens were provided by the study author

Methodological quality

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

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

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Armand 2011	This was a case control study that compared Xpert MTB/RIF assay with an in-house IS6110-based real-time PCR using TaqMan probes (IS6110-TaqMan assay) for TB detection
Blakemore 2011	This was a technical paper that compared bacterial load quantitation determined by Xpert with the load determined by conventional quantitative methods
Causse 2011	This study evaluated Xpert for the diagnosis of extrapulmonary TB
Friedrich 2011a	This study evaluated Xpert for the diagnosis of pleural TB, a form of extrapulmonary TB

(Continued)

Hillemann 2011	This study evaluated Xpert for the diagnosis of extrapulmonary TB
Lawn 2011b	This was a narrative review that covered the development, technical details, and diagnostic accuracy of Xpert in adults and children
Ligthelm 2011	This study evaluated Xpert for the diagnosis of TB lymphadenitis, a form of extrapulmonary TB
Nicol 2011	This study evaluated Xpert for the diagnosis of TB in children
Vadwai 2011	This study evaluated Xpert for the diagnosis of extrapulmonary TB
Van Rie 2010	This was a review that covered technical details of Xpert and the test's potential value as a point-of-care test

Characteristics of ongoing studies [ordered by study ID]

Dheda 2012

Trial name or title	Multicentre randomised control trial of point-of-treatment (Clinic-based) Xpert MTB/RIF assay
Target condition and reference standard(s)	TB: reference standard: MGIT 960
Index and comparator tests	Xpert MTB/RIF assay and smear microscopy
Starting date	7 July 2011
Contact information	Keertan.Dheda@uct.ac.za; Jonny.Peter@uct.ac.za
Notes	RCT to assess the impact of Xpert on time-to-treatment and TB-related patient morbidity in primary care clinics. Identifier: NCT01554384

Luetkemeyer 2012

Trial name or title	Evaluation of Xpert MTB/RIF assay for the rapid identification of TB and TB rifampin resistance in HIV-infected and HIV-uninfected pulmonary tuberculosis suspects
Target condition and reference standard(s)	TB: reference standard: MGIT culture
Index and comparator tests	Xpert MTB/RIF assay
Starting date	24 April 2012
Contact information	Jay (John) Dwyer jdwyer@php.ucsf.edu

Luetkemeyer 2012 (Continued)

Notes	Cohort study of diagnostic accuracy of Xpert in HIV-infected and HIV-uninfected patients suspected of having pulmonary TB. Identifier: NCT01587469
-------	--

Peter 2012

Trial name or title	A randomised control trial of sputum induction, and new and emerging technologies in a high HIV prevalence primary care setting
Target condition and reference standard(s)	TB; liquid culture
Index and comparator tests	Xpert MTB/RIF assay
Starting date	August 2009
Contact information	Jonny.Peter@uct.ac.za
Notes	RCT to evaluate sputum induction for TB diagnosis in a primary care clinic for adults suspected of having TB. Identifier: NCT01545661

DATA

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

Tests. Data tables by test

Test	No. of studies	No. of participants
1 TB detection, all studies	27	7816
2 Add on	23	5719
3 Smear positive	24	1735
4 Smear negative	25	5878
5 HIV positive	11	1557
6 HIV negative	13	1981
7 TB detection, condition of specimen	24	7453
8 TB detection, specimen preparation	27	7816
9 TB prevalence	27	7816
10 Income status	27	7816
11 Rifampicin resistance	20	2340
12 RIF resistance prevalence	20	2340

Test 1. TB detection, all studies.

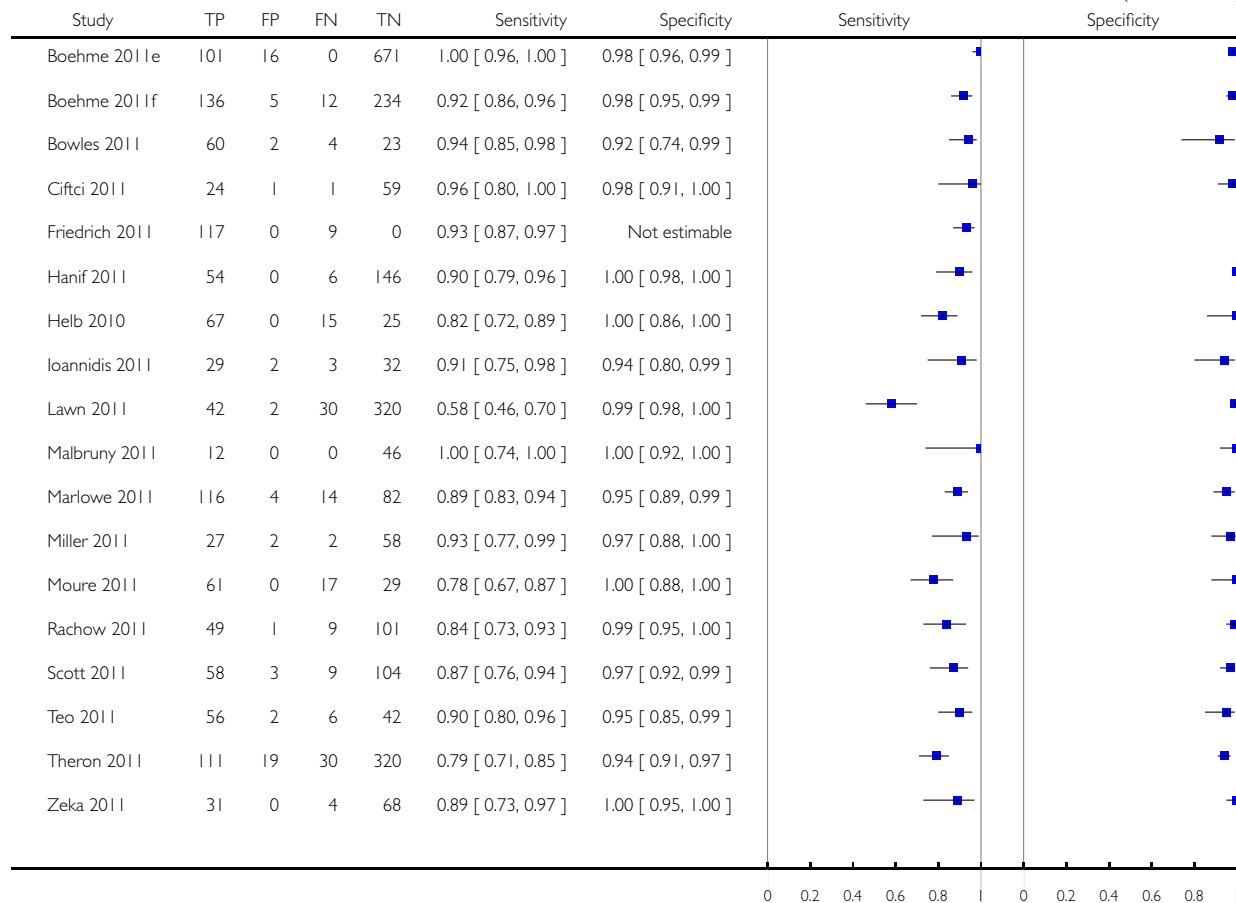
Review: Xpert174 MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults

Test: 1 TB detection, all studies

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Boehme 2010a	123	1	24	68	0.84 [0.77, 0.89]	0.99 [0.92, 1.00]		
Boehme 2010b	201	0	8	101	0.96 [0.93, 0.98]	1.00 [0.96, 1.00]		
Boehme 2010c	136	1	10	185	0.93 [0.88, 0.97]	0.99 [0.97, 1.00]		
Boehme 2010d	36	3	7	215	0.84 [0.69, 0.93]	0.99 [0.96, 1.00]		
Boehme 2010e	179	0	8	35	0.96 [0.92, 0.98]	1.00 [0.90, 1.00]		
Boehme 2011a	203	4	26	303	0.89 [0.84, 0.92]	0.99 [0.97, 1.00]		
Boehme 2011b	171	3	6	825	0.97 [0.93, 0.99]	1.00 [0.99, 1.00]		
Boehme 2011c	201	2	32	669	0.86 [0.81, 0.90]	1.00 [0.99, 1.00]		
Boehme 2011d	121	0	24	144	0.83 [0.76, 0.89]	1.00 [0.97, 1.00]		

(Continued ...)

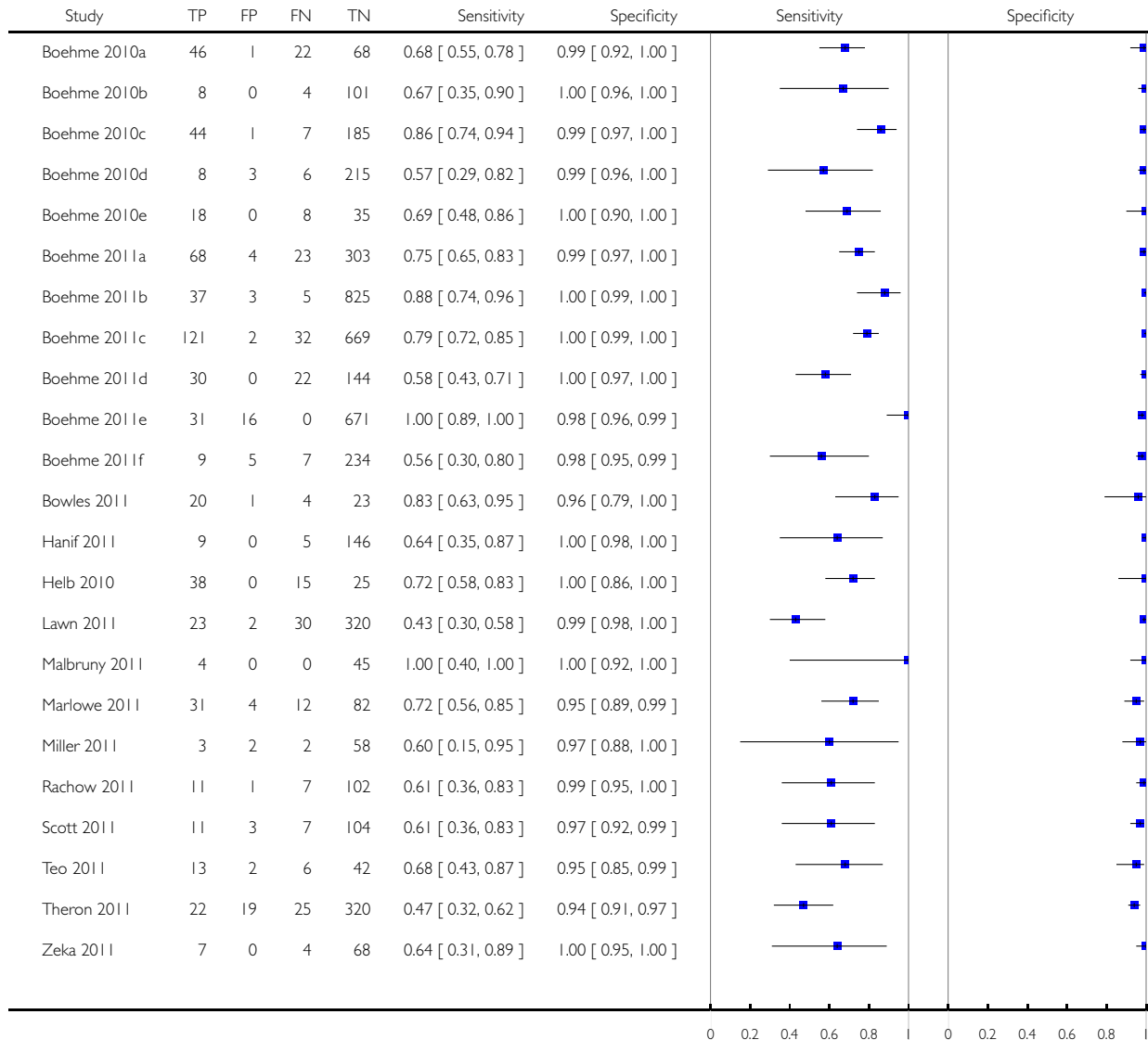
(... Continued)



Test 2. Add on.

Review: Xpert174 MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults

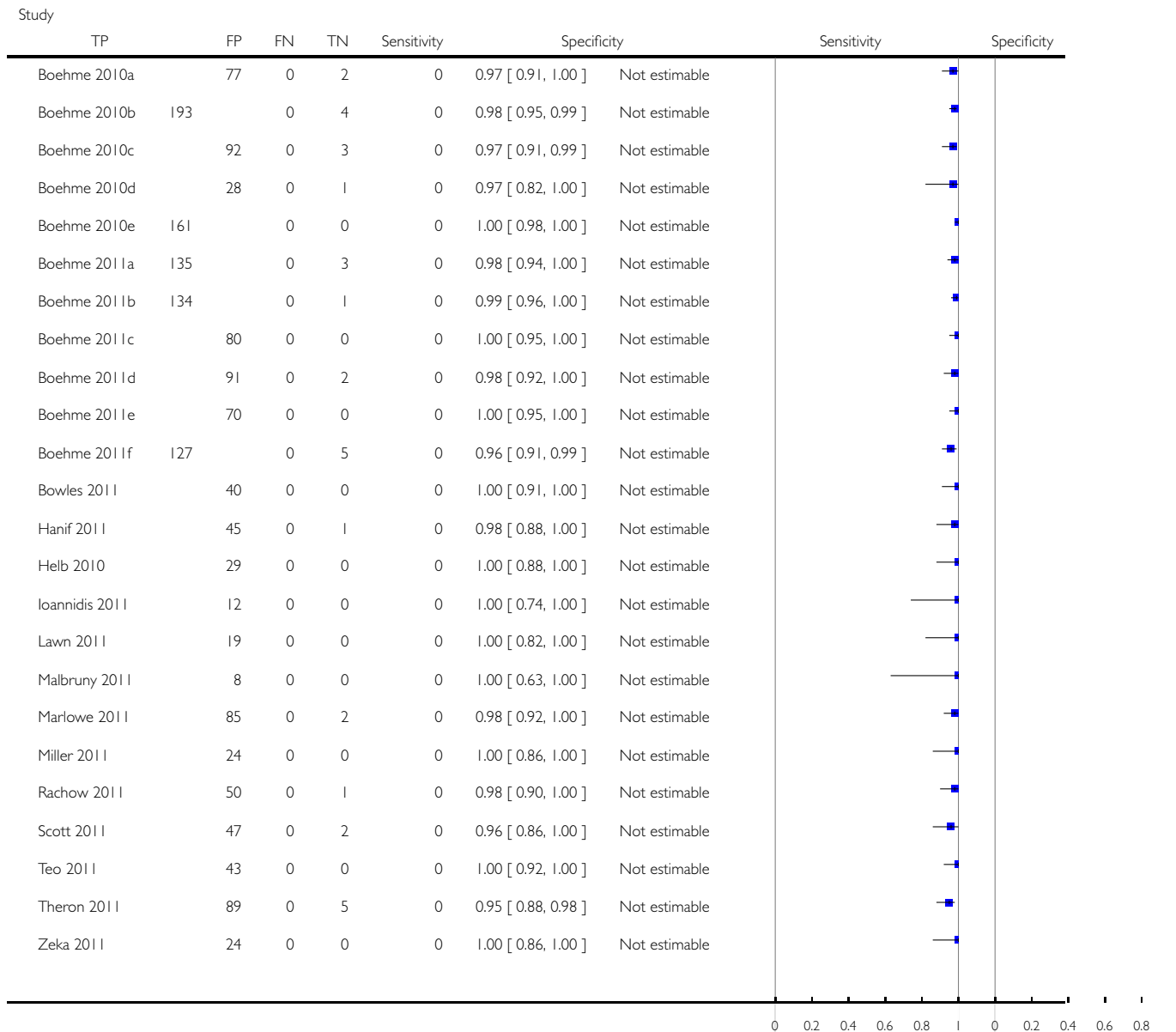
Test: 2 Add on



Test 3. Smear positive.

Review: Xpert174 MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults

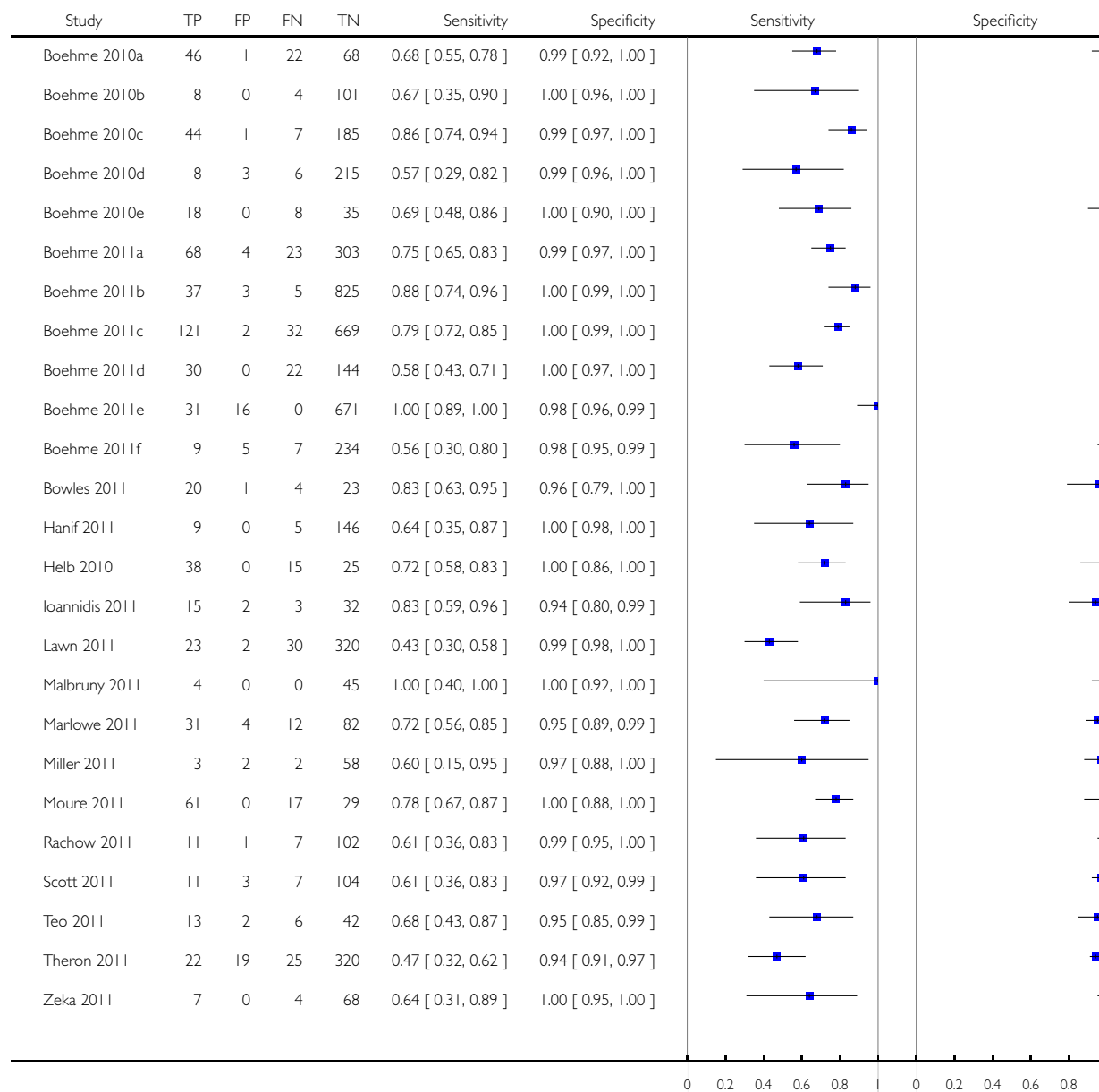
Test: 3 Smear positive



Test 4. Smear negative.

Review: Xpert174 MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults

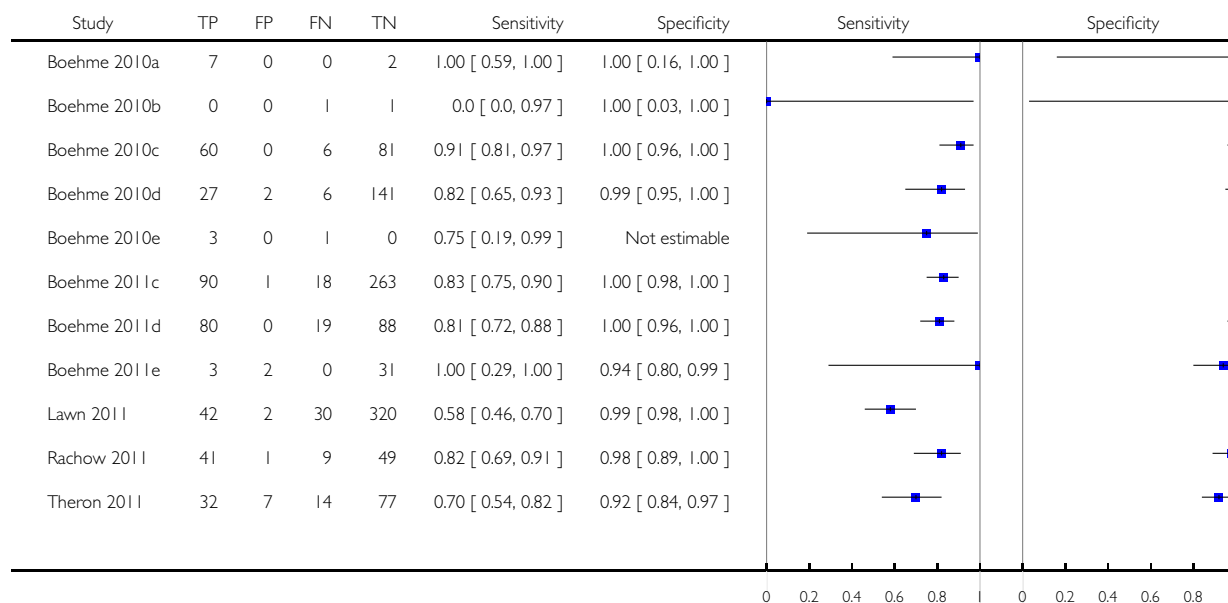
Test: 4 Smear negative



Test 5. HIV positive.

Review: Xpert174 MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults

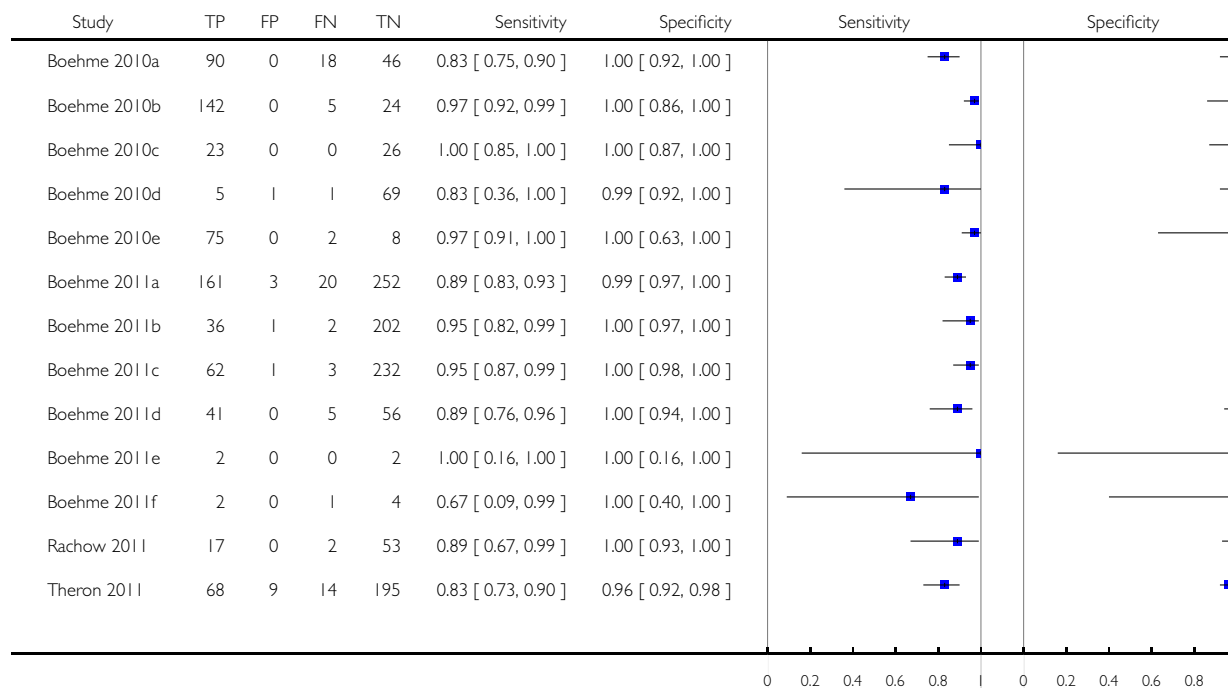
Test: 5 HIV positive



Test 6. HIV negative.

Review: Xpert174 MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults

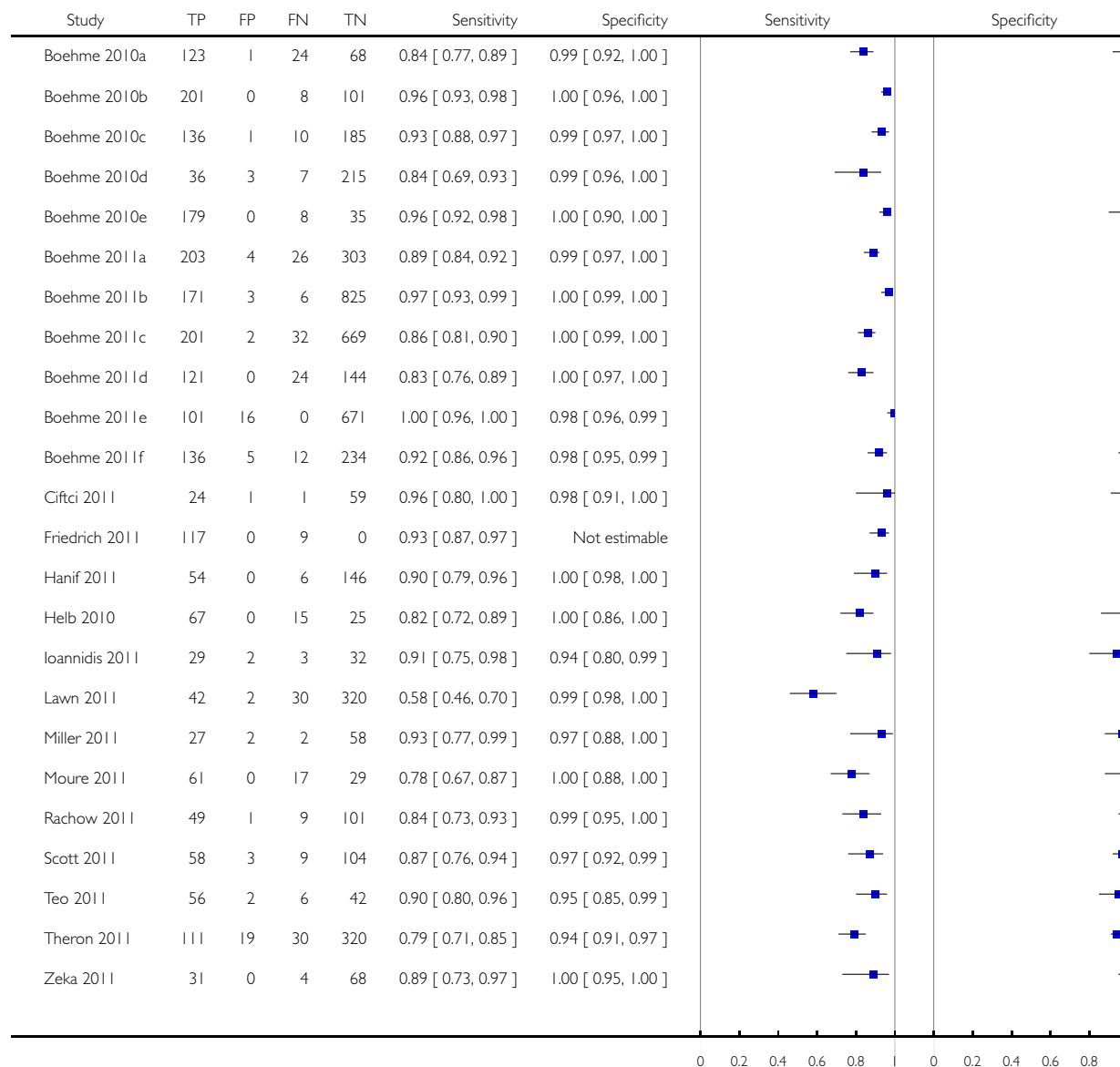
Test: 6 HIV negative



Test 7. TB detection, condition of specimen.

Review: Xpert174 MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults

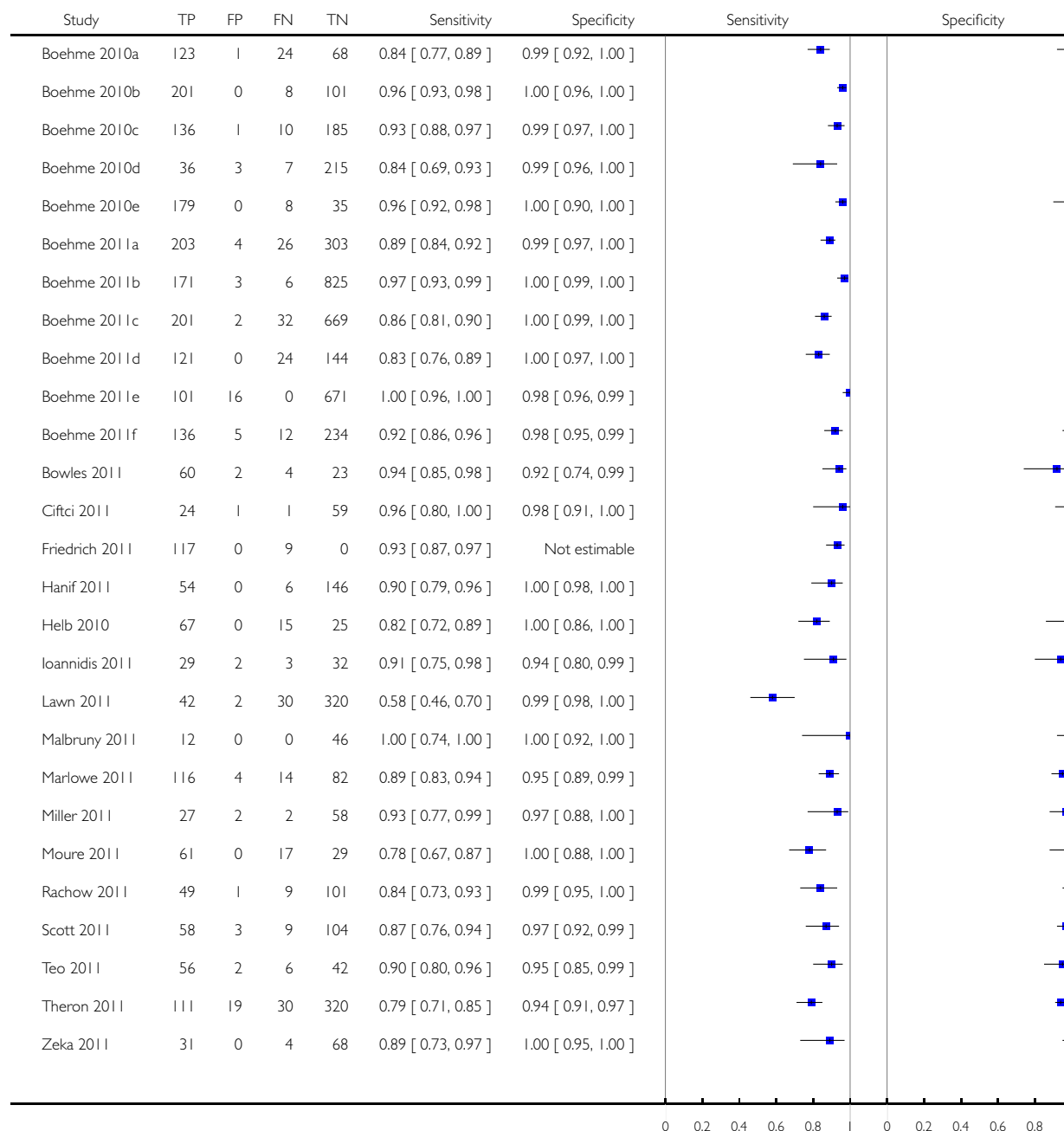
Test: 7 TB detection, condition of specimen



Test 8. TB detection, specimen preparation.

Review: Xpert174 MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults

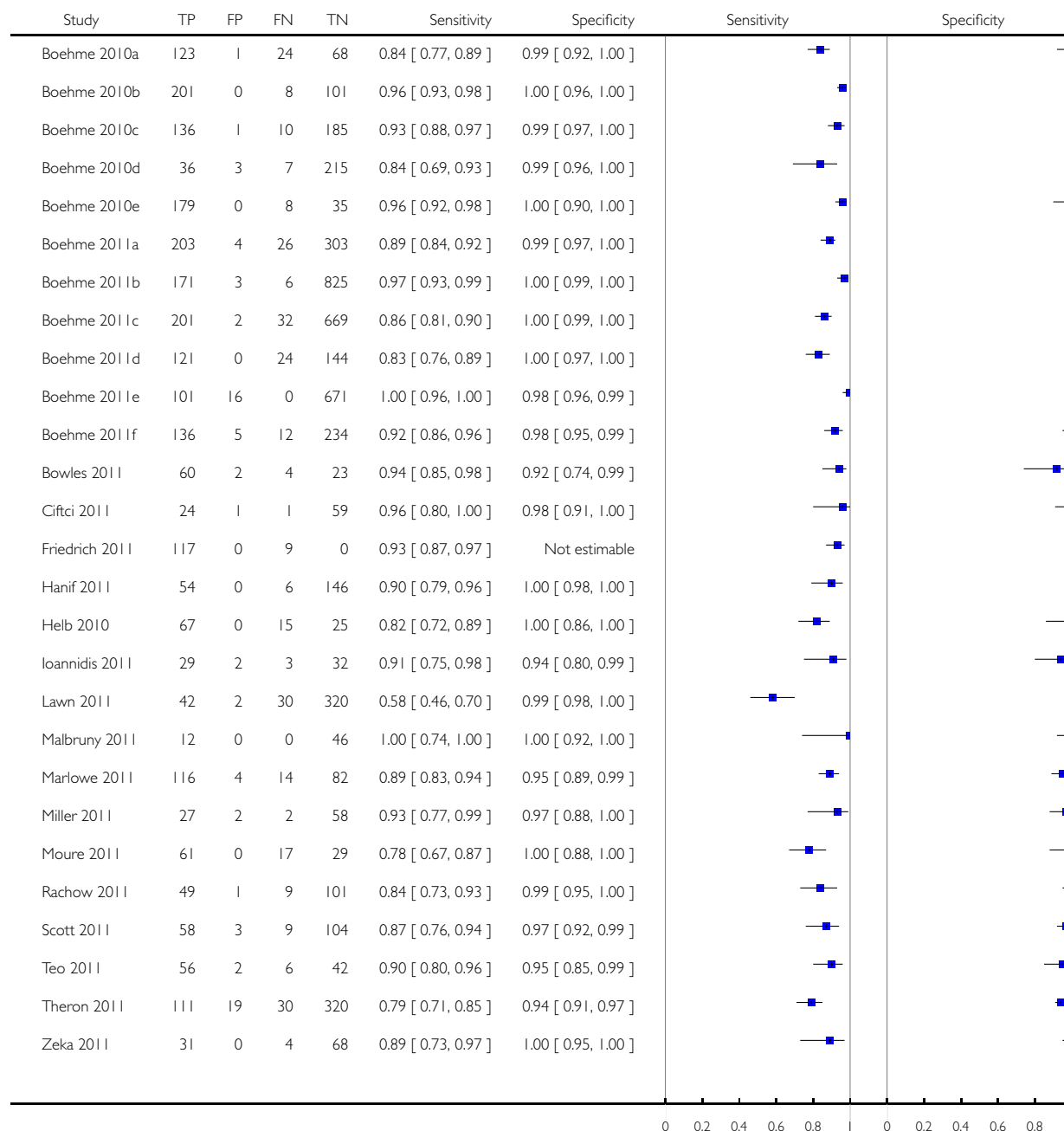
Test: 8 TB detection, specimen preparation



Test 9. TB prevalence.

Review: Xpert174 MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults

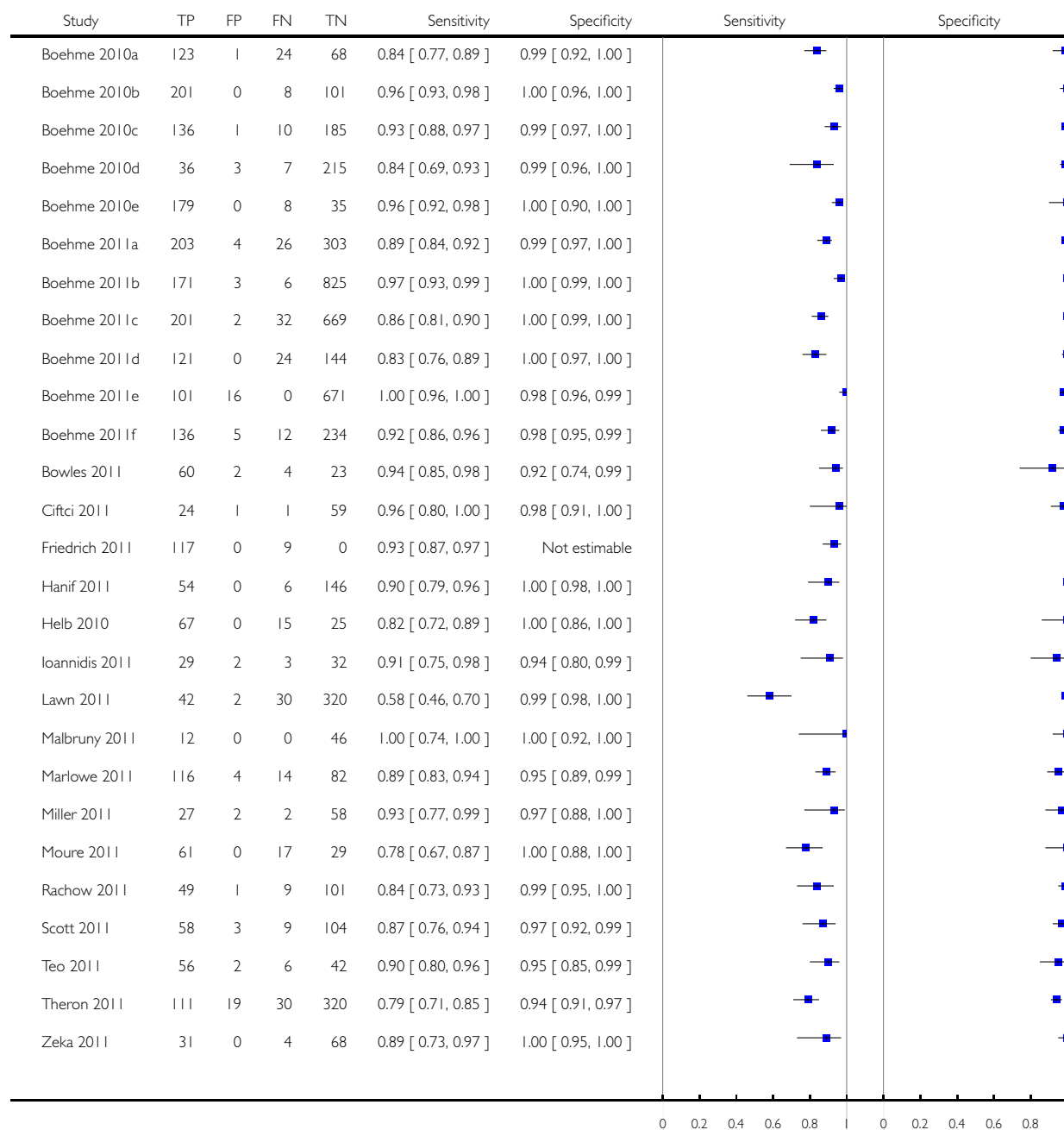
Test: 9 TB prevalence



Test 10. Income status.

Review: Xpert174 MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults

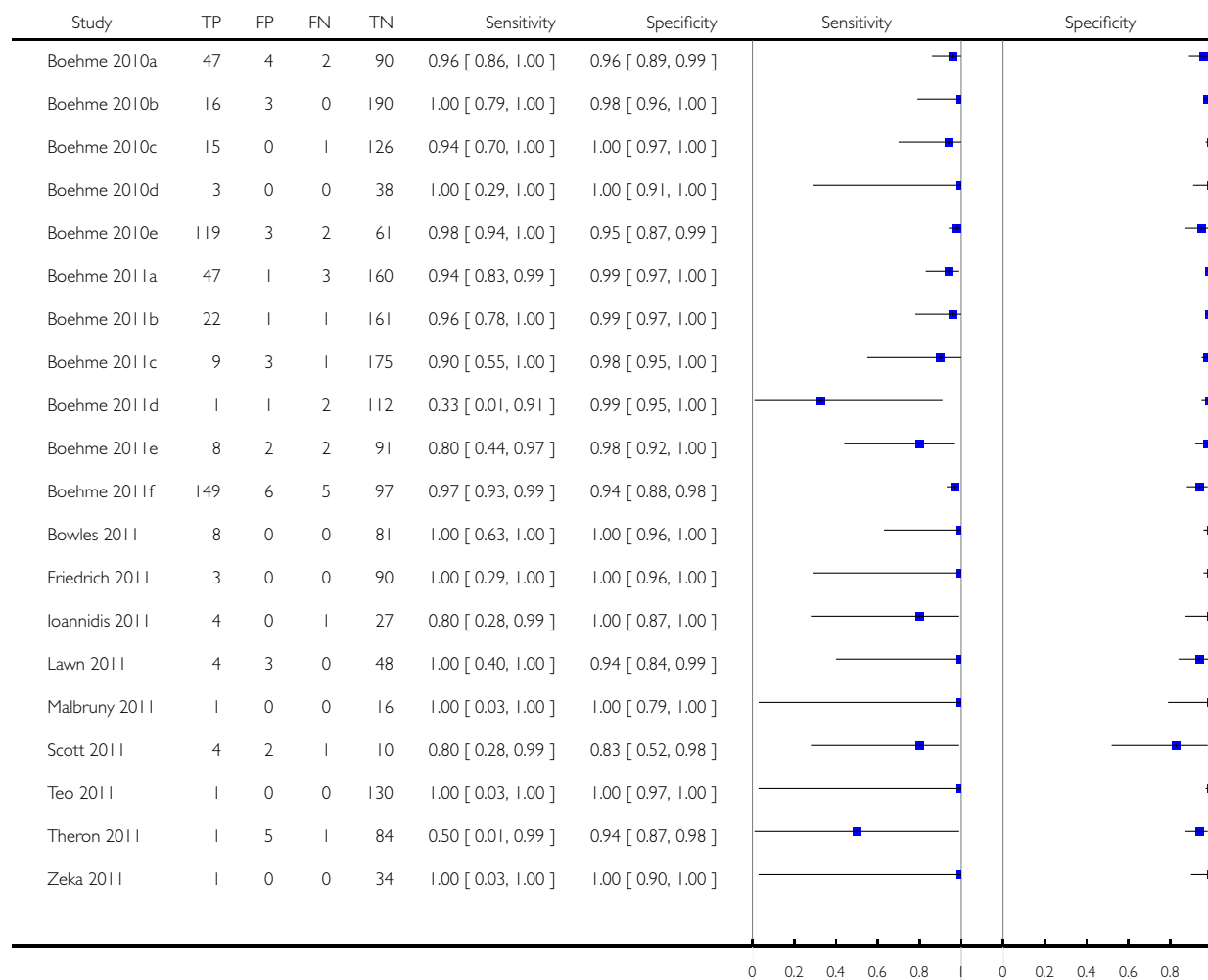
Test: 10 Income status



Test 11. Rifampicin resistance.

Review: Xpert174 MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults

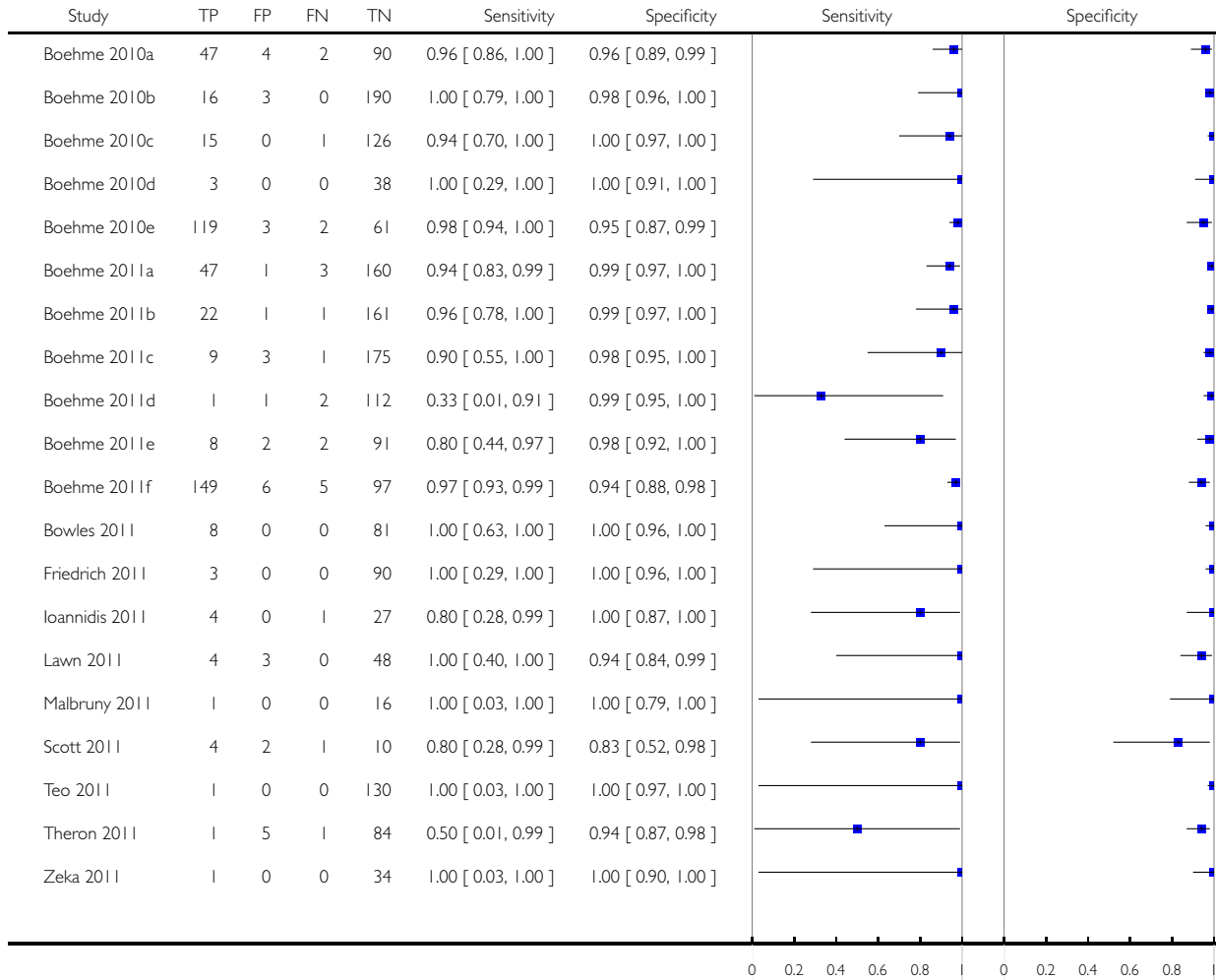
Test: 11 Rifampicin resistance



Test 12. RIF resistance prevalence.

Review: Xpert174 MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults

Test: 12 RIF resistance prevalence



ADDITIONAL TABLES

Table 1. Xpert MTB/RIF assay for detection of TB and rifampicin resistance

Type of analysis (Number of studies)	Pooled sensitivity Median (95% credible interval)	Pooled specificity Median (95% credible interval)	Predicted sensitivity Median (95% credible interval)	Predicted specificity Median (95% credible interval)

Table 1. Xpert MTB/RIF assay for detection of TB and rifampicin resistance (Continued)

Xpert used as an initial test for TB detection replacing microscopy (15)*	88% (83, 92)	98% (97, 99)	88% (66, 97)	98% (92, 100)
Xpert used as an add-on test for TB detection following a negative smear microscopy result (14)	67% (58, 74)	98% (97, 99)	66% (40, 86)	98% (93, 100)
Xpert used as an initial test for rifampicin resistance detection replacing conventional drug susceptibility testing as the initial test (11)	94% (87, 97)	98% (97, 99)	94% (75, 99)	98% (91, 100)

*Three studies that preferentially enrolled smear-positive or smear-negative patients were excluded

Table 2. Impact of covariates on heterogeneity of Xpert sensitivity and specificity for TB detection

Covariate	Sensitivity Median (95% credible interval)	Specificity Median (95% credible interval)
Smear status		
Smear +	98% (97, 99)	***
Smear -	68% (59, 75)	98% (97, 99)
Difference (Smear+ minus Smear-)	31% (23, 39)	**
P (Smear+ > Smear-)	1.00	**
HIV status		
HIV+	80% (67, 88)	97% (93, 99)
HIV-	89% (81, 94)	99% (96, 99)
Difference (HIV+ minus HIV-)	-9% (-22, 3)	-1% (-5, 2)
P (HIV+ > HIV-)	0.06	0.21
Condition of specimen		

Table 2. Impact of covariates on heterogeneity of Xpert sensitivity and specificity for TB detection (Continued)

Fresh	88% (80, 93)	99% (98, 100)
Frozen	85% (77, 91)	97% (95, 99)
Difference (Fresh minus Frozen)	3% (-7, 13)	2% (0.1, 4)
P (Fresh > Frozen)	0.73	0.98
Specimen preparation		
Unprocessed	92% (87, 96)	99% (97, 99)
Processed	85% (79, 90)	98% (96, 99)
Difference (Unprocessed minus Processed)	7% (0.2, 14)	0.8% (-1, 3)
P (Unprocessed > Processed)	0.98	0.84
TB prevalence		
High (> 30%)	89% (84, 93)	98% (96, 99)
Low (\leq 30%)	86% (77, 92)	99% (97, 99)
Difference (High minus Low)	3% (-5, 12)	-0.4% (-2, 1)
P (High > Low)	0.80	0.29
Country income level		
High-income	92% (86, 96)	98% (95, 99)
Low- and middle-income	85% (79, 90)	99% (97, 99)
Difference (High-income minus Low- and middle-income)	6% (-1, 14)	-1% (-3, 1)
P (High-income > Low- and middle-income)	0.96	0.23

P = probability

Table 3. Impact of covariates on heterogeneity of Xpert sensitivity and specificity for rifampicin resistance detection

Covariate	Sensitivity Median (95% credible interval)	Specificity Median (95% credible interval)
Parameter value (delta cycle threshold cutoff)		
Parameter value 5	94% (86, 97)	98% (96, 99)
Parameter value 3.5	96% (81, 100)	100% (98, 100)
Difference (Parameter value 5 minus Parameter value 3)	-3% (-10, 11)	-2% (-4, -0.3)
P (Parameter value 5 > Parameter value 3)	0.26	0.01
Rifampicin resistance prevalence		
High (> 15%)	94% (85, 98)	98% (95, 99)
Low (≤ 15%)	93% (80, 99)	98% (96, 99)
Difference (High minus Low)	0.4% (-8, 13)	-0.4% (-4, 2)
P (High > Low)	0.54	0.34

P = probability

Table 4. Selected patient-important outcomes as reported in the included studies

Study and year of publication	Time to TB detection	Time to detection of rifampicin resistance	Time to treatment initiation
Boehme 2011a; Boehme 2011b; Boehme 2011c; Boehme 2011d; Boehme 2011e; Boehme 2011f	Median (IQR) Xpert: 0 days (0, 1) Smear: 1 day (0, 1) Solid culture: 30 days (23, 43) Liquid culture: 16 days (13-21)	Median (IQR) Xpert: 1 day (0, 1) Line probe assay (direct testing): 20 days (10, 16) Phenotypic DST: 106 days (30, 124)	Median (IQR) Smear-, culture+ TB Before Xpert introduced: 56 days (39, 81) After Xpert introduced: 5 days (2, 8)
Helb 2010	Xpert (1 sample): 1 hour 55 minutes Xpert (8 samples processed together): 2 hours		
Lawn 2011	Median* (IQR) Xpert: 4 days (3, 6)	Xpert: mean 2 days MTBDRplus assay (with pos-	

Table 4. Selected patient-important outcomes as reported in the included studies (Continued)

	Smear: 3 days (2, 5) Liquid culture (smear+): 12 days (10,14) Liquid culture (smear-): 20 days (17, 27)	itive culture isolate): mean 21 days Phenotypic DST (liquid culture): mean 40 days	
Marlowe 2011	Xpert: hands-on time was 5 minutes; run time was less than 2 hours		
Miller 2011	Xpert: hands-on time was 15 minutes; run time was 113 minutes		
Moure 2011	Xpert: total time of 2 hours		
Rachow 2011	Xpert: within two hours		
Zeka 2011**	Xpert (routine practice): 3-24 hours Liquid culture: 19 days mean (range 3-42 days)		

*Delays between sputum collection and results being available to the clinic

**Times provided for both pulmonary and extrapulmonary specimens jointly; DST, drug susceptibility testing; IQR, interquartile range

APPENDICES

Appendix I. Detailed search strategies

Search strategy Medline (OVID) and Embase (OVID)

1. (tuberculosis or TB).tw

limit 1 to yr="2007 -Current"

2. Mycobacterium tuberculosis/

limit 2 to yr="2007 -Current"

3. Tuberculosis, Multidrug-Resistant/ or Tuberculosis/ or Tuberculosis, Pulmonary/

limit 3 to yr="2007 -Current"

4. 1 or 2 or 3

5. (Xpert or GeneXpert or cepheid or(near* patient)). tw.

limit 4 to yr="2007 -Current"

4 and 5

Search strategy Web of Knowledge (SCI-expanded, SSCI, Conference Proceedings science, BIOSIS previews)

(tuberculosis OR TB OR mycobacterium) (topic) AND (Xpert OR Genexpert OR cepheid) (topic)

Search strategy LILACS

(tuberculosis OR TB OR mycobacterium) (Words) AND (xpert OR Genexpert OR Cepheid) (Words)

Search strategy SCOPUS

(tuberculosis OR TB OR mycobacterium) (title, abstract, keywords) AND (xpert OR Genexpert OR Cepheid) (title, abstract, keywords)

Appendix 2. Data extraction form

ID	
ID substudy (for study centres: a, b, c, etc)	
First Author	
Corresponding author & email	
Was author contacted?	1 - Yes 2 - No If yes, dates(s)
Title	
Year (of publication)	
Year (study start date)	
Language	1 - English 2 - Other If other, specify:
Was the study conducted without industry sponsorship?	1 - Yes 2 - No 9 - Unk/NR
If industry sponsorship was present, select one item from the list	Select one: answers ordered from least to most industry involvement: Donation of Xpert® for use in study Xpert® at a special preferred price Receipt of educational support, grants, or speaking fees Financial relationship - author is employee/consultant/stockholder Involvement in design, analysis, or manuscript production
For TB detection, what reference standard(s) was used?	1 - Solid Culture (specify 1a) 2 - Liquid Culture (specify 2a) 3 - Both Solid & Liquid Culture (specify 1a & 2a) 9 - Unk/NR 1a - Solid Culture

(Continued)

	<p>LJ 7H10 7H11 Other 2a - Liquid Culture MGIT 960 Bactec 460 Other</p>
For rifampicin resistance detection, what reference standard(s) was used?	<p>1 - Solid Culture (specify 1a) 2 - Liquid Culture (specify 2a) 3 - Both Solid & Liquid Culture (specify 1a & 2 a) 9 - Unk/NR</p> <p>1a - Solid Culture LJ 7H10 7H11 Other Specify method, eg, proportion 2a - Liquid Culture MGIT 960 Bactec 460 Other</p>
Clinical setting; describe as written in the paper	<p>1 - Outpatient 2- Inpatient 3 - Both out- and in-patient 4 - Other, specify 5 - Laboratory 9 - Unk/NR</p> <p>Describe as in paper:</p>
Laboratory services level Specify type of laboratory	<p>1- Central (Reference) 2 - Intermediate (Regional) 3 - Peripheral (Microscopy centre, health clinic, provincial hospital) 4- Other, specify</p>
Was Xpert run outside of a laboratory, eg, clinic?	<p>1 - Yes 2 - No</p>
Country where study was conducted	
Country World Bank Classification	<p>1 - Middle/Low 2 - High 3 - Both middle/low and high</p>

(Continued)

Study design	1 - Randomized Trial 2 - Cross-sectional 3 - Cohort 4 - Other, specify 9 - Unk/NR If other, specify:
Participant selection	1 - Consecutive 2 - Random 3 - Convenience 7 - Other 9 - NR/Unclear
Direction of study data collection	1 - Prospective 2 - Retrospective 9 - Unk/NR
Comments about study design	
Number after screening by exclusion & inclusion criteria	----- 9 - Unk/NR
Number included in analysis (# screened - # withdrawals)	----- 9 - Unk/NR
Unit of analysis	1 - One specimen per patient 2 - Multiple specimens per patient 3 - Unknown number of specimens per patient 9 - NR/Unclear Describe as in paper, if unclear:
Prior testing by microscopy for triage	1 - Yes 2 - No 9 - Unk/NR
Did the study include patients with previous TB history?	1 - Yes 2 - No 9 - Unk/NR
If so, what is the percentage?	% Specify numerator/denominator
HIV status of participants	0 - HIV - 1 - HIV + 2 - Both HIV +/- 9 - Unk/NR

(Continued)

If HIV-positive participants included, what is the percentage?	% Specify numerator/denominator
Specimen collection (may include expectorated sputum, induced sputum, bronchial alveolar lavage (BAL), tracheal aspiration)	1 - All expectorated 2 - All induced 3 - All BAL 4 - Multiple types 5 - Other 9 - Unk/NR If 4 or 5, describe and record numbers:
Were Xpert sample and culture obtained from same specimen?	1 - Yes 2 - No 9 - Unk/NR
Number of cultures used to exclude TB	1 - One 2 - Two 3 - Three 4 - Four 5 - Other, specify 9 - Unk/NR Specify, if > 4: NOTES:
Pre-treatment processing procedure for Xpert	1 - None 2 - NALC-NaOH 3 - NaOH (Petroff) 4 - Other 9 - Unk/NR
Was microscopy used	1 - Yes 2 - No 9 - Unk/NR
Type of microscopy used	1 - Ziehl-Neelsen 2 - FM 9 - Unk/NR
Smear type	1 - Direct 2 - Concentrated (processed) 9 - Unk/NR
Minimum number of sputum specimens used to determine smear positivity	1 - One 2 - Two 3 - Three 4 - >3 9 - Unk/NR

(Continued)

How was a positive smear defined? (if guideline referenced, look up guideline)	≥ ___ bacilli per ___ high power fields 9 - Unk/NR * complete both fields
For Xpert specimen, what was the condition of the specimen when tested?	1 - Fresh 2 - Frozen 9 - Unk/NR
If fresh, specify:	1 - Tested after storage at room temperature or refrigerated w/in 48 hours of collection 2 - Tested after storage at room temperature or refrigerated > 48 hours of collection 9 - Unk/NR
If frozen, specify:	1 - Tested after frozen < 1 year of storage 2 - Tested frozen ≥ 1 year of storage 9 - Unk/NR
Version of software for test interpretation	1 - Version 1 2 - Version 2 3 - Version 3 4 - Version 4 9 - Unk/NR
Enter percentage contaminated cultures, if provided: # of contaminated culture Total # cultures performed	----- 9 - Unk/NR
Were indeterminate results reported for Xpert for TB detection?	1 - Yes 2 - No 9 - Unk/NR
Were indeterminate results reported for Xpert for rifampicin resistance detection?	1 - Yes 2 - No 9 - Unk/NR
Were patient important outcomes evaluated?	1 - Yes 2 - No 9 - Unk/NR
Time to diagnosis	Xpert: Culture: 9 - Unk/NR
Time to treatment initiation	Xpert: Culture: 9 - Unk/NR

(Continued)

Other patient outcomes	Specify:
------------------------	----------

TABLES

TB detection, all studies		Definite TB		
		Yes	No	Total
Xpert result	Positive			
	Negative			
	Total			
	Indeterminate			

TB Detection, smear positive		Definite TB		
		Yes	No	Total
Xpert result	Positive			
	Negative			
	Total			
	Indeterminate			

TB detection, smear negative		Confirmed TB		
		Yes	No	Total
Xpert result	Positive			
	Negative			
	Total			
	Indeterminate			

TB detection, HIV-positive		Definite TB		
		Yes	No	Total
Xpert result	Positive			
	Negative			
	Total			
	Indeterminate			

TB detection, HIV-negative		Definite TB		
		Yes	No	Total
Xpert result	Positive			
	Negative			
	Total			
	Indeterminate			

RIF resistance detection		Confirmed rifampicin resistance		
		Yes	No	Total
Xpert result	Yes (resistant)			
	No (susceptible)			
	Total			
	Indeterminate			

Discrepant analysis		Comments:
Criteria	Number	
Xpert+/culture - baseline		
Deemed TB after further evaluation		
Percent found to be TB on discrepant analysis		

* Follow-up test included, circle all that apply

** Repeat culture, DNA sequencing, GenoType® MTBDRplus test, other, describe

Microscopy		Definite TB		
		Yes	No	Total
Microscopy result	Positive			
	Negative			
	Total			

Appendix 3. Rules for QUADAS-2

Domain 1 Patient Selection:

Risk of Bias: Could the selection of patients have introduced bias?

Signalling *question 1*: Was a consecutive or random sample of patients enrolled? We scored 'yes' if the study enrolled a consecutive or random sample of eligible patients; 'no' if the study selected patients by convenience; and 'unclear' if the study did not report the manner of patient selection or we could not tell.

Signalling *question 2*: Was a case-control design avoided? Studies using a case-control design were not included in the review because this study design, especially when used to compare results in severely ill patients with those in relatively healthy participants, may lead to overestimation of accuracy in diagnostic studies. We scored 'yes' for all studies.

Signalling *question 3*: Did the study avoid inappropriate exclusions? We scored 'yes' if the study included both smear-positive and smear-negative patients; 'no' if the study included only smear-positive patients; and 'unclear' if we could not tell.

Risk of Bias was scored as 'low concern' if selection was done in a random or consecutive manner and the study was not limited to smear-positive patients; 'high concern' if selection was by convenience or the study included only smear-positive patients; and 'unclear concern' if the manner of participant selection was unclear and no clinical information was provided.

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

We were interested in how Xpert performed in patients suspected of having pulmonary TB or MDR-TB who were evaluated as they would be in settings of intended use, ie in basic laboratories or primary health facilities. We scored 'low concern' if Xpert was evaluated in

provincial hospitals or primary health clinics and 'high' concern if Xpert was evaluated in reference laboratories. We judged applicability to be of 'unclear concern' if Xpert was evaluated in a basic laboratory and the study did not provide any clinical information about the participants.

Domain 2: Index Test

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

Signalling *question 1: Were the index test results interpreted without knowledge of the results of the reference standard?* We answered this question 'yes' for all studies because Xpert test results were automatically generated and the user was provided with printable test results. Thus, there is no room for subjective interpretation of test results.

Signalling *question 2: If a threshold was used, was it prespecified?* The threshold was prespecified in all versions of Xpert. We answered this question 'yes' for all studies.

For risk of bias, we scored 'low concern' for all studies.

Applicability: Are there concerns that the index test, its conduct, or its interpretation differ from the review question? Variations in test technology, execution, or interpretation may affect estimates of the diagnostic accuracy of a test. However, we judged these issues to be of 'low concern' for all studies in this review.

Domain 3: Reference Standard

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

We considered this domain separately for the reference standard for TB detection and the reference standard for rifampicin resistance.

Signalling *question 1: Is the reference standard likely to correctly classify the target condition?* For pulmonary TB: although culture is not 100% accurate, it is considered to be the gold standard for TB diagnosis. For rifampicin resistance: similarly, although drug susceptibility testing by conventional phenotypic methods is not 100% accurate, it is considered to be the gold standard. We answered this question 'yes' for all studies.

Signalling *question 2: (TB) Were the reference standard results interpreted without knowledge of the results of the index test?* We scored 'yes' if the reference test provided an automated result (eg MGIT 960), blinding was explicitly stated, or it was clear that the reference standard was performed at a separate laboratory and/or performed by different people. We scored 'no' if the study stated that the reference standard result was interpreted with knowledge of the Xpert test result. We scored 'unclear' if we could not tell.

Signalling *question 3: (Rifampicin resistance) Were the reference standard results interpreted without knowledge of the results of the index test?* We added a signalling question for rifampicin resistance because judgments might differ for TB and for rifampicin resistance, the two target conditions. We scored 'yes' if the reference test provided an automated result (eg MGIT 960), blinding was explicitly stated, or it was clear that the reference standard was performed at a separate laboratory and/or performed by different people. We scored 'no' if the study stated that the reference standard result was interpreted with knowledge of the Xpert test result. We scored 'unclear' if we could not tell.

For risk of bias, we scored 'low concern' if the reference standard result was automated, blinding explicitly stated, or the reference standard performed at a separate laboratory. We scored 'high concern' if the study explicitly stated the result of the reference standard was interpreted with knowledge of the Xpert test result. We scored 'unclear concern' if we could not tell. We reported results for both TB and rifampicin resistance.

Applicability: Are there concerns that the target condition as defined by the reference standard does not match the question? We judged applicability to be of 'low concern' for all studies for both pulmonary TB and rifampicin resistance.

Domain 4: Flow and Timing

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

Signalling *question 1: Was there an appropriate interval between the index test and reference standard?* In the majority of included studies, we expected specimens for Xpert and culture to be obtained at the same time when patients were suspected of having TB. However, even if there were a delay of several days or weeks between index test and reference standard, TB is a chronic disease and we considered misclassification of disease status to be unlikely. We answered this question 'yes' for all studies.

Signalling *question 2: Did all patients receive the same reference standard?* We answered this question 'yes' for all studies as an acceptable reference standard (either solid or liquid culture) was specified as a criterion for inclusion in the review. However, we acknowledged that it is possible that some specimens could undergo solid culture and others liquid culture. This could potentially result in variations in accuracy, but we thought the variation would be minimal.

Signalling *question 3: Were all patients included in the analysis?* We determined the answer to this question by comparing the number of patients enrolled with the number of patients included in the two-by-two tables.

For risk of bias, we scored 'low concern' if the number of participants enrolled was clearly stated and corresponded to the number presented in the analysis or if exclusions were adequately described. We scored 'high concern' if there were participants missing or excluded from the analysis and there was no explanation given; and 'unclear concern' if not enough information was given to assess

whether participants were excluded from the analysis; usually this meant that the number of participants originally enrolled in the study was not explicitly stated.

Appendix 4. Statistical appendix

Bayesian bivariate hierarchical model

The Bayesian bivariate hierarchical model used for the meta-analyses is summarized below. The hierarchical framework took into account heterogeneity between studies and also between centres within two of the largest studies. The model was derived as an extension of previously described models (Chu 2009; Reitsma 2005). A WinBUGS program to fit this model is provided below. Three independent, dispersed sets of starting values were used to run separate chains. The Gelman-Rubin statistic within the WinBUGS program was used to assess convergence. No convergence problems were observed. The first 3,000 iterations were treated as burn-in iterations and dropped. Summary statistics were obtained based on a total of 15,000 iterations resulting from the three separate chains.

Notation: From the j^{th} centre in the i^{th} study we extracted the cross-tabulation between the index and reference tests TP_{ij} , FP_{ij} , TN_{ij} , FN_{ij} . The sensitivity in ij^{th} study is denoted by S_{ij} and the specificity by SP_{ij} . We denote the Binomial probability distribution with sample size N and probability p as $\text{Binomial}(p, N)$, the Bivariate Normal probability distribution with mean vector μ and variance-covariance matrix Σ as $\text{BVN}(\mu, \Sigma)$, the univariate Normal distribution with mean m and variance s by $N(m, s)$ and the Uniform probability distribution between a and b by $\text{Uniform}(a, b)$.

Likelihood [Figure 14](#):

Figure 14. Bayesian bivariate hierarchical model, likelihood

Centre-level:

For studies with only 1 centre:

$$TP_{i1} \sim \text{Binomial}(S_i, TP_{i1} + FN_{i1}), TN_{i1} \sim \text{Binomial}(SP_i, TN_{i1} + FP_{i1})$$

For multicentre studies:

$$TP_{ij} \sim \text{Binomial}(S_{ij}, TP_{ij} + FN_{ij}), TN_{ij} \sim \text{Binomial}(SP_{ij}, TN_{ij} + FP_{ij})$$

$$\begin{pmatrix} \text{logit}(S_{ij}) \\ \text{logit}(SP_{ij}) \end{pmatrix} \sim \text{BVN}(l_i, \Sigma_i),$$

$$\text{where } l_i = \begin{pmatrix} \text{logit}(S_i) \\ \text{logit}(SP_i) \end{pmatrix} \text{ and } \Sigma_i = \begin{pmatrix} \sigma_{i1}^2 & k_i \sigma_{i1} \sigma_{i2} \\ k_i \sigma_{i1} \sigma_{i2} & \sigma_{i2}^2 \end{pmatrix}$$

Study-level:

$$\begin{pmatrix} \text{logit}(S_i) \\ \text{logit}(SP_i) \end{pmatrix} \sim \text{BVN}\left(\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, T = \begin{pmatrix} \tau_1^2 & \rho \tau_1 \tau_2 \\ \rho \tau_1 \tau_2 & \tau_2^2 \end{pmatrix}\right)$$

The pooled sensitivity is given by $1/1+\exp(-\mu_1)$ and pooled specificity as $1/1+\exp(\mu_2)$.

Prior distributions [Figure 15](#):

Figure 15. Bayesian bivariate hierarchical model, prior distributions

$$\mu_1 \text{ and } \mu_2 \sim N(0, 100)$$

$$k_i \text{ and } \rho \sim U(-1, 1)$$

$$\frac{1}{\sigma_1^2}, \frac{1}{\sigma_2^2}, \frac{1}{\tau_1^2} \text{ and } \frac{1}{\tau_2^2} \sim \text{Gamma}(\text{shape}=2, \text{rate}=0.5)$$

Addition of covariates:

To examine the impact of a dichotomous covariate (Z) on the pooled sensitivity and specificity parameters, we expressed the logit(sensitivity) and logit(specificity) as linear functions of Z as follows:

$$\mu_1 = a_1 + b_1 Z \text{ and } \mu_2 = a_2 + b_2 Z$$

Prior distributions were placed over the coefficients in the linear function: a_1 and $a_2 \sim N(0, 4)$ and b_1 and $b_2 \sim N(0, 1.39)$ (Buzoianu 2008).

```

-----
# WinBUGS PROGRAM FOR ESTIMATING A BIVARIATE HIERARCHICAL META-ANALYSIS MODEL
# FOR SENSITIVITY AND SPECIFICITY ALLOWING FOR HETEROGENEITY BETWEEN STUDIES
# AND HETEROGENEITY BETWEEN CENTRES WITHIN TWO OF THE STUDIES (BOEHME 2010 and 2011)
model {
##### BOEHME 2010 #####
  for(j in 1:5) {

    logit(se.q[j])<-q1[j,1]
    logit(sp.q[j])<-q1[j,2]

    q1[j,1:2]~ dnorm(l[1,1:2], T1[1:2,1:2])

    pos1[j]<-TP1[j]+FN1[j]
    neg1[j]<-TN1[j]+FP1[j]
    TP1[j] ~ dbin(se.q[j],pos1[j])
    FP1[j] ~ dbin(sp.q[j],neg1[j])
  }
  T1[1:2,1:2]<-inverse(SIGMA1[1:2,1:2])

# Between-centre variance-covariance matrix for Boehme 2010
SIGMA1[1,1] <- sigma1[1]*sigma1[1]
SIGMA1[2,2] <- sigma1[2]*sigma1[2]
SIGMA1[1,2] <- k1*sigma1[1]*sigma1[2]
SIGMA1[2,1] <- k1*sigma1[1]*sigma1[2]

prec1[1] ~ dgamma(2,0.5)
prec1[2] ~ dgamma(2,0.5)

```



```

k1 ~ dunif(-1,1)
sigma1[1]<-pow(prec1[1],-0.5)
sigma1[2]<-pow(prec1[2],-0.5)

# Overall sens/spec across centres in Boehme 2010
se[1]<-1/(1+exp(-l[1,1]))
sp[1]<-1/(1+exp(l[1,2]))

l[1,1:2] ~ dnorm(mu[1:2], T[1:2,1:2])

##### BOEHME 2011 #####
for(j in 1:6) {

  logit(se.r[j])<- r1[j,1]
  logit(sp.r[j])<- r1[j,2]

  r1[j,1:2]~ dnorm(l[2,1:2], T2[1:2,1:2])

  pos2[j]<-TP2[j]+FN2[j]
  neg2[j]<-TN2[j]+FP2[j]
  TP2[j] ~ dbin(se.r[j],pos2[j])
  FP2[j] ~ dbin(sp.r[j],neg2[j])

}
T2[1:2,1:2]<-inverse(SIGMA2[1:2,1:2])

# Between-centre variance-covariance matrix for Boehme 2011
SIGMA2[1,1] <- sigma2[1]*sigma2[1]
SIGMA2[2,2] <- sigma2[2]*sigma2[2]
SIGMA2[1,2] <- k2*sigma2[1]*sigma2[2]
SIGMA2[2,1] <- k2*sigma2[1]*sigma2[2]

prec2[1] ~ dgamma(2,0.5)
prec2[2] ~ dgamma(2,0.5)
k2 ~ dunif(-1,1)
sigma2[1]<-pow(prec2[1],-0.5)
sigma2[2]<-pow(prec2[2],-0.5)

# Overall sens/spec across centres in Boheme 2011
se[2]<-1/(1+exp(-l[2,1]))
sp[2]<-1/(1+exp(l[2,2]))

l[2,1:2] ~ dnorm(mu[1:2], T[1:2,1:2])

##### SINGLE CENTRE STUDIES #####
for(i in 3:15) {

  logit(se[i]) <- l[i,1]
  logit(sp[i]) <- l[i,2]

  pos[i]<-TP[i]+FN[i]
  neg[i]<-TN[i]+FP[i]
  TP[i] ~ dbin(se[i],pos[i])

```

```

        FP[i] ~ dbin(sp[i],neg[i])

        l[i,1:2] ~ dnorm(mu[1:2], T[1:2,1:2])
    }
##### HYPER PRIOR DISTRIBUTIONS #####

    mu[1] ~ dnorm(0,0.25)
    mu[2] ~ dnorm(0,0.25)

    T[1:2,1:2]<-inverse(TAU[1:2,1:2])

    # Between-study variance-covariance matrix
    TAU[1,1] <- tau[1]*tau[1]
    TAU[2,2] <- tau[2]*tau[2]
    TAU[1,2] <- rho*tau[1]*tau[2]
    TAU[2,1] <- rho*tau[1]*tau[2]

    tau[1]<-pow(prec[1],-0.5)
    tau[2]<-pow(prec[2],-0.5)

    # prec is the between-study precision in the logit(sensitivity) and logit(specificity)
    prec[1] ~ dgamma(2,0.5)
    prec[2] ~ dgamma(2,0.5)
    rho ~ dunif(-1,1)

    # Pooled sensitivity and specificity
    Pooled_S<-1/(1+exp(-mu[1]))
    Pooled_C<-1/(1+exp(mu[2]))

    # Predicted sensitivity and specificity in a new study
    l.new[1:2] ~ dnorm(mu[],T[,])
    sens.new <- 1/(1+exp(-l.new[1]))
    spec.new <- 1/(1+exp(l.new[2]))
}
##### DATA #####
# DATA WAS READ FROM THREE SEPARATE FILES

# DATA 1 - BOEHME 2010
TP1[]  FP1[]  FN1[]  TN1[]
123    1      24     68
201    0      8      101
136    1      10     185
36     3      7      215
179    0      8      35
END

#row 1 : Azerbaijan
#row 2 : Peru
#row 3 : South Africa, Cape Town
#row 4 : South Africa, Durban
#row 5 : India
#####

```

DATA 2 - FROM BOEHME 2011

TP2[]	FP2[]	FN2[]	TN2[]
203	4	26	303
171	3	6	825
201	2	32	669
121	0	24	144
101	16	0	671
136	5	12	234

END

#Boheme 2011

#row 1 : Azerbaijan

#row 2 : Peru

#row 3 : South Africa

#row 4 : Uganda

#row 5 : India

#row 6 : The Philippines

#####

DATA 3 - FROM BOEHME 2011

TP[]	FP[]	FN[]	TN[]
NA	NA	NA	NA
NA	NA	NA	NA
60	2	4	23
24	1	1	59
54	0	6	146
67	0	15	25
42	2	30	320
12	0	0	46
116	4	14	82
27	2	2	58
49	1	9	101
58	3	9	104
56	2	6	42
111	19	30	320
31	0	4	68

END

#row 1 : Boheme 2010

#row 2 : Boheme 2011

#row 3 : Bowles 2011

#row 4 : Cifci 2011

#row 5 : Hanif 2011

#row 6 : Helb 2010 a

#row 7 : Lawn 2011

#row 8 : Malbruny 2011

#row 9 : Marlowe 2011

#row 10 : Miller 2011

#row 11 : Rachow 2011

#row 12 : Scott 2011

#row 13 : Teo 2011

#row 14 : Theron 2011

#row 15 : Zeka 2011

#####

CONTRIBUTIONS OF AUTHORS

MP had the original idea for the review. KRS and MP wrote the protocol with input from HS and ND. LAK drafted the search strategy. KRS and HS reviewed articles for inclusion and extracted data. KRS, IS, and ND analysed the data. KRS, MP, and ND interpreted the analyses. KRS drafted the manuscript. ND drafted the statistical analysis section and Appendix 4. KRS, LAK, CCB, MP, and ND provided critical revisions to the manuscript.

DECLARATIONS OF INTEREST

The editorial base for the Cochrane Infectious Diseases Group (CIDG) is funded by the UK Department for International Development (DFID) for the benefit of developing countries. CIDG provided funding in part for this review. KRS serves as Co-ordinator of the Evidence Synthesis and Policy Subgroup of Stop TB Partnership's New Diagnostics Working Group. KRS received funding to carry out the review from CIDG and McGill University. MP is a recipient of a New Investigator Award from the Canadian Institutes of Health Research (CIHR) and a salary award from Fonds de recherche du Québec - Santé. MP serves as an external consultant for the Bill & Melinda Gates Foundation. CCB is employed by the Foundation for Innovative New Diagnostics (FIND) and has conducted studies and published on Xpert MTB/RIF as part of a collaborative project between FIND, a Swiss non-profit, Cepheid, a US company, and academic partners. The product developed through this partnership was developed under a contract that obligated FIND to pay for development costs and trial costs and that obligated Cepheid to make the test available at specified preferential pricing to the public sector in developing countries. The authors have no financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the review apart from those disclosed.

SOURCES OF SUPPORT

Internal sources

- Liverpool School of Tropical Medicine, UK.

External sources

- McGill University, Canada.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

In the protocol we stated that we would extract data on industry sponsorship. However, we became aware that the Foundation for Innovative New Diagnostics (FIND) had negotiated a special price for the assay for TB endemic countries. As the majority (78%) of the included study centres were located in TB endemic countries, we assumed Xpert had been purchased at the negotiated price. Therefore, we did not consider the included studies to be sponsored by industry.

We compared the accuracy of Xpert for TB detection in high-income versus low- and middle-income countries. This comparison was not mentioned in the protocol. NTMs were not mentioned in the protocol. We summarized separately data for NTM by determining the percent of false-positive Xpert results in samples that grew NTMs.

We stated we would discuss the consequences when an indeterminate test result was considered to be a (false) true negative result (may lead to missed/delayed diagnosis, with potential for increased morbidity, mortality, and TB transmission), or considered to be a (false) true positive result (may lead to unnecessary treatment with adverse events and increased anxiety). Since the rate of indeterminate results was very low, we did not discuss these consequences.

Exploration of different reference standards, culture-confirmed, clinical, etc, while an interesting and important question was beyond what we could carryout in an already complex review, with two review questions and multiple factors that could affect the results (condition of specimen, income status, clinical subgroups, etc).

We performed additional sensitivity analyses for studies that did not clearly report the reason for testing and clinical information about patients and for studies that did not explicitly report patient age.

We initially used QUADAS, as mentioned in the protocol, but while we were preparing the review, we received a communication advising use of QUADAS-2 for all future reviews. As we had received training in QUADAS-2, we decided to use QUADAS-2 for the current review.