



# Xpert MTB/RIF assay for the diagnosis of extrapulmonary tuberculosis: a systematic review and meta-analysis

Claudia M. Denkinger<sup>1,2</sup>, Samuel G. Schumacher<sup>2</sup>, Catharina C. Boehme<sup>4</sup>, Nandini Dendukuri<sup>2,3</sup>, Madhukar Pai<sup>2,3</sup> and Karen R. Steingart<sup>5</sup>

**Affiliations:** <sup>1</sup>Division of Infectious Disease, Beth Israel Deaconess Medical Centre, Boston, MA, USA. <sup>2</sup>McGill International TB Centre, and Dept of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, QC, Canada. <sup>3</sup>Respiratory Epidemiology and Clinical Research Unit, Montreal Chest Institute, Montreal, QC, Canada. <sup>4</sup>Foundation for Innovative New Diagnostics, Geneva, Switzerland. <sup>5</sup>Portland, OR, USA.

**Correspondence:** Claudia M. Denkinger, McGill International TB Centre, Dept of Epidemiology and Biostatistics, 3650 Saint-Urbain, K-3.09, Montreal, QC H2X 2P4, Canada.  
E-mail: claudia.denkinger@mail.mcgill.ca

**ABSTRACT** Xpert MTB/RIF (Cepheid, Sunnyvale, CA, USA) is endorsed for the detection of pulmonary tuberculosis (TB). We performed a systematic review and meta-analysis to assess the accuracy of Xpert for the detection of extrapulmonary TB.

We searched multiple databases to October 15, 2013. We determined the accuracy of Xpert compared with culture and a composite reference standard (CRS). We grouped data by sample type and performed meta-analyses using a bivariate random-effects model. We assessed sources of heterogeneity using meta-regression for predefined covariates.

We identified 18 studies involving 4461 samples. Sample processing varied greatly among the studies. Xpert sensitivity differed substantially between sample types. In lymph node tissues or aspirates, Xpert pooled sensitivity was 83.1% (95% CI 71.4–90.7%) versus culture and 81.2% (95% CI 72.4–87.7%) versus CRS. In cerebrospinal fluid, Xpert pooled sensitivity was 80.5% (95% CI 59.0–92.2%) against culture and 62.8% (95% CI 47.7–75.8%) against CRS. In pleural fluid, pooled sensitivity was 46.4% (95% CI 26.3–67.8%) against culture and 21.4% (95% CI 8.8–33.9%) against CRS. Xpert pooled specificity was consistently >98.7% against CRS across different sample types.

Based on this systematic review, the World Health Organization now recommends Xpert over conventional tests for diagnosis of TB in lymph nodes and other tissues, and as the preferred initial test for diagnosis of TB meningitis.



@ERSpublications

Xpert MTB/RIF can improve the diagnosis of lymph node TB and TB meningitis <http://ow.ly/uFvjZ>

This article has supplementary material available from [erj.ersjournals.com](http://erj.ersjournals.com)

Received: Jan 09 2014 | Accepted after revision: March 08 2014 | First published online: April 2 2014

**Support statement:** Development and publication of this manuscript was made possible with financial support from the US Agency for International Development. M. Pai is supported by the European and Developing Countries Clinical Trials Partnership (EDCTP-TBNEAT grant) and the Fonds de recherche du Québec – Santé (FRQS). N. Dendukuri is supported by a Chercheur Boursier salary award from the FRQS. C.M. Denkinger is supported by a Richard Tomlinson Fellowship at McGill University and a fellowship of the Burroughs–Wellcome Fund from the American Society of Tropical Medicine and Hygiene. S.G. Schumacher is supported by the Quebec Respiratory Health Training Program. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Conflict of interest:** Disclosures can be found alongside the online version of this article at [erj.ersjournals.com](http://erj.ersjournals.com)

Copyright ©ERS 2014

## Introduction

Worldwide, extrapulmonary tuberculosis (EPTB) accounts for ~25% of all TB cases, and even higher percentages in HIV-infected individuals and children [1–3]. Existing tests for the diagnosis of EPTB are limited in accuracy and time to diagnosis, and often require invasive procedures and special expertise. For pleural TB, culture of pleural fluid has low sensitivity (on average, 30–50%). For lymph node TB, culture of an aspirate has a sensitivity of 60–70% [4]. Culture specificity is 100% if the presence of *Mycobacterium tuberculosis* complex is confirmed with antigen tests or nucleic acid amplification tests (NAATs). Often, biopsy with culture and histopathological examination is necessary to achieve a diagnosis. For TB meningitis, the yield of culture is even lower (on average, 30%), although repeat examination of the cerebrospinal fluid (CSF) or the use of NAATs may increase sensitivity and is associated with high specificity (>98%) [5, 6].

The Xpert MTB/RIF assay (hereafter referred to as Xpert; Cepheid, Sunnyvale, CA, USA) is a rapid, automated molecular test with high accuracy for pulmonary TB detection (sensitivity 89%, specificity 99%) [7]. While Xpert has been approved for TB detection in sputum by regulatory agencies, Xpert for TB detection in nonrespiratory specimens is considered “off-label” use [8–10]. However, given the limitations of available tests for EPTB detection, Xpert has been evaluated in several studies.

This systematic review, commissioned by the World Health Organization (WHO) to inform the recent update of the WHO policy on Xpert, aimed to assess the diagnostic accuracy of Xpert for TB detection in nonrespiratory samples in adults and children [11].

## Methods

We developed a protocol before commencing the review, following standard guidelines [12, 13].

### Search strategy

We searched MEDLINE, EMBASE, the Cochrane Infectious Diseases Group Specialized Register and Web of Knowledge for articles between the January 1, 2007 and October 15, 2013. We reviewed reference lists of all included articles. In addition, we assessed the metaRegister of Controlled Trials and the WHO Clinical Trials Registry Platform. We also attempted to contact all authors who had published abstracts on the topic at the American Thoracic Society Conference, the European Congress of Clinical Microbiology and Infectious Diseases and the Union World Conference on Lung Health (up to 2012), and other experts in the field of TB diagnostics to identify additional studies. No language restriction was applied. The key search terms used were: “Xpert”, “GeneXpert”, “Cepheid”, “tuberculosis” and “*Mycobacterium tuberculosis*” (further details are provided in the online supplementary material).

### Inclusion criteria

We included full-text, peer-reviewed, cross-sectional studies, cohort studies, randomised controlled trials and case-control studies that used Xpert for detecting TB in nonrespiratory samples and compared it to a defined reference standard. We included studies on adults and children (aged <15 years) from all settings and countries with ≥10 nonrespiratory samples of predefined sample types (*i.e.* pleural fluid, lymph node aspirate or tissue, or CSF).

### Reference standard

The reference standards were mycobacterial culture or a composite reference standard (CRS) defined by the study authors of the individual studies. The CRS might have included a NAAT (other than Xpert), histology, smear, biochemical testing results, presenting signs and symptoms or a response to treatment with anti-TB therapy in addition to culture (table 1).

### Study selection and data extraction

Two review authors (C.M. Denkinger and S.G. Schumacher) independently assessed all articles for inclusion in the systematic review, and extracted data on study methodology, characteristics and test accuracy using a standardised extraction form (tables 1 and 2; extraction form available in the online supplementary material). We extracted data on sample processing at the study level, although some steps (*e.g.* homogenisation) might apply to only certain sample types (*e.g.* tissue). We contacted authors when information was not reported in the paper. Any differences in the study selection and data extraction process between the two review authors were resolved by discussion with a third review author (K.R. Steingart).

TABLE 1 Study characteristics

Study	First author [ref.]	Year	Country	Setting	Blinding <sup>#</sup>	Patient selection method	Prospective enrolment	HIV %	Children %	Culture reference standard	CRS	Total samples n	Specimen type (samples n)
1	AL-ATEAH [14]	2012	Saudi Arabia	Tertiary care centre	Yes	Consecutive	Yes	0	2.9	MGIT/LJ	None	67	Pleural fluid (13) Lymph node (8) CSF (14)
2	ARMAND [15]	2011	France	Tertiary care centre	Yes	Convenience	No	2	2	MGIT/LJ	None	56	Pleural fluid (8) Lymph node (20) CSF (5)
3	CAUSSE [16]	2011	Spain	Reference laboratory	Yes	Consecutive	Yes	0	15	MGIT/LJ	None	340	Pleural fluid (34) Lymph node (80) CSF (50)
4	CHRISTOPHER [17]	2013	India	Tertiary care centre	Yes	Consecutive	Yes	0	0	MGIT	Histopathology/smear+culture on pleural biopsy/biochemical analysis of fluid	96	Pleural fluid (96)
5	FRIEDRICH [18]	2011	South Africa	Tertiary care centre	Yes	Consecutive	Yes	28	0	MGIT	Histopathology/improvement on ATT/positive culture from any other site	24	Pleural fluid (24)
6	HANIF [19]	2011	Kuwait	Reference laboratory	No	Consecutive	Yes	0	0	MGIT/Bactec 460/LJ	None	29	Pleural fluid (11) Lymph node (9) CSF (5)
7	HILLEMANN [20]	2011	Germany	Reference laboratory	Yes	Consecutive	Yes	Unknown	5	MGIT/LJ	None	495	Pleural fluid (111) Lymph node (65) CSF (19)
8	LIGTHELM [21]	2011	South Africa	Tertiary care centre	Yes	Consecutive	Yes	18.8	4.2	MGIT	Histopathology	48	Lymph node (48)
9	MALBRUNY [22]	2011	France	Tertiary care centre	No	Convenience	Yes	Unknown	6.8	MGIT	None	124	Pleural fluid (12) Lymph node (23) CSF (15)
10	MOURE [23]	2012	Spain	Tertiary care centre	Yes	Convenience	No	11.2	0	MGIT/LJ	None	149	Pleural fluid (34) Lymph node (38) CSF (14) CSF (386)
11	NHU [24]	2013	Vietnam	Tertiary care centre	No	Consecutive	Yes	34.8	0	MGIT	Clinical symptoms	386	CSF (151)
12	PATEL [25]	2013	South Africa	Tertiary care centre	Yes	Consecutive	Yes	87	1.5	MGIT	NAAT	151	CSF (151)
13	PORCEL [26]	2013	Spain	Tertiary care centre	No	Consecutive	Yes	0	0	LJ (confirmed Mycobacterium tuberculosis) LJ (confirmed M. tuberculosis)	Histopathology/improvement on ATT/smear/biochemical profile/positive sputum	67	Pleural fluid (67)
14	SAFIANOWSKA [27]	2012	Poland	Tertiary care centre	No	Consecutive	Yes	0	0	MGIT/LJ	None	68	Pleural fluid (32) Lymph node (2)
15	TORTOLI [28]	2012	Italy	Tertiary care centre	Yes	Convenience	No	10	33.5	MGIT/LJ	Histopathology/improvement on ATT	1491	Pleural fluid (330) Lymph node (118) CSF (138)
16	VADWAI [29]	2011	India	Tertiary care centre	Yes	Consecutive	Yes	3	2.8	MGIT/LJ	Histopathology/improvement on ATT/radiology	342	Pleural fluid (29) Lymph node (188) Lymph node (344)
17	VAN RIE [30]	2013	South Africa	Tertiary care centre	Yes	Consecutive	Yes	100	0	MGIT	Histopathology/smear with species identification	353	Lymph node (344)
18	ZEKA [31]	2011	Turkey	Tertiary care centre	Yes	Consecutive	No	1.4	13.3	Bactec 460/LJ	Histopathology/radiology	175	Pleural fluid (56) Lymph node (26) CSF (31)

Xpert MTB/RIF is manufactured by Cepheid (Sunnyvale, CA, USA), and the Mycobacteria Growth Indicator Tube (MGIT) and Bactec 460 are manufactured by BD (East Rutherford, NJ, USA). LJ: Löwenstein-Jensen culture; CRS: composite reference standard; NAAT: nucleic acid amplification test; ATT: antituberculous therapy; NAAT: nucleic acid amplification test; CSF: cerebrospinal fluid. #: this item refers to whether the reference standard results were interpreted without knowledge of the index test results; the Xpert result, being automated, was considered blinded to the reference standard result in all studies.

TABLE 2 Processing of samples

Study	First author [ref.]	Year	Specimen condition	Homogenisation	Digestion/decontamination	Concentration (volume mL)	Resuspension (volume mL)	SR/sample ratio	Sample processing	Input volume mL
1	AL-ATEAH [14]	2012	Fresh	Mechanical	NALC/NaOH	Yes	Yes (2.5)	2/1		1
2	ARMAND [15]	2011	Frozen	Mechanical	NALC/NaOH	Yes (variable)	No	3/1		2
3	CAUSSE [16]	2011	Fresh	No	NALC/NaOH	Yes (variable)	No	2/1		1
4	CHRISTOPHER [17]	2013	Fresh	No	No	No	No	2/1		2
5	FRIEDRICH [18]	2011	Fresh	No	No	Yes (50)	Yes (2)	2/1		>2
6	HANIF [19]	2011	Fresh	No	No	No	No	2/1		2
7	HILLEMANN [20]	2011	Fresh	Mechanical (for tissue)	NALC/NaOH	Yes	Yes (1.5)	3/1		2
8	LIGTHELM [21]	2011	Fresh	No	No	No	No	2/1	If sample <0.7 mL, PBS added to increase volume to 0.7 mL	2
9	MALBRUNY [22]	2011	Fresh and frozen	Mechanical	NALC/NaOH	Yes (variable)	No	3/1	All samples concentrated except for CSF	0.5–1
10	MOURE [23]	2012	Frozen	Mechanical	NALC/NaOH	Yes (variable)	Yes (2)	2/1	3 mL added to cartridge; 1 mL sediment, 2 mL SR	3
11	NHU [24]	2013	Fresh	No	No	Yes (5–8)	Yes (0.7)	2/1	200 µL deposit was taken and 500 µL PBS added	2
12	PATEL [25]	2013	Frozen	No	No	No	No	2/1		2
13	PORCEL [26]	2013	Frozen	No	No	Yes (5)	Yes (1)	2/1		2
14	SAFIANOWSKA [27]	2012	Fresh	Mechanical	NALC/NaOH	Yes (see sample processing)	Yes (2.5)	3/1	Centrifugation, followed by additional washing step with 50 mL distilled water and repeat centrifugation	2
15	TORTOLI [28]	2012	Frozen	Mechanical	NALC/NaOH	Yes (variable)	Yes (2)	2/1	Decontamination only for nonsterile samples	2
16	VADWAI [29]	2011	Fresh	Mechanical (for tissue)	No	No	No	2/1	SR/sample 2/1 except for CSF, which was usually <1 mL and increased to 2 mL with SR	2
17	VAN RIE [30]	2013	Fresh	No	No	No	Yes (4)	2/1		2
18	ZEKA [31]	2011	Frozen	No	No	No	No	3/1		2

SR: sample reagent; NALC: N-acetyl-L-cysteine; CSF: cerebrospinal fluid.

### Assessment of methodological quality

We grouped studies according to the type of reference standard (culture *versus* CRS) and assessed study quality separately for the two groups using QUADAS-2, a validated tool for diagnostic studies [32]. The QUADAS-2 protocol is available in the online supplementary material.

### Statistical analysis and data synthesis

We performed descriptive analyses using STATA 12 (STATA Corporation, College Station, TX, USA). For each study, we calculated Xpert sensitivity and specificity along with 95% confidence intervals, compared with culture or CRS, and generated forest plots to display sensitivity and specificity estimates using Review Manager 5.2 (Nordic Cochrane Centre, Copenhagen, Denmark).

### Imperfect reference standard

In diagnostic accuracy studies, an imperfect reference standard may lead to a misclassification of samples [33, 34]. Culture is an imperfect reference standard for EPTB due to the paucibacillary nature of the disease. Assuming that Xpert correctly identifies TB in a sample with a negative culture, the result would appear to be a false positive, leading to an underestimation of Xpert's true specificity. A CRS that classifies TB based on a positive result of one out of several tests or clinical components may sometimes reclassify false positives of Xpert (identified as non-TB using culture) as true positives (TB cases) and thus lead to an increased (*i.e.* more accurate) estimate of Xpert specificity. However, a CRS itself may have reduced specificity that could result in apparent false-negative Xpert results, leading to an underestimation of Xpert's true sensitivity [35, 36]. Therefore, a comparison of the accuracy estimates based on these two reference standards, culture and CRS, should provide a plausible range for sensitivity and specificity.

We excluded noninterpretable test results (*i.e.* invalid or erroneous results) from the analyses for determination of sensitivity and specificity [37].

### Meta-analysis

We assessed heterogeneity in the forest plots (by visually examining the confidence intervals of individual studies) and in summary plots (by examining the width of the prediction region, with a wider prediction region suggesting more heterogeneity). We expected heterogeneity in terms of the sample types. Therefore, we pre-specified subgroups by sample type: pleural fluid, lymph node samples (tissue and aspirate combined; reporting in studies did not allow separating tissue samples from aspirates) and CSF. We used a bivariate random-effects model and carried out meta-analyses using the `metandi` command in STATA [38]. A meta-analysis for a predefined sample type was only carried out if at least four studies were available [39].

Several studies did not contribute to both sensitivity (no true positives or false negatives) and specificity, but only to specificity. In such cases, we performed a univariate random-effects meta-analysis of the specificity estimates separately, so as to make complete use of the available data. We compared the specificity estimate from the univariate analysis with that from a bivariate analysis of the subset of studies that provided both sensitivity and specificity.

### Meta-regression

We anticipated additional heterogeneity with respect to sample processing methods, the condition of samples, laboratory level and HIV prevalence within the predefined subgroups. Therefore, we chose to use a bivariate meta-regression model (command: `mvmeta`) in STATA [40] under the assumption that the pooled sensitivity and specificity were different in each subgroup, but not the between-study variance-covariance matrix. We performed the meta-regression assessing only studies with a culture reference standard because the number of studies using CRS was limited. We also presumed that the effect of the covariate would not differ between the different reference standards. We focused on three categorical covariates (further subgroup analysis was not feasible because of the limited number of studies): concentration step (yes or no), condition of sample (fresh *versus* frozen) and HIV prevalence ( $>10$  or  $\leq 10\%$ ).

### Sensitivity analysis

In addition, we performed two sensitivity analyses by limiting inclusion in the meta-analysis to: 1) studies in which patients were selected consecutively; or 2) studies that did not use a case-control design.

We chose not to carry out formal assessment of publication bias using methods such as funnel plots or regression tests because such techniques are not considered to be valid for diagnostic accuracy reviews [13]. The sponsors of the study had no role in study design, data collection, data analysis, data interpretation or writing of the report.

## Results

We identified 18 studies (fig. 1) [14–31] that included 4461 samples. TB prevalence (based on culture) ranged from 0% to 81%. All studies were written in English. Eight (44%) studies were conducted in low/middle-income countries (table 1). Six studies did not include any HIV-positive patients and for two studies, HIV status was unknown [20, 32]. One study only included HIV-positive patients [30]. The percentages of HIV-positive patients included in remaining studies ranged from 1% to 87% of the study population (table 1). 10 studies included children, with percentages ranging from 2% to 34%.

The median number of samples per study was 137 (interquartile range 67–342). Seven studies included only one sample type (e.g. pleural fluid only) [17, 18, 21, 25, 26, 30, 39] (table 1). The remainder of the studies included different sample types in varying percentages. Six studies used archived frozen samples, 11 used fresh samples, and one used both fresh and frozen samples (table 2).

With respect to specimen processing, the studies varied widely (table 2). Only four (25%) studies used the protocol (i.e. volume of sample included and sample reagent to sample ratio) recommended by the manufacturer [36] for unprocessed sputum samples [17, 19, 29, 25] (table 2). Most of the studies that included a mechanical homogenisation step (40%) also performed a decontamination procedure with *N*-acetyl-L-cysteine and sodium hydroxide solution. 11 (55%) studies reported a concentration step (table 2). The sample reagent/sample volume ratio also varied. Five (25%) studies used a ratio of 3/1, while the remainder of studies used 2/1. Of the eight studies that used a digestion/decontamination step, four studies used a ratio of 2/1 (table 2).

### Methodological quality of studies

The overall methodological quality of the included studies using a culture reference standard is summarised in figure 2 (additional information on the quality assessment for each study individually and using a culture reference standard is given in the online supplementary material). The majority of studies collected data prospectively (n=14; 78%) (table 1) and only three studies used a case-control design [15, 23, 26]. All studies were performed either in tertiary care centres or reference laboratories.

We considered that differences in sample processing might affect estimates of the diagnostic accuracy of Xpert to varying degrees (table 2). In particular, we were concerned that the mechanical homogenisation step could be a source of variation in test accuracy (n=8, 40%) for two reasons: procedural differences might affect the quantity of sample particles in the sample volume and the particles could clog the cartridge valves leading to noninterpretable results. We also considered the possibility that the reference standard could introduce bias due to misclassification of participants (fig. 2) [34, 35].

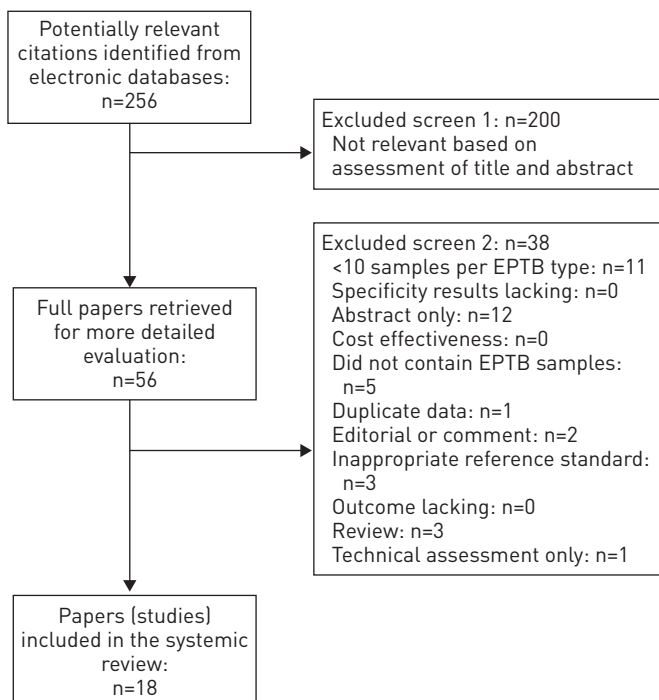


FIGURE 1 Flow diagram of studies in the review. EPTB: extrapulmonary tuberculosis.

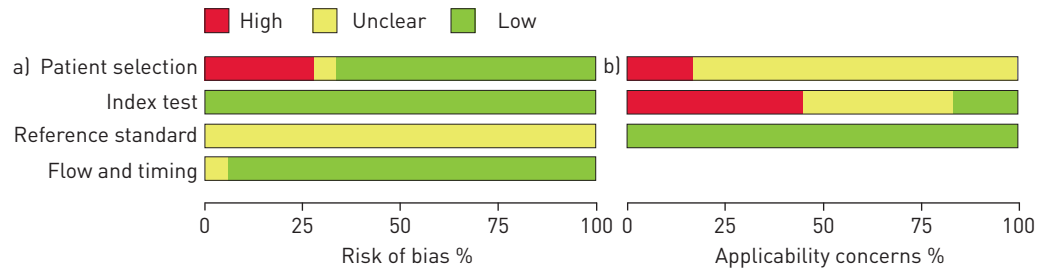


FIGURE 2 a) Risk of bias and b) applicability concerns as percentages across the included studies using a culture-based reference standard for tuberculosis detection.

**Detection of EPTB**

Studies were very heterogeneous, particularly among smear-negative samples; therefore, combining studies to obtain accuracy estimates of Xpert for EPTB (all forms combined) was not considered to be meaningful. For smear-positive samples (506 samples), a univariate analysis yielded a sensitivity of 97.4% (95% CI 95.5–99.3%) across sample types. Data were too limited to estimate specificity.

We focused the remainder of the analysis on predefined subgroups of sample types (*i.e.* pleural fluid, lymph node aspirate or tissue, and CSF) to account for the heterogeneity. Data for smear status of samples were not available for the individual sample types. Therefore, samples included in the subgroups may be smear-positive, negative or unknown.

**Detection of lymph node TB**

For studies [14–16, 19–23, 27–31] that evaluated lymph node biopsy or fine-needle aspirate using a culture reference standard (13 total, 10 with >10 samples; 955 samples, 362 culture-positive), Xpert sensitivity ranged from 50% to 100% (fig. 3a). Pooled sensitivity across studies was 83.1% (95% CI 71.4–90.7%) and

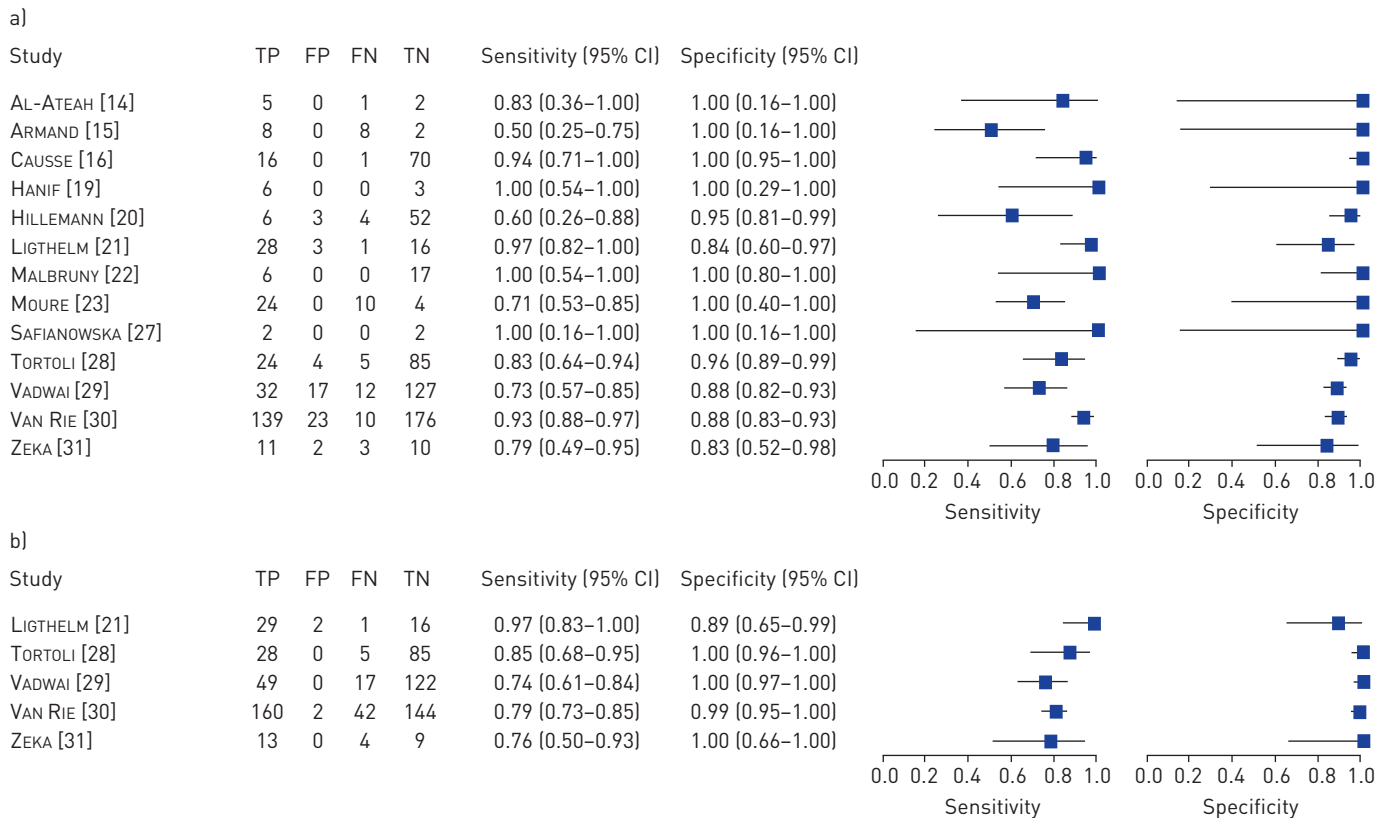


FIGURE 3 Forest plot of Xpert sensitivity and specificity for tuberculosis detection in lymph node samples (tissue or aspirate) with a) culture reference standard and b) composite reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. TP: true positive; FP: false positive; FN: false negative; TN: true negative.



pooled specificity was 93.6% (95% CI 87.9–96.8%). Only two studies reported any noninterpretable results for Xpert: 1.4% (five out of 353) and 10% (two out of 20) [15, 30].

Five studies [21, 28, 29–31] assessed Xpert in lymph node samples against a CRS (fig. 3b). Pooled sensitivity was 81.2% (95% CI 72.4–87.7%). As expected, the specificity was improved, at 99.1% (95% CI 94.5–99.9%), in comparison to the culture reference standard.

Studies that used fresh samples showed a higher sensitivity (86.4%, 95% CI 75.7–97.1%) than those that used frozen samples (74.0%, 95% CI 56.5–91.5%); however, the precision of these estimates was low as data were limited. Only three studies included >10% HIV-positive patients. Accuracy estimates did not differ substantially between these studies and others that included <10% HIV patients.

**Detection of pleural TB**

14 studies [14–20, 22, 23, 26–29, 31] (841 samples in total, 92 culture positive) evaluated Xpert in pleural fluid versus culture. Xpert sensitivity varied widely (0–100%) (fig. 4a). The outliers at the lower and upper ends of the range were studies with few culture-confirmed TB cases. The pooled sensitivity was 46.4% (95% CI 26.3–67.8%) and the pooled specificity was 99.1% (95% CI 95.2–99.8%).

In comparison with the bivariate analysis, a univariate analysis for specificity achieved a similar estimate (98.9%) with a slightly narrower confidence interval (97.9–99.8%). One study reported noninterpretable results for Xpert: 5.4% (six out of 111) [20].

Six studies (598 samples) [17, 18, 26, 28, 29, 31] evaluated Xpert in pleural fluid versus CRS. Only a univariate analysis was feasible given the limited data. Compared with the pooled estimate with culture as the reference standard, the CRS subgroup yielded a lower sensitivity (21.4%, 95% CI 8.8–33.9%) with a slightly higher specificity of 100% (95% CI 99.4–100%) (fig. 4b).

Sensitivity was increased in studies with a low rate of HIV co-infection (49.5% compared with 40.6% in studies with >10% HIV) and in studies that used a concentration step (49.1% compared with 41.6%

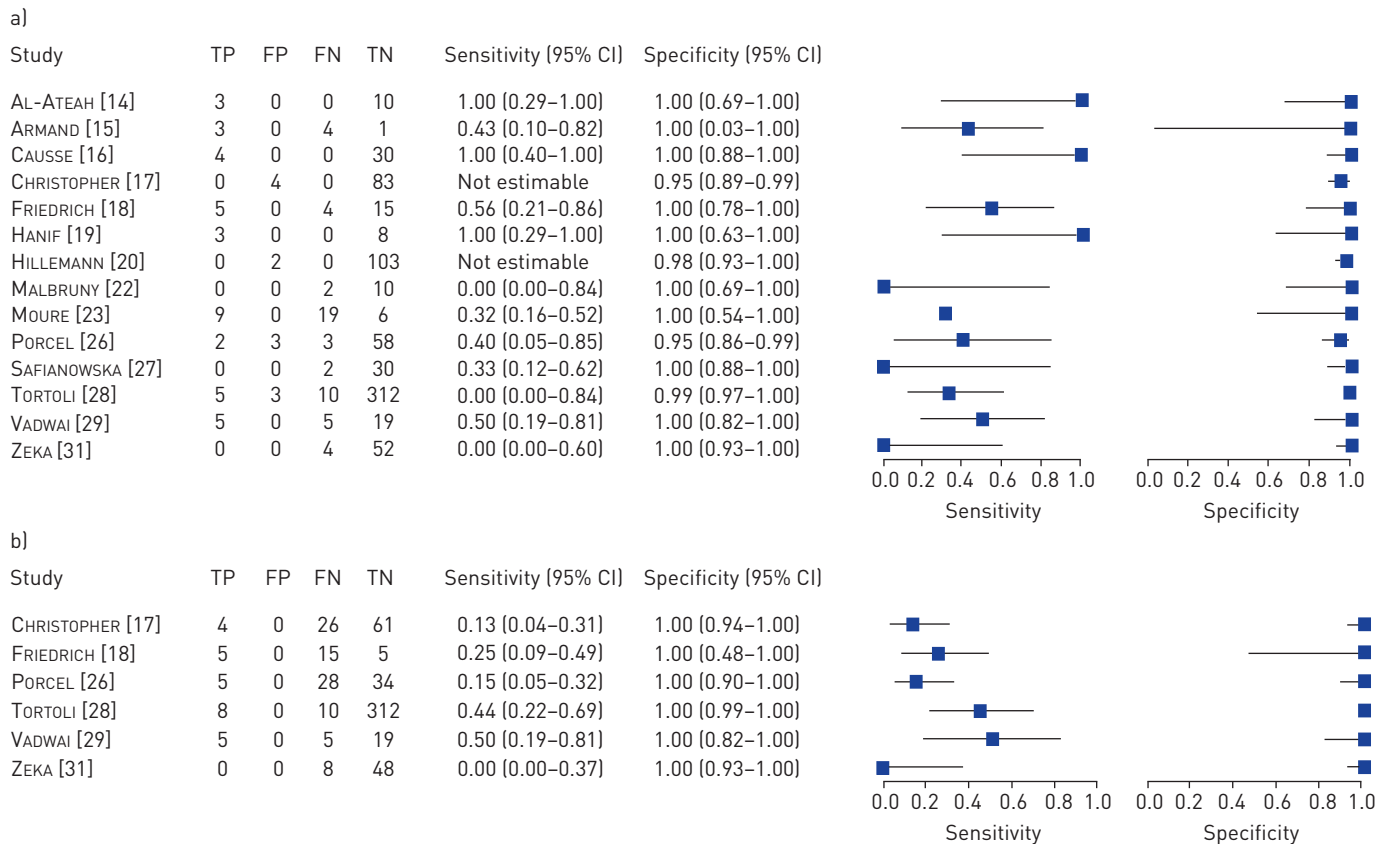


FIGURE 4 Forest plot of Xpert sensitivity and specificity for tuberculosis detection in pleural fluid with a) culture reference standard and b) composite reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. TP: true positive; FP: false positive; FN: false negative; TN: true negative.



without concentration step). A higher sensitivity was also observed for fresh samples (59.0%, 95% CI 41.0–76.9%) compared with frozen samples (31.4%, 95% CI 18.9–43.9%); however, the estimates for all of these subgroup analyses were imprecise and the confidence intervals were wide and overlapping.

#### Detection of TB meningitis

13 studies (839 samples; 10 with >10 samples, 159 culture positive) evaluated Xpert in CSF against culture [14–16, 19, 20, 22–25, 27–29, 31]. Sensitivity varied widely (51–100%), with the study by VADWAI *et al.* [29] (sensitivity none out of 196 samples, three false negatives) considered an outlier (fig. 5a). Pooled sensitivity was 80.5% (95% CI 59.0–92.2%) and pooled specificity was 97.8% (95% CI 95.2–99.0%). Noninterpretable results for Xpert were reported in three studies, with only one study [28] having >2% (*i.e.* 3.6%).

Five studies (711 samples) that assessed Xpert in CSF samples *versus* a CRS found variable sensitivity (20–86%) (fig. 5b) [24, 25, 28–29, 31]. Pooled sensitivity was 62.8% (95% CI 47.7–75.8%) and pooled specificity was 98.8% (95% CI 95.7–100%).

Prevalence of HIV and the condition of the specimen did not have an effect on Xpert sensitivity and specificity in CSF. However, a concentration step in the processing of the sample (table 2) appeared to enhance the sensitivity of Xpert (84.2% (95% CI 78.3–90.1%) *versus* 51.3% (95% CI 35.5–67.1%) for unconcentrated samples; specificity 98.0% (95% CI 96.7–99.2%) *versus* 94.6% (95% CI 90.9–98.2%) for unconcentrated samples).

Sensitivity analyses across all samples types did not substantially affect the results (online supplementary material).

#### Discussion

Our systematic review demonstrated that Xpert sensitivity for TB detection in nonrespiratory samples varied widely across different sample types. While Xpert is a highly sensitive diagnostic for TB detection in lymph node samples and moderately sensitive for the detection of TB meningitis, our results show lower

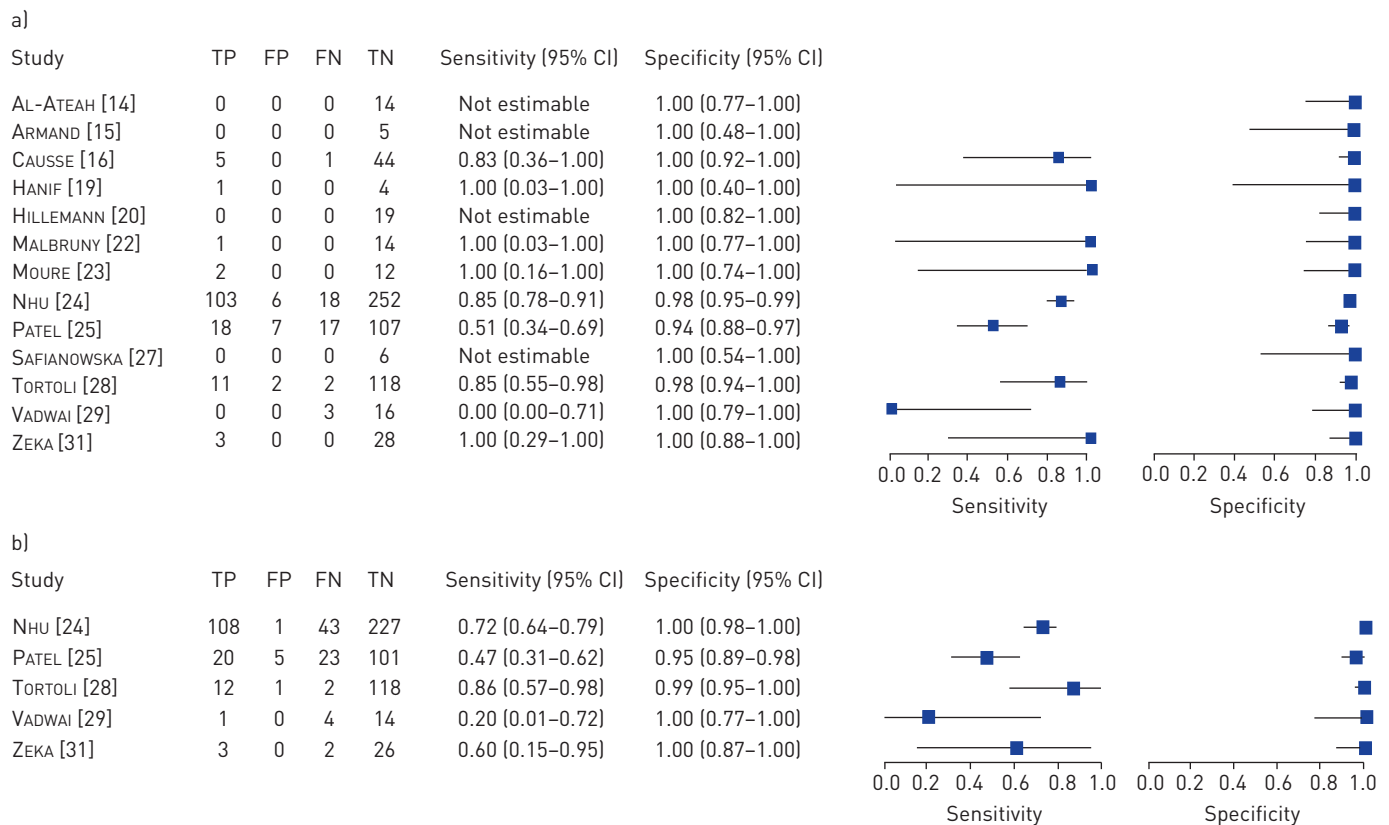


FIGURE 5 Forest plot of Xpert sensitivity and specificity for tuberculosis detection in cerebrospinal fluid with a) culture reference standard and b) composite reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. TP: true positive; FP: false positive; FN: false negative; TN: true negative.

sensitivity for testing pleural fluid. Figure 6 shows the pooled sensitivity estimates for different sample types (pooled specificity estimates are online supplementary material).

The high sensitivity (97.4%, 95% CI 95.5–99.3%) of Xpert in smear-positive samples across sample types and the low proportion of noninterpretable results (1.2%) support the use of the test in nonrespiratory samples in principle. The poor sensitivity of Xpert in pleural fluid is probably due to the paucibacillary nature of the disease and the fact that not pleural fluid but rather pleural biopsy is the sample of choice for the diagnosis of pleural TB (as has been described for culture) [41]. The presence of PCR inhibitors, either in the pleural fluid itself or from blood contamination of the sample, could be considered as well [42, 43]. Where resources are available, Xpert on pleural fluid could still be considered in the work-up of pleural TB as it has higher sensitivity than smear and provides a more rapid diagnosis than culture and histology.

Prior data have suggested a potential role for NAATs in the diagnosis of TB from CSF and lymph node samples, and the results for Xpert here confirm these findings [4, 6]. Interestingly, we observed that a concentration step for CSF increases Xpert sensitivity with unchanged specificity, probably by increasing the bacillary load in the cartridge input volume. While Xpert does not reach the sensitivity of culture, it could improve the diagnosis of CSF and lymph node TB in places where culture or other diagnostic tests are not available or where a rapid diagnosis of TB is necessary (as might be the case for TB meningitis).

The combined confidence intervals for sensitivity and specificity of Xpert *versus* culture and CRS (which includes culture) provide a range in which the “true” sensitivity and specificity are likely to fall. Future analyses could employ statistical models, such as latent class models, that incorporate knowledge about the imperfect accuracy of reference standards to provide a single plausible estimate for the accuracy of Xpert [36].

Strengths of our review include the use of a standard protocol, strict inclusion criteria, standardised data extraction, independent reviewers, a bivariate random-effects model for meta-analysis and pre-specified subgroups to account for heterogeneity. This data set involved comprehensive searching to identify studies as well as repeated correspondence with study authors to obtain additional data on the studies.

Our review also had several limitations. We acknowledge that we may have missed some studies despite the comprehensive search. In addition, the meta-analysis was limited by the small number of studies for the different sample types, particularly those using a CRS. In addition, low event rates (*i.e.* confirmed TB cases) limited the precision of our sensitivity estimates. Furthermore, sample processing was highly variable across and within studies, as there was no recommendation available on how to process nonrespiratory samples from the manufacturer or the WHO. Also, the CRS differed between studies. Because of this heterogeneity, the pooled estimates must be interpreted with caution.

The pooled sensitivity and specificity estimates in our meta-analysis might be overly optimistic for at least three reasons. First, the quality of some studies might have suffered from a lack of a representative patient spectrum (*e.g.* studies using convenience sampling). Second, all of the studies were performed in tertiary care centres or reference laboratories. Third, publication bias must be considered [44]. However, 44% of studies were performed in low-prevalence settings. It is likely that patients in these settings present earlier with more paucibacillary disease, which might result in a decreased sensitivity of Xpert [45].

Given the limited data for the sample types and the large variation in sample processing, a detailed investigation of the best sample processing was not possible in this review. An optimised processing

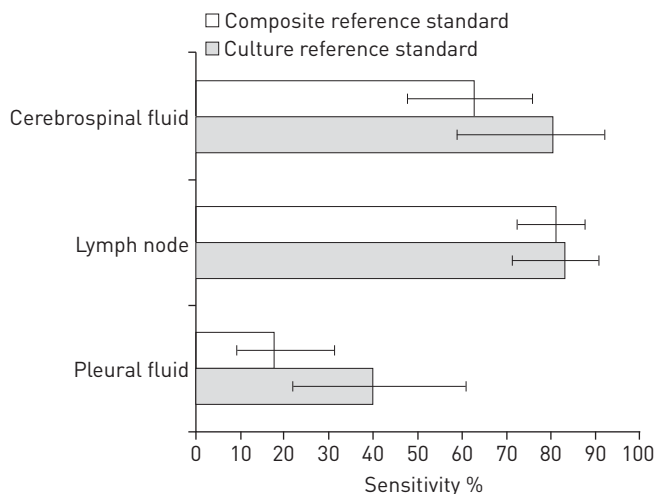


FIGURE 6 Pooled sensitivity estimates across sample types.

procedure might also need to be different for different sample types and might further improve Xpert performance. We therefore would encourage studies to focus on optimisation of sample preparation. Furthermore, additional research is needed on the diagnostic accuracy of Xpert on samples other than those assessed in this review (e.g. blood), and on the impact on patient-important outcomes.

Our review findings have informed an updated WHO policy on Xpert for EPTB [11]. WHO now recommends Xpert over conventional tests for the diagnosis of TB in lymph nodes and other tissues, and as the preferred initial test for the diagnosis of TB meningitis. A draft technical advisory document on the standard operating procedure for sample processing is now available from the WHO [46].

### Acknowledgements

We wish to thank the following individuals for their contributions to this project: Vittoria Lutje (Liverpool School of Tropical Medicine, Liverpool, UK) and Matteo Zignol (WHO, Geneva, Switzerland), and all the authors of the studies analysed.

### References

- 1 World Health Organization. Global tuberculosis report 2012, Geneva, WHO, 2012.
- 2 Jones BE, Young SM, Antoniskis D, *et al.* Relationship of the manifestations of tuberculosis to CD4 cell counts in patients with human immunodeficiency virus infection. *Am Rev Respir Dis* 1993; 148: 1292–1297.
- 3 Nelson LJ, Wells CD. Global epidemiology of childhood tuberculosis. *Int J Tuberc Lung Dis* 2004; 8: 636–647.
- 4 Dinnes J, Deeks J, Kunst H, *et al.* A systematic review of rapid diagnostic tests for the detection of tuberculosis infection. *Health Technol Assess* 2007; 11: 1–196.
- 5 Marais S, Thwaites G, Schoeman JF, *et al.* Tuberculous meningitis: a uniform case definition for use in clinical research. *Lancet Infect Dis* 2010; 10: 803–812.
- 6 Pai M, Flores LL, Pai N, *et al.* Diagnostic accuracy of nucleic acid amplification tests for tuberculous meningitis: a systematic review and meta-analysis. *Lancet Infect Dis* 2003; 3: 633–643.
- 7 Steingart KR, Schiller I, Horne DJ, *et al.* Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database Syst Rev* 2014; 1: CD009593.
- 8 World Health Organization. Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF system, Geneva, WHO, 2011.
- 9 US Food and Drug Administration. Press release: FDA permits marketing of first U.S. test labeled for simultaneous detection of tuberculosis bacteria and resistance to the antibiotic rifampin. [www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm362602.htm](http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm362602.htm) Date last updated: July 25, 2013.
- 10 Weyer K, Mirzayev F, Migliori GB, *et al.* Rapid molecular TB diagnosis: evidence, policy making and global implementation of Xpert MTB/RIF. *Eur Respir J* 2013; 42: 252–271.
- 11 World Health Organization. Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF system for the diagnosis of pulmonary and extrapulmonary TB in adults and children. Geneva, WHO, 2013.
- 12 Leeflang MM, Deeks JJ, Gatsonis C, *et al.* Systematic reviews of diagnostic test accuracy. *Ann Intern Med* 2008; 149: 889–897.
- 13 Macaskill P, Gatsonis C, Deeks JJ, *et al.* Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version. London, The Cochrane Collaboration, 2010.
- 14 Al-Ateah SM, Al-Dowaidi MM, El-Khizzi NA. Evaluation of direct detection of *Mycobacterium tuberculosis* complex in respiratory and non-respiratory clinical specimens using the Cepheid Gene Xpert® system. *Saudi Med J* 2012; 33: 1100–1105.
- 15 Armand S, Vanhuls P, Delcroix G, *et al.* Comparison of the Xpert MTB/RIF test with an IS6110-TaqMan real-time PCR assay for direct detection of *Mycobacterium tuberculosis* in respiratory and nonrespiratory specimens. *J Clin Microbiol* 2011; 49: 1772–1776.
- 16 Causse M, Ruiz P, Gutierrez-Aroca JB, *et al.* Comparison of two molecular methods for rapid diagnosis of extrapulmonary tuberculosis. *J Clin Microbiol* 2011; 49: 3065–3067.
- 17 Christopher DJ, Schumacher SG, Michael JS, *et al.* Performance of Xpert MTB/RIF on pleural tissue for the diagnosis of pleural tuberculosis. *Eur Respir J* 2013; 42: 1427–1429.
- 18 Friedrich SO, von Groote-Bidlingmaier F, Diacon AH. Xpert MTB/RIF assay for diagnosis of pleural tuberculosis. *J Clin Microbiol* 2011; 49: 4341–4342.
- 19 Hanif SN, Eldeen HS, Ahmad S, *et al.* GeneXpert® MTB/RIF for rapid detection of *Mycobacterium tuberculosis* in pulmonary and extra-pulmonary samples. *Int J Tuberc Lung Dis* 2011; 15: 1274–1275.
- 20 Hillemann D, Ruesch-Gerdes S, Boehme C, *et al.* Rapid molecular detection of extrapulmonary tuberculosis by the automated GeneXpert MTB/RIF System. *J Clin Microbiol* 2011; 49: 1202–1205.
- 21 Ligthelm LJ, Nicol MP, Hoek KG, *et al.* Xpert MTB/RIF for rapid diagnosis of tuberculous lymphadenitis from fine-needle-aspiration biopsy specimens. *J Clin Microbiol* 2011; 49: 3967–3970.
- 22 Malbruny B, Le Marrec G, Courageux K, *et al.* Rapid and efficient detection of *Mycobacterium tuberculosis* in respiratory and non-respiratory samples. *Int J Tuberc Lung Dis* 2011; 15: 553–555.
- 23 Moure R, Martin R, Alcaide F. Effectiveness of an integrated real-time PCR method for detection of the *Mycobacterium tuberculosis* complex in smear-negative extrapulmonary samples in an area of low tuberculosis prevalence. *J Clin Microbiol* 2012; 50: 513–515.
- 24 Nhu NT, Heemskerk D, Thu do DA, *et al.* Evaluation of GeneXpert MTB/RIF for diagnosis of tuberculous meningitis. *J Clin Microbiol* 2014; 52: 226–233.
- 25 Patel VB, Theron G, Lenders L, *et al.* Diagnostic accuracy of quantitative PCR (Xpert MTB/RIF) for tuberculous meningitis in a high burden setting: a prospective study. *PLoS Med* 2013; 10: e1001536.
- 26 Porcel JM, Palma R, Valdes L, *et al.* Xpert® MTB/RIF in pleural fluid for the diagnosis of tuberculosis. *Int J Tuberc Lung Dis* 2013; 17: 1217–1279.

- 27 Safianowska A, Walkiewicz R, Nejman-Gryz P, *et al.* Zastosowanie wybranych komercyjnych testów molekularnych w mikrobiologicznej diagnostyce gruźlicy [Two selected commercially based nucleic acid amplification tests for the diagnosis of tuberculosis]. *Pneumonol Alergol Pol* 2012; 80: 6–12.
- 28 Tortoli E, Russo C, Piersimoni C, *et al.* Clinical validation of Xpert MTB/RIF for the diagnosis of extrapulmonary tuberculosis. *Eur Respir J* 2012; 40: 442–447.
- 29 Vadwai V, Boehme C, Nabeta P, *et al.* Xpert MTB/RIF: a new pillar in diagnosis of extrapulmonary tuberculosis? *J Clin Microbiol* 2011; 49: 2540–2545.
- 30 Van Rie A, Page-Shipp L, Mellet K, *et al.* Diagnostic accuracy and effectiveness of the Xpert MTB/RIF assay for the diagnosis of HIV-associated lymph node tuberculosis. *Eur J Clin Microbiol Infect Dis* 2013; 32: 1409–1415.
- 31 Zeka AN, Tasbakan S, Cavusoglu C. Evaluation of the GeneXpert MTB/RIF assay for rapid diagnosis of tuberculosis and detection of rifampin resistance in pulmonary and extrapulmonary specimens. *J Clin Microbiol* 2011; 49: 4138–4141.
- 32 Whiting PF, Rutjes AW, Westwood ME, *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155: 529–536.
- 33 Valenstein PN. Evaluating diagnostic tests with imperfect standards. *Am J Clin Pathol* 1990; 93: 252–258.
- 34 Naaktgeboren CA, Bertens LC, Smeden M, *et al.* Value of composite reference standards in diagnostic research. *BMJ* 2013; 347: f5605.
- 35 Reitsma JB, Rutjes AW, Khan KS, *et al.* A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *J Clin Epidemiol* 2009; 62: 797–806.
- 36 Dendukuri N, Schiller I, Joseph L, *et al.* Bayesian meta-analysis of the accuracy of a test for tuberculous pleuritis in the absence of a gold standard reference. *Biometrics* 2012; 68: 1285–1293.
- 37 Cepheid. Xpert MTB/RIF. [http://tbevidence.org/documents/rescentre/sop/XpertMTB\\_Broch\\_R9\\_EU.pdf](http://tbevidence.org/documents/rescentre/sop/XpertMTB_Broch_R9_EU.pdf)
- 38 Reitsma JB, Glas AS, Rutjes AW, *et al.* Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; 58: 982–990.
- 39 Harbord R. metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. *STATA J* 2009; 9: 211–229.
- 40 White IR. Multivariate random-effects meta-regression: Updates to mvmeta. *STATA J* 2011; 11: 255–270.
- 41 Porcel JM. Tuberculous pleural effusion. *Lung* 2009; 187: 263–270.
- 42 Pai M, Flores LL, Hubbard A, *et al.* Nucleic acid amplification tests in the diagnosis of tuberculous pleuritis: a systematic review and meta-analysis. *BMC Infect Dis* 2004; 4: 6.
- 43 Woods GL. Molecular techniques in mycobacterial detection. *Arch Pathol Lab Med* 2001; 125: 122–126.
- 44 Lijmer JG, Mol BW, Heisterkamp S, *et al.* Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999; 282: 1061–1066.
- 45 Sohn H, Aero AD, Menzies D, *et al.* Xpert MTB/RIF testing in a low TB incidence, high-resource setting: limitations in accuracy and clinical impact. *Clin Infect Dis* 2014; 58: 970–976.
- 46 World Health Organization. Standard Operating Procedure (SOP): Specimen processing of CSF, lymph nodes and other tissues for Xpert MTB/RIF. [www.ghdonline.org/uploads/GeneXpert\\_SOP\\_Xpert\\_processing\\_EPTB\\_specimens\\_DRAFT.pdf](http://www.ghdonline.org/uploads/GeneXpert_SOP_Xpert_processing_EPTB_specimens_DRAFT.pdf) Date last accessed: February 24, 2014.