CADTH

CADTH RAPID RESPONSE REPORT:
SUMMARY WITH CRITICAL APPRAISAL

Rapid and Simultaneous Tuberculosis and Antibiotic Susceptibility Testing for the Diagnosis of Pulmonary Tuberculosis and Rifampicin Resistance: A Review of Diagnostic Accuracy

Service Line:	Rapid Response Service
Version:	1.0
Publication Date:	December 10, 2020
Report Length:	30 Pages

CADTH

Authors: Srabani Banerjee, Melissa Severn

Cite As: Rapid and Simultaneous Tuberculosis and Antibiotic Susceptibility Testing for the Diagnosis of Pulmonary Tuberculosis and Rifampicin Resistance: A Review of Diagnostic Accuracy. Ottawa: CADTH; 2020 Dec. (CADTH rapid response report: summary with critical appraisal).

ISSN: 1922-8147 (online)

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

Questions or requests for information about this report can be directed to **Requests@CADTH.ca**

Abbreviations

CI	confidence interval
CrI	credible interval
PTB	pulmonary tuberculosis
TB	tuberculosis

Context and Policy Issues

Tuberculosis (TB) is one of the top 10 causes of death, worldwide.[1] Worldwide 10 million people developed TB in 2017, (i.e.,133 cases per 100,000).[1-3] However, only 6.4 million cases were reported to WHO, indicating that 36% of cases were undiagnosed or not reported.[2] According to a 2017 estimate, the rate of active TB in Canada was 4.9 per 100,000 population. The rate was highest among Canadian-born Indigenous Peoples (21.5 per 100,000 population)[2].

TB is caused by the bacterium, *Mycobacterium tuberculosis*. It is a communicable disease and can spread from person to person through air (such as when a person with TB coughs).[1,3] TB mainly affects the lungs (pulmonary tuberculosis [PTB]) but can also affect other organs and is referred to as extrapulmonary TB (such as TB meningitis, and TB lymph nodes).[1] TB is a curable disease, if diagnosed early and treated.

The conventional diagnostic approach for individuals with presumptive pulmonary TB includes smear microscopy, followed by the culture-based method to confirm the diagnosis and for drug susceptibility testing.[4] The culture-based method is considered the gold standard. However, it takes two to six weeks to get the culture results and an additional three or more weeks for conventional multi drug resistance testing.[5] Hence there is potential for treatment being delayed if waiting for confirmatory culture results. Also, if treatment is started based on less reliable test results there is potential for initiation of treatment that was necessary. Treatment options for TB include medications such as isoniazide, rifampicin, ethambutol and pyrazinamide.[6] Drug resistant TB continues to be a public health concern. Globally in 2019, close to half a million people developed rifampicin-resistant TB, and 78% of them had multidrug-resistant TB[1] Considering these issues there is a need for a test that is rapid and can simultaneously diagnose TB and antibiotic resistance.

The Xpert MTB/RIF assay is a relatively new test that is rapid (takes less than two hours) and can simultaneously detect *Mycobacterium tuberculosis* and rifampicin resistance. It is a nucleic acid amplification test that requires a disposable cartridge and the GeneXpert Instrument system. Sputum sample collected from the patient suspected of TB, is mixed with a reagent that is provided with the assay, and then the cartridge containing the mixture is placed in the GeneXpert Instrument. From this point onwards the process is fully automated. Technical training to run the test is minimal.[5] Subsequently, to enhance the performance of Xpert MTB/RIF assay, the Xpert MTB/RIF Ultra (Xpert Ultra) assay was developed. Xpert Ultra assay uses a newly developed cartridge and an updated software with the same instrument.[3]

This report is an upgrade from a CADTH report (Summary of Abstracts)[7] published in June 2020. The aim of this current report is to review the evidence regarding the rapid and simultaneous tuberculosis and antibiotic susceptibility testing for PTB and rifampicin

resistance. This report will be followed by a second report on the cost-effectiveness rapid and simultaneous tuberculosis and antibiotic susceptibility testing for the diagnosis of pulmonary tuberculosis and rifampicin resistance. Additionally, these reports are components of a larger CADTH Condition Level Review on tuberculosis. A condition level review is an assessment that incorporates all aspects of a condition, from prevention, detection, treatment, and management. For more information on CADTH's Condition Level Review of tuberculosis, please visit the project page (https://www.cadth.ca/tuberculosis).

Research Questions

1. What is the diagnostic accuracy of rapid and simultaneous tuberculosis and antibiotic susceptibility testing for the detection of pulmonary tuberculosis compared to smear microscopy?
2. What is the diagnostic accuracy of rapid and simultaneous tuberculosis and antibiotic susceptibility testing for the detection of pulmonary tuberculosis compared to mycobacterial cultures?
3. What is the diagnostic accuracy of rapid and simultaneous tuberculosis and antibiotic susceptibility testing for the detection of rifampicin resistance compared to culture-based drug susceptibility testing?

Key Findings

A total of six systematic reviews were identified.

Three systematic reviews reported on the diagnostic accuracy of Xpert with respect to smear microscopy status. There was variability in the estimates of sensitivities and specificities in both the smear positive and smear negative subgroups in the individual studies included in these three systematic reviews, One systematic review reported that the sensitivity with Xpert was 36% to 44% higher than the sensitivity with smear microscopy.

Five systematic reviews reported on the diagnostic accuracy of Xpert and/or Xpert Ultra compared with mycobacterial culture test. The sensitivities of Xpert ranged between 62% and 85%; and the specificities ranged between 98% and 99% (based on four systematic reviews). The sensitivities of Xpert Ultra ranged between 64% and 100%; and specificities ranged between 96% and 100% (based on two systematic reviews).

Six systematic reviews reported on the diagnostic accuracy of Xpert or both Xpert and Xpert Ultra for the detection of rifampicin resistance compared to culture-based drug susceptibility testing. For rifampicin resistance detection with Xpert, the sensitivities ranged between 83% and 100%, and the specificities ranged between 97% and 100% (based on six systematic reviews). For rifampicin resistance detection with Xpert Ultra the sensitivities ranged between 93% and 95%, and the specificities ranged between 98% and 99% (based on two systematic reviews.

The findings need to be interpreted in the light of limitations (such as variations in specimen types and settings; limited evidence for Xpert Ultra; and lack of generalizability with respect to the Canadian setting).

Methods

Literature Search Methods

This report is an update of a literature search strategy developed for a previous CADTH report.[7] For the current report, a limited literature search was conducted on key resources including MEDLINE, the Cochrane Library, University of York Centre for Reviews and Dissemination (CRD) databases, Canadian and major international health technology agencies, as well as a focused internet search. Methodological filters were use to limit search results to health technology assessments, systematic reviews, meta-analyses, or network meta-analyses, any types of clinical trials or observational studies, and economic studies. The initial search was limited to English-language documents published between January 1, 2015 and June 9, 2020. For the current report, database searches were rerun on October 28, 2020 to capture any articles published since the initial search date. The search of major health technology agencies was also updated to include documents published since June 2020.

Selection Criteria and Methods

One reviewer screened citations and selected studies. In the first level of screening, titles and abstracts were reviewed and potentially relevant articles were retrieved and assessed for inclusion. The final selection of full-text articles was based on the inclusion criteria presented in Table 1.

Table 1: Selection Criteria

Population	People of any age with presumptive pulmonary tuberculosis, presumptive rifampicin-resistant tuberculosis or presumptive multi-drug resistant tuberculosis
Intervention	Automated real-time nucleic acid amplification test for rapid and simultaneous detection of tuberculosis and rifampicin resistance (e.g., Xpert)
Comparator	Q1. smear microscopy (e.g., Acid-fast bacilli (AFB) Q2. Liquid or solid mycobacterial cultures Q3. Culture-based drug susceptibility testing
Outcomes	Q1-3. Diagnostic accuracy (e.g., sensitivity, specificity, TB diagnosis, diagnosis of rifampicin resistance, true positive rate, false positive rate)
Study Designs	Health Technology Assessments and Systematic reviews

AFB = acid-fast bacilli; TB = tuberculosis

Exclusion Criteria

Articles were excluded if they did not meet the selection criteria outlined in Table 1, they were duplicate publications, or were published prior to 2015. As there were a large number of systematic reviews and primary studies, primary studies were not included; only health technology assessments and systematic reviews were included. Studies on extrapulmonary TB (such as meningitis TB and lymph node TB) included in the included systematic reviews were not considered.

Critical Appraisal of Individual Studies

The included publications were critically appraised by one reviewer using the following tool as a guide: A MeaSurement Tool to Assess systematic Reviews 2 (AMSTAR 2)[8] for

systematic reviews. Summary scores were not calculated for the included studies; rather, the strengths and limitations of each included publication were described narratively.

Summary of Evidence

Quantity of Research Available

A total of 272 citations were identified in the literature search. Following screening of titles and abstracts, 269 citations were excluded and three potentially relevant reports from the electronic search were retrieved for full-text review. Six potentially relevant publications were retrieved from the previous literature search for full-text review. Of these nine potentially relevant articles, three publications were excluded for various reasons, and six publications met the inclusion criteria and were included in this report. These comprised one health technology assessment[4] and five systematic reviews.[3,9-12] Appendix 1 presents the PRISMA[13] flowchart of the study selection.

Summary of Study Characteristics

One systematic review within a health technology assessment (HTA) report[4] and five systematic reviews.[3,9-12] were identified. For systematic reviews [4,9,10] that had a broad objective, only the studies that were relevant for this report were considered here. Pooled results from the meta-analyses could not be included for some systematic reviews[9,10] as primary studies that were not relevant for this report were included in the pooling. There was some overlap in the studies included in the systematic reviews hence it should be noted there is double counting of studies and findings from the systematic reviews are not exclusive. Xpert MTB/RIF or Gene Xpert will be referred to as Xpert in the text, and Xpert MTB/RIF Ultra as Xpert Ultra. Additional details regarding the characteristics of included publications are provided in Appendix 2.

Study Design

The systematic review within a HTA report[4] included six relevant studies (study type not reported) published between 2010 and 2013. The systematic review by Kay et al.[9] included 28 studies (1 randomized controlled trial [RCT], 14 cohort studies and 13 cross-sectional studies) published between 2011 and 2019. The systematic review by Zhang et al.[10] included eight studies (5 prospective, and 3 retrospective) published between 2017 and 2019.The systematic review by Horne et al.[3] included 95 studies (2 RCTs and 93 cross-sectional) published between 2011 to 2017. The systematic review by Detjen et al.[11] included 15 studies (study type not reported) published between 2011 and 2014. The systematic review by Wang et al.[12] included 11 studies (2 prospective, and 9 cross-sectional) published between 2011 and 2014.

Country of Origin

The first author of the systematic review (within the HTA)[4] was from the UK; and the included studies were conducted in Republic of Korea, Russia, South Africa, and Zambia (1 each) and two studies were multinational.

The first author of the systematic review by Kay et al.,[9] was from the US and included studies conducted in Bangladesh, France, Gambia, Malawi, Vietnam, Zambia, Zimbabwe and multi-national (1 each), India (2), Kenya (2), Tanzania (3), Uganda (3), and South Africa (10).

The first author of the systematic review by Zhang et al.[10] was from China, and the included studies were conducted in Tanzania (1), Switzerland (1), France (2), South Africa (2), and multinational (2).

The first author of the systematic review by Horne et al.[3] was from the UK, and the included studies were conducted in Botswana, Cambodia, Canada, Chile, Cote d'Ivoire, France, Ghana, Japan, Kenya, Lithuania, Nepal, New Zealand, Peru, Poland, Rwanda, Spain, Sudan, Turkey, and Vietnam (1 each); Egypt, Kyrgyzstan, Malawi, Russia, Singapore, and Zambia (2 each); India, Tanzania, Thailand, and Zimbabwe (3 each); Ethiopia (4); China, and US (5 each); Uganda (6); Republic of Korea, and multi-national (7 each).

The first author of the systematic review by Detjen et al.[11] was from the US, and the included studies were conducted in Zambia, Spain, Bangladesh, Malawi, Vietnam, China, Tanzania, Uganda, Italy, and multinational (1 each), and South Africa (5).

The first author of the systematic review by Wang et al.[12] was from China, and the included studies were conducted in Vietnam, Tanzania, Uganda, Zambia, multinational (1 each), China (2) and South Africa (4).

Patient Population

The number of participants in the individual studies included in the systematic reviews ranged between 50 and 1442 in one systematic review,[9] between 33 and 1753 in the second systematic review,[10] between 58 and 6648 in the third systematic review,[3] between 20 and 930 in the fourth systematic review,[11] and between 62 and 6648 in the fifth systematic review.[4] In the sixth systematic review[12] the number of specimens ranged between 73 and 930 in the individual included studies. Three systematic reviews[3,9,11] involved children, one systematic review involved adults,[3] one systematic review[4] included mostly adults (for a few of the included studies the age was not specified) and one systematic review[10] involved both children and adults. HIV status was variable among the individual studies or was not reported for some of the studies.

Interventions and Comparators

The index test was Xpert[3,4,9-12] or Xpert Ultra.[3,9-11]

The reference standard test was a mycobacterial culture (either liquid or solid of various types such as mycobacteria growth indicator tube [MGIT], Lowenstein-Jensen [LJ], 7H11, and Ogawa). For rifampicin-resistance testing the reference standard was drug sensitivity testing (details were not reported).

Outcomes

Outcomes reported included sensitivity,[3,4,9-12] specificity,[3,4,9-12] positive likelihood ratio,[12] negative likelihood ratio,[12] positive predictive value,[9] negative predictive value,[9] and diagnostic odds ratio.[4,12] Specificity and sensitivity data are presented in the main text; and all outcome data are available in Appendix 4.

Summary of Critical Appraisal

An overview of the critical appraisal of the included publications is summarized below. Additional details regarding the strengths and limitations of included publications are provided in Appendix 3.

Overall, the included systematic reviews were well conducted. In all the included systematic reviews[3,4,9-12] the objective was stated, multiple databases were searched, article selection was described and was conducted independently by two reviewers, data extraction was conducted independently by two reviewers and lists of included articles were presented and study characteristics were described. Quality assessment was conducted in duplicate by two reviewers in three systematic reviews,[3,9,12] and in three systematic reviews[4,10,11] it was unclear if quality assessment was done in duplicate. In all the systematic reviews, the QUADAS tool was used to assess quality. In the systematic review by Kay et al.[9] the authors reported that the certainty of evidence (based on Grading Recommendations Assessment, Development and Evaluation [GRADE]) for Xpert was moderate for both sensitivity and specificity; and for Xpert Ultra, the certainty of evidence was low for sensitivity and high for specificity. In the systematic review by Horne et al.[3] the authors reported that the certainty of evidence for Xpert was high for both sensitivity and specificity; and for Xpert Ultra was moderate for both sensitivity and specificity. In the systematic review by Zhang et al.[10] the risk of bias in terms of flow and timing; and patient selection was high or unclear, and in terms of applicability of the assay there was no apparent concern. In the systematic review by Detjen et al.[11] most of the studies seemed to have low risk of bias with respect to patient selection, further details of quality assessment were not presented. In the systematic review by Drobniewski et al.[4] the risk of bias in terms of flow and timing; index test; and reference standard were generally unclear, and in terms of applicability of the assay there was no apparent concern. In the systematic review by Wang et al [12] the risk of bias (in terms of patient selection, index test, reference standard, and flow and timing) appeared to be generally low. In all the systematic reviews meta-analyses were conducted and seemed appropriate. In four systematic reviews[4,10-12] the authors appear to have no conflicts of interest. In the fifth systematic review [9] the majority of the authors had no conflicts of interest; one author received from the manufacturer a discount for Xpert for a TB case finding program which is less likely to be a major concern. In the sixth systematic review[3] the majority of the authors had no conflicts of interest; one author had received financial support from a non-profit agency (that collaborates, with several organizations including the manufacturer), and one author was employed by this agency; hence the potential for bias cannot be entirely ruled out.[4]

Summary of Findings

The main findings are summarized below. Details of the study findings and authors' conclusions are presented in Appendix 4. Pooled estimates from meta-analyses, when available are presented and the number of studies contributing to the pooled estimate are indicated in parenthesis. When pooled estimates were not available, the estimates from the individual studies are presented as a range of estimates and the number of studies is indicated in parenthesis. It is worth noting that there was some overlap in the studies included in the systematic reviews hence findings from the systematic reviews are not exclusive.

Diagnostic Accuracy of Rapid and Simultaneous Tuberculosis and Antibiotic Susceptibility Testing for the Detection of Pulmonary Tuberculosis and Rifampicin Resistance

Detection of pulmonary tuberculosis compared to smear microscopy

Three systematic reviews[3,9,11] reported on diagnostic accuracy of Xpert with respect to smear microscopy status (i.e., smear positive or smear negative).

Children:

In the systematic review by Kay et al.[9] the pooled sensitivity of Xpert in smear-positive sputum specimens was 97.8%; 95%CI, 91.6% to 99.4%, with the sensitivities for the individual studies ranging between 93% and 100% (11 studies). In this systematic review, the specificity of Xpert in smear-positive sputum specimens ranged from 67% to 100% with the lower and upper bounds of the 95% CI ranging between 3% and 42%, and between 98% and 100%, respectively (4 studies). In the systematic review by Detjen et al.[11] the sensitivity of Xpert in smear-positive sputum samples ranged between 92% and 100%, with lower bound of the 95% CI ranging between 3% and 85%, and the upper bound of the 95% CI being 100% (3 studies). Also, the sensitivity of Xpert in smear-positive gastric fluid ranged between 92% and 100%, with the lower bound of the 95% CI ranging between 3% to 78%, and the upper bound of the 95% CI being 100% (6 studies). In this systematic review, the specificity of Xpert in smear-positive sputum samples was 86%; 95% CI, 42% to 100% (1 study); and in smear-positive gastric fluid was 96%; 95% CI, 78% to 100% (1 study).

In the systematic review by Kay et al.[9] the pooled sensitivity of Xpert in smear-negative sputum specimens was 58.9%%; 95% CI, 45.6% to 71.0% with the sensitivities for the individual studies ranging between 22% and 100%; and the pooled specificity of Xpert in smear-negative sputum specimens was 99.1%; 95% CI, 97.1% to 99.7%, with the specificities for the individual studies ranging between 86% and 100% (12 studies). In the systematic review by Detjen et al.[11] the sensitivity of Xpert in smear-negative sputum specimens ranged between 25% and 100%, with lower bound of the 95% CI ranging between 3% and 46%, and the upper bound of the 95% CI ranging between 58% to 100%, (11 studies). In this systematic review, the specificity of Xpert in smear-negative sputum ranged between 93% and 100%, with lower bound of the 95% CI ranging between 81% and 98%, and the upper bound of the 95% CI ranging between 99% to 100%, (11 studies).

In the systematic review by Detjen et al.[11] the sensitivity with Xpert was 36% to 44% higher than the sensitivity with smear microscopy.

Adults:

In the systematic review by Horne et al.[3] the pooled sensitivity of Xpert in smear-positive specimens was 98%; 95% CrI, 97 % to 98% (45 studies) with the sensitivities for the individual studies ranging between 75% and 100%; the pooled specificity could not be determined as in many studies the values of true negatives were zero. In this systematic review, the sensitivity of Xpert in smear-negative specimens was 67%; 95% CrI, 62% to 72% with sensitivities for individual studies ranging between 28% and 100%; and the specificity was 98%; CrL, 98% to 99% with specificities for individual studies ranging between 57% and 100% (45 studies).

In summary, in three systematic reviews[3,9,11] the sensitivities and specificities of Xpert were reported for the smear positive subgroups and the smear negative subgroups. In the smear positive subgroups, the sensitivities and specificities in the individual studies ranged from 75% to 100%, and from 67% to 100% respectively. In the smear negative subgroups, the sensitivities and specificities in the individual studies ranged from 22% to 100%, and from 57% to 100%. Only one systematic review[11] reported that the sensitivity with Xpert was 36% to 44% higher than the sensitivity with smear microscopy.

Detection of pulmonary tuberculosis compared to mycobacterial cultures

Five included systematic reviews[3,9-12] reported on diagnostic accuracy of Xpert compared to mycobacterial culture . Three systematic reviews[3,9,10] also reported on the diagnostic accuracy of Xpert Ultra.

Children:

In the systematic review by Kay et al.[9] with sputum specimens the pooled sensitivity of Xpert was 65%; 95% CI, 55% to 73%; and the pooled specificity was 99%, 95% CI, 98% to 100% (23 studies). Similar trends for sensitivity and specificity were found with gastric aspirate specimens, stool specimens, and nasopharyngeal specimens; and also with Xpert Ultra.

In the systematic review by Detjen et al.[11] the pooled sensitivity of Xpert using sputum specimens was 62%; 95% CrI, 51% to 73%; and the pooled specificity was 98%; 95% CrI, 97% to 99% (12 studies) Similar trends for sensitivity and specificity were found with gastric aspirate specimens.

In the systematic review by Wang et al.[12] the pooled sensitivity of Xpert was 65%; 95% CI, 61% to 69%, (11 studies); and the specificity was 99%, 95% CI, 98% to 99%, (11 studies).

In the systematic review by Zhang et al.[10] the sensitivities with Xpert Ultra were 75%; 95% CI, 64% to 85% in one study and 64%, 95% CI, 44% to 81% in another study. In this systematic review, the specificities with Xpert Ultra were 97%; 95% CI,94% to 99 % in one study; and 100%; 95% CI, 97% to 100% in another study.

Adults:

In the systematic review by Zhang et al.[10] the sensitivities of Xpert Ultra ranged between 82% and 100%, with the lower bound of 95% CI ranging between 65% and 93%, and the upper bound of 95% CI ranging between 91% and 100%, (6 studies). In this systematic review, the specificities of Xpert Ultra ranged between 96% and 100%, with the lower bound of 95% CI ranging between 88% and 94%, and the upper bound of 95% CI ranging between 97% and 100% (5 studies). This systematic review also reported on Xpert. For Xpert, the sensitivities were 93%; 95% CI, 80 to 98 in one study, and 95%; 95%CI, 91% to 98%. For Xpert the specificities were 98%; 95% CI, 93% to100% in one study, and 98%; 95% CI, 96%-99% in another study.

In the systematic review by Horne et al.[3] the pooled sensitivity of Xpert was 85%; 95% CrI, 82% to 88%; and the pooled specificity of Xpert was 98%; 95% CrI, 97% to 98% (70 studies). In this systematic review, the sensitivity with Xpert Ultra was 95%; 95% CI, 90% to 98%; and specificity was 98%; 95% CI, 97% to 99% (1 study).

In summary, the sensitivities of Xpert ranged between 62% and 85%; and the specificities ranged between 98% and 99%. The sensitivities of Xpert Ultra ranged between 64% and 100%; and specificities ranged between 96% and 100%.

Detection of rifampicin resistance compared to culture-based drug sensitivity testing

All six included systematic reviews[3,4,9-12] reported on detection of rifampicin resistance with Xpert compared to culture-based drug sensitivity testing.

Children:

In the systematic review by Kay et al.[9] for rifampicin resistance the sensitivities ranged between 83% and 100% with the lower bound of 95% CI ranging between 16% and 59% and the upper bound of 95% CI being 100% (3 studies). In this systematic review for rifampicin resistance the specificities ranged between 97% and 100%. with the lower bound of 95% CI ranging between 72% and 95% and the upper bound being 100% (3 studies).

In the systematic review by Detjen et al.[11] for rifampicin resistance the pooled sensitivity of Xpert was 86%; 95% CrI, 53% to 98% ; and the pooled specificity was 98%; 95% CrI, 94% to 100% (3 studies).

In the systematic review by Wang et al.[12] for rifampicin resistance the pooled sensitivity of Xpert was 94%; 95% CI, 73% to 100%; and the pooled specificity was 99%, 95% CI, 98% to 99%, (5 studies).

Adults:

In the systematic review by Horne et al.[3] for rifampicin resistance the pooled sensitivity of Xpert was 96%; 95% CrI, 94% to 97%; and the pooled specificity was 98%; 95% CrI, 98% to 99%, (48 studies). In this systematic review, the sensitivity of Xpert Ultra was 95%; 95% CI, 90% to 98%; and specificity was 98%; 95% CI, 97% to 99% (1 study). Of note, in this same study, for Xpert the sensitivity was 95%; 95% CI, 93% to 100%; and specificity was 98%, 95% CI, 96% to 99%.

In the systematic review by Zhang et al.[10] for rifampicin resistance the sensitivity of Xpert was 93%, 95% CI, 80% to 98% in one study, and 95%, 95% CI, 91% to 98%, in another study; and the specificity was 98%; 95% CI, 93% to 100% in one study and 98%; 95% CI, 96% to 99% in another study. The sensitivities and specificities for Xpert Ultra followed a similar trend.

In the systematic review by Drobniewski[4] for rifampicin resistance the pooled sensitivity of Xpert was 97%; 95% CI, 94% to 99%; and the pooled specificity was 98%; 95% CI, 98% to 99%, (6 studies [Population was adults in 4 studies and not reported in 2 studies]).

In summary, for rifampicin resistance for Xpert the sensitivities ranged between 83% and 100%, and the specificities ranged between 97% and 100%. For rifampicin resistance, for Xpert Ultra the sensitivities ranged between 93% and 95%, and the specificities ranged between 98% and 99%.

Limitations

There was variability in the sensitivity and specificity estimates in the individual studies, likely because of heterogeneity among the individual studies. The tests were conducted using various types of specimens (such as sputum, respiratory, gastric aspirates, and nasopharyngeal) which may impact the findings. Also, factors such as patient age group, health status of the patient, and setting where the test is performed could impact findings. Of note, categorization of the adult group varied, some systematic reviews on adults included studies that included participants greater than or equal to 14 years, 15 years or 16 years. It is however unclear whether including the ages 14 years up to 18 years in the adult group would impact the findings for the adult group. It appears from the results reported in the Summary of Findings section, that the estimates of sensitivities of Xpert were generally lower for the pediatric population compared to that for the adult population.

Comparison of the sensitivity and specificity of Xpert and Xpert Ultra were based on the assumption that the culture reference standard is 100% sensitive and specific. However it has been reported that the culture reference standard may not be the most accurate one to detect paucibacillary TB in children.[9]

There was some overlap in the studies included in the systematic reviews hence results are not exclusive. In some systematic reviews, pooled results for sensitivity and specificity, include studies that were not relevant for this report hence for those systematic reviews the pooled results could not be used. Instead, for sensitivities and specificities, ranges of values from the individual studies were reported.

The studies included in the systematic reviews were conducted in various countries with majority of studies being conducted in African countries. There was one included study that was conducted in Canada. Hence the generalizability of the findings to the Canadian setting may not be possible.

Conclusions and Implications for Decision or Policy Making

A total of six systematic reviews were identified, these comprised, one systematic review within a health technology assessment (HTA) report[4] and five systematic reviews.[3,9-12]

In three systematic reviews[3,9,11] the sensitivities and specificities of Xpert were reported for the smear positive subgroups and the smear negative subgroups. In the smear positive subgroups, the sensitivities and specificities in the individual studies ranged from 75% to 100%, and from 67% to 100% respectively. In the smear negative subgroups, the sensitivities and specificities in the individual studies ranged from 22% to 100%, and from 57% to 100% respectively. Only one systematic review[11] reported that the sensitivity with Xpert was 36% to 44% higher than the sensitivity with smear microscopy.

Five systematic reviews[3,9-12] reported on the diagnostic accuracy of Xpert and/or Xpert Ultra compared with mycobacterial culture test. The sensitivities of Xpert ranged between 62% and 85%; and the specificities ranged between 98% and 99% (based on four systematic reviews[3,9,11,12]). The sensitivities of Xpert Ultra ranged between 64% and 100%; and specificities ranged between 96% and 100% (based on two systematic reviews[3,10]).

Six systematic reviews[3,4,9-12] reported on the diagnostic accuracy of Xpert or both Xpert and Xpert Ultra for the detection of rifampicin resistance compared to culture-based drug susceptibility testing. For rifampicin resistance detection with Xpert, the sensitivities ranged between 83% and 100%, and the specificities ranged between 97% and 100% (based on six systematic reviews[3,4,9-12]). For rifampicin resistance detection with Xpert Ultra the sensitivities ranged between 93% and 95%, and the specificities ranged between 98% and 99% (based on two systematic reviews[3,10]).

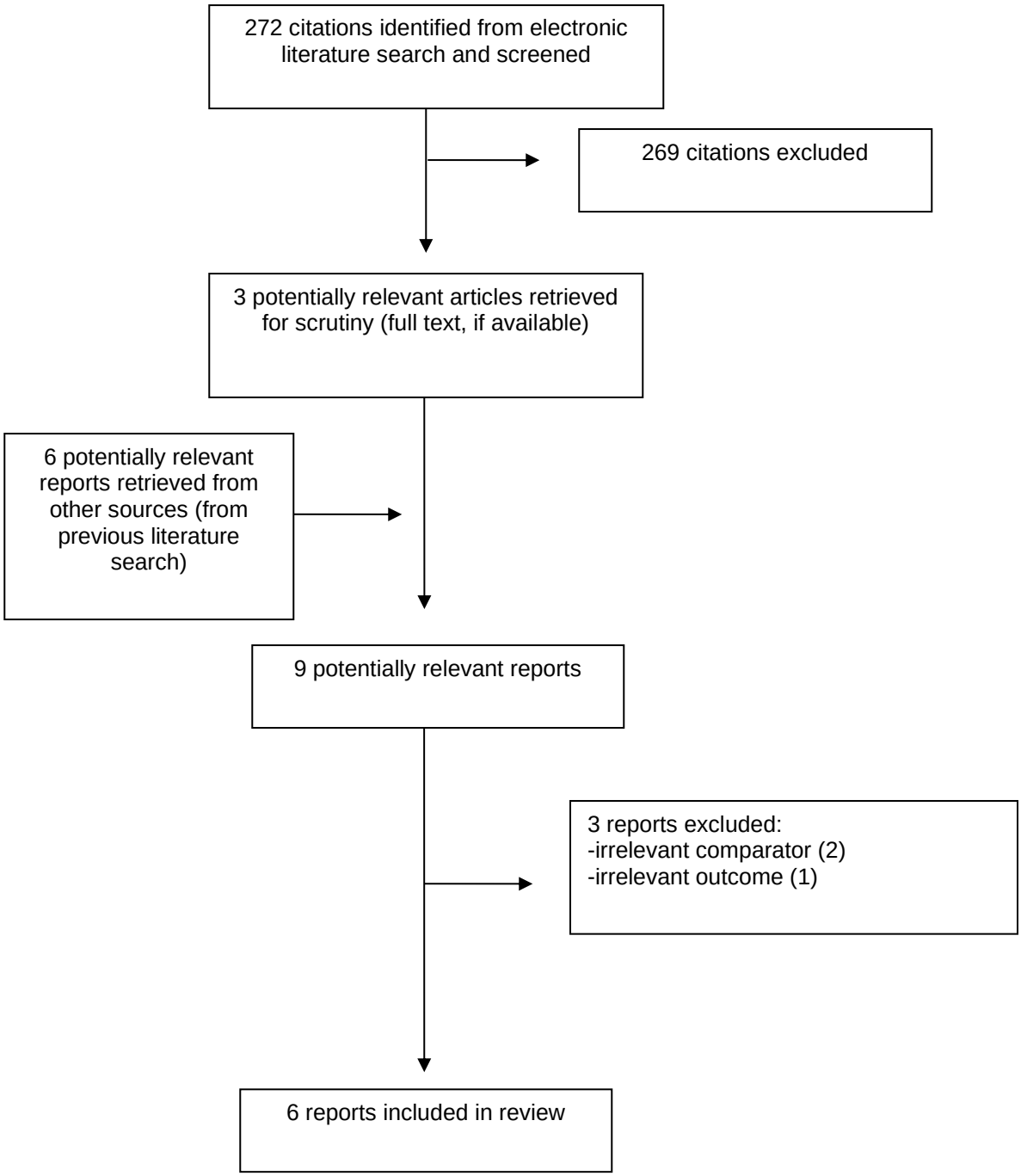
Findings need to be interpreted in the light of limitations. In some instances, ranges for sensitivities and specificities reported included pooled as well as individual study data (depending on the type of data available). Findings may be impacted by variations in specimen types, population subgroups, and settings. Generalizability to the Canadian setting is unclear. There was limited amount of evidence available for Xpert Ultra.

Factors such as acceptability, accessibility, and affordability need to be considered when implementing these diagnostic tests. Further research, to determine the level of accuracy of the test with respect to different specimen types, patient subgroups, and settings would be useful.

References

1. WHO. Global Tuberculosis Report: Executive Summary. 2020: https://www.who.int/docs/default-source/documents/tuberculosis/execsumm-11nov2020.pdf?sfvrsn=e1d925f_4 Accessed 17 Nov 2020.
2. Government of Canada. Tuberculosis: Monitoring. https://www.canada.ca/en/public-health/services/diseases/tuberculosis/surveillance.html Accessed 17 November 2020.
3. Horne DJ, Kohli M, Zifodya JS, et al. Xpert MTB/RIF and Xpert MTB/RIF Ultra for pulmonary tuberculosis and rifampicin resistance in adults. *The Cochrane database of systematic reviews.* 2019;6(6):Cd009593.
4. Drobniewski F, Cooke M, Jordan J, et al. Systematic review, meta-analysis and economic modelling of molecular diagnostic tests for antibiotic resistance in tuberculosis. *Health technology assessment (Winchester, England).* 2015;19(34):1-188, vii-viii.
5. CDC. A New Tool to Diagnose Tuberculosis: The Xpert MTB/RIF assay. https://www.cdc.gov/tb/publications/factsheets/pdf/xpertmtb-rifassayfactsheet_final.pdf Accessed 18 November 2020.
6. MayoClinic. Tuberculosis https://www.mayoclinic.org/diseases-conditions/tuberculosis/diagnosis-treatment/drc-203T51256 Accessed 28 October 2020.
7. Kumar D, Brett K, Argaez C. Rapid and Simultaneous Tuberculosis and Antibiotic Susceptibility Testing for the Diagnosis of Pulmonary Tuberculosis and Rifampicin Resistance: Diagnostic Accuracy and Cost-Effectiveness. *CADTH rapid response report: summary of abstracts*. Ottawa: CADTH; 2020: https://www.cadth.ca/sites/default/files/pdf/htis/2020/RB1499%20GeneXpert%20Final.pdf Accessed 26 October 2020.
8. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ.* 2017;358:j4008.
9. Kay AW, Gonzalez Fernandez L, Takwoingi Y, 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;8:CD013359.
10. 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 : IJID : official publication of the International Society for Infectious Diseases.* 2020;90:35-45.
11. Detjen AK, DiNardo AR, Leyden J, et al. Xpert MTB/RIF assay for the diagnosis of pulmonary tuberculosis in children: a systematic review and meta-analysis. *The Lancet Respiratory medicine.* 2015;3(6):451-461.
12. Wang XW, Pappoe F, Huang Y, et al. Xpert MTB/RIF Assay for Pulmonary Tuberculosis and Rifampicin Resistance in Children: a Meta-Analysis. *Clinical laboratory.* 2015;61(11):1775-1785.
13. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol.* 2009;62(10):e1-e34.

Appendix 1: Selection of Included Studies



Appendix 2: Characteristics of Included Publications

Table 2: Characteristics of Included Systematic Reviews

Study citation, country, funding source	Study designs and numbers of primary studies included	Population characteristics	Intervention and comparator(s)	Clinical outcomes, length of follow-up
Kay,[9] 2020, US. Funding: Received support from Liverpool School of Tropical Medicine, UK; Texas Children's Hospital, US; Thrasher Foundation, US; DFID, UK; and USAID, US.	Systematic review and meta-analysis. Number of studies: 28 studies (1 RCT, 14 cohort studies and 13 cross-sectional studies); conducted in Bangladesh (1), France (1), Gambia (1), India (2), Kenya (2), Malawi (1), S.Africa (10), Tanzania (3), Uganda (3), Uganda & Tanzania (1), Vietnam (1), Zambia (1), and Zimbabwe (1) Inclusion criteria: RCT, cohort studies, or cross-sectional studies; Children of age range: 0 to 14 years; target conditions: PTB, EPTB (lymph node TB or meningitis TB) or rifampicin resistance; index test: MTB/RIF Xpert or Xpert Ultra; and reference test: culture (solid or liquid medium) or composites reference standard (based on culture or on a clinical features). Exclusion criteria: Case-control studies or case-reports were excluded. Aim: to assess the diagnostic accuracy of	Children with presumptive PTB. N = 6812 (in 25 studies for Xpert MTB/RIF) N = 697 (in 3 studies for Xpert Ultra) The number of participants ranged between 50 and 1442in the individual included studies. Age (months): Median ages ranged from 10m to 127 m (21 studies), mean ages from 12m to 88m (3 studies), up to 180m (1 study), 20% less than 60m and 80% older than 60m (1 study), and NR (1 study). HIV positive status: 0% (2 studies), 2% to 54% (21 studies), and NR (5 studies)	Index test: Xpert MTB/RIF or Ultra Reference standard : culture	Sensitivity, and specificity

Study citation, country, funding source	Study designs and numbers of primary studies included	Population characteristics	Intervention and comparator(s)	Clinical outcomes, length of follow-up
	Xpert and Xpert Ultra for diagnosing (1) PTB, (2) TB meningitis, and (3) lymph node TB in children presumed to have TB *Note:* This systematic review included 49 studies. Of these studies, 28 studies were relevant for this report and are described here.			
Zhang,[10] 2020, China. Funding: from National Natural Science Foundation of China.	Systematic review and meta-analysis. Number of studies = 8 (retrospective [3], and prospective [5]; blinded [4], not blinded [2], and NR [2]) Studies were conducted in Tanzania (1), Switzerland (1), France (2), South Africa (2), and multinational (2) Inclusion: Studies including adults and/or children, diagnosis of TB (PTB or EPTB) Exclusion: Animal studies, reviews, correspondences, commentaries, interim analyses, case reports, and editorial 19.5 to 62s. Aim: to assess the diagnostic accuracy of Xpert Ultra for the detection of TB in adults and children	Adults and/or children with presumptive PTB. Number of participants ranged between 33 and 1753 in the individual included studies Age: adults (6 studies) and children (2 studies. HIV positive status: 0% to 62% (5 studies), NR (3 studies)	Index test: Xpert Ultra Reference standard: culture Culture: MGIT and LJ, or MGIT	Sensitivity, and specificity

Study citation, country, funding source	Study designs and numbers of primary studies included	Population characteristics	Intervention and comparator(s)	Clinical outcomes, length of follow-up
	Note: This systematic review included 16 studies. Of these studies, 8 studies were relevant for this report and are described here.			
Horne,[3] 2019, US. Funding: Received support from Liverpool; DFID, UK; and USAID, US.	Systematic review and meta-analyses. Number of studies: 95 (2 RCTS and 93 cross-sectional studies). Studies were conducted in Botswana, Cambodia, Canada, Chile, Cote d'Ivorie, France, Ghana, Japan, Kenya, Lithuania, Nepal, New Zealand, Peru, Poland,Rwanda, Spain, Sudan, Turkey, and Vietnam (1 each); Egypt, Kyrgyzstan, Malawi, Russia, Singapore, and Zambia (2 each); India, Tanzania, Thailand, and Zimbabwe (3 each); Ethiopia (4); China, and US (5 each); Uganda (6); Republic of Korea, and multi-national (7 each). Inclusion criteria: RCT, cohort studies, or cross-sectional studies; adults (≥ 15 years) with presumptive PTB, rifampicin-resistant TB, or MDRTB; Exclusion criteria: Case-control studies	Adults for the diagnosis of PTB N = 42,091 (for the individual studies numbers of participants ranged between 58 and 6648) Age (years): median values within 25 to 64 (53 studies), mean values within 33 to 65 (20 studies), range values within 15 to 83 (6 studies), adult or ≥ 18 (8 studies), ≥ 15 (6 studies), ≥14 (1 study), and majority between 16 to 30 (1study) HIV positive status: 0% (4 studies), 0.1% to 1% (7 studies), 2% to 100% (55 studies), and NR (29 studies)	Index test: Xpert MTB/RIF Reference standard: culture (95 studies). Index test: Xpert Ultra Reference standard: culture	Sensitivity, and specificity

Study citation, country, funding source	Study designs and numbers of primary studies included	Population characteristics	Intervention and comparator(s)	Clinical outcomes, length of follow-up
	or case-reports were excluded. Aim: to assess the diagnostic accuracy of Xpert and Xpert Ultra in detecting PTB and RIF resistance in adults			
Detjen,[11] 2015, US Funding: WHO, Global TBProgram of Texas Children's Hospital	Systematic review and meta-analyses. Number of studies: 15 (study type not reported). Studies were conducted in Zambia, Spain, Bangladesh, Malawi, Vietnam, China, Tanzania, Uganda, Italy, multinational (1 each), and S.Africa (5). Inclusion criteria: RCTs, cohort studies and cross-sectional studies; published, in press, or unpublished studies; children of age 0 to 15 years with presumed PTB. Exclusion criteria: case-control studies, case reports, and studies only available as abstracts. Aim: To assess the sensitivity and specificity of Xpert MTB/RIFassay compared with microscopy in the diagnosis of PTB in children	Children with presumed PTB. Number of participants ranged between 20 and 930 in the individual included studies. Age (months): median values within 13 to 106 (9 studies), mean values within 13 to 91 (5 studies), and NR (1 study). HIV positive status: 0% (1 study), 10% to 54% (10 studies), x10% to x% (x studies), and NR (4 studies)	1) Index test: Xpert MTB/RIF, Reference standard: culture. 2) Index test: smear microscopy, Reference standard: culture	Sensitivity, and specificity

Study citation, country, funding source	Study designs and numbers of primary studies included	Population characteristics	Intervention and comparator(s)	Clinical outcomes, length of follow-up
Drobniewski,[4] 2015, UK Funding: provided by the Health Technology program of the National Institute for Health Research	Health technology assessment which includes a section on a systematic review and meta-analyses. Number of studies: 6 (study type: NR) Studies were conducted in Republic of Korea, Russia, S.Africa, and Zambia (1 study each) and 2 studies were multinational. Aim: To assess the accuracy of molecular tests used to detect drug resistance in MTB *Note:* This systematic review included 56 studies. Of these studies, 6 studies were relevant for this report and are described here.	Participants with presumptivef PTB N = 9372 (for the individual studies numbers of participants ranged between 62 and 6648) Age: adults (≥16 years) for 4 studies and NR for 2 studies. HIV status: HIV negative (1 study), both HIV positive and HIV negative (3 studies, and NR (2 studies	Index test: GeneXpert Reference standard: culture	Sensitivity, specificity, and diagnostic odd's ratio.
Wang,[12] 2015, China Funding: National Natural Science Foundation of China	Systematic review and meta-analyses. Number of studies: 11 (2 prospective and 9 cross-sectional; participants were consecutively or randomly enrolled) Studies were conducted in Vietnam, Tanzania, Uganda, Zambia, multinational (1 each), China (2) and S.Africa (4). Inclusion criteria: prospective or cross-sectional studies	Children with presumptive PTB Number of included specimens ranged between 73 and 930 for the individual included studies Age: NR HIV status: 5 of the 11 studies included both HIV positive and HIV negative participants; NR for the remaining 6 studies	Index test: GeneXpert Reference standard: culture (in 3 studies it was culture plus clinical TB) Accuracy of Xpert MIB/RIF for detecting RIF resistance was determined as compared to phenotypic drug susceptibility testing	Sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, and diagnostic odd's ratio.

Study citation, country, funding source	Study designs and numbers of primary studies included	Population characteristics	Intervention and comparator(s)	Clinical outcomes, length of follow-up
	Exclusion criteria: case-control studies and studies with data available from abstracts only . Aim: to assess accuracy of Xpert MTB/RIF assay for diagnosing PTB and RIF-resistance in children			

DFID = Department for International Development; EPTB = extrapulmonary tuberculosis; LJ = Lowenstein-Jensen culture; m = month; MDR-TB = multi-drug resistant TB; MGIT = Mycobacteria Growth Indicator Tube; MTB = mycobacterium tuberculosis; MTB/RIF = mycobacterium tuberculosis complex and resistance to rifampicin; PTB = pulmonary tuberculosis; RCT = randomized controlled trial; RIF = rifampicin; TB = tuberculosis; USAID = United States Agency for International Development.

Appendix 3: Critical Appraisal of Included Publications

Table 3: Strengths and Limitations of Systematic Reviews Using AMSTAR 2[8]

Strengths	Limitations
Kay,[9] 2020, US	
• The objective was clearly stated • Study selection was described, and a flow chart was presented • Multiple databases searched up to 29 April 2019 (Medline [from 1966], EMBASE [from 1974], CINAHL [from 1982], Scopus [from 1970], Science Citation Index [from 1900]), Clinical trials registries were searched Also, reference list of irrelevant reviews and studies were searched. Researchers and experts in the field were also contacted. • Study selection was described, and a flow chart was presented • A list of included studies was provided • A list of excluded studies was provided • Article selection was done by two reviewers • Data extraction was done by two reviewers • Quality assessment was done by two reviewers using QUADAS 2. Certainty of evidence was assessed using GRADE. For Xpert MTB/RIF certainty is moderate; for Xpert Ultra the certainty is variable (low to moderate) • Characteristics of the included studies were presented • Meta-analysis was conducted and appeared to be appropriate. • Conflicts of interest were declared and for most authors did not appear to be of concern, except one author had received Xpert MTB/RIF Ultra cartridges at a discounted price from the manufacturer.	• Publication bias was not formally examined. However, the authors reported that publication bias was not of serious concern as the literature search was comprehensive and additionally there was extensive outreach to TB researchers to identify studies
Zhang,[10] 2020, China	
• The objective was clearly stated • Study selection was described, and a flow chart was presented • Multiple databases searched from inception to 20 May 2019 (Medline, EMBASE, Web of Science, Scopus, and Cochrane Central Register of Controlled Trials. Manual search of references cited in selected articles and systematic reviews. No language restrictions • Study selection was described, and a flow chart was presented • A list of included studies was provided • Article selection was done by two reviewers • Data extraction was done by two reviewers • Quality assessment was done using QUADAS 2. Risk of bias in terms of flow and timing; and patient selection was high or unclear. There appeared to be no concerns in terms of applicability of the assay • Characteristics of the included studies were presented • Publication bias was examined using the Funnel plot and the possibility of publication bias appeared to be low.	• A list of excluded studies was not provided • Unclear if quality assessments were done in duplicate • As pooled results from meta-analyses included studies relevant for this report as well as studies not relevant for this report, the pooled results could not be presented. Instead results of individual studies were presented.

Strengths	Limitations
• Meta-analysis was conducted and appeared to be appropriate. • The authors reported that there were no conflicts of interest.	
Horne,[3] 2019, US.	
• The objective was clearly stated • Study selection was described, and a flow chart was presented • Multiple databases searched up to 11 October 2018 (Medline, EMBASE, Science Citation Index, Web of Science, Latin American Caribbean Health Sciences Literature). Clinical trials registries were searched. No language restrictions. • Study selection was described, and a flow chart was presented • A list of included studies was provided • A list of excluded studies was provided • Article selection was done by two reviewers • Data extraction was done by two reviewers • Quality assessment was done by two reviewers using QUADAS 2. Certainty of evidence was assessed using GRADE; the authors reported high certainty for Xpert MTB/RIF and moderate certainty for Xpert Ultra. • Characteristics of the included studies were presented • Meta-analysis was conducted and appeared to be appropriate. In some instances a Bayesian approach was used and credible intervals were reported. • Conflicts of interest were declared. Most authors had no conflicts of interest. One of the authors had collaborations with several organizations including the manufacture. Also the manufacturer provided the test at a preferential price.	• Publication bias was not formally examined using techniques such as Funnel plots or regression tests. The authors reported that such techniques were not useful in the case of diagnostic test accuracy studies. Further, since Xpert MTB/RIF and Xpert Ultra are produced by one manufacturer and there has been considerable attention and scrutiny, the authors believe reporting bias is likely minimal.
Detjen,[11] 2015, US	
• The objective was clearly stated • Study selection was described, and a flow chart was presented • Multiple databases searched up to 6 January 2015 (Medline, Scopus). Reference list of included studies and review articles were also searched for additional studies. Study authors and researchers were also contacted. There were no language restrictions. • Study selection was described, and a flow chart was presented • A list of included studies was provided • Article selection was done by two reviewers • Data extraction was done by two reviewers • Quality assessment was conducted using QUADAS 2. The authors reported that most studies seemed to have a low risk of bias for patient selection as participants were recruited consecutively. Further details of the quality assessment were not presented. • Characteristics of the included studies were presented	• A list of excluded studies was not provided • Unclear if quality assessment was done in duplicate • The authors mentioned that publication bias was not assessed as the methods were not applicable for studies of diagnostic accuracy.

Strengths	Limitations
• Meta-analysis was conducted and appeared to be appropriate. • The authors reported that there were no conflicts of interest	
Drobniewski,[4] 2015, UK	
• The objective was clearly stated • Study selection was described, and a flow chart was presented • Multiple databases searched from 1 January 2000 to 15 Aug 2013 (Medline, PubMed, EMBASE, BIOSIS, Web of Science, and grey literature [SIGLE]) • Study selection was described, and a flow chart was presented • A list of included studies was provided • Article selection was done by two reviewers • Data extraction was done by two reviewers • Quality assessment was done by two reviewers using QUADAS 2. Risk of bias in terms of flow and timing; index test; and reference standard were generally unclear. There appeared to be no concerns in terms of applicability of the assay • Characteristics of the included studies were presented • Meta-analysis was conducted and appeared to be appropriate. In some instances a Bayesian approach was used and credible intervals were reported. • The authors reported that there was no statistically significant (P <0.05) evidence of publication bias as determined by Egger's test and Begg's test • Conflicts of interest were declared and seemed less likely to be an issue	• A list of excluded studies was not provided
Wang,[12] 2015, China	
• The objective was clearly stated • Study selection was described, and a flow chart was presented • Multiple databases searched up to 28 October 2014 (PubMed and Science Direct). Authors of studies were also contacted. • Study selection was described, and a flow chart was presented • A list of included studies was provided • Article selection was done by two reviewers • Data extraction was done by two reviewers • Quality assessment was done using QUADAS 2. Risk of bias (in terms of patient selection, index test, reference standard, and flow and timing) was generally low. Applicability concerns. in terms of patient selection was generally low but applicability concerns in terms of index test and reference standard were variable • Characteristics of the included studies were presented • Meta-analysis was conducted and appeared to be appropriate.	• A list of excluded studies was not provided • It was unclear if quality assessment was done in duplicate

Strengths	Limitations
• The authors reported that publication bias was explored using Funnel plots, Egger's tests, and Begg's test and was less likely to be of concern; results were not presented. • The authors reported that there were no conflicts of interest.	

AMSTAR 2 = A MeaSurement Tool to Assess systematic Reviews 2; GRADE = Grading of Recommendations, Assessment, Development and Evaluations; MTB/RIF = mycobacterium tuberculosis complex and resistance to rifampicin; QUADAS 2 = Quality Assessment of Diagnostic Accuracy Studies 2

Appendix 4: Main Study Findings and Authors' Conclusions

Table 4: Summary of Findings Included Systematic Reviews

Main study findings	Authors' conclusion
Kay,[9] 2020, US	
Xpert MTB/RIF sensitivity and specificity for pulmonary tuberculosis in children using culture as reference standard (23 studies, 6612 participants). *Sputum specimens (23 studies, 6703 participants)* Pooled Sensitivity (95% CI): 64.6% (55.3% to 72.9%). Pooled Specificity (95% CI): 99.0%,(98.1% to 99.5%) PPV (95% CI): 88.2% (79.6% to 93.5%) NPV (95% CI): 96.2 (95.1 to 97.0) *Gastric aspirate specimens (14 studies, 3482 participants)* Pooled Sensitivity: 73.0%; 95% CI, 52.9% to 86.7%. Pooled Specificity: 98.1%; 95% CI, 95.5% to 99.2%. PPV (95% CI): 81.0 (65.5 to 90.6) NPV (95% CI): 97.0 (94.5 to 98.4) *Stool specimen (11 studies, 1592 participants)* Pooled Sensitivity: 61.5%; 95% CI, 44.1% to 76.4%. Pooled Specificity: 98.5%; 95% CI, 97.0% to 99.2%. PPV (95% CI): 81.7 (72.2 to 88.5) NPV (95% CI): 95.8 (93.8 to 97.3) *Nasopharyngeal specimens (4 studies, 1125 participants)* Pooled Sensitivity: 45.7%; 95% CI, 27.6% to 65.1%. Pooled Specificity: 99.6%; 95% CI, 98.9% to 99.8%. PPV (95% CI): 92.6 (81.1 to 97.3) NPV (95% CI): 94.3 (92.0 to 95.9) **Xpert MTB/RIF sensitivity and specificity for pulmonary tuberculosis in children using culture as reference standard in various subgroups.** *Sputum specimens, smear-positives (11 studies, 91 participants)* Sensitivity %, (95% CI): 97.8 (91.6 to 99.4). Specificity %, (95% CI): 67 to 100 (lower bound: 3 to 2, upper bound: 98 to100), (estimable in 4 of the 11 studies) PPV (95% CI): NR NPV (95% CI): NR (Note: The authors conducted a univariate meta-analysis for this analysis group because in many studies, few or zero false-positive and true-negative values were reported). *Sputum specimens, smear-negatives (112 studies, 3118 participants)* Sensitivity %, (95% CI): 58.9 (45.6 to 71.0). Specificity %, (95% CI): 99.1 (97.1 to 99.7) PPV (95% CI): 88.4 (68.8 to 96.3) NPV (95% CI): 95.6 (94.0 to 96.8). *Sputum specimens, HIV-positive (10 studies, 642 participants)* Sensitivity %, (95% CI): 72.2 (59.9 to 81.8). Specificity %, (95% CI): 99.4 (97.2 to 99.9). PPV (95% CI): 93.2 (74.0 to 98.5) NPV (95% CI): 97.0 (95.5 to 97.9) *Sputum specimens, HIV-negative (12 studies, 2784 participants)*	"We found Xpert MTB/RIF sensitivity to vary by specimen type, with gastric aspirate specimens having the highest sensitivity followed by sputum and stool, and nasopharyngeal specimens the lowest; specificity in all specimens was > 98%. [….]. Xpert MTB/RIF was accurate for detection of rifampicin resistance." (p. 3)[9]

Main study findings	Authors' conclusion
Sensitivity %, (95% CI): 54.3 (43.5 to 64.7). Specificity %, (95% CI): 99.3 (98.1 to 99.7). PPV (95% CI): 89.7 (80.5 to 94.9) NPV (95% CI): 95.1 (93.9 to 96.2) *Gastric aspirate specimens, HIV-positive (3 studies, 634 participants) s* Sensitivity %, (95% CI): 73.3 (54.9 to 86.1). Specificity %, (95% CI): 98.5 (97.1 to 99.2). PPV (95% CI): 84.1 (72.7 to 91.3) NPV (95% CI): 97.1 (93.8 to 98.4) *Stool specimens, HIV-positive (4 studies, 526 participants)* Sensitivity %, (95% CI): 69.8 (56.3 to 80.6). Specificity %, (95% CI): 98.6 (96.1 to 99.5). PPV (95% CI): 84.7 (66.2 to 94.0) NPV (95% CI): 98.6 (96.1 to 99.5) **Xpert Ultra sensitivity and specificity for pulmonary tuberculosis in children using culture as reference standard.** *Sputum specimens (3 studies, 697 participants)* Pooled Sensitivity: 72.8%; 95% CI, 64.7% to 79.6%. Pooled Specificity: 97.5%; 95% CI, 95.8% to 98.5%. PPV (95% CI): 76.4 (65.6 to 84.6) NPV (95% CI): 97.7 (95.9 to 97.7) *Gastric aspirate specimens: No studies identified*. *Stool specimens: No studies identified.* *Nasopharyngeal specimens (1 study, 195 participants)* Sensitivity: 45.7%; 95% CI, 28.9% to 63.3%. Specificity: 97.5%; 95% CI, 93.7% to 99.3% PPV (95% CI): 67.0 (42.0 to 85.1) NPV (95% CI): 94.1 (92.2 to 95.6) **Xpert MTB/RIF sensitivity and specificity for rifampicin resistance pulmonary tuberculosis in children using culture as reference standard.** (Results from 3 individual studies involving a total of 180 participants) Sensitivities (95% CI): 1.00 (0.59 to 1.00); 0.83 (0.36 to 1.00); and 1.00 (0.16 to 1.00). Specificities (95% CI): 1.00 (0.72 to 1.00); 0.99 (0.95 to 1.00); and 0.97 (0.87 to 1.00) Pooled results were not presented.	
Zhang,[10] 2020, China	
Diagnosis of PTB in adults, using Xpert Ultra. The sensitivities varied between 0.82 and 1.00 with the lower bound of 95% CI varying between 0.65 and 0.93 and the upper bound of 95% CI varying between 0.91 and 1.00; (from 6 studies). The specificities varied between 0.96 and 1.00, with the lower bound of 95% CI varying between 0.88 and 0.94 and the upper bound of 95% CI varying between 0.97 to 1.00 (from 5 studies and not estimable in 1 study). **Diagnosis of PTB in children, using Xpert Ultra.**	"As a rapid and highly sensitive test for the detection of TB and simultaneous detection of RIF resistance, Xpert Ultra exhibits a viable alternative in sensitivities in both pulmonary TB (PTB) and extrapulmonary TB (EPTB), which was proved to be higher than Xpert in the comparative analysis, and also shows a good

Main study findings	Authors' conclusion
The sensitivities (95% CI) were 0.75 (0.64 to 0.85 in one study, and 0.64 (0.44 to 0.81) in the second study. The specificities (95% CI) were 0.97 (0.94 to 0.99) in one study and 1.00 (0.97 to 1.00) in the second study. **Diagnosis of RIF resistance in adults, using Xpert Ultra or Xpert** *Xpert Ultra:* The sensitivities (95% CI) were 0.93 (0.80 to 0.98) in one study, and 0.95 (0.90 to 0.98) in the second study. The specificities (95% CI) were 0.99 (0.94 to 1.00) in one study and 0.98 (0.97 to 0.99) in the second study. *Xpert:* The sensitivities (95% CI) were 0.93 (0.80 to 0.98) in one study, and 0.95 (0.0.91 to 0.98) in the second study. The specificities (95% CI) were 0.98 (0.93 to 1.00) in one study and 0.98 (0.96 to 0.99) in the second study.	performance in the detection of RIF resistance" (p.35).[10]
Horne,[3] 2019, US.	
Xpert MTB/RIF sensitivity and specificity for pulmonary tuberculosis in adults using culture as reference standard (70 studies, 37,237 participants). Median pooled sensitivity (95% CrI): 85% (82% to 88%). Median pooled specificity (95% CrI): 98% (97% to 98%) **Subgroups: Xpert MTB/RIF sensitivity and specificity for pulmonary tuberculosis in adults using culture as reference standard.** *Smear-positive (45 studies, 4064 participants)* Median pooled sensitivity (95% CrI): 98% (97% to 98%) Median pooled specificity (95% CrI): could not be determined *Smear-negative (45 studies; 18,962 participants)* Median pooled sensitivity (95% CrI): 67% (62% to 72%) Median pooled specificity (95% CrI): 98% (98% to 99%) *HIV-positive (14 studies; 4,664 participants)* Median pooled sensitivity (95% CrI): 81% (75% to 86%) Median pooled specificity (95% CrI): 98% (97% to 99%) *HIV-negative (14 studies; 3,866 participants)* Median pooled sensitivity (95% CrI): 88% (83% to 92%) Median pooled specificity (95% CrI): 98% (97% to 99%) **Xpert Ultra sensitivity and specificity for pulmonary tuberculosis in adults using culture as reference standard** (1 study, 1439 participants) Sensitivity (95% CrI): 95% (90% to 98%) Specificity (95% CrI): 98% (97% to 99%) **Rifamine resistance: Xpert MTB/RIF sensitivity and specificity for rifampicin resistance pulmonary tuberculosis in adults using culture as reference standard** (48 studies, 8020 participants). Median pooled sensitivity (95% CrI): 96% (94% to 97%) Median pooled specificity (95% CrI): 98% (98% to 99%)	"We found Xpert MTB/RIF to be sensitive and specific for diagnosing PTB and rifampicin resistance, consistent with findings reported previously. Xpert MTB/RIF was more sensitive for tuberculosis in smear-positive than smear-negative participants and HIV-negative than HIV positive participants. Compared with Xpert MTB/RIF, Xpert Ultra had higher sensitivity and lower specificity for tuberculosis and similar sensitivity and specificity for rifampicin resistance (1 study). Xpert MTB/RIF and Xpert Ultra provide accurate results and can allow rapid initiation of treatment for multidrug-resistant tuberculosis." (p. 2)[3].

Main study findings	Authors' conclusion
Rifamine resistance: Xpert Ultra sensitivity and specificity for rifampicin resistance pulmonary tuberculosis in adults using culture as reference standard (1 study, 551 specimens). Sensitivity (95% CI): 95% (90% to 98%) Specificity (95% CI): 98% (97% to 99%) (Note: For this same study, the sensitivity (95% CI) and specificity (95% CI) for Xpert MTB/RIF were 95% (91% to 98%), and 98% (96% to 99%) Note: The authors used a Bayesian approach. They reported that they used a sample from the posterior distribution to obtain various descriptive statistics of interest. They estimated the median pooled sensitivity and specificity and their 95% CrI; considering that the posterior distribution of some parameters may be skewed so median values would be a better point estimate of the parameter than the mean in such cases	
Detjen,[11] 2015, US	
Sensitivity and specificity of Xpert (against culture reference standard) to detect PTB in children using various types of specimen *Sputum expectorated and/or induced (12 studies)* Sensitivity (95% CI): ranged between 0.25 and 1.00, with lower bound of the 95% CI ranging between 0.16 and 0.64, with the upper bound of the 95% CI ranging between 0.63 and 1.00 Specificity (95% CI): ranged between 0.93 and 1.00, and the lower and upper bounds of the 95% CI ranging between 0.81 to 0.98 and between 0.96 and 1.00, respectively. *Gastric fluid (7 studies)* Sensitivity (95% CI): ranged between 0.40 and 1.00, with lower bound of the 95% CI ranging between 0.12 and 0.54, and the upper bound of the 95% CI ranging between 0.68 and 1.00 Specificity (95% CI): ranged between 0.93 and 1.00,with the lower and upper bounds of the 95% CI ranging between 0.79 to 0.98, and between 0.98 and 1.00. *Nasopharyngeal aspirate (2 studies)* Sensitivity (95% CI): 0.44 (0.33 to 0.55), and 0.30 (0.15 to 0.43). Sensitivity (95% CI): 1.00 (0.99 to 1.00), and 0.99 (0.97 to 1.00) **Sensitivity and specificity of Xpert (against culture reference standard) to detect PTB in children using various types of specimen and by smear status** *In smear positive children using sputum expectorated and/or induced (11 studies)* Sensitivity (95% CI): ranged between 0.92 and 1.00, with lower bound of the 95% CI ranging between 0.03 and 0.85, and the upper bound of the 95% CI being 1.00; (Of the 11 studies, sensitivity was not estimable in 3 studies). Specificity (95% CI): 0.86 (0.42 to 1.00); (Of the 11 studies, specificity was not estimable in 10 studies). *In smear positive children using gastric fluid (7 studies)* Sensitivity (95% CI): ranged between 0.92 and 1.00, with lower bound of the 95% CI ranging between 0.03 and 0.78, and the upper bound of the 95% CI being 1.00; (Of the 7 studies sensitivity was not estimable in 1 study). Specificity (95% CI): 0.96 (0.78 to 1.00); (Of the 7 studies specificity was not estimable in 6 studies). *In smear negative children using sputum expectorated and/or induced (11 studies)* Sensitivity (95% CI): ranged between 0.25 and 1.00, with lower bound of the 95% CI ranging between 0.03 and 0.46, and the upper bound of the 95% CI ranging between 0.58 to 1.00. Specificity (95% CI): ranged between 0.93 and 1.00, with lower bound of the 95% CI ranging between 0.81 and 0.98, and the upper bound of the 95% CI ranging between 0.99 to 1.00	"Compared with microscopy, Xpert offers better sensitivity for the diagnosis of pulmonary tuberculosis in children and its scale-up will improve access to tuberculosis diagnostics for children. Although Xpert helps to provide rapid confirmation of disease, its sensitivity remains suboptimum compared with culture tests. A negative Xpert result does not rule out tuberculosis. Good clinical acumen is still needed to decide when to start antituberculosis therapy and continued research for better diagnostics is crucial." (p. 2)[11]

Main study findings	Authors' conclusion
In smear negative children using gastric fluid (8 studies) Sensitivity (95% CI): ranged between 0.40 and 1.00, with lower bound of the 95% CI ranging between 0.05 and 0.43, and the upper bound of the 95% CI ranging between 0.65 and 1.00. Specificity (95% CI): ranged between 0.70 and 1.00, with lower bound of the 95% CI ranging between 0.63 and 0.99, and the upper bound of the 95% CI ranging between 0.76 and 1.00. **Pooled median sensitivity and specificity of Xpert (against culture reference standard) to detect PTB in children using various types of specimen** *Expectorated induced sputum (12 studies, 2380 children):* Pooled median sensitivity (95% CrI): 62% (51 to 73). Pooled median specificity (95% CrI): 98% (97 to 99). *Gastric fluid (7 studies, 1319 children)* Pooled median sensitivity (95% CrI): 66% (51 to 81). Pooled median specificity (95% CrI): 98% (96 to 99). **Pooled median sensitivity and specificity of smear microscopy (against culture reference standard) to detect PTB in children using various types of specimen** *Expectorated induced sputum (12 studies, 2380 children):* Pooled median sensitivity (95% CrI): 26% (14 to 39). Pooled median specificity (95% CrI): 100% (99 to 100). *Gastric fluid (7 studies, 1319 children)* Pooled median sensitivity (95% CrI): 22% (12 to 35). Pooled median specificity (95% CrI): 99% (97 to 100). **Sensitivity for Xpert was 36% to 44% higher than that for smear microscopy** **Pooled median sensitivity of Xpert (against culture reference standard) to detect PTB in children using expectorated or induced sputum by smear and HIV status** *Smear-positive status, and HIV positive (6 studies, 25 children)* Pooled median sensitivity (95% CrI): 97% (87 to 100), *Smear-positive status, and HIV negative (7 studies, 41 children)* Pooled median sensitivity (95% CrI): 94% (83 to 99). *Smear-negative status, and HIV positive (7 studies, 36 children)* Pooled median sensitivity (95% CrI): 60% (40 to 77). *Smear-negative status, and HIV negative (7 studies, 125 children)* Pooled median sensitivity (95% CrI): 44% (30 to 59). **Rifamin-resistance (3 studies, 176 children)** Pooled sensitivity (95% CrI): 86% (53 to 98). Pooled specificity (95% CrI): 98% (94 to 100) Note: The authors used a Bayesian approach. They estimated the median pooled sensitivity and specificity and their 95% CrI. The reason for reporting point estimates as median values and not mean values was not presented.	
Drobniewski,[4] 2015, UK	
Rifamine resistance: GeneXpert sensitivity and specificity for rifampicin resistance pulmonary tuberculosis	"Rapid molecular tests such as the manual line probe assays (LPAs) and automated

Main study findings	Authors' conclusion
(in adults [4 studies], NR [2 studies]) using culture as reference standard (6 studies, 9372 participants) Pooled sensitivity (95% CI): 96.8% (94.2% to 99.4%) Pooled specificity (95% CI): 98.4% (97.8% to 99.0%) Diagnostic odd's ratio (95% CI): 1209 (446 to 3276), P <0.05	GeneXpert are able to identify rifampicin resistance (and isoniazid resistance for some LPAs) with promising levels of specificity and are almost as sensitive as microbiological culture, but produce results more quickly (within 1 day of the sample being obtained)." (p. xxviii)[4] Note: LPA and isoniazid resistance were not relevant for this report and hence are not described here
Wang,[12] 2015, China	
Pooled results for Xpert MTB/RIF for pulmonary tuberculosis in children using culture as reference standard (11 studies; specimen types: sputum, respiratory, nasopharyngeal, gastric aspirate, BALF [not explained], and body fluid) Pooled sensitivity (95% CI): 65% (61% to 69%), heterogeneity, I^2 = 78.8%. Pooled specificity (95% CI): 99% (98.0% to 99%), heterogeneity, I^2 = 46.9%. Pooled positive likelihood ratio (95% CI): 43.89 (30.31 to 63.55), heterogeneity, I^2 =.27.3% Pooled negative likelihood ratio (95% CI): 0.31 (0.24 to 0.41), heterogeneity, I^2 =.81.4 Pooled diagnostic odd's ratio (95% CI): 164.99 (111.89 to 240.64), heterogeneity, I^2 =.1.4%. **Pooled results for Xpert MTB/RIF for pulmonary tuberculosis in children using culture as reference standard in various subgroups.** *HIV positive (5 studies; specimen types: sputum, nasopharyngeal, and gastric aspirate)* Pooled sensitivity (95% CI): 84% (74% to 92%); heterogeneity, I^2 = 62.2%. Pooled specificity (95% CI): 99% (97% to 99%); heterogeneity, I^2 = 67.2%. *HIV negative (5 studies; specimen types: sputum, nasopharyngeal, and gastric aspirate)* Pooled sensitivity (95% CI): 65% (59% to 70%); heterogeneity, I^2 =45.3%. Pooled specificity (95% CI): 98% (97% to 99%); heterogeneity, I^2 = 85.5%. **Pooled results for Xpert MTB/RIF for detecting RIF resistance in children** (5 studies, specimen types: sputum, BALF [not explained], and respiratory) Pooled sensitivity (95% CI): 94% (73% to 100%); heterogeneity, I^2 = 0% Pooled specificity (95% CI): 99% (98% to 100%), heterogeneity, I^2 = 74.0%	"The Xpert MTB/RIF is sensitive and specific for diagnosing paediatric pulmonary TB. It is also effective in detecting rifampicin resistance. It can, therefore, be used as an initial diagnostic tool." (p. 1775)[12]

CI = confidence interval; CrI = credible interval; MTB/RIF = mycobacterium tuberculosis complex and resistance to rifampicin; NPV = negative predictive value; NR = not reported; PPV = positive predictive value; PTB = pulmonary tuberculosis; RIF = rifampicin; TB = tuberculosis.