Cochrane Library

Cochrane Database of Systematic Reviews

Low-complexity automated nucleic acid amplification tests for extrapulmonary tuberculosis and rifampicin resistance in adults and adolescents (Review)

Kohli M, Inbaraj LR, Salomon A, Scandrett K, Korobitsyn A, Ismail N, Srinivasalu VA, Daniel J, Steingart KR, Takwoingi Y

Kohli M, Inbaraj LR, Salomon A, Scandrett K, Korobitsyn A, Ismail N, Srinivasalu VA, Daniel J, Steingart KR, Takwoingi Y.
Low-complexity automated nucleic acid amplification tests for extrapulmonary tuberculosis and rifampicin resistance in adults and adolescents.
Cochrane Database of Systematic Reviews 2025, Issue 8. Art. No.: CD012768.
DOI: 10.1002/14651858.CD012768.pub4.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Review]

Low-complexity automated nucleic acid amplification tests for extrapulmonary tuberculosis and rifampicin resistance in adults and adolescents

Mikashmi Kohli[1], Leeberk Raja Inbaraj[2], Angela Salomon[3], Katie Scandrett[4], Alexei Korobitsyn[5], Nazir Ismail[5,6], Vignes Anand Srinivasalu[7], Jefferson Daniel[8], Karen R Steingart[9], Yemisi Takwoingi[4]

[1]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada. [2]Clinical Research, IMCR-Indian Council of Medical Research, Chennai, India. [3]Department of Family Medicine, McGill University, Montreal, Canada. [4]Department of Applied Health Sciences, University of Birmingham, Birmingham, UK. [5]Global Tuberculosis Programme, World Health Organization, Geneva, Switzerland. [6]Clinical Microbiology and Infectious Diseases, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. [7]Clinical Research, ICMR-National Institute for Research in Tuberculosis, Chennai, India. [8]Pulmonary Medicine, Christian Medical College, Vellore, India. [9]Honorary Research Fellow (retired), Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, UK

Contact: Mikashmi Kohli, mikashmikohli@gmail.com.

Editorial group: Cochrane Central Editorial Service.
Publication status and date: New search for studies and content updated (no change to conclusions), published in Issue 8, 2025.

Citation: Kohli M, Inbaraj LR, Salomon A, Scandrett K, Korobitsyn A, Ismail N, Srinivasalu VA, Daniel J, Steingart KR, Takwoingi Y. Low-complexity automated nucleic acid amplification tests for extrapulmonary tuberculosis and rifampicin resistance in adults and adolescents. *Cochrane Database of Systematic Reviews* 2025, Issue 8. Art. No.: CD012768. DOI: 10.1002/14651858.CD012768.pub4.

ABSTRACT

Background

Low-complexity automated nucleic acid amplification tests **(**LC-aNAATs) are molecular World Health Organization (WHO)-recommended rapid diagnostic tests widely used for simultaneous detection of *Mycobacterium tuberculosis* complex and rifampicin resistance in sputum. To extend our previous review on extrapulmonary tuberculosis, we performed this update to inform a WHO policy update.

Objectives

To estimate the diagnostic accuracy of LC-aNAATs for extrapulmonary tuberculosis and rifampicin resistance in adults and adolescents with presumptive extrapulmonary tuberculosis.

Search methods

We searched the Cochrane Central Register of Controlled Trials, MEDLINE, Embase, Science Citation Index, Latin American Caribbean Health Sciences Literature, Scopus, ClinicalTrials.gov, the WHO International Clinical Trials Registry Platform, the International Standard Randomized Controlled Trial Number Registry, and ProQuest, up to 11 October 2023, without language restriction. A WHO public call for data was made between 30th November 2023 and 15th February 2024 to identify unpublished studies.

Selection criteria

We included cross-sectional and cohort studies using non-respiratory specimens and eight forms of extrapulmonary tuberculosis: tuberculous meningitis and pleural, lymph node, bone or joint, genitourinary, peritoneal, pericardial, and disseminated tuberculosis. Reference standards were culture and a study-defined composite reference standard (tuberculosis detection); and phenotypic drug

susceptibility testing with or without genotypic drug susceptibility testing (rifampicin resistance detection). Index tests included Xpert Ultra, Truenat assays, STANDARD M10, and Iron qPCR.

Data collection and analysis

Two review authors independently extracted data and assessed the risk of bias and applicability using the QUADAS-2 tool. For tuberculosis detection, we performed separate analyses by specimen type and reference standard using the bivariate model to estimate summary sensitivity and specificity with 95% confidence intervals (CIs). Based on a pre-defined condition, based on sample sizes and type of technology for performing class-based analysis, data for Truenat MTB Plus were not included in the meta-analyses for LC-aNAATs. Hence, we present results for Xpert Ultra and Truenat MTB Plus separately. We assessed the certainty of evidence using the GRADE approach.

Main results

We included 37 unique studies where 36 studies evaluated Xpert Ultra and three studies evaluated Truenat MTB plus. We found no eligible studies for the other index tests. Overall, the risk of bias was low for patient selection, index test, and flow and timing domains. For the reference standard, the risk of bias for included studies was low (75%) or unclear (25%). Applicability for the patient selection domain was unclear for most studies because we were unsure of the clinical settings, and the applicability concern was low for most studies for the reference standard domain.

Cerebrospinal fluid

Xpert Ultra (16 studies)

Xpert Ultra summary sensitivity and specificity (95% CI) against a microbiological reference standard were 88.2% (83.7 to 91.6) (287 participants; high-certainty evidence) and 96.0% (86.8 to 98.9) (1397 participants; moderate-certainty evidence).

Truenat MTB Plus (2 studies)

There were not enough data to meta-analyze, and we have provided descriptive results for Truenat MTB Plus. The sensitivities in these two studies ranged from 95% to 100% while the specificities ranged from 55% to 100% against a microbiological reference standard. The sensitivity was 78.7% (70 to 86) and the specificity was 100% (91 to 100) against a composite reference standard from a single study.

Pleural fluid

Xpert Ultra (13 studies)

Xpert Ultra summary sensitivity and specificity against a microbiological reference standard were 74.0% (60.8 to 83.9; 264 participants; low-certainty evidence) and 88.1% (78.8 to 93.6; 777 participants; very low-certainty evidence).

Truenat MTB Plus (1 study)

The sensitivity was 100% (2.5 to 100) and specificity was 100% (95.3 to 100) against a microbiological reference standard.

Lymph node aspirate

Xpert Ultra (6 studies)

Xpert Ultra summary sensitivity and specificity (95% CI) against a composite reference standard were 71.3% (64.3 to 77.4) (243 participants; moderate-certainty evidence) and 97.4% (82.3 to 99.7) (218 participants; very low-certainty evidence).

Truenat MTB Plus (1 study)

The sensitivity and specificity were 77.1% (66 to 86) and 100% (88 to 100), respectively, against a microbiological reference standard. The sensitivity was 100% (81 to 100) and specificity was 56% (45 to 67) against a composite reference standard.

Rifampicin resistance

Xpert Ultra (13 studies)

Xpert Ultra summary sensitivity and specificity were 100.0% (93.4 to 100.0; 54 participants; high-certainty evidence) and 99.4% (92.1 to 100.0; 392 participants; high-certainty evidence).

Authors' conclusions

LC-aNAATs are helpful in diagnosing extrapulmonary tuberculosis. Sensitivity varies across different extrapulmonary specimens, while for most specimens specificity is high, the tests rarely yielding a positive result for people without tuberculosis. For tuberculous meningitis, Xpert Ultra had high sensitivity against culture. Xpert Ultra also had high sensitivity and specificity for rifampicin resistance. Future research

should acknowledge the concern associated with culture as a reference standard in paucibacillary specimens and consider ways to address this limitation. Additionally, there is a critical need for robust evidence on other technologies within the LC-aNAAT class.

Funding

Funded by the WHO Global Tuberculosis Program.

Registration

This is an update to the published review "Xpert MTB/RIF Ultra and Xpert MTB/RIF assays for extrapulmonary tuberculosis and rifampicin resistance in adults" via doi: 10.1002/14651858.CD012768.pub3.

PLAIN LANGUAGE SUMMARY

How accurate are low-complexity automated rapid molecular tests for diagnosing tuberculosis outside the lungs (extrapulmonary tuberculosis) and rifampicin resistance in adults and adolescents?

Key messages

- Low-complexity rapid molecular tests can help identify people with extrapulmonary tuberculosis and rifampicin resistance.
- These tests can accurately identify tuberculosis in cerebrospinal fluid, pleural fluid, pleural tissue, synovial, peritoneal and pericardial fluid.

Why is using low-complexity rapid molecular tests for extrapulmonary tuberculosis important?

Tuberculosis is one of the top 10 causes of death worldwide. Tuberculosis mainly affects the lungs but can also occur elsewhere in the body (extrapulmonary). Quick and accurate tests help people start treatment sooner, which saves lives. Low-complexity automated nucleic acid amplification tests (LC-aNAATs) are rapid tests that give results in about two hours, unlike culture that takes weeks. They can also detect resistance to important antibiotics for tuberculosis, like rifampicin.

What is the aim of this review?

To update evidence on how well LC-aNAATs detect extrapulmonary tuberculosis and rifampicin resistance in adults and adolescents.

What did we do?

LC-aNAATs are World Health Organization-recommended rapid molecular diagnostic tests for diagnosing tuberculosis and rifampicin resistance. We combined study results to find out:

- sensitivity for tuberculosis detection: proportion of people with tuberculosis correctly diagnosed as having tuberculosis;
- specificity for tuberculosis detection: proportion of people without tuberculosis correctly identified as not having tuberculosis;
- sensitivity for rifampicin resistance detection: proportion of people with rifampicin resistance correctly diagnosed as being rifampicin resistant;
- specificity for rifampicin resistance detection: proportion of people with rifampicin susceptibility correctly identified as being rifampicin susceptible.

We assessed LC-aNAAT results against a microbiological and a composite reference standard (neither is a perfect reference standard because extrapulmonary tuberculosis has fewer bacteria).

What are the main results of this review?

Thirty-seven studies tested lymph node, pleural, and cerebrospinal fluid, and other specimens from people presumed to have extrapulmonary tuberculosis. We found data for 2 LC-aNAATs (Xpert Ultra and Truenat MTB Plus), but could only pool the data for Xpert Ultra to produce summary estimates which are presented below.

For every 1000 people tested, if 100 had tuberculosis:

cerebrospinal fluid (16 studies)

The sensitivity was 88% with a specificity of 96% against a microbiological reference standard. This means that 124 people would test positive in total, of which 36 would be without tuberculosis (false positive); also, 876 people would test negative in total, of which 12 would have tuberculosis (false negative).

pleural fluid (13 studies)

The sensitivity was 74% with a specificity of 88% against a microbiological reference standard. This means that 181 people would test positive in total, of which 107 would be without tuberculosis (false positive); also, 819 people would test negative in total, of which 26 would have tuberculosis (false negative).

lymph node aspirate (6 studies)

The sensitivity was 71% with a specificity of 97% against a composite reference standard. This means that 94 people would test positive in total, of which 23 would be without tuberculosis (false positive); also, 906 people would test negative in total, of which 29 would have tuberculosis (false positive).

rifampicin resistance (13 studies)

The sensitivity was 100% with a specificity of 99% against a microbiological reference standard. This means 105 people would test positive in total for resistance, of which 5 would be without resistance (false positive); also 895 people would test negative for resistance in total.

Who do the results of this review apply to?

People with presumed extrapulmonary tuberculosis.

How confident are we in our results?

We are fairly confident about LC-aNAAT in cerebrospinal fluid and less confident about lymph node aspirate and pleural fluid for Xpert Ultra because our question was to understand how these tests would perform in routine settings. However, most studies for lymph node aspirate and pleural fluid did not always report the settings. Their results also differed a lot between studies and some of the studies were very small. We are less confident about Truenat MTB plus, as there were few studies and few people tested. Both reference standards are imperfect, which may affect accuracy estimates.

What are the implications of this review?

LC-aNAATs may be helpful in diagnosing extrapulmonary tuberculosis, though sensitivity varies across different extrapulmonary specimens. While for most specimens, specificity is high, the test rarely yields a positive result for people without tuberculosis. LC-aNAATs had high sensitivity for tuberculous meningitis and high sensitivity and specificity for rifampicin resistance. However, there is a need for more data on other tests within the same technology class.

What are the limitations of the evidence?

As culture is not a perfect reference standard for this form of tuberculosis, multiple cultures per specimen could help strengthen the reference standard. However, not all studies reported the number of cultures per specimen, which is a limitation in this review.

How up-to-date is this review?

10 October 2023. A World Health Organization public call for data was made between 30 November 2023 and 15 February 2024 to identify unpublished studies.

SUMMARY OF FINDINGS

Summary of findings 1. LC-aNAATs in cerebrospinal fluid

Participants: people presumed to have tuberculous meningitis

Prior testing: people who received LC-aNAATs may first have undergone a health examination (history and physical examination) and possibly received a chest radiograph

Role: initial test, replacement for usual practice

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

Index tests: LC-aNAATs (Xpert Ultra only)*

Reference standard: solid or liquid culture

Studies: cross-sectional studies

Limitations: participants were evaluated exclusively as inpatients at a tertiary care center, or, if the clinical setting was not reported, LC-aNAAT was performed at a reference laboratory rather than at primary care facilities and local hospitals. However, we would expect TB meningitis patients to be evaluated in tertiary care settings due to the nature of this form of tuberculosis.

LC-aNAAT summary sensitivity (95% CI): 0.88 (95% CI 0.84 to 0.92)| summary specificity: 0.96 (95% CI 0.87 to 0.99)

Test result	Number of results per 1000 patients tested (95% CI)			Number of participants (studies)	Certainty of the Evidence (GRADE)
	Prevalence 2.5%	Prevalence 10%	Prevalence 20%		
True positives	22 (21 to 23)	88 (84 to 92)	176 (167 to 183)	287 (16)	⨁⨁⨁⨁ **High**
False negatives	3 (2 to 4)	12 (8 to 16)	24 (17 to 33)		
True negatives	936 (846 to 964)	864 (781 to 890)	768 (694 to 791)	1397 (16)	⨁⨁⨁◯ **Moderate**[a,b]
False positives	39 (11 to 129)	36 (10 to 119)	32 (9 to 106)		

Abbreviations: CI: confidence interval; LC-aNAAT: low-complexity automated nucleic acid amplification tests
*Only Xpert Ultra was included for this SoF, as Truenat data were insufficient to be included in the class-based analyses.
Explanations
[a]Although the specificity ranged from 50% to 100%, we could explain some of the inconsistency based on whether decontaminated specimens were used for culture inoculation, which could lead to loss of viable bacteria in the process, thus causing a false negative on the culture and decreasing the specificity.
[b]The very wide 95% CI around false positives may lead to different decisions, depending on which confidence limits are assumed. We downgraded by one level for imprecision. We downgraded one level only, keeping in mind that culture is an imperfect reference standard and might lead to decreased specificity.
GRADE certainty of the evidence
High: we are very confident that the true effect lies close to that of the estimate of the effect.

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

Summary of findings 2. LC-aNAATs in pleural fluid

Participants: people presumed to have pleural tuberculosis

Prior testing: people who received LC-aNAATs may first have undergone a health examination (history and physical examination) and received a chest radiograph

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

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

Index tests: LC-aNAATs (Xpert Ultra only)*

Reference standard: solid or liquid culture

Studies: cross-sectional studies

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

LC-aNAAT summary sensitivity (95% CI): 0.74 (95% CI 0.61 to 0.84) |summary specificity: 0.88 (95% CI 0.79 to 0.94)

Xpert Ultra result	1000 people tested for TB using Xpert Ultra (95% CI)			Number of participants (studies)	Certainty of the evidence (GRADE)
	Prevalence of 2.5%	Prevalence of 10%	Prevalence of 20%		
True-positives (patients with pleural TB)	19 (15 to 21)	74 (61 to 84)	148 (122 to 168)	264 (13)	⊕⊕◯◯ **Low**[a,b]
False-negatives (patients incorrectly classified as not having pleural TB)	6 (4 to 10)	26 (16 to 39)	52 (32 to 78)		
True-negatives (patients without pleural TB)	859 (768 to 913)	793 (709 to 842)	705 (630 to 749)	777 (13)	⊕◯◯◯ **Very low**[b,c,d]
False-positives (patients incorrectly classified as having pleural TB)	116 (62 to 207)	107 (58 to 191)	95 (51 to 170)		

Abbreviations: CI: confidence interval; LC-aNAAT: low-complexity automated nucleic acid amplification tests; TB: tuberculosis
*Only Xpert Ultra was included for this SoF, as Truenat data were insufficient to be included in the class-based analyses.
Explanations
[a]The sensitivity of the included studies varied from 0% to 100%. We could not completely explain this heterogeneity in sensitivity estimates. We downgraded by one level for inconsistency.
[b]We were interested in how LC-aNAATs performed in patients presumed to have extrapulmonary tuberculosis who were evaluated as they would be in routine practice. However, most studies did not report information on the clinical setting. We downgraded by one level for indirectness.
[c]The specificity of the included studies varied from 57% to 100%. We could not completely explain this heterogeneity. We downgraded by one level for inconsistency.
[d]The very wide 95% CIs around true negatives and false positives could lead to different decisions, depending on which confidence limits are assumed. We downgraded by one level for imprecision.
GRADE certainty of the evidence
High: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.
The results presented in this table should not be interpreted in isolation from results of the individual included studies contributing to each summary test accuracy measure.

Summary of findings 3. LC-aNAATs in lymph node aspirate

Participants: people presumed to have lymph node tuberculosis

Prior testing: people who received LC-aNAATs may first have undergone a health examination (history and physical examination) and possibly received a chest radiograph

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

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

Index tests: LC-aNAATs (Xpert Ultra only)*

Reference standard: composite reference standard

Studies: cross-sectional studies

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

LC-aNAATs summary sensitivity (95% CI): 0.71 (95% CI 0.64 to 0.77)|summary specificity: 0.97 (95% CI 0.82 to 1.00)

Xpert Ultra result	1000 people tested for TB using Xpert Ultra (95% CI)			Number of participants (studies)	Certainty of the evidence (GRADE)
	Prevalence of 2.5%	Prevalence of 10%	Prevalence of 20%		
True-positives (patients with lymph node TB)	18 (16 to 19)	71 (64 to 77)	143 (129 to 155)	243	⊕⊕⊕◯

False-negatives (patients incorrectly classified as not having lymph node TB)	7 (6 to 9)	29 (23 to 36)	57 (45 to 71)	(6)	**Moderate**[a]
True-negatives (patients without lymph node TB)	950 (801 to 972)	877 (740 to 897)	779 (658 to 798)	218 (6)	⊕◯◯◯ **Very low**[a,b,c,d]
False-positives (patients incorrectly classified as having lymph node TB)	25 (3 to 174)	23 (3 to 160)	21 (2 to 142)		

Abbreviations: CI: confidence interval; LC-aNAAT: low-complexity automated nucleic acid amplification tests; TB: tuberculosis
*Only Xpert Ultra was included for this SoF, as Truenat data were insufficient to be included in the class-based analyses.
[a]For indirectness, regarding applicability, for the patient selection domain, we considered most studies to have unclear concerns. We were interested in how LC-aNAATs performed in patients presumed to have extrapulmonary tuberculosis who were evaluated as they would be in routine practice. However, none of the studies reported this information. We downgraded by one level for indirectness.
[b]The composite reference standard was defined by the primary study authors and, therefore, was not uniform. We downgraded by one level for risk of bias.
[c]The specificity ranged from 71% to 100%, and we were unable to explain this heterogeneity. We downgraded by one level for inconsistency.
[d]The very wide 95% CI for true negatives and false positives may lead to different decisions depending on which confidence limits are assumed. We downgraded by one level for imprecision.
GRADE certainty of the evidence
High: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.
The results presented in this table should not be interpreted in isolation from results of the individual included studies contributing to each summary test accuracy measure.

Summary of findings 4. LC-aNAATs for rifampicin resistance

Participants: people with tuberculosis detected by LC-aNAATs

Role: initial test**,** replacement test for standard practice, which includes culture-based drug susceptibility testing or line probe assay

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

Index tests: LC-aNAATs (Xpert Ultra only)*

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

Studies: cross-sectional studies

LC-aNAAT summary sensitivity (95% CI): 1.00 (95% CI 0.93 to 1.00)|summary specificity: 0.99 (95% CI 0.92 to 1.00)

Xpert Ultra result	1000 people tested for rifampicin resistance using Xpert Ultra (95% CI)	Number of participants (studies)	Certainty of the evidence (GRADE)

	Prevalence of 2%	Prevalence of 10%	Prevalence of 15%		
True-positives (patients correctly classified as rifampicin resistant)	20 (19 to 20)	100 (93 to 100)	150 (140 to 150)	54 (13)	⊕⊕⊕⊕ High
False-negatives (patients incorrectly classified as rifampicin susceptible)	0 (0 to 1)	0 (0 to 7)	0 (0 to 10)		
True-negatives (patients correctly classified as rifampicin susceptible)	974 (903 to 980)	895 (829 to 900)	845 (783 to 850)	392 (13)	⊕⊕⊕⊕ High
False-positives (patients incorrectly classified as rifampicin resistant)	6 (0 to 77)	5 (0 to 71)	5 (0 to 67)		

Abbreviations: CI: confidence interval; LC-aNAAT: low-complexity automated nucleic acid amplification tests; TB: tuberculosis
*Only Xpert Ultra was included for this SoF, as Truenat data were insufficient to be included in the class-based analyses.
GRADE certainty of the evidence
High: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.
The results presented in this table should not be interpreted in isolation from results of the individual included studies contributing to each summary test accuracy measure.

BACKGROUND

Tuberculosis causes tremendous suffering worldwide and has surpassed HIV/AIDS as the world's leading infectious cause of death. The World Health Organization (WHO) estimates that globally in 2023, 10.8 million (95% uncertainty interval [UI]: 10.1 to 11.7 million) people became ill with tuberculosis. In 2023, around 1.09 million HIV-negative people died from tuberculosis and 161,000 HIV-positive people died from tuberculosis [1]. Globally, extrapulmonary tuberculosis accounted for 20% of the 8.2 million cases of tuberculosis notified in 2023 [1]. Among countries in the European Union, extrapulmonary tuberculosis was responsible for 19% of all notified cases (range: 6% to 44%) [2]. The WHO estimates that from 2000 to 2019 more than 60 million lives were saved by diagnosing and treating tuberculosis. However, the COVID-19 pandemic resulted in disruptions in tuberculosis diagnosis and treatment, reversing gains made over recent years.

A large retrospective analysis from China found that, of 19,279 hospitalized tuberculosis patients, around 33% had extrapulmonary tuberculosis [3]. The number of people affected by extrapulmonary tuberculosis is likely to be higher, given that (according to the WHO), extrapulmonary tuberculosis is notified as pulmonary tuberculosis when the two forms exist together, and diagnosing extrapulmonary tuberculosis is challenging, as described below. Additionally, extrapulmonary tuberculosis accounts for an increasing proportion of tuberculosis cases in some countries, in part because of host and genetic considerations, and the association of extrapulmonary tuberculosis and HIV [4, 5, 6, 7]. Based on surveillance and epidemiological data, extrapulmonary tuberculosis affects a greater proportion of children than adults [8].

Drug-resistant tuberculosis is a serious threat to global health. In 2023, 175,923 people were diagnosed and treated for multidrug-resistant or rifampicin-resistant tuberculosis (MDR/RR-TB), which is only 44% of the total estimated people who developed MDR/RR-TB [1]. For the purpose of surveillance and treatment, drug-resistant tuberculosis is classified as rifampicin-resistant tuberculosis (RR-TB), MDR-TB, and extensively drug-resistant tuberculosis. MDR-TB is defined as resistance to at least isoniazid and rifampicin, the two most important first-line anti-tuberculosis drugs. Extensively drug-resistant (XDR-TB) tuberculosis is defined as MDR-TB plus resistance to at least one fluoroquinolone (levofloxacin or moxifloxacin) and to at least one other Group A drug (bedaquiline or linezolid). Ten countries (China, the Democratic People's Republic of Korea, India, Indonesia, Myanmar, Nigeria, Pakistan, the Philippines, Ukraine, and Vietnam) account for 70% of the estimated number of people who develop MDR/RR-TB each year globally [1].

In 2014, the World Health Assembly unanimously approved the WHO End TB Strategy, a 20-year strategy devised to end the global tuberculosis epidemic [9]. Early diagnosis of tuberculosis, including universal drug susceptibility testing (DST) and systematic screening of contacts and high-risk groups, is a part of pillar one of the strategy.

Target condition being diagnosed

Extrapulmonary tuberculosis

Tuberculosis is caused by infection with *Mycobacterium tuberculosis* (*M tuberculosis*) bacteria. Tuberculosis predominantly affects the lungs (pulmonary tuberculosis). Extrapulmonary tuberculosis refers to tuberculosis in parts of the body other than the lungs. Extrapulmonary tuberculosis is known to affect virtually every part of the body, with lymph nodes and the pleura being the most common sites [10]. Although active pulmonary tuberculosis is transmissible by droplets spread by coughing, extrapulmonary tuberculosis is thought to result from hematogenous spread (spread by way of the bloodstream) from an initial lung infection and is not infectious. Extrapulmonary tuberculosis can occur alone or together with pulmonary tuberculosis.

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

Diagnosis of extrapulmonary tuberculosis is challenging for several reasons. Many forms of extrapulmonary tuberculosis require invasive diagnostic sampling; gathering adequate specimens can pose a risk of harm to the patient and can be costly. Most forms of extrapulmonary tuberculosis are paucibacillary (tuberculosis disease caused by a small number of bacteria), making diagnosis by various tests less sensitive. Culture, for example, has reduced sensitivity in paucibacillary disease. In addition, culture takes several weeks for results and requires a highly-equipped laboratory. Limitations are also associated with histology, which relies on highly-trained operators, and characteristic morphology is shared with other diseases. As a result of these difficulties, the diagnosis of extrapulmonary tuberculosis is often made on the grounds of clinical suspicion alone, and many people receive the wrong diagnosis, leading to unnecessary tuberculosis treatment or poor outcomes from untreated extrapulmonary tuberculosis.

Tuberculosis treatment regimens must contain multiple drugs to which the organisms are sensitive to cure tuberculosis and avoid selection for drug resistance. WHO tuberculosis treatment guidelines recommend the same drug regimens for extrapulmonary and pulmonary disease, with notable mention of tuberculous meningitis and bone or joint tuberculosis, for which longer treatment regimens are recommended [11, 12, 13]. For patients with tuberculous meningitis or tuberculous pericarditis, the use of adjuvant corticosteroid therapy is recommended in addition to appropriate tuberculosis treatment regimens [11, 13]. Other tuberculosis treatment guidelines include India [14], and those issued by the American Thoracic Society, the Centers for Disease Control and Prevention (CDC), and the Infectious Diseases Society of America [15]. The WHO currently recommends three categories of regimen for drug-resistant tuberculosis: a short oral regimen consisting of bedaquiline, pretomanid, linezolid and moxifloxacin (BPaLM) for people with MDR/RR-TB and without moxifloxacin for people with pre XDR-TB; an all-oral short regimen for people with MDR/RR-TB; and longer regimens of 18 to 20 months that may include injectable drugs. The shorter 6-month regimen is recommended for individuals aged 14 years and older with MDR/RR-TB or pre-XDR-TB.

Rifampicin resistance

Rifampicin inhibits bacterial DNA-dependent RNA polymerase, encoded by the RNA polymerase gene (*rpoB*) [16]. Resistance to this drug has been associated mainly with mutations in a limited

region of the *rpoB* gene [17]. Rifampicin resistance may occur alone or in association with resistance to isoniazid and other drugs. In settings with a high burden of MDR-TB, the presence of rifampicin resistance alone may serve as a proxy for MDR-TB [18]. People with drug-resistant tuberculosis can transmit the infection to others.

Index test(s)

The WHO groups individual tests with similar characteristics and performance into one of three classes: low-complexity automated nucleic acid amplification tests (LC-aNAATs); low-complexity manual NAATs (LC-mNAATs), and moderate-complexity automated NAATs (MC-aNAATs). The classes are defined by the type of technology (e.g. automated or reverse hybridization NAATs), the complexity of the test for implementation (e.g. low, moderate, or high - considering the requirements of infrastructure, equipment and technical skills of laboratory staff) and the target conditions (e.g. diagnosis of tuberculosis, and detection of resistance to first-line or second-line drugs) [19]. The following four index tests in the LC-aNAATs class were considered in this review.

Xpert MTB/RIF Ultra

Xpert MTB/RIF Ultra (Xpert Ultra, Cepheid Inc, subsidiary of Danaher Corporation, Sunnyvale, USA) is a nucleic acid amplification test (NAAT) (i.e. molecular test) used for diagnosing tuberculosis and rifampicin-resistant tuberculosis. Xpert Ultra cartridges are used with the GeneXpert system [20, 21]. Xpert Ultra is able to detect both *M tuberculosis* complex and rifampicin resistance within 90 minutes of starting the test, with minimal hands-on technical time. Unlike conventional NAATs, Xpert Ultra's sample processing and polymerase chain reaction (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 assays' sample reagent, used to liquefy sputum, has potent tuberculocidal (the ability to kill tuberculosis bacteria) properties and so largely eliminates biosafety concerns during the test procedure [22]. Except as described below for Xpert Ultra trace call results, a single Xpert Ultra run will provide both detection of tuberculosis and detection of rifampicin resistance. One cannot deselect testing for rifampicin resistance and only run the assay for tuberculosis detection.

In order to overcome the limitations with Xpert MTB/RIF of low sensitivity in paucibacillary specimens, Cepheid developed Xpert Ultra, a re-engineered assay that uses a newly-developed cartridge, but may be run on the same device after a software upgrade. To improve sensitivity for tuberculosis detection, Xpert Ultra incorporates two different multi-copy amplification targets and a larger DNA reaction chamber than Xpert MTB/RIF [23]. A laboratory study reported that the limit of detection (the lowest number of colony-forming units (CFUs) per sample that can be reproducibly distinguished from negative samples with 95% confidence) using Xpert Ultra improved to 15.6 CFU/mL of sputum compared to 112.6 CFU/mL for Xpert MTB/RIF [24]. Xpert Ultra has added a new result category, 'trace call', that corresponds to the lowest bacillary load for *M tuberculosis* detection [23]. This new category is reported as 'MTB trace DETECTED'. Interpreting a trace call result requires a reassessment of clinical symptoms and history of prior tuberculosis. No rifampicin resistance results are available (indeterminate) for people with trace results. As with Xpert MTB/RIF [25], Xpert Ultra detects both live and dead bacteria.

To address limitations in rifampicin resistance detection, Xpert Ultra uses melting temperature-based analysis, in lieu of real-time PCR analysis with Xpert MTB/RIF. Melting temperature-based analysis allows Xpert Ultra to better distinguish resistance-conferring mutations from silent mutations with improved diagnostic accuracy for rifampicin resistance detection [26].

For sputum specimens, the test procedure may be used either directly on raw sputum specimens or sputum pellets created after decontaminating and concentrating the sputum [27]. In both cases, the test material is combined with the assay sample reagent (sodium hydroxide and isopropanol), mixed by hand or vortex, and incubated at room temperature for 15 minutes. After the incubation step, 2 mL of the treated specimen is transferred to the cartridge and the run is initiated [28]. According to the manufacturer, as with Xpert MTB/RIF, Xpert Ultra may be used with fresh sputum specimens, which may be either unprocessed sputum or processed sputum sediments. The sample reagent:sample volume ratio is 2:1 for unprocessed sputum and 3:1 for sputum pellets. The manufacturer does not specifically mention the use of Xpert Ultra with frozen specimens [20, 21]. Xpert Ultra using the GeneXpert sytem requires an uninterrupted and stable electrical power supply, temperature control, and yearly calibration of the cartridge modules [29]. Like previous Xpert cartridge generations, Xpert Ultra can be performed by operators with minimal technical expertise [30]. The time to run the assay is shorter for Xpert Ultra (around 65 to 87 minutes) than Xpert MTB/RIF [26]. Currently, the manufacturer has made no claim for the use of Xpert Ultra in non-sputum specimens [21]. However, there is a standard operating procedure provided by the WHO for processing non-sputum specimens [31].

Truenat MTB assays

Truenat assays, developed by Molbio diagnostics in Bangalore, India, include Truenat MTB, Truenat MTB Plus, and Truenat MTB-RIF Dx. The Truenat and Xpert assays can both detect dead and live bacilli in the test sample. Penn Nicholson and colleagues reported that Truenat assays were non-inferior to Xpert assays [32]. Truenat MTB targets the ribonucleoside-diphosphate reductase B single-copy gene (*nrdB*), and the target of Truenat MTB Plus is the *nrdZ* gene and the multicopy insertion sequence IS*6110* for identifying *M tuberculosis*. Truenat MTB is a quantitative test that gives actual colony-forming units (CFUs) per milliliter count, while Truenat MTB Plus is semi-quantitative and gives four grades (high, medium, low, and very low) based on CFUs, but does not specify the actual count [33, 34]. Both assays have similar run time and shelf life. Truenat MTB-RIF Dx targets the *rpoB* gene (RNA polymerase gene's beta subunit) for detecting rifampicin resistance. Once Truenat MTB plus is positive, Truenat MTB-RIF Dx is performed as a follow-on test for drug resistance detection.

STANDARD M10, SD Biosensor, Republic of Korea

The STANDARD M10 is an in vitro molecular diagnosis system that automatically extracts and amplifies nucleic acids (DNA, RNA) from the collected samples, and analyzes the data from amplification in real-time, by combining the processes of isothermal amplification and real-time PCR in a single platform using all-in-one cartridge systems. There are two variations currently available for tuberculosis:

- STANDARD™ M10 MTB/NTM; and
- STANDARD™ M10 MDR-TB.

These assays are intended for use with the STANDARD M10 system [35].

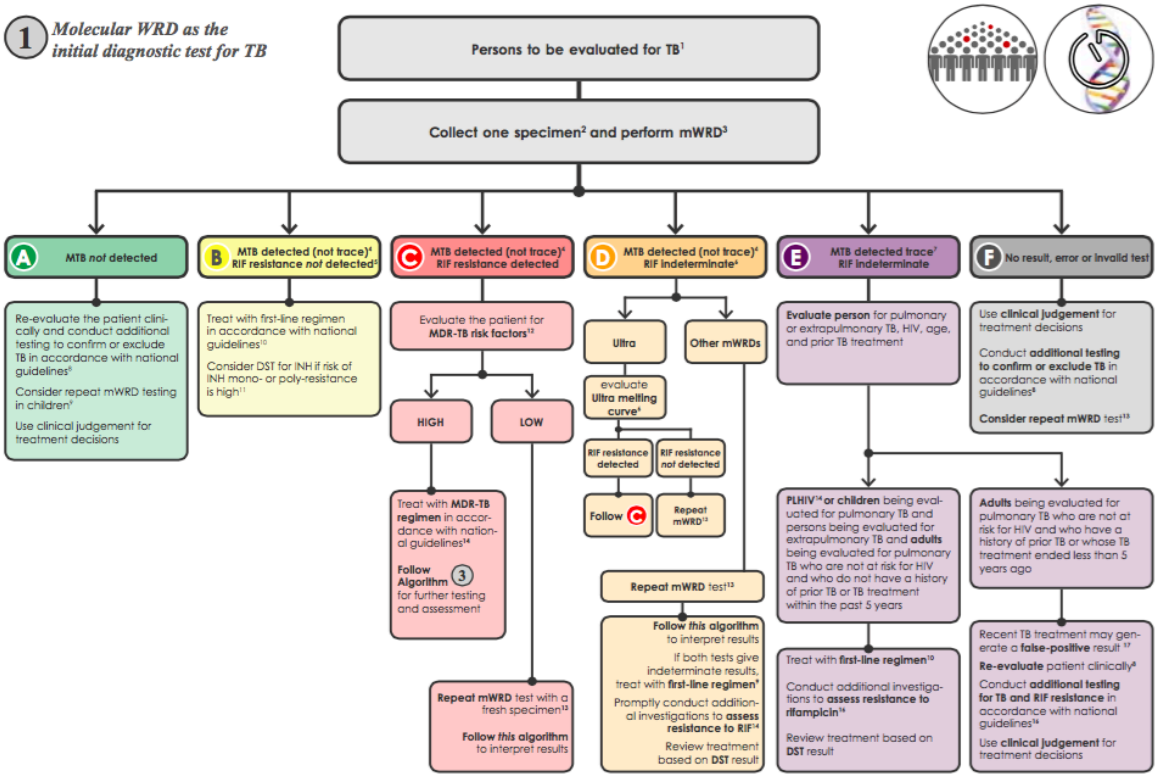
IRON qPCR, Bioneer, Republic of Korea

IRON-qPCR ™ is a point-of-care molecular diagnostics system which combines nucleic acid extraction and real-time PCR. It does syndromic diagnostics, i.e. the process of using a single cartridge test to detect > 10 gene targets which cause overlapping signs and symptoms. IRON-qPCR ™ has high multiplexing capacity of up to 40 targets in a single test and provides results in 40 minutes. The IRON qPCR RFIA kit is a cartridge-based assay that detects *M tuberculosis* and resistance to 11 drugs: rifampicin, isoniazid, fluoroquinolones (levofloxacin, moxifloxacin, gatifloxacin and ofloxacin), aminoglycosides (kanamycin, amikacin, capreomycin), prothionamide, and ethionamide. It uses IRON qPCR, which is a single instrument for extraction and amplification and detection. Two modules are available per instrument [36]. IRON qPCR RFIA is not commercially available.

Clinical pathway

LC-aNAATs such as Xpert Ultra and Truenat MTB assays are used for the diagnosis of extrapulmonary tuberculosis and rifampicin resistance. Figure 1 shows the clinical pathway and presents the context in which molecular WHO-recommended diagnostic tests (mWRDs) might be used [37]. The target conditions are extrapulmonary tuberculosis, which includes several forms (e.g. tuberculous meningitis, pleural tuberculosis) and rifampicin resistance.

Figure 1. The clinical pathway describes how patients might present and the point in the pathway at which they would be considered for testing with low-complexity automated NAATs. This algorithm for the use of a molecular WHO-recommended rapid diagnostic (WRD) comes from the WHO operational handbook on tuberculosis [37]. Copyright © [2020] [World Health Organization]: reproduced with permission. Abbreviations:
DST: drug susceptibility testing
INH: isoniazid
MDR-TB: multidrug-resistant TB
MTB: Mycobacterium tuberculosis
PLHIV: people living with HIV
RIF: rifampicin
TB: tuberculosis
WRD: WHO-recommended rapid diagnostic



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

We have described in Table 1 the signs and symptoms of the forms of extrapulmonary tuberculosis included in this review. The clinician should take a careful history, noting a history of tuberculosis exposure, prior tuberculosis disease, and medical conditions that increase the risk for tuberculosis disease (e.g. HIV, diabetes mellitus, low body weight). In comparison with HIV-negative people, HIV-positive people have higher rates of extrapulmonary tuberculosis or mycobacteremia (tuberculosis bloodstream infection). People with HIV with signs or symptoms of extrapulmonary tuberculosis should have specimens taken from the suspected site(s) of involvement to increase the likelihood of tuberculosis diagnosis. Tuberculous meningitis is the most severe form of tuberculosis. In tuberculous meningitis, diagnosis is often delayed, with appalling consequences for patients. For all forms of extrapulmonary tuberculosis, patients may be evaluated in primary- or secondary-care settings. However, if more complex or invasive tests are needed, patients may be referred to a tertiary medical center [10, 38, 39].

The downstream consequences of testing include the following.

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

Prior test(s)

People presumed to have tuberculosis could undergo prior screening by modalities like symptomatic screening, chest X-ray, C-reactive protein, cytology/histopathology for extrapulmonary specimens, etc. These help triage people for downstream diagnostic testing. Smear microscopy could also be considered as prior testing before participants are included or diagnosed further.

For the purpose of this review, we recorded and judged whether participants were included in the study based on prior testing, especially if these tests could over or under-estimate the diagnostic accuracy of LC-aNAATs.

Role of index test(s)

We were interested in the following roles for testing:

I. LC-aNAATs for detection of extrapulmonary tuberculosis

An index test used as an initial test replacing usual practice (including conventional microscopy, culture, or histopathology) for the diagnosis of extrapulmonary tuberculosis in adults and adolescents with presumptive extrapulmonary tuberculosis [19]. An initial test does not mean that other tests will follow.

II.LC-aNAATs for detection of rifampicin resistance

An index test used as an initial test replacing culture and phenotypic DST for the diagnosis of rifampicin-resistant tuberculosis in adults and adolescents with presumptive extrapulmonary tuberculosis [19].

As mentioned, in high MDR-TB settings, the presence of rifampicin resistance alone may serve as a proxy for MDR-TB. LC-aNAATs do not eliminate the need for subsequent culture and phenotypic DST, which are required to monitor treatment progress and to detect resistance to drugs other than rifampicin.

Alternative test(s)

For a comprehensive review of new tests not yet in widespread use, we refer the reader to these references [40, 41, 42].

Smear microscopy (light microscopy (Ziehl-Neelsen), fluorescence microscopy, or light-emitting diode (LED) fluorescence microscopy) is the examination of smears for acid-fast bacilli (tuberculosis bacteria) under a microscope. Around 5000 to 10,000 organisms per mL must be present in the specimen for tuberculosis bacteria to be visible by microscopy [43]. For extrapulmonary tuberculosis, microscopy can be performed on fluid or tissue specimens from sites of disease involvement; for example, in cerebrospinal fluid (CSF) in presumptive tuberculous meningitis or in lymph node tissue in presumptive lymph node tuberculosis. For most extrapulmonary sites, because there are usually few organisms, the sensitivity of smear microscopy is generally low. Ranges from studies, some with selected cases, are quoted here: 0% to 10% in pleural fluid; 14% to 39% in pleural tissue; 2% to 30% in CSF; < 5% in peritoneal fluid; and 0% to 42% in pericardial fluid. In contrast, the specificity of smear microscopy tends to be quite high, as can be seen in pulmonary tuberculosis (≥ 90%) [41, 44].

Mycobacterial culture is a method used to grow bacteria on nutrient-rich media. In comparison with microscopy, a positive culture requires only around 100 organisms per mL and therefore can detect lower numbers of tuberculosis bacteria [43]. Additionally, culture is essential for species identification and DST. However, culture takes several weeks and requires a highly-equipped laboratory. Culture has reduced sensitivity in paucibacillary disease (reference standards have included culture from a different specimen, such as sputum, smear microscopy, NAATs, presence of granulomatous inflammation, clinical criteria, imaging studies, and response to anti-tuberculosis therapy, done alone or in various combinations): CSF 45% to 70%; pleural fluid 23% to 58%; urine 80% to 90%; peritoneal tuberculosis 45% to 69%; pericardial tuberculosis 50% to 65% [41]; lymph node tuberculosis (excisional biopsy) 18% to 93%; and lymph node tuberculosis (fine-needle aspirate) 10% to 67% [45].

Histological examination involves examination of tissue specimens under a microscope. Diagnosis of extrapulmonary tuberculosis by histological examination is based on finding acid-fast bacilli and granulomatous inflammation, frequently with caseous (cheese-like) necrosis (necrotizing granulomas). The sensitivity of histology has been reported to vary for different forms of extrapulmonary tuberculosis (reference standards have included smear microscopy, culture, NAATs, clinical criteria, and imaging studies, done alone or in various combinations): 59% to 88% for lymph node tuberculosis (excisional biopsy) [45]; 69% to 97% in pleural tissue (closed pleural biopsy); 86% to 94% in urological tissue; 60% to 70% in endometrial curettage; 79% to 100% in peritoneal biopsy; and 73% to 100% in pericardial tissue [41]. Sensitivity has also been observed to vary for different diagnostic techniques. Diacon and colleagues found thoracoscopy to be more sensitive (sensitivity of 100%) than closed-needle biopsy (sensitivity of 66%) for establishing a diagnosis of pleural tuberculosis (reference standards have included microscopy smear, culture, or presence of granulomatous inflammation with caseous necrosis) [46]. Specificity has been observed to be low because of the presence of granulomas in other diseases, both infectious and non-infectious [41], although the presence of 'necrotizing' granulomatous inflammation increases specificity [47]. Histological examination carries the additional concern that invasive procedures that are complex and costly may be required to obtain the necessary specimens [4].

Cytopathological examination of fluid specimens (such as pleural and peritoneal fluid) may be performed, first to exclude cancer, and then to obtain material for additional analyses, such as measurement of levels of adenosine deaminase and free interferon-gamma (IFN-γ) and cell counts [41, 48]. Advantages of these tests are that they are rapid and simple and can be performed in most clinical laboratories [49]. In pleural, pericardial, and peritoneal fluid, a predominance of lymphocytes, especially in the absence of mesothelial cells, is highly suggestive of tuberculosis [48]. However, in HIV-positive people, this pattern may not be observed [48]. Adenosine deaminase, an enzyme involved in purine metabolism, has been extensively studied for its potential role in the diagnosis of pleural tuberculosis, peritoneal tuberculosis, and tuberculous meningitis [41]. IFN-γ is released after it is sensitized by T cells in response to specific *M tuberculosis* antigens. An evidence synthesis using GRADE provides the following recommendations:

- "...cell counts and chemistries be performed on amenable fluid specimens (including pleural, cerebrospinal, ascitic, and joint fluid) collected from sites of presumed extrapulmonary TB (conditional recommendation, very low-quality evidence);
- ...adenosine deaminase levels be measured, rather than not measured, on fluid collected from patients with presumed pleural TB, TB meningitis, peritoneal TB, or pericardial TB (conditional recommendation, low-quality evidence);
- ...free IFN-γ levels be measured, rather than not measured, on fluid collected from patients with presumed pleural TB or peritoneal TB (conditional recommendation, low-quality evidence)" [41].

NAATs including LC-aNAATs and LC-mNAATs are molecular techniques that can detect small quantities of genetic material (DNA or RNA) from micro-organisms, such as *M tuberculosis.* The key advantage of NAATs is that they are rapid diagnostic tests, potentially providing results in a few hours. This is a particularly important feature of the test in life-threatening forms of extrapulmonary tuberculosis, such as tuberculous meningitis. 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 laboratory) and are used routinely in high-income countries for tuberculosis detection. In-house PCR is widely used in low-income countries because these tests are less expensive than commercial kits. A systematic review found that NAATs have relatively low sensitivity for extrapulmonary tuberculosis but high specificity (e.g. for tuberculous meningitis, for pleural TB), indicating that these tests cannot be used reliably to rule out tuberculosis [49]. An evidence synthesis reported sensitivities of 72% to 88% in lymph node tissue, 28% to 81% in pleural fluid, 90% in pleural tissue, and 31% to 56% in CSF. Specificity ranged from 90% to 100% [41]. A systematic review of 24 studies estimated a summary sensitivity of 77% (95% CI: 68 to 85) and summary specificity of 99% (95% CI: 96 to 100) for TB- LAMP, a LC-mNAAT test. However, the review did not report the findings based on different sites of extrapulmonary tuberculosis and also included in-house assays [50].

Alternative molecular methods for DST include the commercial line-probe assays, GenoType MTBDR*plus* assay (MTBDR*plus,* Hain LifeScience, Nehren, Germany), and the Nipro NTM+MDRTB detection kit 2 (Nipro, Tokyo, Japan), which detect the presence of mutations associated with drug resistance to isoniazid and rifampicin [51]. MTBDR*plus* is the most widely studied line-probe assay. Advantages of line-probe assays are that they can provide a result for the detection of tuberculosis and drug resistance in one to two days. Drawbacks are that line-probe assays are expensive and need to be used in intermediate and central laboratories [42]. The WHO recommends that for persons with a sputum smear-positive specimen or a cultured tuberculosis isolate, commercial molecular line-probe assays may be used as the initial test instead of phenotypic culture-based DST to detect resistance to rifampicin and isoniazid [19]. Other molecular assays for the detection of tuberculosis and resistance to rifampicin and isoniazid along with instruments are in development [52].

Alere Determine™ TB LAM Ag (AlereLAM) Alere Inc, (Waltham, USA) is a commercially-available point-of-care test for tuberculosis disease (pulmonary and extrapulmonary tuberculosis). The test detects lipoarabinomannan (LAM), a component of the bacterial cell wall, which is present in the urine of some people with tuberculosis. AlereLAM is performed by placing urine on one end of a test strip, with results appearing as a band on the strip if tuberculosis is present. The test is simple, requires no special equipment, and shows results in 25 minutes. This urine test has potential advantages over sputum-based testing due to ease of sample collection. The accuracy of urinary LAM detection is improved among people living with HIV with advanced immunosuppression [53]. In two randomized trials, the use of Alere LAM in adult inpatients living with HIV was shown to reduce mortality [54, 55]. Based on evidence from the randomized trials and a Cochrane Review [53], the WHO currently recommends that AlereLAM should be used to assist in the diagnosis of active tuberculosis in adults, adolescents, and children living with HIV [19]. The key change from the WHO 2015 guidelines is broadening the indication for the use of LF-LAM among inpatients with HIV with signs and symptoms of active tuberculosis (pulmonary and extrapulmonary); the test is now recommended for all such patients, irrespective of their CD4 count. The WHO issued a rapid communication in September

2024 regarding concurrent use of a molecular test on respiratory samples and LF-LAM on urine for the diagnosis of tuberculosis in adults and adolescents with HIV. Use of concurrent tests has improved accuracy compared with a single test and has moderate cost requirements [19].

Fujifilm SILVAMP TB LAM (FujiLAM, co-developed by FIND, Geneva, Switzerland and Fujifilm, Tokyo, Japan) is a new, urine-based, point-of-care test for tuberculosis diagnosis in people living with HIV. In an individual participant data meta-analysis that included five cohorts of people living with HIV, FujiLAM was found to have superior sensitivity, 70.7% (95% CI 59.0% to 80.8%), compared to AlereLAM sensitivity of 42.3% (31.7% to 51.8%), against a microbiological reference standard; FujiLAM had lower specificity, 90.9% (87.2 to 93.7), compared to AlereLAM specificity of 95.3% (92.2 to 97.7) [56]. There has been additional evidence generation on FujiLAM [57], however, owing to a lot of variation for this test [58], manufacturers have been working on redesigning it and there has been no policy guidance on this test.

Rationale

LC-aNAATs are rapid tests that may provide benefits for patients (earlier diagnosis and the opportunity to begin earlier, appropriate treatment), especially in high tuberculosis-burden countries.

Since 2010, the WHO has recommended the use of Xpert MTB/RIF as the preferred initial diagnostic test for people thought to have MDR-TB or HIV-associated tuberculosis (strong recommendation, moderate-certainty evidence) [59]. In 2013, the WHO expanded the recommendations, stating that Xpert MTB/RIF may be used rather than conventional microscopy and culture as the initial diagnostic test in all adults suspected of having tuberculosis (conditional recommendation acknowledging resource implications, high-quality evidence) [60]. The 2013 recommendations extended to the diagnosis of several forms of extrapulmonary tuberculosis, including tuberculous meningitis and lymph nodes and other tissues. In addition, the WHO recommended that, following an Xpert MTB/RIF test that demonstrates rifampicin resistance, subsequent DST (e.g. using a line-probe assay for second-line drugs) remains essential to detecting resistance to drugs other than rifampicin [60]. In 2017, based on a non-inferiority analysis of Xpert Ultra compared with Xpert MTB/RIF [61], the WHO stated that recommendations on the use of Xpert MTB/RIF also apply to the use of Xpert Ultra as the initial diagnostic test for all adults and children with signs and symptoms of tuberculosis [23].

In December 2019, the WHO convened a Guideline Development Group (GDG) to update the recommendations on the use of molecular assays intended as initial tests for the diagnosis of pulmonary and extrapulmonary tuberculosis and rifampicin resistance. In 2020, the WHO recommended class-based tests. Currently, low-complexity automated NAATs (LC-aNAATs) such as Xpert Ultra and Truenat MTB assays are widely available and recommended by the WHO as initial tests for the diagnosis of tuberculosis as they detect both tuberculosis disease and rifampicin resistance. We performed this systematic review on the diagnostic accuracy of LC-aNAATs for the detection of extrapulmonary tuberculosis in adults and adolescents to inform the 2024 update of the WHO policy guideline on rapid NAATs for tuberculosis detection. The GDG meeting was convened in May 2024 to update the previous WHO policy. This review is an update of a Cochrane review first published in 2018 [62], and previously updated in 2021 [63]. The review was updated to inform WHO policy guideline development in 2024. Previous reviews only included Xpert MTB/RIF and Xpert Ultra; however, this review update was intended to include newer studies for Xpert Ultra and search for more technologies that fit into this class of LC-aNAATs.

The Background and Methods sections of this review include some text that overlaps with some of our other Cochrane Reviews for LC-aNAATs for diagnosing tuberculosis [64, 65, 66, 67].

OBJECTIVES

To estimate the diagnostic accuracy of LC-aNAATs for:
a) extrapulmonary tuberculosis by site of disease; and
b) rifampicin resistance in adolescents and adults with presumptive extrapulmonary tuberculosis.

Presumptive tuberculosis refers to an individual who presents with symptoms or signs suggestive of tuberculosis.

Secondary objectives

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

For potential sources of heterogeneity, for extrapulmonary tuberculosis, we included smear status, HIV status, and prevalence of extrapulmonary tuberculosis. For cerebrospinal fluid (CSF), we considered the presence of a concentration step and specimen volume.

METHODS

Criteria for considering studies for this review

Types of studies

We included cross-sectional and cohort studies using non-respiratory specimens to estimate the accuracy of LC-aNAATs against a microbiological or composite reference standard for tuberculosis, and culture-based drug susceptibility testing or whole genome sequencing for rifampicin resistance. We included the following common forms of extrapulmonary tuberculosis: tuberculous meningitis and pleural, lymph node, bone or joint, genitourinary, peritoneal, pericardial, and disseminated tuberculosis. We excluded studies that evaluated the index tests in gastric fluid, as this specimen is used most often to investigate pulmonary tuberculosis in children. We also excluded stool specimens because tuberculosis bacteria may be swallowed and passed into stool as a marker of pulmonary tuberculosis. We only included studies that reported data comparing the index test(s) to an acceptable reference standard from which we could extract true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) values. We excluded case-control studies (i.e. multiple-group studies with two or more sets of eligibility criteria) and case reports.

Participants

We included studies where at least 85% of the participants enrolled were adults and adolescents aged 10 years or older with presumptive extrapulmonary tuberculosis from all settings and countries. We excluded studies where we could not disaggregate data on adults from those in children and studies where we could not tell the age of the participants enrolled. Restricting the age

group to adults and adolescents differs from the original review, where we also included children [62]. The update in 2021 also did not include children and was reflected clearly in the 2021 update. We did this because children are now included in a separate Cochrane Review, *Xpert MTB/RIF and Xpert MTB/RIF Ultra assays for active tuberculosis and rifampicin resistance in children* [65] (Supplementary material 12). Adults and adolescents were defined as people aged 10 years and above and were included in this review.

We included non-respiratory specimens (such as CSF, pleural fluid, lymph node aspirate or tissue). We excluded sputum and other respiratory specimens, such as fluid obtained from bronchial alveolar lavage and tracheal aspiration. As we anticipated finding many studies, we set a bar to exclude smaller studies to reduce unnecessary work. We therefore required studies to provide data for at least five specimens for a given form of extrapulmonary tuberculosis included in the review. We excluded studies evaluating the use of LC-aNAATs to diagnose relapse of previously-treated extrapulmonary tuberculosis to avoid the selection bias that may arise by limiting to a group that is already at elevated risk of extrapulmonary tuberculosis. We attempted to identify studies that included participants who were not taking anti-tuberculosis drugs or had taken anti-tuberculosis drugs for fewer than seven days.

We have tried to eliminate stigmatizing language, for example, by changing ‘suspected tuberculosis' to ‘presumptive tuberculosis'. Whenever possible, we extracted data per participant rather than per specimen. For most studies, the number of specimens was the same as the number of participants.

Index tests

We considered the following LC-aNAATs: Xpert Ultra; Truenat (Truenat MTB Plus and Truenat MTB-RIF Dx); STANDARD M10, SD Biosensor; and IRON-qPCR, BIONEER.

Index test results are automatically generated (i.e. there is a single threshold), and the user is provided with a printable test result as follows.

Xpert Ultra

- MTB (M tuberculosis) DETECTED HIGH; RIF (rifampicin) Resistance DETECTED;
- MTB DETECTED MEDIUM; RIF Resistance DETECTED;
- MTB DETECTED LOW; RIF Resistance DETECTED;
- MTB DETECTED VERY LOW; RIF Resistance DETECTED;
- MTB DETECTED HIGH; RIF Resistance NOT DETECTED;
- MTB DETECTED MEDIUM; RIF Resistance NOT DETECTED;
- MTB DETECTED LOW; RIF Resistance NOT DETECTED;
- MTB DETECTED VERY LOW; RIF Resistance NOT DETECTED;
- MTB DETECTED HIGH; RIF Resistance INDETERMINATE;
- MTB DETECTED MEDIUM; RIF Resistance INDETERMINATE;
- MTB DETECTED LOW; RIF Resistance INDETERMINATE;
- MTB DETECTED VERY LOW; RIF Resistance INDETERMINATE;
- MTB Trace DETECTED; RIF Resistance INDETERMINATE;
- INVALID (the presence or absence of MTB cannot be determined);
- ERROR (the presence or absence of MTB cannot be determined);
- NO RESULT (the presence or absence of MTB cannot be determined).

Xpert Ultra incorporates a semi-quantitative classification for results: trace, very low, low, moderate, and high. ‘Trace' corresponds to the lowest bacterial burden for detection of *M tuberculosis* [24]. We considered a trace result to mean MTB (*M tuberculosis*) DETECTED. However, no rifampicin-resistance result was available for participants with trace results because the trace sample is always reported as 'INDETERMINATE' for rifampin resistance [20]. Additionally, a rifampicin resistance or susceptible result can only be determined if the person has tuberculosis.

Truenat MTB Plus

We did not include Truenat MTB in this review because the WHO received official information from Molbio Diagnostics in 2024 that the Truenat MTB assay will no longer be sold in the international market. The results of Truenat MTB plus are categorized as follows.

1. MTB (M tuberculosis) DETECTED HIGH; RIF (rifampicin) Resistance DETECTED;
2. MTB DETECTED MEDIUM; RIF Resistance DETECTED;
3. MTB DETECTED LOW; RIF Resistance DETECTED;
4. MTB DETECTED VERY LOW; RIF Resistance DETECTED;
5. MTB DETECTED HIGH; RIF Resistance NOT DETECTED;
6. MTB DETECTED MEDIUM; RIF Resistance NOT DETECTED;
7. MTB DETECTED LOW; RIF Resistance NOT DETECTED;
8. MTB DETECTED VERY LOW; RIF Resistance NOT DETECTED.

If results are positive from Truenat MTB plus, as a follow-on test, Truenat MTB-RIF Dx is performed and the results are shown as “Rif Resistance Detected” if mutations are detected or “Rif Resistance Not Detected” if mutations are not detected. For failed runs, it is shown as “Indeterminate” or “Error”.

STANDARD M10 SD, SD Biosensor

These assays provide qualitative results within 80 minutes from sputum or sputum sediment specimens.

1. STANDARD™ M10 MTB/NTM assay aids in the simultaneous detection of *M tuberculosis* complex (MTBC) and non-tuberculous mycobacteria (NTM) DNA.
2. STANDARD™ M10 MDR-TB assay aids in the simultaneous detection of *M tuberculosis* and drug-resistance against rifampicin (RIF) and isoniazid (INH).

IRONqPCR, Bioneer

This assay provides qualitative results for TB and drug resistance detection. The system has the capacity for simultaneous detection of *Mycobacterium tuberculosis* and detection of resistance to rifampicin (RIF), fluoroquinolones (FQs; including moxifloxacin (MOX)), isoniazid (INH), and aminoglycosides (AGs). It can test two samples simultaneously within 30 minutes.

Target conditions

The target conditions were extrapulmonary tuberculosis and rifampicin resistance. We included eight common forms of extrapulmonary tuberculosis and considered these subcategories of the target condition as separate diagnostic classifications [2, 10, 68].

- Tuberculous meningitis;
- Pleural tuberculosis;

- Lymph node tuberculosis;
- Genitourinary tuberculosis;
- Bone or joint tuberculosis;
- Peritoneal tuberculosis;
- Pericardial tuberculosis;
- Disseminated tuberculosis.

Table 1 lists the forms of extrapulmonary tuberculosis and specimens used for diagnosis in the review. We excluded less common forms, such as cutaneous tuberculosis, ocular tuberculosis, female genital tuberculosis, and tuberculosis of the breast, ear, and paranasal sinuses [10].

Reference standards

Detection of extrapulmonary tuberculosis

We included two reference standards.

- Solid or liquid mycobacterial culture (microbiological reference standard)
  - 'Tuberculosis' was defined as a positive *M tuberculosis* culture
  - 'Not tuberculosis' was defined as a negative *M tuberculosis* culture;
- Composite reference standard
  - 'Tuberculosis' was defined as a positive *M tuberculosis* culture or positive composite reference test
  - 'Not tuberculosis' was defined as a negative *M tuberculosis* culture and a negative composite reference test.

The composite reference standard is based on the results of microbiological tests, culture or NAATs other than the index tests mentioned above: imaging studies, histology, and clinical characteristics, and includes at least one component test that is positive, according to the definition of the primary study authors.

For pleural tuberculosis, we defined the composite reference standard as the presence of granulomatous inflammation or a positive culture. We proposed this definition because we found evidence to support the inclusion of histopathological examination in the definition. Around 60% of patients undergoing pleural biopsy will show granulomatous inflammation [43]. A prospective cohort study of participants with clinical and radiological findings consistent with pleural tuberculosis [69] found that histological examination of tissue obtained from pleural biopsy had a higher diagnostic yield (78%; 66/84) than that of culture (62%; 52/84).

Culture is considered the best reference standard for tuberculosis. However, culture may lead to the misclassification of some cases of extrapulmonary tuberculosis as 'not tuberculosis', owing to the paucibacillary nature of the disease. This means that culture may have low sensitivity for extrapulmonary tuberculosis overall and, further, that culture sensitivity may differ for different forms of extrapulmonary tuberculosis [41]. This misclassification by culture may lead to biased estimates (overestimation or underestimation) of the diagnostic accuracy of the index tests.

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

In contrast to culture, a composite reference standard that includes culture, other tests, and clinical characteristics may correctly classify index test results as TPs (instead of as FPs with respect to culture), especially in people with paucibacillary disease in whom culture may be negative. However, because of the uncertainties that surround a clinical diagnosis of tuberculosis and, in some instances, the conditional dependence of the index tests and other tests in the composite reference standard (for example, for most of these tests, detection of tuberculosis depends on bacillary load), a reference standard that uses additional tests and clinical characteristics (in culture-negative people) may incorrectly classify people without tuberculosis as having tuberculosis [70]. An additional challenge with including a composite reference standard is that the definition of the composite reference standard may vary across studies, making it difficult to interpret the accuracy estimates.

Thus, both reference standards, culture and composite, are imperfect and may affect accuracy estimates.

Detection of rifampicin resistance

The reference standard was culture-based DST, also known as phenotypic DST (pDST) using solid or liquid media, as recommended by the WHO [19, 71] with or without whole genome sequencing, also known as genotypic drug susceptibility testing (gDST).

Search methods for identification of studies

We attempted to identify all relevant studies, regardless of language or publication status (published, unpublished, in press, or ongoing). We monitored abstracts to see if these studies were published during the time we performed the review. We also reached out to authors and other researchers in this area to ensure no relevant studies are missed.

We did not use generative AI in the search process.

We updated the search terms as indicated in Supplementary material 1, where we added new index test terms as this review update includes index tests other than Xpert Ultra as well.

Electronic searches

The Cochrane Infectious Diseases Group Information Specialist (VL) performed literature searches on 11 October 2023, without language restrictions, in the following databases, using the search terms described in Supplementary material 1:

- Ovid MEDLINE(R) ALL (from 1946 to 11 October 2023);
- Embase (Ovid, from 1947 to 11 October 2023);
- Science Citation Index Expanded (SCI-EXPANDED; Web of Science, from 1900 to 11 October 2023);

- BIOSIS previews (Web of Science, from1926 to 11 October 2023);
- Scopus (Elsevier), from 1970 to 11 October 2023);
- Cochrane Central Register of Controlled Trials (CENTRAL) published in the Cochrane Library, Issue 10 of 12, October 2023;
- WHO Global Index Medicus (accessed 11 October 2023).

We also searched ClinicalTrials.gov (clinicaltrials.gov/) and the WHO Clinical Trials Registry platform (www.who.int/trialsearch) on 18 October 2023, to identify ongoing trials. These searches were updated by the addition of index test names to capture other tests included in this class of technology.

Searching other resources

TThe authors examined the reference lists of included articles and relevant review articles identified through the electronic searches. VL also searched for relevant dissertations in ProQuest Dissertations & Theses A&I on 11 October 2023. The authors searched information on ongoing and unpublished studies from manufacturers through a WHO public call; and from experts working on new diagnostics for TB such as STOP TB Partnership's New Diagnostic Working Group and FIND (the global alliance for diagnostics).

We did not run another literature search to keep this review update aligned with the WHO policy update. As there were no ongoing or studies awaiting classification, no new studies were added after this initial search was performed in October 2023. However, there was a WHO public call for data till 15 February 2024 to identify unpublished and upcoming studies. We included one unpublished study that came through this WHO call.

Data collection and analysis

Selection of studies

We used Covidence to manage the selection of studies [72]. Two review authors (MK, LRI) independently scrutinized titles and abstracts identified by electronic literature searching to identify potentially eligible studies. We selected any citation identified by either review author as potentially eligible for full-text review. The same review authors independently assessed full-text papers for study eligibility using predefined inclusion and exclusion criteria, and resolved any discrepancies by discussion. We recorded all studies excluded after full-text assessment and their reasons for exclusion. We illustrated the study selection process in a PRISMA diagram [73, 74].

Data extraction and management

Using a previously-developed form (Supplementary material 6), two review authors (MK, LRI) worked independently to extract data on the following characteristics.

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

We classified a country's income status as either low- and middle-income or high-income, according to the World Bank List of Economies [75].

We extracted TP, FP, FN, and TN values for the following specimens: CSF, pleural fluid and tissue, lymph node aspirate and tissue (the latter specimen acquired by surgical biopsy), bone or joint aspirate and tissue, urine, peritoneal fluid and tissue, pericardial fluid and tissue, and blood. We extracted these values for each of the specimen types separately. For example, we used one 2 × 2 table for lymph node aspirate, and another 2 × 2 table for lymph node tissue. In situations in which a participant contributed more than one specimen but of different types, we extracted data for all specimens. When a study included data for both raw specimens and concentrated sediment involving the same participants, we preferentially extracted data for raw specimens, except in the case of CSF, for which we extracted data for concentrated sediment as recommended by the WHO [31]. We extracted accuracy data according to the defined reference standards (see Reference standards). We did not encounter any situations in which a subset of participants in a study received the reference standard, but others did not. Hence, there was no need to make corrections for verification bias in the statistical analysis [76].

In most studies, the number of specimens was the same as the number of participants. However, in some studies, the number of specimens exceeded the number of participants or study authors reported only the number of specimens.

We contacted authors of primary studies for missing data or clarifications. We entered all data into Microsoft Excel 2019 [77]. Primary authors were contacted twice with a one-month difference.

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

Assessment of methodological quality

We used the QUADAS-2 tool, tailored to this review, to assess the quality of the included studies (Supplementary material 7) [78]. QUADAS-2 consists of four domains: patient selection, index test, reference standard, and flow and timing. We assessed all domains for risk of bias and the first three domains for concerns about applicability. Two review authors (MK and LRI)

independently completed QUADAS-2 and resolved disagreements through discussion. We present the results of this quality assessment in text, tables, and graphs.

We modified QUADAS-2 as follows.

Participant selection domain, applicability: for tuberculous meningitis, owing to the severity of the illness, we judged 'low concern' if participants were evaluated as inpatients at tertiary care centers. In the original review, we judged tertiary care to be a setting of high concern.

Reference standard domain: we clarified that CSF, pleural fluid, and lymph node aspirates are usually considered to be sterile, and standards specify that these specimens may be placed directly into the culture medium. Overly processing specimens may lead to false-negative cultures. We scored 'yes' if studies did not use N-acetyl-L-cysteine-sodium hydroxide for processing sterile specimens and 'unclear' if studies used N-acetyl-L-cysteine-sodium hydroxide.

Investigations of heterogeneity: for specimen volume, we restricted this analysis to CSF because it was most clinically meaningful. For other fluid specimen types, the manufacturer's instructions for sputum were usually followed. In terms of the WHO standard operating procedure for lymph node tissue, we did not investigate this further because 80% (8/10) of the included studies followed the WHO recommendations. In performing the review, it became clear that, because a homogenization step is part of the WHO standard operating procedure for preparing tissue specimens, there was no need to perform an additional separate analysis to confirm the presence of a homogenization step. We removed the condition of the specimen (fresh or frozen) from the analysis, because we identified only six studies in the current review that used frozen specimens, and we had already performed an analysis of this possible source of heterogeneity for the Cochrane Review on Xpert MTB/RIF for pulmonary tuberculosis [79].

Statistical analysis and data synthesis

We performed descriptive analyses of the characteristics of the included studies using Stata 18.0 [80], and we present key study characteristics in the Supplementary material 2 table. We used data reported in the TP, FP, FN, and TN formats to calculate sensitivity and specificity estimates and 95% confidence intervals (CIs) for individual studies. We present individual study results graphically by plotting the estimates of sensitivity and specificity (and their 95% CIs) in forest plots using Review Manager (RevMan) web [81].

To assess the diagnostic accuracy of the index tests, we estimated summary sensitivity and specificity and corresponding 95% CIs using a bivariate random-effects model for meta-analysis [82,83]. The bivariate model allowed us to calculate the summary 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. In addition, we determined predictive values at a pretest probability of 10%, a value suggested by the WHO.

If bivariate meta-analysis was not possible due to sparse data, or if there was little or no variability in estimates of sensitivity and specificity across studies, we fitted univariate random-effects logistic regression models separately for sensitivity and specificity, as appropriate [84, 85]. We only conducted meta-analyses if there were three or more studies for each form of extrapulmonary tuberculosis. In addition to the number of studies, we also included the condition that the number of TB cases should be more than 30. If meta-analysis was not possible, we provided ranges of sensitivity and specificity.

All analyses were conducted using Stata 18.0 and meta-analysis models were fitted using the 'metandi' user-written function or the 'meqrlogit' command. Where sensitivity was 100% in all studies and univariate logistic regression analysis failed to converge, we calculated the summary estimate by summing the number of TPs across studies and calculating an exact (Clopper-Pearson [86]) 95% binomial CI.

We performed separate analyses grouped by type of extrapulmonary specimen (e.g. CSF, pleural fluid, peritoneal fluid), rather than estimating summary accuracy for all forms of extrapulmonary tuberculosis combined, because we considered the former approach to be the most clinically meaningful. In addition, we performed separate analyses using the reference standard.

For the accuracy of LC-aNAATs for detection of rifampicin resistance, we included participants who: (1) were culture-positive; (2) were tuberculosis-positive by the index test; and (3) had a valid result for rifampicin resistance, detected or not detected (susceptible) by the index test.

- Sensitivity = Index test rifampicin resistance detected/ phenotypic DST* resistant.
- Specificity = Index test rifampicin resistance not detected/ phenotypic DST* susceptible.

*With or without genotypic DST

For the detection of rifampicin resistance, when a study included multiple types of specimens, we based our analysis on all available data in the study, including data for specimens that we did not include in the primary analyses for the detection of extrapulmonary tuberculosis. For example, if a study provided data for several specimen types combined (e.g. all tissue specimens) and we could not disaggregate the data for a specific specimen type, we included all data (for all tissue specimens) in the analysis for rifampicin resistance detection. We did this because we did not expect the accuracy of rifampicin resistance detection to vary by specimen type. We used the bivariate random-effects model to estimate summary sensitivity and specificity.

For rifampicin resistance, we planned to assess the impact of the prevalence of rifampicin resistance on accuracy estimates, but we had insufficient data for this analysis.

Analyses performed in this review have been provided in Supplementary material 4 and the data package used in this review has been provided in Supplementary material 5.

Class-level analyses

As this review is about a class of technology, we developed criteria to assess when the test results could be combined into one class to provide summary accuracy estimates. These conditions are

provided in Table 2. Based on these conditions, Truenat MTB Plus did not meet the criteria to be included in a class-based analysis.

Approach to inconclusive index test results

The proportion of inconclusive (non-determinate) rates for the detection of pulmonary tuberculosis is the number of tests classified as 'invalid', 'error', or 'no result' divided by the total number of index tests performed. The proportion of inconclusive (indeterminate) rates for detection of rifampicin resistance is the number of tests classified as 'MTB DETECTED, Rif (rifampicin) resistance INDETERMINATE' divided by the total number of index test-positive results.

Investigations of heterogeneity

Initially, we investigated heterogeneity through visual examination of forest plots of sensitivities and specificities and through visual examination of the ROC space of the raw data. We used the summary ROC plots to visually assess heterogeneity; however, these plots were not included in the review as the forest plots provide a clearer and comprehensive visualization of individual study estimates and the uncertainty around the estimates. When data allowed, we evaluated potential sources of heterogeneity using subgroup analyses and bivariate meta-regression. We included the following covariate:

- HIV status;
- For tuberculous meningitis, the concentration step used for preparing specimens (yes or no);
- CSF specimen volume.

We had planned to investigate smear status, history of tuberculosis, and whether WHO standard procedures for preparing tissue specimens were followed. However, we had insufficient data to do this.

Sensitivity analyses

We did not perform sensitivity analyses.

Assessment of reporting bias

We did not perform a formal assessment of publication bias using methods such as funnel plots or regression tests because such techniques have not been helpful for diagnostic test accuracy studies [87].

Summary of findings and assessment of the certainty of the evidence

We assessed the certainty of evidence using the GRADE approach for diagnostic studies [88, 89, 90, 91]. As recommended, we rated the certainty of evidence as either high (not downgraded), moderate (downgraded by one level), low (downgraded by two levels), or very low (downgraded by more than two levels) based on five domains: risk of bias, indirectness, inconsistency, imprecision, and publication bias. For each outcome, the certainty of evidence started as high when there were high-quality studies (cross-sectional or cohort studies) that enrolled participants with diagnostic uncertainty. If we found a reason for downgrading, we used our judgment to classify the reason as either serious (downgraded by one level) or very serious (downgraded by two levels). If we downgraded by two levels for one of the five domains and downgraded by one level for any other of the four domains, this could lead to very low certainty of evidence. Two review authors discussed judgments and applied GRADE in the following way [92, 93, 94].

- *Assessment of risk of bias*. We used QUADAS-2 to assess the risk of bias.
- *Indirectness*. We assessed indirectness in relation to the population (including disease spectrum), setting, interventions, and outcomes (accuracy measures). We also used the prevalence of the target condition as a guide to whether there was indirectness in the population.
- *Inconsistency*. GRADE recommends downgrading for unexplained inconsistency in sensitivity and specificity estimates. We carried out prespecified analyses to investigate potential sources of heterogeneity and downgraded when we could not explain inconsistency in the accuracy estimates.
- *Imprecision*. We considered a precise estimate to be one that would allow a clinically meaningful decision. We considered the width of the CI and asked, 'Would we make a different decision if the lower or upper boundary of the CI represented the truth?' In addition, we worked out projected ranges for TP, FN, TN, and FP for a given prevalence of tuberculosis and made judgments on imprecision from these calculations.
- *Publication bias*. We rated publication bias as undetected (not serious) for several reasons: the comprehensiveness of the literature search and extensive outreach to tuberculosis researchers to identify studies; the presence only of studies that produced precise estimates of high accuracy despite small sample size; and our knowledge of studies that were conducted but not published.

For the summary of findings tables for CSF and pleural fluid, we provide evidence using a microbiological reference standard, which is considered the best reference standard for tuberculosis [41]. For lymph node aspirate, we provide evidence using a composite reference because, based on findings from the original review [62], we believe a composite reference standard is preferable for estimating accuracy.

Prevalences used in these tables were finalized after conversations and consensus from the WHO. These were decided based on the range of prevalences observed in different high and low-burden settings for these forms of extrapulmonary tuberculosis. The domains mentioned above helped in reaching the certainty of evidence which are provided as explanatory notes for each table. Based on the meta-analyzed results and prevalences, we provided the TP, FN, TN, and FP values that would be observed in a cohort of 1000 people.

RESULTS

Results of the search

In this review update, we identified 575 records for screening. After removing 351 duplicates, 224 abstracts were assessed. Of these, 75 full-text articles were reviewed against our inclusion criteria. We excluded 49 studies for reasons including: not evaluating the index test(s), insufficient data to construct 2 × 2 tables, inappropriate reference standard, lack of data by specimen type, absence of extrapulmonary specimens, duplicate data, case-control designs, inclusion of children, screening study designs, case report formats, or fewer than five specimens for a given type.

From our previous review of Xpert Ultra, 11 additional studies were included, bringing the total to 37 studies in this update. Among these, 36 evaluated Xpert Ultra and three evaluated Truenat MTB Plus. Two Truenat MTB Plus studies (Sharma 2023 [95], Sharma 2021 [96]) also assessed Xpert Ultra. No eligible studies were found for Truenat MTB-RIF Dx. Two studies on Iron qPCR and Standard M10 were identified but excluded — one was an ongoing trial on pulmonary TB, and the other was a case-control study presented only as a conference abstract.

In total, 37 studies (comprising 39 datasets) met the inclusion criteria. Of these, all but one study (Boloko 2022 [97]) were included in the quantitative analyses.

Figure 2 shows the flow of studies in the review. We recorded the excluded studies and the reasons for their exclusion in Supplementary material 3.

Figure 2. Flow of studies through the screening and selection process.

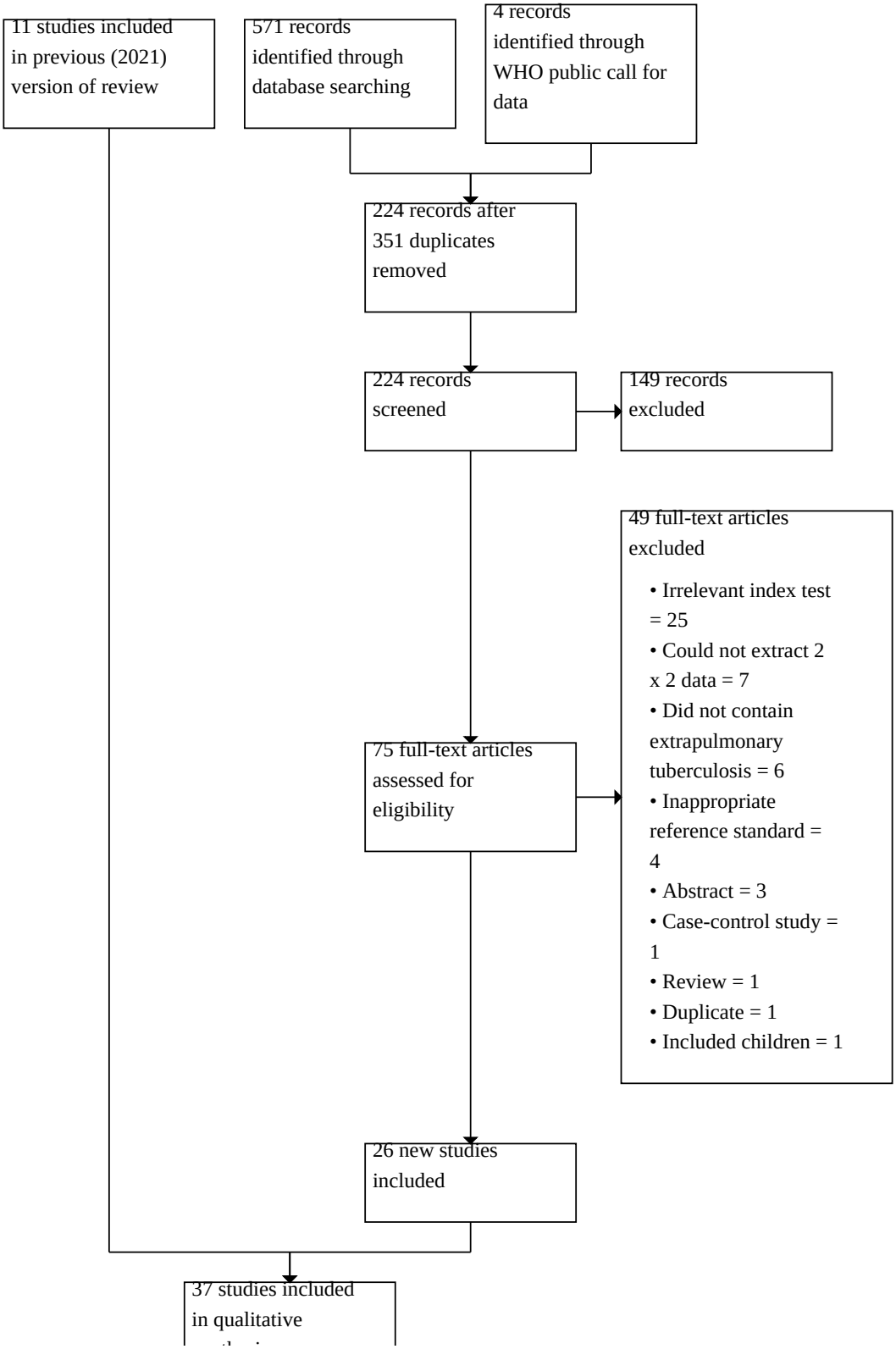
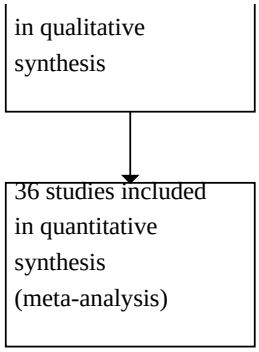


Figure 2. (Continued)



Methodological quality of included studies

Studies evaluating LC-aNAATs for detection of extrapulmonary tuberculosis

Figure 3 shows the risk of bias and applicability concerns for each of the 37 studies included for tuberculosis detection. Also, we separately present risk of bias and applicability concerns for studies evaluating different forms of extrapulmonary tuberculosis (Supplementary material 8; Supplementary material 9; and Supplementary material 10). Table 3 provides summary characteristics of all included studies.

Figure 3. Risk of bias and applicability concerns summary of LC-aNAATs for tuberculosis detection: review authors' judgments about each domain for each included study ***Note: most boxes are blank for Index test: Truenat as there are only three studies that included this index test.***

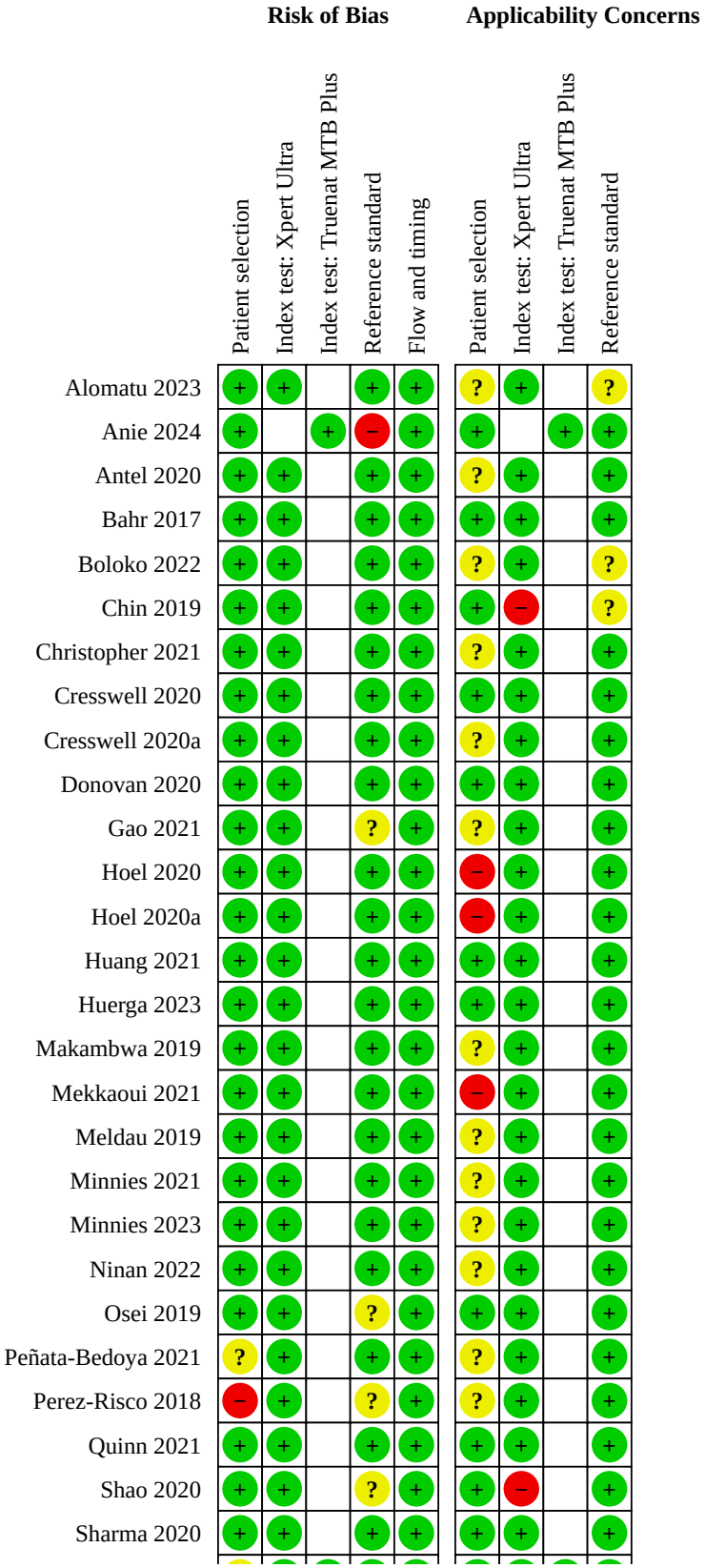
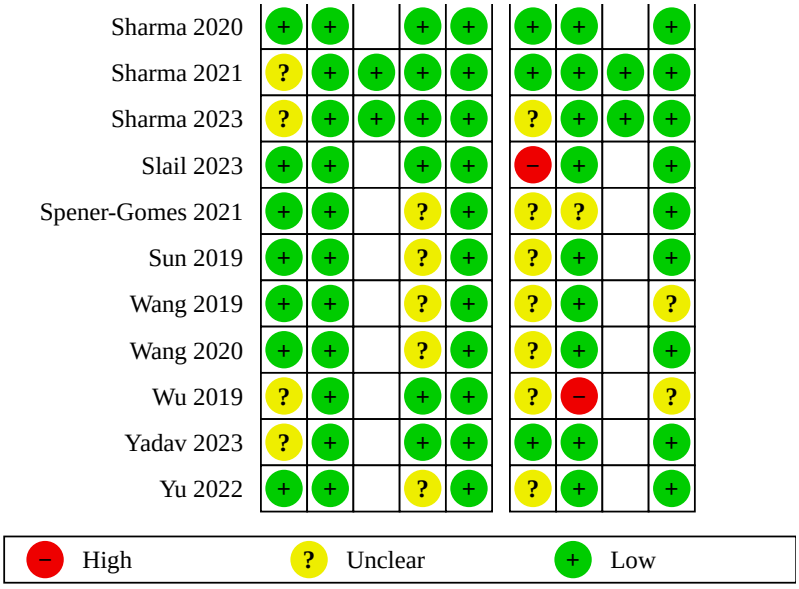


Figure 3. (Continued)



In the patient selection domain, we thought that 31 studies (84%) had low risk of bias, and five studies (13%) had unclear risk of bias because the manner of patient selection was unclear or non-consecutive or random (Peñata-Bedoya 2021 [98]; Sharma 2021; Sharma 2023; Wu 2019 [99]; Yadav 2023 [100]) and high risk of bias due to inappropriate exclusions (Perez-Risco 2018 [101]). Regarding applicability (patient characteristics and setting), we thought that 13 studies (35%) had low concerns because participants were evaluated in local hospitals or primary health settings or, in the case of tuberculous meningitis, tertiary centers (Anie 2024 [102]; Bahr 2017 [103]; Chin 2019 [104]; Cresswell 2020 [105]; Donovan 2020 [106]; Huang 2021 [107]; Huerga 2023 [108]; Osei 2019 [109]; Quinn 2021 [110]; Shao 2020 [111]; Sharma 2020 [112]; Sharma 2021; Yadav 2023). Four studies (11%) had high concerns because participants were evaluated exclusively as inpatients at a tertiary care center (Hoel 2020 [113]; Hoel 2020a [114]; Mekkaoui 2021 [115]; Slail 2023 [116]); and 20 (54%) studies had unclear concerns because we could not identify the clinical setting (Alomatu 2023 [117]; Antel 2020 [118]; Boloko 2022; Christopher 2021 [119]; Cresswell 2020a [120]; Gao 2021 [121]; Makambwa 2019 [122]; Meldau 2019 [123]; Minnies 2021 [124]; Minnies 2023 [125]; Ninan 2022 [126]; Peñata-Bedoya 2021; Perez-Risco 2018; Sharma 2023; Spener-Gomes 2021 [127]; Sun 2019 [128]; Wang 2019 [129]; Wang 2020 [130]; Wu 2019; Yu 2022 [131]).

In the index test domain, we judged that all studies had low risk of bias because the results of the index tests are automatically generated, the user is provided with printable test results, and the test threshold is prespecified. Regarding applicability, we judged that 32 studies (86%) had low concerns and three studies (8%) had high concerns because the index test was not performed according to WHO standard operating procedures (Chin 2019; Shao 2020; Wu 2019).

In the reference standard domain, 27 studies (73%) had low risk of bias because the results of the reference standard were interpreted without knowledge of the results of the index test and only non-sterile specimens were decontaminated. One study (3%) had a high risk of bias because the results of the reference standard were interpreted with knowledge of the results of the index test (Anie 2024). Nine studies (24%) had unclear risk of bias for the following reasons: three studies did not report whether there was blinding of the reference standard (Perez-Risco 2018; Wang 2019; Wang 2020), and six studies decontaminated specimens generally considered to be sterile (Gao 2021; Osei 2019; Shao 2020; Spener-Gomes 2021; Sun 2019; Yu 2022).

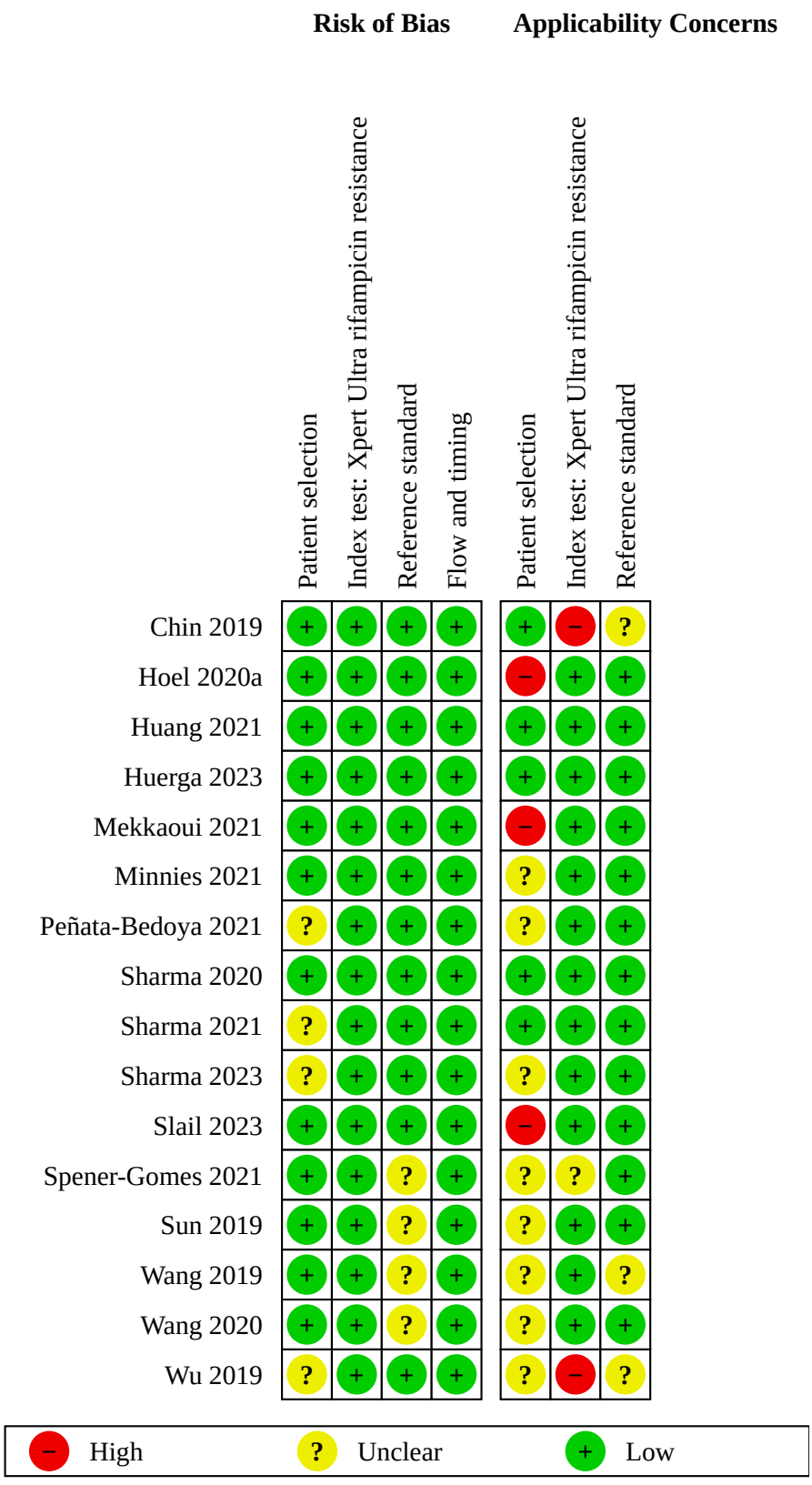
Regarding the applicability of the reference standard, we judged that 32 studies (86%) had low concerns because these studies performed a test to identify *M tuberculosis* species (speciation) and five studies (14%) had unclear concerns because we could not tell whether the study performed speciation (Alomatu 2023; Boloko 2022; Chin 2019; Wang 2019; Wu 2019).

In the flow and timing domain, we considered all studies to have low risk of bias, noting that all participants were accounted for in the analysis.

Studies evaluating LC-aNAATs for detection of rifampicin resistance

Figure 4 shows risk of bias and applicability concerns for each of the 16 studies included for rifampicin resistance detection.

Figure 4. Risk of bias and applicability concerns summary of LC-aNAATs for detection of rifampicin resistance: review authors' judgments about each domain for each included study



In the patient selection domain, we judged that four studies (25%) had unclear risk of bias (Peñata-Bedoya 2021; Sharma 2021; Sharma 2023; Wu 2019) as the manner of patient selection was unclear, while most (75%) of the studies had low risk of bias. For applicability, we thought that three studies (19%) had high concerns as the studies were conducted in low tuberculosis-burden countries (Hoel 2020a; Mekkaoui 2021; Slail 2023). Eight studies (50%) had unclear concerns because we could not identify the details of the clinical setting (Minnies 2021; Peñata-Bedoya 2021; Sharma 2023; Spener-Gomes 2021; Sun 2019; Wang 2019; Wang 2020; Wu 2019).

In the index test domain, we judged that all studies had low risk of bias because the results of the index tests are automatically generated, the user is provided with printable test results, and the test threshold is prespecified. For applicability, two studies (12%) had high concerns because fewer than 50% of the specimen types in these studies were processed according to WHO recommendations (Chin 2019; Wu 2019). One study [127] had unclear applicability concerns in the index test domain as it was not clear if the specimens processed the specimens according to the WHO recommendations.

In the reference standard domain, four studies (60%) had unclear risk of bias as it was unclear whether blinding of the reference standard was performed (Spener-Gomes 2021; Sun 2019; Wang 2019; Wang 2020). For the applicability of the reference standard, we judged that all studies had low concerns because detection of rifampicin resistance occurs only when the *M tuberculosis* target is present within the specimen.

In the flow and timing domain, we considered all studies to have low risk of bias, noting that all participants were accounted for in the analysis.

Findings

The 37 studies were conducted in 13 different countries. Most of the studies were conducted in China (n = 8), India (n = 10), South Africa (n = 9), and Uganda (n = 6). Of the 37 studies, 31 (84%) took place in high tuberculosis-burden countries and 22 (59%) in high-tuberculosis/HIV-burden countries. Most studies performed the index tests and culture on the same specimen type, except for one study (Boloko 2022) in which Xpert Ultra was performed on blood and culture was performed on sputum. Most studies did not report the exact number of cultures used to confirm a diagnosis of tuberculosis, but it is likely that many studies used a single culture.

We contacted the primary authors of six studies to request 2 × 2 tables by specimen type, details on blinding of the reference standard, and demographic information on the included populations. For two studies (Sharma 2020, Sharma 2021), the authors could not provide demographic data because the specimens were received directly by the laboratory without accompanying participant information. However, all authors were able to provide data on 2 x 2 tables, reference standards, and blinding.

We present key characteristics of the included studies in Supplementary material 2.

I. Detection of extrapulmonary tuberculosis

Both studies on Xpert Ultra and Truenat MTB plus are categorized as LC-aNAAT.

Xpert Ultra: of the 36 studies, the number of studies evaluating different specimens was as follows: tuberculous meningitis (CSF), 16 studies; pleural tuberculosis (pleural fluid), 13 studies; lymph node tuberculosis (lymph node aspirate), nine studies; genitourinary tuberculosis (urine), seven studies; bone or joint tuberculosis (bone or joint aspirate), five studies; pericardial tuberculosis (pericardial fluid), five studies; peritoneal tuberculosis (peritoneal fluid), three studies and disseminated tuberculosis (blood), one study. Several studies included more than one specimen.

Truenat MTB plus: of the three studies, the number of studies evaluating different specimens was as follows: tuberculous meningitis (CSF), two studies; lymph node tuberculosis (lymph node aspirate), three studies; bone or joint tuberculosis, one study, and peritoneal tuberculosis, one study.

Table 4 presents summary sensitivity and specificity estimates and predictive values by reference standard for all forms of extrapulmonary tuberculosis and specimen types included in the review.

For the class-level analysis of each specimen type across different technologies, we could only include studies from Xpert Ultra based on the criteria in Table 2. Therefore, the summary estimates in Table 4 were obtained only from studies of Xpert Ultra. We present the findings in detail below for each specimen type together with statements about the certainty of the evidence. For Truenat MTB Plus, we present data separately in Supplementary material 11.

A: LC-aNAAT testing in cerebrospinal fluid for tuberculous meningitis

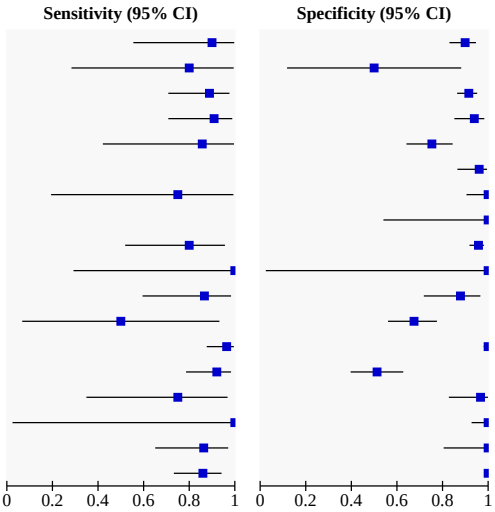
Microbiological reference standard

Eighteen studies (16 Xpert Ultra; 2 Truenat MTB plus) evaluated Xpert Ultra in cerebrospinal fluid (CSF) specimens against culture. LC-aNAAT sensitivity ranged from 50% to 100% and specificity ranged from 50% to 100% (Figure 5). Chin 2019 reported the lowest specificity (50%). In this study, the investigators inoculated uncentrifuged CSF, which could have led to lower culture positivity, thus resulting in a higher number of false positives. Perez-Risco 2018 (specificity 100%) contributed only one participant to this analysis. Only 16 Xpert Ultra studies contributed to the meta-analysis and summary sensitivity and specificity (95% CI) were 88.2% (83.7 to 91.6) and 96.0% (86.8 to 98.9) (1684 participants; Table 4, Summary of findings 1).

Figure 5. Forest plots of Xpert Ultra sensitivity and specificity in cerebrospinal fluid by reference standard and in people living with HIV. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive

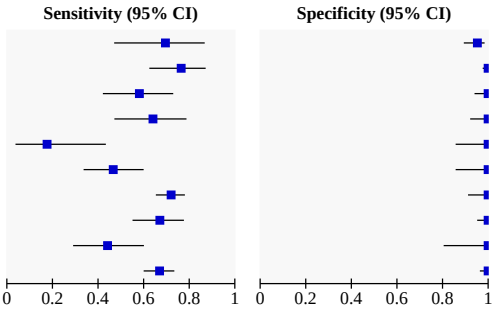
Cerebrospinal fluid, Xpert Ultra, culture

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bahr 2017	9	12	1	107	0.90 [0.55, 1.00]	0.90 [0.83, 0.95]
Chin 2019	4	3	1	3	0.80 [0.28, 0.99]	0.50 [0.12, 0.88]
Cresswell 2020	24	15	3	162	0.89 [0.71, 0.98]	0.92 [0.86, 0.95]
Donovan 2020	20	4	2	62	0.91 [0.71, 0.99]	0.94 [0.85, 0.98]
Huang 2021	6	19	1	58	0.86 [0.42, 1.00]	0.75 [0.64, 0.84]
Mekkaoui 2021	0	2	0	49	Not estimable	0.96 [0.87, 1.00]
Ninan 2022	3	0	1	37	0.75 [0.19, 0.99]	1.00 [0.91, 1.00]
Osei 2019	0	0	0	6	Not estimable	1.00 [0.54, 1.00]
Peñata-Bedoya 2021	12	8	3	181	0.80 [0.52, 0.96]	0.96 [0.92, 0.98]
Perez-Risco 2018	3	0	0	1	1.00 [0.29, 1.00]	1.00 [0.03, 1.00]
Quinn 2021	13	4	2	29	0.87 [0.60, 0.98]	0.88 [0.72, 0.97]
Shao 2020	2	26	2	54	0.50 [0.07, 0.93]	0.68 [0.56, 0.78]
Sharma 2020	54	0	2	188	0.96 [0.88, 1.00]	1.00 [0.98, 1.00]
Sharma 2021	35	38	3	40	0.92 [0.79, 0.98]	0.51 [0.40, 0.63]
Slail 2023	6	1	2	29	0.75 [0.35, 0.97]	0.97 [0.83, 1.00]
Spener-Gomes 2021	1	0	0	49	1.00 [0.03, 1.00]	1.00 [0.93, 1.00]
Wang 2019	19	0	3	17	0.86 [0.65, 0.97]	1.00 [0.80, 1.00]
Yadav 2023	43	0	7	250	0.86 [0.73, 0.94]	1.00 [0.99, 1.00]



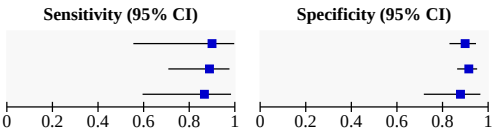
Cerebrospinal fluid, Xpert Ultra, composite reference standard

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bahr 2017	16	5	7	101	0.70 [0.47, 0.87]	0.95 [0.89, 0.98]
Cresswell 2020	39	0	12	153	0.76 [0.63, 0.87]	1.00 [0.98, 1.00]
Donovan 2020	25	0	18	60	0.58 [0.42, 0.73]	1.00 [0.94, 1.00]
Huang 2021	25	0	14	45	0.64 [0.47, 0.79]	1.00 [0.92, 1.00]
Ninan 2022	3	0	14	24	0.18 [0.04, 0.43]	1.00 [0.86, 1.00]
Shao 2020	28	0	32	24	0.47 [0.34, 0.60]	1.00 [0.86, 1.00]
Sharma 2020	147	0	57	40	0.72 [0.65, 0.78]	1.00 [0.91, 1.00]
Sharma 2021	49	0	24	75	0.67 [0.55, 0.78]	1.00 [0.95, 1.00]
Wang 2019	19	0	24	17	0.44 [0.29, 0.60]	1.00 [0.80, 1.00]
Yadav 2023	134	0	66	100	0.67 [0.60, 0.73]	1.00 [0.96, 1.00]



Cerebrospinal fluid, Xpert Ultra, culture, people living with HIV

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bahr 2017	9	12	1	107	0.90 [0.55, 1.00]	0.90 [0.83, 0.95]
Cresswell 2020	24	15	3	162	0.89 [0.71, 0.98]	0.92 [0.86, 0.95]
Quinn 2021	13	4	2	29	0.87 [0.60, 0.98]	0.88 [0.72, 0.97]



Composite reference standard

LC-aNAAT summary sensitivity and specificity (95% CI) were 60.3% (50.9 to 69.0) and 99.2% (98.1 to 99.7) from 10 Xpert Ultra studies (1397 participants; Table 4, Figure 5).

Investigations of heterogeneity

Xpert Ultra in people living with HIV

We identified three studies (Bahr 2017; Cresswell 2020; Quinn 2021) that provided information for Xpert Ultra accuracy in people living with HIV, against a microbiological reference standard. There were no Truenat studies which provided this information. LC-aNAAT summary sensitivity and specificity (95% CI) were 88.5% (76.6 to 94.7) and 90.6% (86.9 to 93.3) (3 studies; 381 participants; Figure 5).

Specimen concentration

Xpert Ultra

We found that concentrating CSF improved both Xpert Ultra sensitivity and specificity. Xpert Ultra summary sensitivity in concentrated specimens was 92.8% (87.5 to 96.0) (5 studies; 781 participants) versus 81.0% (68.0 to 89.5) (6 studies; 470 participants) in unconcentrated specimens. Xpert Ultra summary specificity in concentrated specimens was 93.6% (70.8 to 98.9) versus 85.8% (68.9 to 94.3) in unconcentrated specimens (Table 5).

Cerebrospinal fluid collection volumes

Xpert Ultra

Eleven studies reported the volume of CSF collected for Xpert Ultra testing, which ranged from 0.8 mL to > 6 mL. Table 6 provides accuracy estimates, CSF volumes and concentration steps. We did not observe any important trends for this analysis.

B: LC-aNAAT testing in pleural fluid for pleural tuberculosis

Microbiological reference standard

Thirteen studies evaluated Xpert Ultra in pleural fluid with respect to culture. Xpert Ultra sensitivity ranged from 0% to 100%, and specificity ranged from 57% to 100% (Figure 6).

Figure 6. Forest plot of Xpert Ultra sensitivity and specificity in pleural fluid and tissue by reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive

Pleural fluid, Xpert Ultra, culture

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Christopher 2021	4	9	2	65	0.67 [0.22, 0.96]	0.88 [0.78, 0.94]
Gao 2021	3	10	3	45	0.50 [0.12, 0.88]	0.82 [0.69, 0.91]
Hoel 2020	0	0	1	12	0.00 [0.00, 0.97]	1.00 [0.74, 1.00]
Makambwa 2019	13	6	22	8	0.37 [0.21, 0.55]	0.57 [0.29, 0.82]
Mekkaoui 2021	6	1	3	67	0.67 [0.30, 0.93]	0.99 [0.92, 1.00]
Minnies 2023	12	9	0	59	1.00 [0.74, 1.00]	0.87 [0.76, 0.94]
Peñata-Bedoya 2021	3	6	0	84	1.00 [0.29, 1.00]	0.93 [0.86, 0.98]
Perez-Risco 2018	10	0	11	3	0.48 [0.26, 0.70]	1.00 [0.29, 1.00]
Slail 2023	26	21	5	127	0.84 [0.66, 0.95]	0.86 [0.79, 0.91]
Spener-Gomes 2021	3	2	0	6	1.00 [0.29, 1.00]	0.75 [0.35, 0.97]
Wang 2019	48	18	11	33	0.81 [0.69, 0.90]	0.65 [0.50, 0.78]
Wang 2020	46	1	9	83	0.84 [0.71, 0.92]	0.99 [0.94, 1.00]
Wu 2019	17	30	6	72	0.74 [0.52, 0.90]	0.71 [0.61, 0.79]

Pleural fluid, Xpert Ultra, composite reference standard

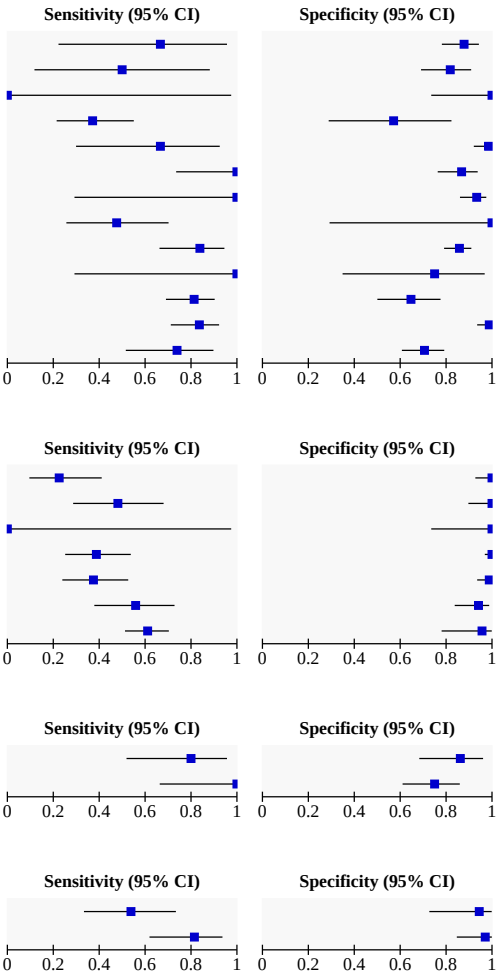
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Christopher 2021	7	0	24	49	0.23 [0.10, 0.41]	1.00 [0.93, 1.00]
Gao 2021	13	0	14	34	0.48 [0.29, 0.68]	1.00 [0.90, 1.00]
Hoel 2020	0	0	1	12	0.00 [0.00, 0.97]	1.00 [0.74, 1.00]
Makambwa 2019	19	0	30	116	0.39 [0.25, 0.54]	1.00 [0.97, 1.00]
Meldau 2019	18	1	30	83	0.38 [0.24, 0.53]	0.99 [0.94, 1.00]
Minnies 2023	19	3	15	48	0.56 [0.38, 0.73]	0.94 [0.84, 0.99]
Wang 2019	66	1	42	22	0.61 [0.51, 0.70]	0.96 [0.78, 1.00]

Pleural tissue, Xpert Ultra, culture

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Christopher 2021	12	4	3	25	0.80 [0.52, 0.96]	0.86 [0.68, 0.96]
Gao 2021	9	13	0	39	1.00 [0.66, 1.00]	0.75 [0.61, 0.86]

Pleural tissue, Xpert Ultra, composite reference standard

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Christopher 2021	14	1	12	17	0.54 [0.33, 0.73]	0.94 [0.73, 1.00]
Gao 2021	22	1	5	33	0.81 [0.62, 0.94]	0.97 [0.85, 1.00]



LC-aNAATs summary sensitivity and specificity (95% CI) were 74.0% (60.8 to 83.9) and 88.1% (78.8 to 93.6) from 13 Xpert Ultra studies (1041 participants; Table 4; Summary of findings 2).

Composite reference standard

Seven studies evaluated Xpert Ultra in pleural fluid with respect to a composite reference standard (Figure 6). Sensitivity ranged from 23% to 61%, and specificity ranged from 94% to 100%.

LC-aNAATs summary sensitivity and specificity (95% CI) were 43.6% (32.8 to 55.0) and 99.2% (95.2 to 99.9) from seven Xpert Ultra studies (667 participants; Table 4).

C: LC-aNAAT testing in pleural tissue for pleural tuberculosis

Microbiological reference standard

We identified two studies evaluating Xpert Ultra in pleural tissue against culture (Christopher 2021; Gao 2021). Meta-analysis was not performed due to paucity of data. The sensitivities of LC-aNAATs ranged from 80% to 100% and specificities from 75% to 86% from two Xpert Ultra studies (105 participants; Table 4; Figure 6).

Composite reference standard

We did not perform meta-analysis against a composite reference standard. The sensitivities of LC-aNAATs ranged from 54% to 81%

and specificities ranged from 94% to 97% from two Xpert Ultra studies (105 participants; Table 4; Figure 6).

D: LC-aNAAT testing in lymph node aspirate for lymph node tuberculosis

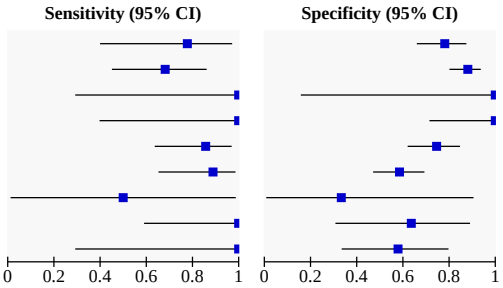
Microbiological reference standard

Nine studies evaluated Xpert Ultra in lymph node aspirate with respect to culture. Xpert Ultra sensitivity ranged from 50% to 100% and specificity ranged from 33% to 100% (Figure 7).

Figure 7. Forest plots of Xpert Ultra sensitivity and specificity in lymph node aspirate by reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive

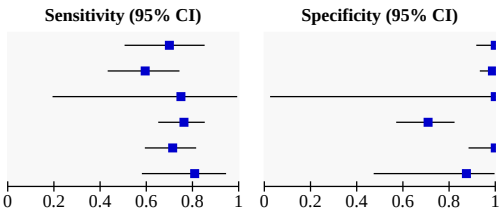
Lymph node aspirate, Xpert Ultra, culture

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Antel 2020	7	14	2	50	0.78 [0.40, 0.97]	0.78 [0.66, 0.87]
Christopher 2021	15	12	7	89	0.68 [0.45, 0.86]	0.88 [0.80, 0.94]
Hoel 2020	3	0	0	2	1.00 [0.29, 1.00]	1.00 [0.16, 1.00]
Hoel 2020a	4	0	0	11	1.00 [0.40, 1.00]	1.00 [0.72, 1.00]
Minnies 2021	18	16	3	47	0.86 [0.64, 0.97]	0.75 [0.62, 0.85]
Sharma 2023	16	34	2	48	0.89 [0.65, 0.99]	0.59 [0.47, 0.69]
Slail 2023	1	2	1	1	0.50 [0.01, 0.99]	0.33 [0.01, 0.91]
Spener-Gomes 2021	7	4	0	7	1.00 [0.59, 1.00]	0.64 [0.31, 0.89]
Yu 2022	3	8	0	11	1.00 [0.29, 1.00]	0.58 [0.33, 0.80]



Lymph node aspirate, Xpert Ultra, composite reference standard

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Antel 2020	21	0	9	43	0.70 [0.51, 0.85]	1.00 [0.92, 1.00]
Christopher 2021	25	1	17	80	0.60 [0.43, 0.74]	0.99 [0.93, 1.00]
Hoel 2020	3	0	1	1	0.75 [0.19, 0.99]	1.00 [0.03, 1.00]
Minnies 2021	58	16	18	39	0.76 [0.65, 0.85]	0.71 [0.57, 0.82]
Sharma 2023	50	0	20	30	0.71 [0.59, 0.82]	1.00 [0.88, 1.00]
Yu 2022	17	1	4	7	0.81 [0.58, 0.95]	0.88 [0.47, 1.00]



The summary sensitivity and specificity (95% CI) of LC-aNAATs were 85.3% (73.4 to 92.4) and 74.1% (63.5 to 82.5) from nine Xpert Ultra studies (445 participants; Table 4).

Composite reference standard

Six studies evaluated Xpert Ultra in lymph node aspirates with respect to a composite reference standard. The sensitivity ranged from 60% to 81% and the specificity ranged from 71% to 100% (Figure 7).

The summary sensitivity and specificity (95% CI) of LC-aNAATs were 71.3% (64.3 to 77.4) and 97.4% (82.3 to 99.7) from six Xpert Ultra studies (461 participants; Table 4; Summary of findings 3). Of note, with a composite reference standard, specificity was higher (100%) than that observed when using culture as the reference standard (74%).

E: LC-aNAAT testing in lymph node biopsies for lymph node tuberculosis

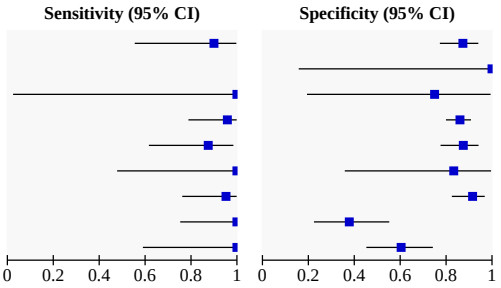
Microbiological reference standard

Nine studies evaluated Xpert Ultra for lymph node biopsies against culture as a reference standard. The sensitivity ranged from 88% to 100% and the specificity ranged from 38% to 100% (Figure 8).

Figure 8. Forest plots of Xpert Ultra sensitivity and specificity in lymph node biopsy by reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive

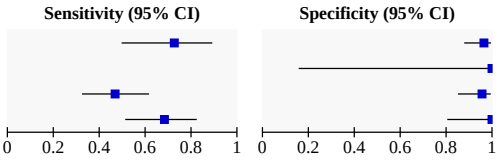
Lymph node biopsy, Xpert Ultra, culture

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Antel 2020	9	9	1	62	0.90 [0.55, 1.00]	0.87 [0.77, 0.94]
Hoel 2020	0	0	0	2	Not estimable	1.00 [0.16, 1.00]
Hoel 2020a	1	1	0	3	1.00 [0.03, 1.00]	0.75 [0.19, 0.99]
Mekkaoui 2021	23	24	1	148	0.96 [0.79, 1.00]	0.86 [0.80, 0.91]
Ninan 2022	14	9	2	63	0.88 [0.62, 0.98]	0.88 [0.78, 0.94]
Peñata-Bedoya 2021	5	1	0	5	1.00 [0.48, 1.00]	0.83 [0.36, 1.00]
Slail 2023	20	6	1	65	0.95 [0.76, 1.00]	0.92 [0.83, 0.97]
Wu 2019	13	23	0	14	1.00 [0.75, 1.00]	0.38 [0.22, 0.55]
Yu 2022	7	19	0	29	1.00 [0.59, 1.00]	0.60 [0.45, 0.74]



Lymph node biopsy, Xpert Ultra, composite reference standard

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Antel 2020	16	2	6	55	0.73 [0.50, 0.89]	0.96 [0.88, 1.00]
Hoel 2020	0	0	0	2	Not estimable	1.00 [0.16, 1.00]
Ninan 2022	23	2	26	44	0.47 [0.33, 0.62]	0.96 [0.85, 0.99]
Yu 2022	26	0	12	17	0.68 [0.51, 0.82]	1.00 [0.80, 1.00]



The summary sensitivity and specificity (95% CI) of LC-aNAATs were 96.5% (84.7 to 99.3) and 79.4% (65.4 to 88.8) from eight Xpert Ultra studies (578 participants; Table 4).

Composite reference standard

Four studies evaluated lymph node biopsies for Xpert Ultra. The sensitivity ranged from 47% to 73% and the specificity ranged from 96% to 100%.

The summary sensitivity and specificity (95% CI) of LC-aNAATs were 61.5% (47.1 to 74.2) and 96.7% (91.5 to 98.7) from three Xpert Ultra studies (229 participants; Figure 8).

F: LC-aNAAT testing in urine for genitourinary tuberculosis

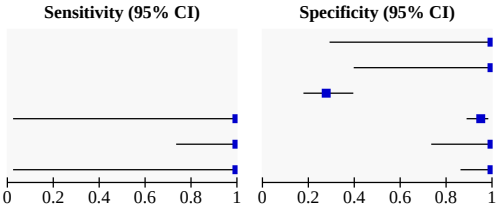
Microbiological reference standard

Six studies evaluated Xpert Ultra in urine against culture as the reference standard. The sensitivity was 100% (3 studies) and specificity ranged from 28% to 100% (6 studies, 232 participants; Figure 9).

Figure 9. Forest plots of Xpert Ultra sensitivity and specificity in urine by reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive

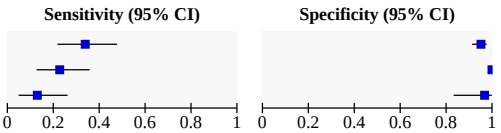
Urine, Xpert Ultra, culture

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Huerga 2023	0	0	0	3	Not estimable	1.00 [0.29, 1.00]
Mekkaoui 2021	0	0	0	4	Not estimable	1.00 [0.40, 1.00]
Osei 2019	0	52	0	20	Not estimable	0.28 [0.18, 0.40]
Peñata-Bedoya 2021	1	5	0	97	1.00 [0.03, 1.00]	0.95 [0.89, 0.98]
Perez-Risco 2018	12	0	0	12	1.00 [0.74, 1.00]	1.00 [0.74, 1.00]
Spener-Gomes 2021	1	0	0	25	1.00 [0.03, 1.00]	1.00 [0.86, 1.00]



Urine, Xpert Ultra, composite reference standard

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Cresswell 2020a	19	10	37	198	0.34 [0.22, 0.48]	0.95 [0.91, 0.98]
Huerga 2023	13	0	44	295	0.23 [0.13, 0.36]	1.00 [0.99, 1.00]
Minnies 2021	6	1	40	30	0.13 [0.05, 0.26]	0.97 [0.83, 1.00]



Due to the low number of specimens, it was not possible to meta-analyze the data.

Composite reference standard

Three studies evaluated Xpert Ultra in urine against a composite reference standard. The sensitivity ranged from 13% to 34% and the specificity ranged from 95% to 100% (Figure 9).

The summary sensitivity and specificity (95% CI) of LC-aNAATs were 23.0% (14.7 to 34.1) and 98.9% (89.7 to 99.9) from three Xpert Ultra studies (693 participants; Table 4).

G: LC-aNAAT testing in bone or joint aspirate for bone or joint tuberculosis

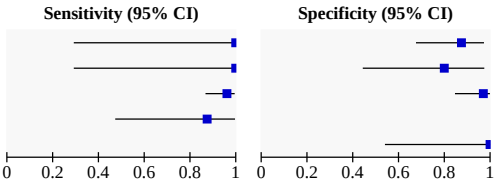
Microbiological reference standard

Five studies evaluated Xpert Ultra in bone or joint aspirate against culture as the reference standard. The sensitivity ranged from 88% to 100% and specificity ranged from 80% to 100% (Figure 10).

Figure 10. Forest plots of Xpert Ultra sensitivity and specificity in bone or joint fluid and tissue by reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive

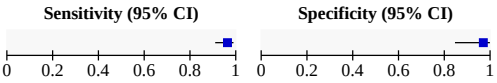
Bone or joint aspirate, Xpert Ultra, culture

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Mekkaoui 2021	3	3	0	21	1.00 [0.29, 1.00]	0.88 [0.68, 0.97]
Slail 2023	3	2	0	8	1.00 [0.29, 1.00]	0.80 [0.44, 0.97]
Sun 2019	50	1	2	33	0.96 [0.87, 1.00]	0.97 [0.85, 1.00]
Perez-Risco 2018	7	0	1	0	0.88 [0.47, 1.00]	Not estimable
Peñata-Bedoya 2021	0	0	0	6	Not estimable	1.00 [0.54, 1.00]



Bone or joint aspirate, Xpert Ultra, composite reference standard

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Sun 2019	107	1	4	33	0.96 [0.91, 0.99]	0.97 [0.85, 1.00]



The summary sensitivity and specificity (95% CI) of LC-aNAATs were 96.6% (87.2 to 99.1) and 91.1% (80.8 to 96.2) from three Xpert Ultra studies (126 participants; Table 4; Figure 10).

Composite reference standard

In bone or joint aspirate, Xpert Ultra sensitivity and specificity against a composite reference standard were 96% (91 to 99) and 97% (85 to 100), (1 study; 145 participants; Figure 10). We did not perform meta-analysis due to lack of sufficient data.

H: LC-aNAAT testing in peritoneal fluid for peritoneal tuberculosis

Microbiological reference standard

Three studies evaluated Xpert Ultra in peritoneal fluid against culture as the reference standard. The sensitivity ranged from 33% to 67% and the specificity ranged from 94% to 100% (3 studies, 69 participants; Figure 11). The data were insufficient to do a meta-analysis.

Figure 11. Forest plot of Xpert Ultra sensitivity and specificity for peritoneal, pericardial and disseminated TB by reference standard. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive

Peritoneal fluid, Xpert Ultra, culture

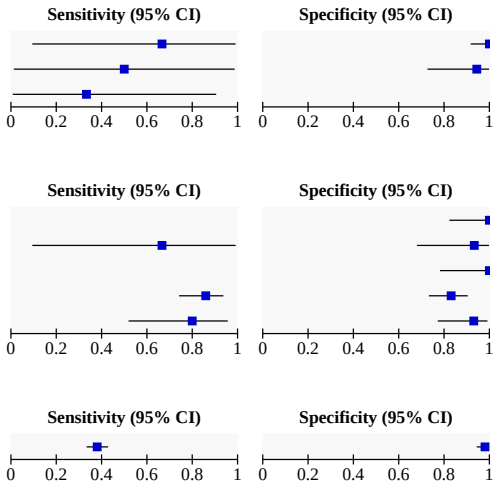
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Peñata-Bedoya 2021	2	0	1	43	0.67 [0.09, 0.99]	1.00 [0.92, 1.00]
Slail 2023	1	1	1	17	0.50 [0.01, 0.99]	0.94 [0.73, 1.00]
Perez-Risco 2018	1	0	2	0	0.33 [0.01, 0.91]	Not estimable

Pericardial fluid, Xpert Ultra, culture

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Mekkaoui 2021	0	0	0	19	Not estimable	1.00 [0.82, 1.00]
Slail 2023	2	1	1	14	0.67 [0.09, 0.99]	0.93 [0.68, 1.00]
Peñata-Bedoya 2021	0	0	0	15	Not estimable	1.00 [0.78, 1.00]
Minnies 2023	49	14	8	69	0.86 [0.74, 0.94]	0.83 [0.73, 0.90]
Alomatu 2023	12	2	3	27	0.80 [0.52, 0.96]	0.93 [0.77, 0.99]

Blood, Xpert Ultra, culture

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Boloko 2022	161	3	262	152	0.38 [0.33, 0.43]	0.98 [0.94, 1.00]



Composite reference standard

We did not identify any studies that evaluated Xpert Ultra in peritoneal fluid against a composite reference standard.

I: LC-aNAAT testing in pericardial fluid for pericardial tuberculosis

Microbiological reference standard

Five studies evaluated Xpert Ultra in pericardial fluid against culture as the reference standard. The sensitivity ranged from 67% to 86% and the specificity ranged from 83% to 100% (Table 4; Figure 11).

The summary sensitivity and specificity (95% CI) of LC-aNAATs were 84.0% (73.9 to 90.7) and 86.6% (79.5 to 91.5) from three Xpert Ultra studies (202 participants; Table 4; Figure 11).

Composite reference standard

We did not identify any studies that evaluated Xpert Ultra in pericardial fluid against a composite reference standard.

J: LC-aNAAT testing in blood for disseminated tuberculosis

Microbiological reference standard

One study evaluated Xpert Ultra in blood against culture as the reference standard. The sensitivity was 38% (33.0 to 43.0) and specificity was 98% (94.0 to 100) (1 study; 578 participants; Figure 11).

Composite reference standard

We did not identify any studies that evaluated Xpert Ultra in blood against a composite reference standard.

II. Detection of rifampicin resistance

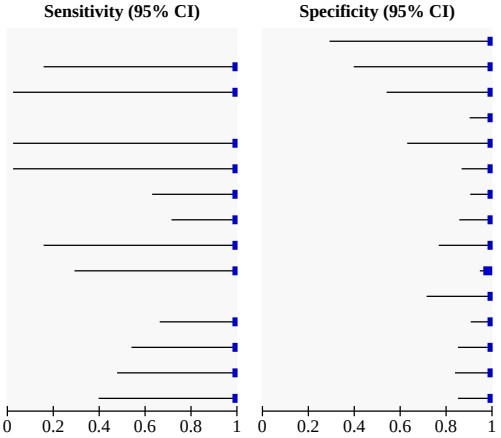
LC-aNAAT testing for rifampicin resistance

Xpert Ultra

Thirteen studies evaluated Xpert Ultra for the detection of rifampicin resistance. Xpert Ultra sensitivity estimates were 100%; specificity varied from 98% to 100% (Figure 12). One study reported zero participants with rifampicin resistance and thus sensitivity was not estimable (Chin 2019). LC-aNAAT summary sensitivity and specificity values (95% CI) were 100.0% (93.4 to 100.0) and 99.4% (92.1 to 100.0) (13 Xpert Ultra studies: 446 participants; Table 4; Summary of findings 4).

Figure 12. Forest plots of Xpert Ultra sensitivity and specificity for rifampicin resistance. The squares represent the sensitivity and specificity of one study, the black line its confidence interval. FN: false-negative; FP: false-positive; TN: true-negative; TP: true-positive

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Chin 2019	0	0	0	3	Not estimable	1.00 [0.29, 1.00]
Hoel 2020a	2	0	0	4	1.00 [0.16, 1.00]	1.00 [0.40, 1.00]
Huerga 2023	1	0	0	6	1.00 [0.03, 1.00]	1.00 [0.54, 1.00]
Mekkaoui 2021	0	0	0	36	Not estimable	1.00 [0.90, 1.00]
Minnies 2021	1	0	0	8	1.00 [0.03, 1.00]	1.00 [0.63, 1.00]
Peñata-Bedoya 2021	1	0	0	26	1.00 [0.03, 1.00]	1.00 [0.87, 1.00]
Sharma 2020	8	0	0	37	1.00 [0.63, 1.00]	1.00 [0.91, 1.00]
Sharma 2021	11	0	0	24	1.00 [0.72, 1.00]	1.00 [0.86, 1.00]
Sharma 2023	2	0	0	14	1.00 [0.16, 1.00]	1.00 [0.77, 1.00]
Slail 2023	3	3	0	160	1.00 [0.29, 1.00]	0.98 [0.95, 1.00]
Spener-Gomes 2021	0	0	0	11	Not estimable	1.00 [0.72, 1.00]
Sun 2019	9	0	0	38	1.00 [0.66, 1.00]	1.00 [0.91, 1.00]
Wang 2019	6	0	0	23	1.00 [0.54, 1.00]	1.00 [0.85, 1.00]
Wang 2020	5	0	0	21	1.00 [0.48, 1.00]	1.00 [0.84, 1.00]
Wu 2019	4	0	0	23	1.00 [0.40, 1.00]	1.00 [0.85, 1.00]



DISCUSSION

Summary of main results

This systematic review update summarizes the current literature and includes 37 unique studies on the accuracy of LC-aNAATs which includes Xpert Ultra and Truenat MTB plus for detection of extrapulmonary tuberculosis and rifampicin resistance. We identified 36 studies evaluating Xpert Ultra, and 3 studies evaluating Truenat MTB plus. We included two reference standards: a microbiological and a composite reference standard, and have stratified all analyses by type of reference standard. Major findings from our review include the following:

- LC-aNAAT sensitivity for extrapulmonary tuberculosis varied across different types of specimens (from 74.0% in pleural fluid to 96.6% in bone or joint fluid) against a microbiological reference standard (Table 4).
- The sensitivity for each form of extrapulmonary tuberculosis decreased when assessed against a microbiological reference standard, while specificity increased when assessed against a composite reference standard (Table 4).
- In cerebrospinal fluid, LC-aNAAT sensitivity and specificity were 88.2% (83.7 to 91.6) and 96.0% (86.8 to 98.9) against a microbiological reference standard (Summary of findings 1).
- In pleural fluid, LC-aNAAT sensitivity and specificity were 74.0% (60.8 to 83.9) and 88.1% (78.8 to 93.6) against a microbiological reference standard (Summary of findings 2).

- In lymph node aspirate, LC-aNAAT sensitivity and specificity were 71.3% (64.3 to 77.4) and 97.4% (82.2 to 99.7) against a composite reference standard (Summary of findings 3).
- For the detection of rifampicin resistance, Xpert Ultra sensitivity and specificity were 100.0% (93.4 to 100.0) and 99.4% (92.1 to 100.0) against culture-based drug susceptibility testing using solid or liquid media (Summary of findings 4).

It is important to note that these summary estimates for LC-aNAAT include studies for Xpert Ultra only. For Truenat MTB plus, the data were insufficient to be included in the class-level analyses (Table 2) and have been provided separately in Supplementary material 11.

LC-aNAAT testing in cerebrospinal fluid

See Summary of findings 1

Results of these studies indicate that, in theory, for a population of 1000 people where 100 have tuberculosis meningitis in culture, 124 would be LC-aNAAT-positive - of these, 36 (29%) would not have tuberculosis (false-positives); and 876 would be LC-aNAAT-negative - of these, 12 (1%) would have tuberculosis (false-negatives).

Rapid diagnosis of tuberculous meningitis is critical so that lifesaving treatment can be started promptly. Around 50% of those affected die or experience disabling consequences [132]. Xpert Ultra was designed to improve tuberculosis detection, in particular, in people with paucibacillary disease. The limit of detection for MTB is lower with Xpert Ultra (16 bacterial colony-forming units (CFU) per mL) than with Xpert MTB/RIF (131 CFU per mL) [24]. In subgroup analyses, we found slightly higher Xpert Ultra accuracy in studies that concentrated the cerebrospinal fluid (CSF): pooled sensitivity of 92.8% in concentrated specimens versus 81.0% in unconcentrated specimens, and pooled specificity of 93.6% in concentrated specimens versus 85.8% in unconcentrated specimens. The Tuberculous Meningitis International Research Consortium has recommended increasing the volume of CSF collected for diagnosis followed by centrifugation as a way of improving Xpert MTB/RIF (now superseded by Xpert Ultra) sensitivity [133]; however, we did not have sufficient data to investigate CSF collection volume.

LC-aNAAT testing in pleural fluid

See Summary of findings 2

Results of these studies indicate that, in theory, for a population of 1000 people where 100 have pleural tuberculosis on culture, 181 would be LC-aNAAT-positive - of these, 107 (59%) would not have tuberculosis (false-positives), and 819 would be LC-aNAAT-negative - of these, 26 (3%) would have tuberculosis (false-negatives).

LC-aNAAT pooled sensitivity in pleural fluid was lower than that of CSF. One reason for the lower sensitivity of LC-aNAAT in pleural fluid could be the paucibacillary nature of pleural tuberculosis. Other possible reasons are contamination of blood or the presence of certain polymerase chain reaction (PCR) inhibitors in the pleural fluid [134, 135]. However, Theron and colleagues found that extrapulmonary specimens showed less evidence of PCR inhibition than pulmonary specimens, with bacterial load being more important for a positive Xpert MTB/RIF result. Given that false-negative results were common (low sensitivity), a negative LC-aNAAT result may not be relied on to exclude tuberculosis.

LC-aNAAT testing in lymph node aspirates

See Summary of findings 3

Results of these studies indicate that, in theory, for a population of 1000 people where 100 have lymph node tuberculosis verified by a composite reference standard, 94 would be LC-aNAAT positive - of these, 23 (24%) would not have tuberculosis (false-positives), and 906 would be LC-aNAAT-negative - of these, 29 (3%) would have tuberculosis (false-negatives).

Regarding Xpert testing for lymph node aspirates, it is important to point out that although tissue biopsy provides material for histological examination which may be of substantial diagnostic value, a fluid specimen may be collected more easily. In addition, fine-needle aspiration of lymph nodes is well-suited for use in resource-limited settings because the procedure is simple, easy to learn, minimally invasive, and inexpensive [136]. Thus, clinicians may want to consider fine-needle aspiration of lymph nodes before surgical biopsy.

We considered several reasons why the specificity of LC-aNAATs in lymph node aspirates against culture would be lower than in other extrapulmonary specimens. Although not always reported, studies may have included participants receiving tuberculosis treatment. We considered the type of culture used in the included studies because liquid culture is more sensitive than solid culture [43]. Although most studies used liquid culture or a combination of solid and liquid culture, culture results may also be negative owing to inefficient specimen collection or errors in sampling, differing bacterial loads, and contamination [136]. Negative culture results in lymph node tuberculosis have previously been reported [45].

Another reason for negative culture results is that there may have been a decrease in live tuberculosis bacteria during processing with N-acetyl-L-cysteine-sodium hydroxide, which is routinely used to homogenize, decontaminate, and liquefy non-sterile specimens, such as sputum, for mycobacterial culture [43]. Harsh decontamination practices have been noted to contribute to false-negative culture results, especially in paucibacillary specimens [137]. Standards specify that "specimens collected from normally sterile sites may be placed directly into the culture medium" [43]. CSF, pleural fluid, and lymph node aspirates are usually considered to be sterile specimens. It is our understanding that some laboratories do decontaminate sterile site specimens as a precaution against non-sterile collection procedures. We did not have sufficient data to further investigate laboratory practices.

In summary, several factors probably contributed to low LC-aNAAT specificity against culture in lymph node aspirates. The 'true' specificity of LC-aNAATs in lymph node aspirates is likely to be higher for the aforementioned reasons. The index test specificity was higher against a composite reference standard, similar to that found in CSF, pleural fluid, and other specimens (Table 4).

LC-aNAAT testing for rifampicin resistance

See Summary of findings 4

Results of these studies indicate that, in theory, for a population of 1000 people where 100 have rifampicin resistance, 105 would be LC-aNAAT-positive (resistant) - of these, five (5%) would not have rifampicin resistance (false-positives); and 895 would be LC-aNAAT-

negative (susceptible) - of these, none (0%) would have rifampicin resistance.

For the detection of rifampicin resistance in extrapulmonary specimens, we found the sensitivity and specificity of LC-aNAATs (100%) to be comparable to estimates in pulmonary specimens [64]. These findings suggest that the use of Xpert Ultra in extrapulmonary specimens could assist in rapid diagnosis of rifampicin-resistant tuberculosis and early initiation of treatment for multidrug-resistant tuberculosis (MDR-TB).

Notably, concerns have been raised about rapid drug susceptibility testing (DST) methods; in particular, the automated mycobacteria growth indicator tube (MGIT) 960 for tuberculosis drug resistance using the recommended critical concentrations [138].

People-important outcomes, such as mortality, are especially relevant to patients, decision-makers, and the wider tuberculosis community. While performing this systematic review, we did not identify direct evidence of studies linking true-positives, false-positives, true-negatives, and false-negatives to people-important outcomes when either Xpert Ultra or Truenat MTB plus was used to diagnose extrapulmonary tuberculosis.

This review represents the most comprehensive review of the diagnostic accuracy of LC-aNAATs for extrapulmonary tuberculosis in adults. These reviews provide evidence that may help countries to make decisions about scaling up the tests for programmatic management of tuberculosis and drug-resistant tuberculosis. Although the information in this review will help to inform such decisions, other factors such as resource requirements and feasibility (including stable electrical power supply, temperature control, and maintenance of the cartridge modules) will also be important considerations.

Strengths and weaknesses of the review

Completeness of evidence

This is a reasonably complete data set. We included any non-English studies that we found from which we could obtain accuracy data. However, we acknowledge that we may have missed some studies despite the comprehensive search and our outreach to investigators. We included eight common forms of extrapulmonary tuberculosis in the review. However, for some of these forms, such as disseminated tuberculosis, data were insufficient to allow us to determine summary accuracy estimates. We did not include less common forms, such as cutaneous tuberculosis, ocular tuberculosis, female genital tuberculosis, and tuberculosis of the breast. Our inclusion criteria, limiting eligibility to adults and adolescents, meant that some of the studies included in our previous reviews were excluded from the current update. We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy (PRISMA-DTA) [74]. To keep this updated review aligned with the WHO policy guidance, we did not re-run another literature search for these technologies, which is a limitation of this updated review. However, we also had a WHO public call for data to include unpublished studies and, therefore, we are confident about the comprehensiveness of this review.

Accuracy of the reference standards used

In a systematic review of diagnostic test accuracy studies, the reference standard is the best available test to determine the presence or absence of the target condition. For the detection of tuberculosis, we used two reference standards: culture and a composite reference standard, both of which are known to be imperfect. While the composite reference standard is designed to have improved accuracy compared to culture alone, it may still lead to biased accuracy estimates of the index test, depending on various factors such as the accuracy of the different components; decision rules for combining them; prevalence of the target condition; and conditional dependence between the components and the index test [139]. Conditional dependence between two imperfect tests arises when both tests make the same false-positive or false-negative errors more often than expected by chance [70]. Hence, conditional dependence may arise between the index test and both reference standards we have used, as they are imperfect. As a consequence, we may over- or underestimate the diagnostic accuracy of the index tests. An additional challenge with including a composite reference standard is that the definition of the composite reference standard may vary across studies, making it difficult to interpret the accuracy estimates.

Several factors may have contributed to false-negative culture results for the accuracy of the reference standard for lymph node aspirate in particular, including inefficient specimen collection and overly harsh decontamination.

Establishing a diagnosis of extrapulmonary tuberculosis would ideally include pursuing the diagnosis of pulmonary tuberculosis as well, because participants with tuberculosis may have both pulmonary and extrapulmonary tuberculosis and the lung may be the only site where the presence of tuberculosis can be established. Because of the difficulties involved in diagnosing HIV-associated tuberculosis, it is recommended that multiple cultures from sputum and other types of specimens be evaluated in people with HIV [53, 140]. Given the limitations in the reference standard, we recommend that future studies consider using liquid culture because this is more sensitive than solid culture, and that researchers obtain multiple specimens for culture to confirm the diagnosis of extrapulmonary tuberculosis [141].

Most studies included in this review used culture-based DST (either Löwenstein-Jensen (LJ) or mycobacteria growth indicator tube (MGIT) 960) as the reference standard for the detection of rifampicin resistance. Concerns have been raised about rapid DST methods, in particular, automated MGIT 960, for tuberculosis drug resistance using the recommended critical concentrations [138].

Quality and quality of reporting of the included studies

The risk of bias was low for the participant selection, index test, and flow and timing domains and was high or unclear for the reference standard domain (most of these studies performed specimen decontamination before culture inoculation). For the applicability domain, most studies had an unclear concern as the testing was done either in reference laboratories or the setting was not clear. Because of this, it was difficult to tell if a given reference laboratory provided services mainly to very sick patients (inpatients in tertiary care) or to all patients, including very sick patients and those with less severe disease (primary, secondary, and tertiary care). A limitation was that several studies included more than one specimen per participant, which artificially inflated

the sample size of the study and may have led to overestimation or underestimation of the accuracy estimates. Additionally, it is important to note that clinical practice may differ in studies where participant inclusion could be based on symptom screening, chest X-ray abnormality, C-reactive protein (CRP) levels or a combination of multiple factors. In general, the studies were fairly well-reported, although we corresponded with the primary study authors to ask for additional data and missing information, where applicable. In several studies, accuracy data by site of extrapulmonary disease were not reported, and, in a minority of studies, blinding was not reported. We strongly encourage the authors of future studies to follow the recommendations provided in the updated Standards for Reporting Diagnostic Accuracy (STARD) statement to improve the quality of reporting [142].

Interpretability of subgroup analyses

We investigated potential sources of heterogeneity in the different extrapulmonary specimens. Importantly, we found slightly higher Xpert Ultra accuracy in studies with concentrated cerebrospinal fluid (CSF) in comparison to unconcentrated specimens. We note that subgroup findings should be interpreted with caution, as there were only three studies and few participants with tuberculous meningitis were included in these analyses.

Comparison with other systematic reviews

We identified one systematic review that estimated the summary accuracy of Xpert Ultra that found, for all forms of extrapulmonary tuberculosis combined, pooled sensitivity and specificity of 85.1% (95% CI 76.7 to 90.8) and 95.7% (95% CI 87.9 to 98.6) (7 studies; 1500 specimens) [143].

Applicability of findings to the review question

For the patient selection domain, most studies (other than the ones evaluating TB meningitis) had high or unclear concern for applicability because either participants were evaluated exclusively as inpatients in tertiary care or we were not sure about the clinical settings. We therefore cannot be sure about the applicability of our findings to primary care. Studies that take place in referral settings may include participants whose conditions are more difficult to diagnose than are seen at lower levels of the health system. However, we recognize that classifying studies as primary, secondary, or tertiary care may not adequately account for differences in the disease spectrum [144]. For the index and reference test domains, most studies had low concern for applicability.

AUTHORS' CONCLUSIONS

Implications for practice

In people presumed to have extrapulmonary tuberculosis, LC-aNAAT may be helpful in confirming the diagnosis. Sensitivity varies across different extrapulmonary specimens; however, for most specimens specificity is high, and the test rarely yields a positive result for people without tuberculosis. For tuberculous meningitis, LC-aNAAT had a high summary sensitivity and specificity against culture. LC-aNAATs had high accuracy for rifampicin resistance.

Implications for research

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

Rapid point-of-care diagnostic tests for extrapulmonary tuberculosis are critically needed. Research groups should focus on developing diagnostic tests and strategies that use readily-available clinical specimens, such as urine, rather than specimens that require invasive procedures for collection. As this is a class-level analysis, it is important that other tests falling into this class by definition should generate robust data to be evaluated and pooled for future reviews.

SUPPLEMENTARY MATERIALS

Supplementary materials are available with the online version of this article: 10.1002/14651858.CD012768.pub3.

Supplementary material 1 Search strategies

Supplementary material 2 Characteristics of included studies

Supplementary material 3 Characteristics of excluded studies

Supplementary material 4 Analyses

Supplementary material 5 Data package

Supplementary material 6 Data extraction form

Supplementary material 7 Rules for QUADAS-2

Supplementary material 8 Risk of bias and applicability concerns table for tuberculous meningitis

Supplementary material 9 Risk of bias and applicability concerns table for pleural tuberculosis

Supplementary material 10 Risk of bias and applicability concerns table for lymph node tuberculosis

Supplementary material 11 Truenat MTB Plus accuracy data

Supplementary material 12 Differences between previous versions and current review

ADDITIONAL INFORMATION

Acknowledgements

The Cochrane Infectious Diseases Group (CIDG) supported the authors in the development of this diagnostic test accuracy (DTA) review update.

We thank Vittoria Lutje, CIDG Information Specialist, for her help designing the literature searches and performing searches. The CIDG editorial base is funded by UK aid from the UK government for the benefit of low- and middle-income countries (project number 300342-104). The views expressed do not necessarily reflect the UK government's official policies.

This review was commissioned by the WHO Global Tuberculosis Programme to inform the guideline development group (GDG) meeting in May 2024, and in part made possible with financial support from the WHO Global Tuberculosis Programme.

We acknowledge Eleanor Ochodo, Centre for Evidence-based Health Care Department of Global Health, Stellenbosch, and Fredrick Haraka, Ifakara Health Institute, Bagamoyo, Tanzania, for the section on people-important outcomes. We thank Hannah Ryan, from the Tropical and Infectious Diseases Unit, Royal Liverpool University Hospital, for her help with the protocol. We thank Chunli Lu, from the Centre for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing, and Marcela Perlwitz, Ivy Tech Community College of Indiana, West Lafayette, for help with translation. We thank all authors of the included studies for providing answers to our questions along with additional data. We thank the referees for their helpful comments.

Editorial and peer-reviewer contributions

The Cochrane Infectious Diseases group supported the authors in the development of this diagnostic test accuracy review update.

The following people conducted the editorial process for this article:

• Sign-off Editors (final editorial decision): Mariska M. G. Leeflang, Amsterdam UMC, University of Amsterdam (DTA); Dr. Michael Eisenhut, Consultant Paediatrician, Luton & Dunstable University Hospital (clinical)

• Managing Editor (provided editorial guidance to authors, edited the article): Ben Ridley, Cochrane Central Editorial Service

• Editorial Assistant (selected peer-reviewers, conducted editorial policy checks, collated peer-reviewer comments and supported editorial team): Leticia Rodrigues, Cochrane Central Editorial Service

• Copy Editor (copy editing and production): Anne Lethaby, Cochrane Central Production Service

• Peer-reviewers (provided comments and recommended an editorial decision): Mariska M. G. Leeflang, Amsterdam UMC, University of Amsterdam (methods); Sofia Tsokani, Methods Support Unit, Cochrane CET (statistical reviewer); Suzanne Avis, Cardiovascular Discovery Group University of Sydney, Australia (search); Heidi Albert, FIND South Africa (clinical); and Brian Duncan (consumer).

Contributions of authors

MK and KRS wrote early drafts of the protocol. MK, KRS, and YT wrote the latest protocol.
MK, AS, LRI, JD, VAS reviewed the studies, extracted accuracy data and assessed methodological quality of the included studies.
MK and KRS tailored QUADAS-2 to the review.
KS and YT performed the statistical analyses.
All review authors interpreted the findings.
All review authors contributed to the final manuscript.

In the previous versions, authors like Claudia M Denkinger (CMD), Keertan Dhedha (KD), and Samuel Schumacher (SGS) contributed methodological advice. CMD and SGS tailored QUADAS-2 to the review. Ian Schiller (IS), Nandini Dendukuri (ND), and Mandy Yao (MY) performed the statistical analyses in the review. CMD, KD, SGS, IS, ND, MY were involved in previous published versions of this review in 2018 and 2021 and are no longer included in the author byline. Some of the content retained in this review reflects their contributions.

Declarations of interest

MK has no conflict of interest.

LRI has no known conflicts of interest. He was one of the authors of an included study (Anie 2024). However, he did not make study eligibility decisions, extract the data, carry out the risk of bias assessment, or perform GRADE assessments for that study.

AS has no conflict of interest.

KS has no conflict of interest.

YT is a Cochrane Editorial Board Member and a Cochrane Infectious Diseases and Diagnostic Test Accuracy Editor. They were not involved with the editorial process of this review.

AK is a WHO staff member in the Global Tuberculosis Programme, which commissioned the 2024 update for tuberculosis molecular diagnostics.

NI is a WHO staff member in the Global Tuberculosis Programme, which commissioned the 2024 update for tuberculosis molecular diagnostics.

VAS has no conflict of interest.

JD has no conflict of interest.

KRS was previously a Cochrane Infectious Diseases and Diagnostic Test Accuracy Editor. They were not involved with the editorial process of this review.

We have no financial involvement with any organization or entity that has a financial interest in, or financial conflict with, the subject matter or materials discussed in the review apart from those disclosed.

All authors received funding from the World Health Organization Global Tuberculosis Programme, Switzerland.

Sources of support

Internal sources

- Liverpool School of Tropical Medicine (LSTM), UK

  LSTM hosts the Cochrane Infectious Diseases Group (CIDG) editorial base, which supported the authors in the development of this review update.

External sources

- World Health Organization (WHO), Switzerland

  This review was commissioned by the WHO Global Tuberculosis Programme to inform guideline development group (GDG) meeting in May 2024, and in part made possible with financial support from the WHO Global Tuberculosis Programme.
- Foreign, Commonwealth, and Development Office (FCDO), Other

Project number 300342-104

Registration and protocol

This review was part of the larger WHO policy recommendations and this is an update to previous reviews. These versions are stated below:

Protocol: Xpert® MTB/RIF assay for extrapulmonary tuberculosis and rifampicin resistance (protocol) (2017); DOI: 10.1002/14651858.CD012768.

Original review: Xpert® MTB/RIF assay for extrapulmonary tuberculosis and rifampicin resistance (2018); https://doi.org/10.1002/14651858.CD012768.pub2

Updated review: Xpert MTB/RIF Ultra and Xpert MTB/RIF assays for extrapulmonary tuberculosis and rifampicin resistance in adults (2021); https://doi.org/10.1002/14651858.CD012768.pub3

We have also described differences in the previous versions of this review and current review in Supplementary material 12.

Data, code and other materials

As part of the published Cochrane Review, the following is made available for download for users of the Cochrane Library:

- *Full search strategies for each database;*
- *Full citations of each unique report for all studies included, ongoing or waiting classification, or excluded at the full text screen, in the final review;*
- *Study data, including study information, study arms, and study results or test data;*
- *Consensus risk of bias assessments and analysis data, including overall estimates and settings, subgroup estimates, and individual data rows.*

Appropriate permissions have been obtained for such use. Analyses and data management were conducted using Cochrane's authoring tool, RevMan, using the inbuilt computation methods. All analyses were conducted using Stata 18.0 and meta-analysis models were fitted using the 'metandi' user-written function or the 'meqrlogit' command. We provide all analyses in Supplementary material 4 and Supplementary material 5. *Template data extraction forms from MS Word documents are available in* Supplementary material 6.

What's new

Date	Event	Description
4 August 2025	New citation required but conclusions have not changed	This is an update to the previous review which will be a new citation.
4 August 2025	New search has been performed	We updated the literature search for Xpert Ultra and included other technologies in this class of LC-aNAATs. We included new studies based on this updated search. The authors have also changed in this review.

History

Protocol first published: Issue 8, 2017
Review first published: Issue 8, 2018

Date	Event	Description
11 January 2021	New search has been performed	We have updated the review with more information. There are no major changes to the conclusions.
11 January 2021	New citation required but conclusions have not changed	We updated the literature search and included 22 new studies.

REFERENCES

1. World Health Organization. Global Tuberculosis Report 2024 (Licence: CC BY-NC-SA 3.0 IGO). Geneva: World Health Organization, 2024.

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

3. Pang Y, An J, Shu W, Huo F, Chu N, Gao M, et al. Epidemiology of extrapulmonary tuberculosis among inpatients, China, 2008-2017. *Emerging Infectious Diseases* 2019;**25**:457-64.

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

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

6. 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.

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

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

9. World Health Organization. The END TB strategy. www.who.int/tb/strategy/end-tb/en/ 2014 (accessed 1 April 2020).

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

11. World Health Organization. Module 4: treatment - drug-susceptible tuberculosis treatment. In: WHO Consolidated Guidelines on Tuberculosis. https://www.who.int/publications/i/item/9789240048126 2022 (accessed 30 Sep 2024).

12. World Health Organization. Module 4: treatment - drug-resistant tuberculosis treatment (2022 update). In: WHO Consolidated Guidelines on Tuberculosis. www.who.int/publications/i/item/9789240063129 2022 (accessed 30 September 2024).

13. World Health Organization. Compendium of WHO guidelines and associated standards: ensuring optimum delivery of the cascade of care for patients with tuberculosis (2nd edition). www.who.int/tb/publications/Compendium_WHO_guidelines_TB_2017/en/ (accessed 2 July 2020).

14. Sharma SK, Ryan H, Khaparde S, Sachdeva KS, Singh AD, Mohan A, et al. Index-TB guidelines: guidelines on extrapulmonary tuberculosis for India. Indian Journal of Medical Research 2017;**145**(4):448-63.

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

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

17. 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.

18. World Health Organization. Rapid implementation of the Xpert MTB/RIF diagnostic test. Technical and operational 'How-to'. Practical considerations. who.int/tb/publications/tb-amplificationtechnology-implementation/en/ 2011 (accessed 2 July 2020).

19. World Health Organization. Module 3: diagnosis – rapid diagnostics for tuberculosis detection. In: WHO Consolidated Guidelines on Tuberculosis. https://www.who.int/publications/i/item/9789240089488 2024 (accessed 2 Oct 2024).

20. Cepheid. Brochure: Xpert® MTB/RIF Ultra. www.cepheid.com/en/tests/Critical-Infectious-Diseases/Xpert-MTB-RIF-Ultra (accessed 26 March 2020).

21. Cepheid. Xpert® MTB/RIF. Two-hour detection of MTB and rifampin resistance mutations. www.cepheid.com/Package%20Insert%20Files/Xpert-MTB-RIF-ENGLISH-Package-Insert-301-1404-Rev-F.pdf (accessed 29 March 2020).

22. 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.

23. World Health Organization. WHO meeting report of a technical expert consultation: non-inferiority analysis of Xpert MTB/RIF Ultra compared to Xpert MTB/RIF. WHO/HTM/TB/2017.04. www.who.int/tb/publications/2017/XpertUltra/en/ 2017 (accessed 2 July 2020).

24. Chakravorty S, Simmons AM, Rowneki M, Parmar H, Cao Y, Ryan J, et al. The new Xpert MTB/RIF Ultra: improving detection of Mycobacterium tuberculosis and resistance to rifampin in an assay suitable for point-of-care testing. *MBio* 2017;**8**(4):e00812-17. [DOI: 10.1128/mBio.00812-17]

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

26. Global Laboratory Initiative. Planning for country transition to Xpert® MTB/RIF Ultra cartridges. www.stoptb.org/wg/gli/assets/documents/gli_ultra.pdf (accessed 2 July 2020).

27. 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.

28. 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.

29. Global Laboratory Initiative. Practical guide to implementing a quality assurance system for Xpert MTB/RIF testing. www.stoptb.org/wg/gli/assets/documents/Xpert-QA-guide-2019.pdf (accessed 7 January 2020).

30. Theron G, Zijenah L, Chanda D, Clowes P, Rachow A, Lesosky M, et al. Feasibility, accuracy, and clinical effect of point-of-care Xpert MTB/RIF testing for tuberculosis in primary-care settings in Africa: a multicentre, randomised, controlled trial. *Lancet* 2014;**383**(9915):424-35.

31. World Health Organization. Xpert MTB/RIF implementation manual. Technical and operational 'how-to' practical considerations. apps.who.int/iris/bitstream/10665/112469/1/9789241506700_eng.pdf 2014 (accessed 2 July 2020).

32. Penn-Nicholson A, Gomathi SN, Ugarte-Gil C, Meaza A, Lavu E, Patel P, et al. A prospective multicentre diagnostic accuracy study for the Truenat tuberculosis assay. *European Respiratory Journal* 2021;**58**(5):2100526.

33. Molbio Diagnostics Private Limited. Truenat® MTB: Chip-based Real Time PCR Test for Mycobacterium tuberculosis. https://www.molbiodiagnostics.com/product_details.php?id=1 (accessed 15 August 2024).

34. Molbio Diagnostics Private Limited. Truenat® MTB Plus: Chip-based Real Time PCR Test for Mycobacterium tuberculosis. https://www.molbiodiagnostics.com/product_details.php?id=49 (accessed 05 August 2024).

35. SD Biosensor. www.sdbiosensor.com/product/product_view?product_no=125 (accessed 05 August 2024).

36. Bioneer South Korea. IRON-qPCR™. https://eng.bioneer.com/20-a-2600.html#producttechnical (accessed 05 August 2024).

37. World Health Organization. Module 3: diagnosis - rapid diagnostics for tuberculosis detection. In: WHO Operational Handbook on Tuberculosis (Licence: CC BY-NC-SA 3.0 IGO). who.int/publications/i/item/who-operational-handbook-on-tuberculosis-module-3-diagnosis---rapid-diagnostics-for-tuberculosis-detection (accessed 6 August 2020).

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

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

40. Branigan D. Pipeline report 2023 tuberculosis diagnostics. https://www.treatmentactiongroup.org/wp-content/uploads/2023/11/2023_pipeline_TB_diagnostics_final.pdf (accessed November 2023).

41. Lewinsohn DM, Leonard MK, LoBue PA, Cohn DL, Daley CL, Desmond E, et al. Official American Thoracic Society/Infectious Diseases Society of America/Centers for Disease Control and Prevention clinical practice guidelines: diagnosis of tuberculosis in adults and children. *Clinical Infectious Diseases* 2017;**64**(2):e1–33. [DOI: 10.1093/cid/ciw694]

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

43. American Thoracic Society, Centers for Disease Control Prevention, Infectious Disease Society of America. Diagnostic standards and classification of tuberculosis in adults and children. *American Journal Respiratory and Critical Care Medicine* 2000;**161**(4 Pt 1):1376-95.

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

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

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

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

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

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

50. Yu G, Shen Y, Zhong F, Ye B, Yang J, Chen G. Diagnostic accuracy of the loop-mediated isothermal amplification assay for extrapulmonary tuberculosis: a meta-analysis. *PLoS One* 2018;**13**(6):e0199290.

51. Nathavitharana RR, Cudahy PG, Schumacher SG, Steingart KR, Pai M, Denkinger CM. Accuracy of line probe assays for the diagnosis of pulmonary and multidrug-resistant tuberculosis: a systematic review and meta-analysis. *European Respiratory Journal* 2017;**49**(1):pii: 1601075.

52. Walzl G, McNerney R, Du Plessis N, Bates M, McHugh TD, Chegou NN, et al. Tuberculosis: advances and challenges in development of new diagnostics and biomarkers. *Lancet Infectious Diseases* 2018;**18**(7):e199-e210.

53. Bjerrum S, Schiller I, Dendukuri N, Kohli M, Nathavitharana RR, Zwerling AA, et al. Lateral flow urine lipoarabinomannan assay for detecting active tuberculosis in people living with HIV. *Cochrane Database of Systematic Reviews* 2019, Issue 10. Art. No: CD011420. [DOI: 10.1002/14651858.CD011420.pub3]

54. Gupta-Wright A, Corbett EL, Van Oosterhout JJ, Wilson D, Grint D, Alufandika-Moyo M, et al. Rapid urine-based screening for tuberculosis in HIV-positive patients admitted to hospital in Africa (STAMP): a pragmatic, multicentre, parallel-group, double-blind, randomised controlled trial. *Lancet* 2018;**392**(10144):292-301. [PMID: 30032978]

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

56. Broger T, Nicol MP, Székely R, Bjerrum S, Sossen B, Schutz C, et al. Diagnostic accuracy of a novel tuberculosis point-of-care urine lipoarabinomannan assay for people living with HIV: a meta-analysis of individual in- and outpatient data. *PLOS Medicine* 2020;**17**(5):e1003113.

57. Huerga H, Bastard M, Lubega AV, Akinyi M, Antabak NT, Ohler L, et al. Novel FujiLAM assay to detect tuberculosis in HIV-positive ambulatory patients in four African countries: a diagnostic accuracy study. *Lancet Global Health* 2023;**11**(1):e126-35. [DOI: 10.1016/S2214-109X(22)00463-6]

58. Székely R, Sossen B, Mukoka M, Muyoyeta M, Nakabugo E, Hella J, et al; FujiLAM Study Consortium. Prospective multicentre accuracy evaluation of the FUJIFILM SILVAMP TB LAM test for the diagnosis of tuberculosis in people living with HIV demonstrates lot-to-lot variability. *PloS One* 2024;**19**(5):e0303846. [DOI: 10.1371/journal.pone.0303846]

59. 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. www.who.int/tb/publications/tb-amplificationtechnology-statement/en/ 2011 (accessed 2 July 2020).

60. 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: policy update. WHO/HTM/TB/2013.14. apps.who.int/iris/handle/10665/112472 2013 (accessed 2 July 2020).

61. Dorman SE, Schumacher SG, Alland D, Nabeta P, Armstrong DT, King B, et al. Xpert MTB/RIF Ultra for detection of Mycobacterium tuberculosis and rifampicin resistance: a prospective multicentre diagnostic accuracy study. *Lancet Infectious Diseases* 2018;**18**(1):76-84.

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

63. Kohli M, Schiller I, Dendukuri N, Yao M, Dheda K, Denkinger CM, et al. Xpert MTB/RIF Ultra and Xpert MTB/RIF assays for extrapulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database of Systematic Reviews* 2021, Issue 1. Art. No: CD012768. [DOI: 10.1002/14651858.CD012768.pub3]

64. Horne DJ, Kohli M, Zifodya JS, Schiller I, Dendukuri N, Tollefson D, et al. Xpert MTB/RIF and Xpert MTB/RIF Ultra for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database of Systematic Reviews* 2019, Issue 6. Art. No: CD009593. [DOI: 10.1002/14651858.CD009593]

65. Kay AW, González Fernández L, Takwoingi Y, Eisenhut M, Vu RD, Steingart KR, et al. Xpert MTB/RIF and Xpert MTB/RIF Ultra assays for active tuberculosis and rifampicin resistance in children. *Cochrane Database of Systematic Reviews* 2020, Issue 8. Art. No: CD013359. [DOI: 10.1002/14651858.CD013359] [PMID: 32853411]

66. Shapiro AE, Ross JM, Schiller I, Kohli M, Dendukuri N, Steingart KR, et al. Xpert MTB/RIF and Xpert Ultra assays for pulmonary tuberculosis and rifampicin resistance in adults irrespective of signs or symptoms of pulmonary tuberculosis. *Cochrane Database of Systematic Reviews* 2020, Issue 7. Art. No: CD013694. [DOI: 10.1002/14651858.CD013694]

67. Vonasek B, Ness T, Takwoingi Y, Kay AW, Wyk SS, Ouellette L, et al. Screening tests for active pulmonary tuberculosis in children. *Cochrane Database of Systematic Reviews* 2020, Issue 7. Art. No: CD013693. [DOI: 10.1002/14651858.CD013693]

68. Centers for Disease Control and Prevention. Reported tuberculosis in the United States, 2018. www.cdc.gov/tb/statistics/reports/2018/table15.htm (accessed 29 June 2020).

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

70. Naaktgeboren CA, Bertens LC, Van Smeden M, De Groot JA, Moons KG, Reitsma JB. Value of composite reference standards in diagnostic research. *BMJ (Clinical Research Ed.)* 2013;**347**:f5605.

71. World Health Organization. Updated interim critical concentrations for first-line and second-line DST (as of May 2012). www.stoptb.org/wg/gli/assets/documents/Updated%20critical%20concentration%20table_1st%20and%202nd%20line%20drugs.pdf 2012 (accessed 2 July 2020).

72. Covidence. Version accessed prior to 16 July 2025. Melbourne: Veritas Health Innovation, 2025. Available at www.covidence.org.

73. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;**372**:PMC8005924. [DOI: 10.1371/journal.pmed1000097]

74. McInnes MD, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T et al; PRISMA-DTA Group. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement [published correction appears in JAMA. 2019 Nov 26;322(20):2026]. *JAMA* 2018;**319**(4):388-96. [DOI: 10.1001/jama.2017.19163]

75. World Bank. World bank list of economies. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups 2023 (accessed July 2024).

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

77. Microsoft Excel. Microsoft Corporation, 2019.

78. 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.

79. Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database of Systematic Reviews* 2014, Issue 1. Art. No: CD009593. [DOI: 10.1002/14651858.CD009593.pub3]

80. Stata. StataCorp, Version Release 18. College Station, TX: StataCorp LLC, 2023.

81. Review Manager (RevMan). Version 8.7.0. The Cochrane Collaboration, 2024. Available at https://revman.cochrane.org.

82. 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.

83. Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *Journal of Clinical Epidemiology* 2006;**59**(12):1331–2.

84. Takwoingi Y, Guo B, Riley RD, Deeks JJ. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. *Statistical Methods in Medical Research* 2017;**26**:1896-911.

85. Takwoingi Y, Dendukuri N, Schiller I, Rücker G, Jones HE, Partlett C, et al. Code for undertaking meta-analysis. In: Deeks JJ, Bossuyt PM, Leeflang MM, Takwoingi Y, editor(s). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. Version 2.0 (updated July 2023), Cochrane 2023. Available from https://www.cochrane.org/authors/handbooks-and-manuals/handbook-systematic-reviews-diagnostic-test-accuracy.

86. Clopper, CJ, & Pearson, E S. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* 1935;**26**:404–413. [DOI: https://doi.org/10.1093/biomet/26.4.404]

87. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and presenting results. In: Deeks JJ, Bossuyt PM, C Gatsonis (editors). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 2010. Available at: methods.cochrane.org/sdt/handbook-dta-reviews.

88. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. *Journal of Clinical Epidemiology* 2011;**64**(4):401–6.

89. GRADEpro GDT. Version accessed 2 July 2020. Hamilton (ON): McMaster University (developed by Evidence Prime), 2020. Available at gradepro.org.

90. Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ (Clinical Research Ed.)* 2008;**336**(7653):1106-10.

91. Schünemann HJ, Mustafa R, Brozek J, Santesso N, Alonso-Coello P, Guyatt G, et al. GRADE Working Group. GRADE Guidelines: 16. GRADE evidence to decision frameworks for tests in clinical practice and public health. *Journal of Clinical Epidemiology* 2016;**76**:89-98. [DOI: 10.1016/j.jclinepi.2016.01.032]

92. Schünemann HJ, Mustafa R, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 1. Study design, risk of bias and indirectness in rating the certainty across a body of evidence for test accuracy. *Journal of Clinical Epidemiology* 2020;**122**:129-41. [DOI: 10.1016/j.jclinepi.2019.12.020]

93. Schünemann HJ, Mustafa R, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 2. Inconsistency, imprecision, publication bias and other domains for rating the certainty of evidence for test accuracy and presenting it in evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology* 2020;**122**:142-52. [DOI: 10.1016/j.jclinepi.2019.12.021]

94. Mustafa RA, El Mikati IK, Murad MH, Hultcrantz M, Steingart KR, Yang B, et al. GRADE guidance 37: rating imprecision in a body of evidence on test accuracy. *Journal of Clinical Epidemiology* 2024;**165**:111189. [DOI: 10.1016/j.jclinepi.2023.10.005]

95. Sharma K, Sharma M, Gupta N, Modi T, Joshi H, Shree R, et al. Determining the diagnostic potential of Truenat MTB Plus for tubercular lymphadenitis and detection of drug resistance and a comparison with GeneXpert Ultra. *Tuberculosis (Edinburgh, Scotland)* 2023;**142**:102379. [DOI: 10.1016/j.tube.2023.102379]

96. Sharma K, Sharma M, Modi M, Singla N, Sharma A, Sharma A, et al. Comparative analysis of Truenat™ MTB Plus and Xpert Ultra in diagnosing tuberculous meningitis. *International Journal of Tuberculosis and Lung Disease* 2021;**25**(8):626-31. [DOI: 10.5588/ijtld.21.0156]

97. Boloko L, Schutz C, Sibiya N, Balfour A, Ward A, Shey M, et al. Xpert Ultra testing of blood in severe HIV-associated tuberculosis to detect and measure Mycobacterium tuberculosis blood stream infection: a diagnostic and disease

biomarker cohort study. *Lancet. Microbe* 2022;**3**(7):e521-e532. [DOI: 10.1016/S2666-5247(22)00062-3]

98. Peñata-Bedoya A, Zuluaga-Avendaño S, Castaño-Sepúlveda T, Bustamante-Mira J, Ospina-Ospina S. Performance of real-time semi-quantitative polymerase chain reaction assay for optimum diagnosis of extrapulmonary tuberculosis and sensitivity to rifampin in a tertiary care center. *Revista de Investigacion Clinica* 2021;**73**(6):354-61. [DOI: 10.24875/RIC.21000040]

99. Wu X, Tan G, Gao R, Yao L, Bi D, Guo Y, et al. Assessment of the Xpert MTB/RIF Ultra assay on rapid diagnosis of extrapulmonary tuberculosis. *International Journal of Infectious Diseases* 2019;**81**:91-6.

100. Yadav B, Sharma M, Singla N, Shree R, Goyal M, Modi T, et al. Molecular diagnosis of tuberculous meningitis: sdaA-based multi-targeted LAMP and GeneXpert Ultra. *Tuberculosis (Edinburgh, Scotland)* 2023;**140**:102339. [DOI: 10.1016/j.tube.2023.102339]

101. Perez-Risco D, Rodriguez-Temporal D, Valledor-Sanchez I, Alcaide F. Evaluation of the Xpert MTB/RIF Ultra assay for direct detection of Mycobacterium tuberculosis complex in smear-negative extrapulmonary samples. *Journal of Clinical Microbiology* 2018;**56**(9):pii: e00659-18.

102. Anie JR, Leeberk RI, Ria CV, Adhin B, Renu M. Diagnostic accuracy of Truenat MTB plus for the detection of pulmonary and extrapulmonary tuberculosis. *Indian Journal of Medical Microbiology* 2024;**51**:100709. [DOI: https://doi.org/10.1016/j.ijmmb.2024.100709]

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

104. Chin JH, Musubire AK, Morgan N, Pellinen J, Grossman S, Bhatt JM, et al. Xpert MTB/RIF Ultra for detection of Mycobacterium tuberculosis in cerebrospinal fluid. *Journal of Clinical Microbiology* 2019;**57**(6):pii: e00249-19.

105. Cresswell FV, Tugume L, Bahr NC, Kwizera R, Bangdiwala AS, Musubire AK, et al. Xpert MTB/RIF Ultra for the diagnosis of HIV-associated tuberculous meningitis: a prospective validation study. *Lancet Infectious Disease* 2020;**20**(3):308-17.

106. Donovan J, Anh Thu DD, Phu NH, Dung VT, Quang TP, Nghia HD, et al. Xpert MTB/RIF Ultra versus Xpert MTB/RIF for the diagnosis of tuberculous meningitis: a prospective, randomised, diagnostic accuracy study. *Lancet Infectious Disease* 2020;**20**:299-307.

107. Huang M, Wang G, Sun Q, Jiang G, Li W, Ding Z, et al. Diagnostic accuracy of Xpert MTB/RIF Ultra for tuberculous meningitis in a clinical practice setting of China. *Diagnostic Microbiology and Infectious Disease* 2021;**100**(1):115306.

108. Huerga H, Bastard M, Lubega AV, Akinyi M, Antabak NT, Ohler L, et al. Novel FujiLAM assay to detect tuberculosis in HIV-positive ambulatory patients in four African countries: a diagnostic accuracy study. *Lancet Global Health* 2023;**11**:e126–e135.

109. Osei SJ, Maphalala N, Malinga LA, Mbelle NM, Maningi NE. A comparative evaluation of the new Genexpert MTB/RIF Ultra and other rapid diagnostic assays for detecting tuberculosis in pulmonary and extra pulmonary specimens. *Scientific Reports* 2019;**9**(1):16587. [DOI: 10.1038/s41598-019-53086-5]

110. Quinn CM, Kagimu E, Okirworth M, Bangdiwala AS, Mugumya G, Ramachandran PS, et al. Fujifilm SILVAMP TB LAM assay on cerebrospinal fluid for the detection of tuberculous meningitis in adults with human immunodeficiency virus. *Clinical Infectious Diseases* 2021;**73**(9):e3428-e3434. [DOI: 10.1093/cid/ciaa1910]

111. Shao L, Qiu C, Zheng L, Yang Y, Yang X, Liang Q, et al. Comparison of diagnostic accuracy of the GeneXpert Ultra and cell-free nucleic acid assay for tuberculous meningitis: a multicentre prospective study. *International Journal of Infectious Diseases* 2020;**98**:441-6. [DOI: 10.1016/j.ijid.2020.06.076]

112. Sharma K, Sharma M, Shree R, Modi M, Goyal M, Narang D, et al. Xpert MTB/RIF ultra for the diagnosis of tuberculous meningitis: a diagnostic accuracy study from India. *Tuberculosis (Edinburgh, Scotland)* 2020;**125**:101990. [DOI: 10.1016/j.tube.2020.101990]

113. Hoel IM, Sviland L, Syre H, Dyrhol-Riise AM, Skarstein I, Jebsen P, et al. Diagnosis of extrapulmonary tuberculosis using the MPT64 antigen detection test in a high-income low tuberculosis prevalence setting. *BMC Infectious Diseases* 2020;**20**:130. [DOI: 10.1186/s12879-020-4852-z]

114. Hoel IM, Syre H, Skarstein I, Mustafa T. Xpert MTB/RIF Ultra for rapid diagnosis of extrapulmonary tuberculosis in a high-income low-tuberculosis prevalence setting. *Scientific Reports* 2020;**10**(1):13959. [DOI: 10.1038/s41598-020-70613-x]

115. Mekkaoui L, Hallin M, Mouchet M, Payen M, Maillart E, Clevenbergh P, et al. Performance of Xpert MTB/RIF Ultra for diagnosis of pulmonary and extra-pulmonary tuberculosis, one year of use in a multi-centric hospital laboratory in Brussels, Belgium. *PloS One* 2021;**16**:e0249734.

116. Slail MJ, Booq RY, Al-Ahmad IH, Alharbi AA, Alharbi SF, Alotaibi MZ, et al. Evaluation of Xpert MTB/RIF ultra for the diagnosis of extrapulmonary tuberculosis: a retrospective analysis in Saudi Arabia. *Journal of Epidemiology and Global Health* 2023;**13**(4):782-93. [DOI: 10.1007/s44197-023-00150-z]

117. Alomatu S, Vasaikar S, Thomas K, Dubula T, Moeketsi K. The efficiency of TB LAM antigen test to Xpert MTB/RIF Ultra test for the diagnosis of tuberculous pericarditis using pericardial fluid samples. *Pathogens* 2023;**12**(9):1175. [DOI: 10.3390/pathogens12091175]

118. Antel K, Oosthuizen J, Malherbe F, Louw VJ, Nicol MP, Maartens G, et al. Diagnostic accuracy of the Xpert MTB/Rif Ultra for tuberculosis adenitis. *BMC Infectious Diseases* 2020;**20**(1):33.

119. Christopher DJ, Coelho V, Ebby GS, Shankar D, Gupta R, Thangakunam B. Incremental yield of Xpert® MTB/RIF Ultra

over Xpert® MTB/RIF in the diagnosis of extrapulmonary TB. *International Journal of Tuberculosis and Lung Disease* 2021;**25**(11):939-44. [DOI: 10.5588/ijtld.21.0280]

120. Cresswell FV, Ellis J, Kagimu E, Bangdiwala AS, Okirwoth M, Mugumya G, et al. Standardized urine-based tuberculosis (TB) screening with TB-lipoarabinomannan and xpert MTB/RIF ultra in Ugandan adults with advanced human immunodeficiency virus disease and suspected meningitis. *Open Forum Infectious Diseases* 2020;**7**(4):ofaa100. [DOI: 10.1093/ofid/ofaa100]

121. Gao S, Wang C, Yu X, Teng T, Shang Y, Jia J, et al. Xpert MTB/RIF Ultra enhanced tuberculous pleurisy diagnosis for patients with unexplained exudative pleural effusion who underwent a pleural biopsy via thoracoscopy: a prospective cohort study. *International Journal of Infectious Diseases* 2021;**106**:370-5. [DOI: 10.1016/j.ijid.2021.04.011]

122. Makambwa E, Maboreke HR, Fadul M, Meldau R, Dhansay M, Esmail A, et al. Clinical characteristics that portend a positive Xpert Ultra test result in patients with pleural tuberculosis. *African Journal of Thoracic and Critical Care Medicine* 2019;**25**(2):10.7196/AJTCCM.2019.v25i2.011. [DOI: 10.7196/AJTCCM.2019.v25i2.011]

123. Meldau R, Randall P, Pooran A, Limberis J, Makambwa E, Dhansay M, et al. Same day tools, including Xpert Ultra and unstimulated IFN-gamma, for the rapid diagnosis of pleural tuberculosis - a prospective observational study. *Journal of Clinical Microbiology* 2019;**57**(9):e00614-19.

124. Minnies S, Reeve BWP, Rockman L, Nyawo G, Naidoo CC, Kitchin N, et al. Xpert MTB/RIF ultra is highly sensitive for the diagnosis of tuberculosis lymphadenitis in a high-HIV setting. *Journal of Clinical Microbiology* 2021;**59**(12):10.1128/jcm.01316-21. [DOI: 10.1128/jcm.01316-21]

125. Minnie S, Theron G. Diagnostic accuracy of Xpert Ultra using multiple extrapulmonary specimens. Author correspondence 2023 to Minnie S, Theron G (data on file).

126. Ninan MM, Rupali P, Varghese GM, Shalini EB, Venkatesh K, Jesudason MR, et al. Xpert Ultra in diagnosing extrapulmonary TB: accuracy and trace calls. *International Journal of Tuberculosis and Lung Disease* 2022;**26**:441-5. [DOI: 10.5588/ijtld.21.0480]

127. Spener-Gomes R, Costa AG, Melo HF, Souza AB, Beraldi-Magalhães F, Jesus JS, et al. Examination of respiratory specimens improves microbiological diagnosis of patients with presumptive extrapulmonary tuberculosis. *International Journal of Infectious Diseases* 2021;**105**:743-5. [DOI: 10.1016/j.ijid.2021.03.022]

128. Sun Q, Wang S, Dong W, Jiang G, Huo F, Ma Y, et al. Diagnostic value of Xpert MTB/RIF Ultra for osteoarticular tuberculosis. *Journal of Infection* 2019;**79**(2):153-8.

129. Wang G, Wang S, Jiang G, Yang X, Huang M, Huo F, et al. Xpert MTB/RIF Ultra improved the diagnosis of paucibacillary tuberculosis: a prospective cohort study. *Journal of Infection* 2019;**78**(4):311-6.

130. Wang G, Wang S, Yang X, Sun Q, Jiang G, Huang M, et al. Accuracy of Xpert MTB/RIF Ultra for the diagnosis of pleural TB in a multicenter cohort study. *Chest* 2020;**157**(2):268-75.

131. Yu X, Zhang T, Kong Y, Wang F, Dong L, Han M, et al. Xpert MTB/RIF Ultra outperformed the Xpert assay in tuberculosis lymphadenitis diagnosis: a prospective head-to-head cohort study. *International Journal of Infectious Diseases* 2022;**122**:741-6. [DOI: 10.1016/j.ijid.2022.07.039]

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

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

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

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

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

137. Foundation for Innovative New Diagnostics (FIND). Report for WHO: a multicentre non-inferiority diagnostic accuracy study of the Ultra assay compared to the Xpert MTB/RIF assay. www.finddx.org/publication/ultra-report/ (accessed 2 July 2020)(1.8):1-82.

138. World Health Organization. Technical report on critical concentrations for drug susceptibility testing of isoniazid and the rifamycins (rifampicin, rifabutin and rifapentine). https://www.who.int/publications/i/item/9789240017283 2021 (accessed 2 July 2024).

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

140. Shah M, Hanrahan C, Wang ZY, Dendukuri N, Lawn SD, Denkinger CM, et al. Lateral flow urine lipoarabinomannan assay for detecting active tuberculosis in HIV-positive adults. *Cochrane Database of Systematic Reviews* 2016, Issue 5. Art. No: CD011420. [DOI: 10.1002/14651858.CD011420.pub2]

141. Drain PK, Gardiner J, Hannah H, Broger T, Dheda K, Fielding K, et al. Guidance for studies evaluating the accuracy of biomarker-based nonsputum tests to diagnose tuberculosis. *Journal of Infectious Diseases* 2019;**220**(Suppl 3):S108-S115.

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

143. Zhang M, Xue M, He JQ. Diagnostic accuracy of the new Xpert MTB/RIF Ultra for tuberculosis disease: a preliminary systematic review and meta-analysis. *International Journal of Infectious Diseases* 2020;**90**:35-45.

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

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

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

147. Gupta RK, Lucas SB, Fielding KL, Lawn SD. Prevalence of tuberculosis in post-mortem studies of HIV-infected adults and children in resource-limited settings: a systematic review and meta-analysis. *AIDS (London, England)* 2015;**29**(15):1987-2002.

ADDITIONAL TABLES

Table 1. Forms of extrapulmonary TB

Form of extrapulmonary TB	Characteristics	Diagnostic specimens and means of collection
Tuberculous meningitis	Tuberculosis of the meninges affects people of all ages but is most common among children and people with untreated HIV infection. In adults, tuberculous meningitis presents with gradual onset of headache, neck stiffness, malaise, and fever, and if untreated can progress to altered sensorium, focal neurological deficits, coma, and death. Young children may present with poor weight gain, low-grade fever, and listlessness. Infants may present with fever, cough (related to the primary pulmonary infection that occurs before tuberculous meningitis develops), change of consciousness at presentation, bulging anterior fontanel, and seizures [132]. Tuberculous meningitis is sometimes associated with a concurrent cerebral tuberculoma, or, more rarely, a tuberculous abscess.	Cerebrospinal fluid, acquired by lumbar puncture with or without radiological guidance; biopsy of tuberculoma, acquired surgically
Pleural tuberculosis, also called TB pleurisy	TB infection of the pleura presents with gradual onset of pleuritic chest pain, shortness of breath, fever, night sweats, and weight loss. Chest X-ray may demonstrate unilateral or occasionally bilateral pleural effusion. The severity of symptoms is highly variable, with many patients experiencing spontaneous resolution of symptoms, while others may develop severe pleural effusions requiring drainage. Pleuro-pulmonary tuberculosis, in which parenchymal lung involvement is visible on a chest X-ray, is associated with higher mortality than isolated pleural infection, which appears to be rarely fatal [145].	Pleural fluid; pleural biopsy, which may be performed via thoracoscopy or percutaneously with Abram's needle, with or without ultrasound guidance
Lymph node tuberculosis, also called TB lymphadenitis	Tuberculosis of the lymph nodes may affect one node or a group of nodes, or multiple groups within a chain. Lymph node tuberculosis is relatively more common among children than adults. The most common presentation is of a single, firm, non-tender enlarged node in the neck, although any lymph node group can be affected. This may be accompanied by fever, weight loss, and night sweats, particularly in people with HIV. Patients with tuberculosis in deep lymph nodes, such as the mediastinal or mesenteric lymph nodes, may present with fever, night sweats, and weight loss, or, more rarely, with symptoms related to compression of adjacent structures. Over time, lymph nodes become fluctuant and may discharge via a sinus to the skin or an adjacent viscus. It should be noted that lymphadenopathy may also be seen in other forms of tuberculosis as part of the immune response, but this is not usually caused by direct infection of the lymph nodes.	Fine-needle aspiration of fluid from affected lymph node, with or without radiological guidance; surgical biopsy of superficial lymph nodes; endoscopic biopsy of deep lymph nodes with ultrasound guidance
Bone or joint tuberculosis	Tuberculosis of bones or joints, or both, causes chronic pain, deformity, and disability, and tuberculosis of the cervical spine can be life-threatening. The usual presenting symptom is pain. Fever and weight loss, with or without signs	Aspiration of joint fluid or periarticular abscesses; percutaneous

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

	of spinal cord compression, may be present. Patients with advanced disease may have severe pain, spinal deformity, paraspinal muscle wasting, and neurological deficits. Children may have failure to thrive and difficulty walking.	computed tomography-guided biopsy of lesions is preferred, but some patients may require open biopsy
Genitourinary tuberculosis	Tuberculosis of the genitourinary tract includes renal tuberculosis and tuberculosis of the reproductive system. Renal tuberculosis presents with flank pain, hematuria, and dysuria. Female genital tuberculosis presents with infertility (and may be otherwise asymptomatic), pelvic pain, and vaginal bleeding. Testicular tuberculosis presents with a scrotal mass and infertility.	Urine; biopsy of affected organs, acquired under radiological guidance or surgically
Pericardial tuberculosis, also called TB pericarditis	Tuberculosis of the pericardium presents with fever, malaise, night sweats, and weight loss. Chest pain and shortness of breath are also commonly experienced symptoms. Pericardial tuberculosis may be associated with pericardial effusion, which can be severe and lead to life-threatening tamponade. Some patients go on to develop pericardial constriction, which can lead to heart failure and death and may require surgical intervention even after mycobacterial cure.	Pericardial fluid acquired by pericardiocentesis; pericardial biopsy, acquired under radiological guidance or surgically
Peritoneal tuberculosis	Tuberculosis of the peritoneum usually presents with pain and abdominal swelling, which may be accompanied by fever, weight loss, and anorexia.	Ascitic fluid acquired by paracentesis; peritoneal biopsy [146]
Disseminated tuberculosis, also called miliary tuberculosis. It has been proposed that the designation 'miliary TB' be restricted to disseminated TB with miliary shadows on chest radiograph [39].	Disseminated tuberculosis involves two or more distinctly separate sites. Manifestations may be varied, ranging from acute fulminant disease to nonspecific symptoms of fever, weight loss, and weakness. HIV-positive people are more likely to have disseminated tuberculosis than HIV-negative people. In a systematic review of the prevalence of tuberculosis in post-mortem evaluations of HIV-positive people among adults, disseminated tuberculosis was found in 88% of tuberculosis cases and was considered the cause of death in 91% of TB cases [147].	Blood; specimens acquired from affected extrapulmonary sites

Abbreviations:
HIV: human immunodeficiency virus
TB: tuberculosis
We adapted the table from [14].

Table 2. Criteria for the inclusion of individual test technologies in LC-aNAAT class

Parameters	Sensitivity	Specificity
Pre-condition	Total number across included studies ≥ 50 TB+ or drug-resistant TB+ (number with TB or drug resistance)	Total number across included studies ≥ 100 TB − or drug-resistant TB− (number with no TB or drug resistance)
Condition 1	The summary estimate of an assay lies within ±5 percentage points of the overall point estimate.	The summary estimate of an assay lies within ±2 percentage points of the overall point estimate.
Condition 2	The summary estimate for an assay lies within 95% CI of the overall point estimate AND	The summary estimate for an assay lies within 95% CI of the overall point estimate AND The summary estimate for an assay lies within ±5 percentage points of the overall summary estimate.

Table 2. Criteria for the inclusion of individual test technologies in LC-aNAAT class *(Continued)*

	The summary estimate for an assay lies within ±10 percentage points of the overall point estimate.

Abbreviations:
CI: confidence interval
TB: tuberculosis

Table 3. Table of summary study characteristics

Author, year	Country	Age	Female	Clinical setting	Target condition	Index test	Number of specimens evaluated
Alomatu 2023	South Africa	Median age: 48 years (IQR: 33.75 to 66)	47.80%	Tertiary care center	Pericardial TB	Xpert Ultra	44
Anie 2024	India	Mean: 52.3 (SD 19.6)	44.40%	Tertiary care hospital	TB meningitis, pleural TB, lymph node TB, genitourinary TB, bone or joint TB, peritoneal TB	Truenat MTB plus	372
Antel 2020	South Africa	Median 37 years (IQR 30 to 49)	55%	Tertiary referral center, inpatients and outpatients	Lymph node TB	Xpert Ultra	99
Bahr 2017	Uganda	Median 32 years (IQR 30 to 34)	45%	Tertiary care center (inpatient)	TB meningitis	Xpert Ultra	129
Boloko 2022	South Africa	Median 36.3 years [interquartile range (IQR) 40 to 44]	52%	Tertiary care center	Disseminated TB	Xpert Ultra	582
Chin 2019	Uganda	Range 20–41 years	Not reported	Tertiary care center (inpatients)	TB meningitis	Xpert Ultra	11
Christopher 2021	India	Mean 47 years, range: 15–83 years	39%	Tertiary care center	Lymph node TB, pleural TB	Xpert Ultra	250
Cresswell 2020	Uganda	Median age 32 years (IQR: 29 to 38)	42.60%	Tertiary care center (inpatients)	TB meningitis	Xpert Ultra	204
Cresswell 2020a	Uganda	Median age 32 years (IQR: 29 to 38)	42.60%	Tertiary care center (inpatients)	Disseminated TB	Xpert Ultra	264
Donovan 2020	Vietnam	Median age 42 (IQR: 31 to 57)	40%	Tertiary care center (inpatients)	TB meningitis	Xpert Ultra	205
Gao 2021	China	Mean 53.5 years [range 16–80]	29.50%	Tertiary TB referral center (Beijing chest hospital)	Pleural TB	Xpert Ultra	61
Hoel 2020	Norway	Not reported	Not reported	Regional tertiary care hospitals	Pleural TB, lymph node TB	Xpert Ultra	288

Table 3. Table of summary study characteristics *(Continued)*

Hoel 2020a	Norway	Not reported	Not reported	Regional tertiary care hospitals	Lymph node TB	Xpert Ultra	86
Huang 2021	China	Not reported	44%	Regional tertiary care hospitals	TB meningitis	Xpert Ultra	84
Huerga 2023	Uganda, Kenya, Mozambique, and South Africa	Median age: symptomatic HIV-positive group: 43 (35–53); asymptomatic HIV-positive group: 37 (30–45)	52.20%	Outpatient clinics attached to HIV and TB referral hospitals and PHCs	Genitourinary TB	Xpert Ultra	351
Makambwa 2019	South Africa	Mean: 45.6	65%	Referral hospital	Pleural TB	Xpert Ultra	49
Mekkaoui 2021	Belgium	Mean age ± SD: 52.46 ± 28.49 years	32.30%	Tertiary care hospital	Lymph node TB, TB meningitis, genitourinary TB	Xpert Ultra	461
Meldau 2019	South Africa	Median 39 years (IQR 28 to 57)	11%	Tertiary care hospital	Pleural TB	Xpert Ultra	149
Minnies 2021	South Africa	Median age: 36 (IQR: 29–46.5)	53%	Tertiary referral clinic	Lymph node TB	Xpert Ultra	135
Minnies 2023	South Africa	Median age: 41 (IQR: 34–53)	42%	Tertiary care hospital	Pleural TB, pericardial TB	Xpert Ultra	270
Ninan 2022	India	Not reported	Not reported	Tertiary care hospital	Lymph node TB, bone or joint TB, TB meningitis	Xpert Ultra	242
Osei 2019	South Africa	Not reported	Not reported	University hospital	TB meningitis, genitourinary TB	Xpert Ultra	78
Peña-ta-Bedoya 2021	Colombia	Range: 15–92 years	37%	University hospital	TB meningitis, pleural TB, Lymph node TB, bone or joint TB, peritoneal TB, pericardial TB	Xpert Ultra	540
Perez-Risco 2018	Spain	> 18 years	Not reported	Laboratory-based evaluation	TB meningitis, pleural TB, genitourinary TB, bone or joint TB	Xpert Ultra	75

Table 3. Table of summary study characteristics *(Continued)*

Quinn 2021	Uganda	Median age: 33 (IQR: 26–40)	36.60%	Referral hospital	TB meningitis	Xpert Ultra	48
Shao 2020	China	Mean: 37.6 (range: 18–69)	52%	University hospital	TB meningitis	Xpert Ultra	84
Sharma 2020	India	Not reported	Not reported	Tertiary care hospital	TB meningitis	Xpert Ultra	244
Sharma 2021	India	Not reported	Not reported	Tertiary care hospital	TB meningitis	Xpert Ultra and Truenat MTB plus	108
Sharma 2023	India	Not reported	Not reported	Tertiary care hospital	Lymph node TB	Xpert Ultra and Truenat MTB plus	100
Slail 2023	Saudi Arabia	32 ± 17.1 years	39.50%	Referral hospital	TB meningitis, lymph node TB, genitourinary TB, bone or joint TB, pleural TB, pericardial TB, peritoneal TB	Xpert Ultra	845
Spen-er-Gomes 2021	Brazil	37 (IQR: 30–43)	38%	University hospital	TB meningitis, lymph node TB, peritoneal TB, pleural TB, genitourinary TB	Xpert Ultra	157
Sun 2019	China	Median 51 years (range 16 to 86)	55%	National TB referral center	Bone or joint TB	Xpert Ultra	166
Wang 2019	China	Range: 15–89 years	Pleural TB: 20%; TB meningitis: 445	National TB referral center	TB meningitis, pleural TB	Xpert Ultra	131
Wang 2020	China	Median 45 years; range: 15 to 89	32.50%	Laboratory-based evalua-tion	Pleural TB	Xpert Ultra	139
Wu 2019	China	> 16 years	32%	Tertiary care hospital	Lymph node TB, pleural TB	Xpert Ultra	119
Yadav 2023	India	Mean ± SD: 37.15 ± 16.27	56%	Tertiary care hospital	TB meningitis	Xpert Ultra	300

Table 3. Table of summary study characteristics *(Continued)*

Yu 2022	China	Median age: 32 (range: 16-89)	58%	Tertiary care hospital	Lymph node TB	Xpert Ultra	106

Abbreviations:
IQR; interquartile range
PHC: primary health care
TB: tuberculosis

Table 4. Summary accuracy of LC-aNAAT for detection of extrapulmonary tuberculosis and rifampicin resistance

EPTB specimen	Reference standard	Number of studies	Number of participants (cases)	Summary sensitivity (95% CI)	Summary specificity (95% CI)	Positive predictive value (95% CI)*	Negative predictive value (95% CI)*
Cerebrospinal fluid	MRS	16	1684 (287)	88.2 (83.7 to 91.6)	96.0 (86.8 to 98.9)	70.9 (41.3 to 89.9)	98.7 (98.0 to 99.1)
Cerebrospinal fluid	CRS	10	1397 (753)	60.3 (50.9 to 69.0)	99.2 (98.1 to 99.7)	89.6 (75.3 to 96.0)	95.7 (94.7 to 96.7)
Pleural fluid	MRS	13	1041 (264)	74.0 (60.8 to 83.9)	88.1 (78.8 to 93.6)	40.8 (24.1 to 59.4)	96.8 (94.8 to 98.1)
Pleural fluid	CRS	7	667 (298)	43.6 (32.8 to 55.0)	99.2 (95.2 to 99.9)	85.4 (42.9 to 97.8)	94.1 (92.7 to 95.2)
Pleural tissue	MRS	2	105 (24)	-	-	-	-
Pleural tissue	CRS	2	105 (53)	-	-	-	-
Lymph node aspirate	MRS	9	445 (89)	85.3 (73.4 to 92.4)	74.1 (63.5 to 82.5)	26.8 (18.3 to 37.0)	97.8 (95.6 to 99.0)
Lymph node aspirate	CRS	6	461 (243)	71.3 (64.3 to 77.4)	97.4 (82.3 to 99.7)	75.6 (28.7 to 96.5)	96.8 (95.4 to 97.5)
Lymph node tissue	MRS	8	578 (97)	96.5 (84.7 to 99.3)	79.4 (65.4 to 88.8)	34.3 (21.4 to 49.5)	99.5 (97.5 to 99.9)
Lymph node tissue	CRS	3	229 (109)	61.5 (47.1 to 74.2)	96.7 (91.5 to 98.7)	67.2 (38.0 to 86.8)	95.8 (94.0 to 97.2)
Urine	MRS	6	232 (14)	-	-	-	-
Urine	CRS	3	693 (159)	23.0 (14.7 to 34.1)	98.9 (89.7 to 99.9)	70.7 (13.7 to 97.4)	92.0 (90.4 to 93.2)
Bone or joint fluid	MRS	3	126 (58)	96.6 (87.2 to 99.1)	91.1 (80.8 to 96.2)	54.8 (33.6 to 74.2)	99.6 (98.3 to 99.9)

Table 4. Summary accuracy of LC-aNAAT for detection of extrapulmonary tuberculosis and rifampicin resistance *(Continued)*

Bone or joint fluid	CRS	1	145 (111)	-	-	-	-
Pericardial fluid	MRS	3	202 (75)	84.0 (73.9 to 90.7)	86.6 (79.5 to 91.5)	41.1 (28.6 to 54.3)	98.0 (96.5 to 98.9)
Peritoneal fluid	MRS	3	69 (8)	-	-	-	-
Blood	MRS	1	578 (423)	-	-	-	-
Rifampicin resistance	MRS	13	446 (54)	100.0 (93.4 to 100.0)	99.4 (92.1 to 100.0)	94.6 (56.7 to 100.0)	100.0 (99.2 to 100.0)

Abbreviations:
CI: confidence interval
CrI: credible interval
CRS: Composite reference standard
CSF: cerebrospinal fluid
EPTB: extrapulmonary tuberculosis
LC-aNAAT: low-complexity automated nucleic acid amplification test
LPA: Line probe assay
MRS: Microbiological Reference Standard
TB: tuberculosis
Studies in this table include estimates from Xpert Ultra only as Truenat MTB plus could not be included in the class-level analyses.
Studies included in the table are limited to those that report data for both sensitivity and specificity; thus, the number of studies (specimens) may differ slightly from those reported in the main text of the review. For tuberculosis detection, the reference standard was a microbiological reference standard and a composite reference standard. For rifampicin resistance detection, the reference standards were culture-based drug susceptibility testing or line probe assay. Pooled sensitivity and pooled specificity are posterior median estimates.
Analyses where univariate/fixed-effect models were used were: CSF for CRS; lymph node biopsy for CRS; urine for CRS; bone or joint fluid for MRS; pericardial fluid for MRS; rifampicin resistance.

Table 5. Impact of concentrating cerebrospinal fluid on LC-aNAAT sensitivity and specificity

Covariate (number of studies, participants)	Pooled sensitivity (95% CrI)	Pooled specificity (95% CrI)
Microbiological reference standard		
Concentrated specimen (5, 781)	92.8 (87.5 to 96.0)	93.6 (70.8 to 98.9)
Unconcentrated specimen (6, 470)	81.0 (68.0 to 89.5)	85.8 (68.9 to 94.3)
Composite reference standard		
Concentrated specimen (5, 828)	70.0 (64.7 to 74.7)	99.9 (81.8 to 100.0)
Unconcentrated specimen (3, 228)	50.9 (41.7 to 60.0)	100.0 (97.4 to 100.0)

Abbreviations:
CrI: credible interval
LC-aNAAT: low-complexity automated nucleic acid amplification test

Table 6. Cerebrospinal fluid starting volumes, concentration step and accuracy estimates

Study	CSF volume	Concentration	Sensitivity (95% CI)	Specificity (95% CI)
Chin 2019	0.8 mL	No	80 (28 to 99)	50 (12 to 88)
Cresswell 2020	Target > 6 mL	Yes	89 (71 to 98)	92 (86 to 95)
Donovan 2020	6 mL	Yes	91 (71 to 99)	94 (85 to 98)
Huang 2021	3–5 mL	No	86 (42 to 100)	75 (64 to 84)
Peñata-Bedoya 2021	< 2 mL	No	80 (52 to 96)	96 (92 to 98)
Quinn 2021	2 mL	No	87 (60 to 98)	88 (72 to 97)
Shao 2020	6 mL	No	50 (7 to 93)	68 (56 to 78)
Sharma 2020	3–5 mL	Yes	96 (88 to 100)	100 (98 to 100)
Sharma 2021	2–3 mL	Yes	92 (79 to 98)	51 (40 to 63)
Wang 2019	> 3 mL	No	86 (65 to 97)	100 (80 to 100)
Yadav 2023	2.5–3.5 mL	Not clear	86 (73 to 94)	100 (99 to 100)

Abbreviations:
CSF: cerebrospinal fluid
CI: confidence interval
Sensitivity and specificity are presented as percentages (95% CI).