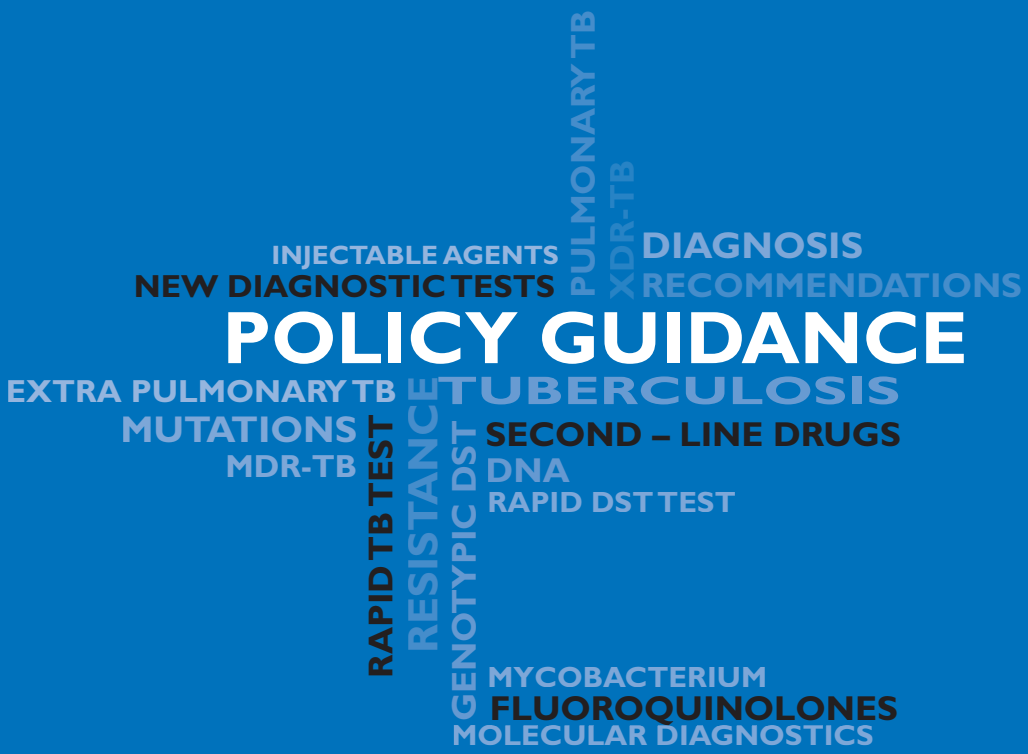
The use of molecular line probe assays for the detection of resistance to second-line anti-tuberculosis drugs



THE END TB STRATEGY

World Health Organization

The use of molecular line probe assays for the detection of resistance to second-line anti-tuberculosis drugs

Policy guidance

WHO Library Cataloguing-in-Publication Data

The use of molecular line probe assays for the detection of resistance to second-line anti-tuberculosis drugs. Policy guidance.

I.World Health Organization.

ISBN 978 92 4 151056 1

Subject headings are available from WHO institutional repository

Designed by GPS Publishing

Printed in Spain

WHO/HTM/TB/2016.07

Contents

Tables

Figures

Abbreviations

AFB	acid-fast bacilli
CI	confidence interval
CRS	composite reference standard
DOI	Declaration of Interests
DST	drug-susceptibility testing
GRADE	Grading of Recommendations Assessment, Development and Evaluation
GDG	Guideline Development Group
HIV	human immunodeficiency virus
MDR-TB	multidrug-resistant tuberculosis
NAAT	nucleic acid amplification test
PCR	polymerase chain reaction
PICO	Population, Intervention, Comparator, Outcome
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
RR-TB	rifampicin-resistant tuberculosis
SLID	Second-line injectable drug
SL-LPA	Second-line lineprobe assays
TB	tuberculosis
USAID	United States Agency for International Development
WHO	World Health Organization
XDR-TB	Extensively drug-resistant tuberculosis

Acknowledgements

This document was prepared by Christopher Gilpin and Alexei Korobitsyn, with input from Karin Weyer, Fuad Mirzayev, Dennis Falzon, Ernesto Jaramillo, Wayne Van Gemert and Matteo Zignol (WHO Global TB Programme), on the basis of consensus agreed at a Guideline Development Group (GDG) meeting convened by WHO in Montreux, Switzerland on 3 March 2015.

The findings and recommendations from the meeting were presented to an External Review Group in March 2016. The External Review Group agreed with the recommendations made by the GDG on the use of molecular line probe assays for the detection of mutations associated with resistance to fluoroquinolones and second-line injectable agents (SLID). This document was finalized following consideration of all comments and suggestions from the participants of the GDG (Annex 1) and the External Review Group (Annex 2).

WHO gratefully acknowledges the contributions of the Chairperson of the GDG (Holger Schünemann), the members of the GDG, and the External Review Group. Karen Steingart and Grant Theron (systematic review authors) are thanked for preparing the systematic reviews and presenting their findings to the members of the GDG.

Funding from the Bill and Melinda Gates Foundation and United States Agency for International Development is gratefully acknowledged through USAID-WHO Consolidated Grant No. GHA-G-00-09-00003 / US 2014-741.

Declarations of Interests

The members of the GDG, the systematic review team and members of the External Review Group completed Declarations of Interests (DOIs). These were reviewed by the WHO Steering Group (Annex 3) prior to the meeting and prior to preparing the policy guidance. The review of each DOI assessed whether an interest had been declared and, if so, whether it was insignificant or potentially significant. If the Steering Group determined that no relevant interest had been declared or such interest was insignificant or minimal, individuals were invited to participate. A summary of DOI statements is provided in Annex 4. None of the DOIs of the GDG members were declared significant or potentially significant. Members of the systematic review team were invited to provide technical input and answer technical questions. These individuals did not participate in the final discussions when recommendations were developed. Also, they were not involved in developing the report of the GDG meeting, nor in preparing the WHO policy guidance.

Executive summary

Background

Genotypic (molecular) methods have considerable advantages for scaling up programmatic management and surveillance of drug-resistant tuberculosis (TB), offering speed of diagnosis, standardised testing, potential for high through-put, and fewer requirements for laboratory biosafety. Molecular tests for detecting drug resistance such as the Genotype® MTBDR*sl* assay, Hain Lifescience, Nehren, Germany (henceforth called MTBDR*sl*) have shown promise for the diagnosis of drug-resistant TB. These tests are rapid (can be performed in a single working day) and detect the presence of mutations associated with drug resistance. MTBDR*sl* belongs to a category of molecular genetic tests called second-line line probe assays (SL-LPA).

MTBDR*sl* (version 1.0) was the first commercial SL-LPA for detection of resistance to second-line TB drugs. In 2015, the manufacturer developed and made commercially available version 2.0 of the MTBDR*sl* assay. Version 2.0 detects the mutations associated with fluoroquinolone and second-line injectable drug (SLID) resistance detected by version 1.0, as well as additional mutations. Once a diagnosis of rifampicin-resistant TB (RR-TB) or multidrug-resistant TB (MDR-TB) has been established, SL-LPA can be used to detect additional resistance to second-line drugs.

The MTBDR*sl* assay incorporates probes to detect mutations within genes (*gyr*A and *rrs* for version 1.0 and, in addition, *gyr*B and the *eis* promoter for version 2.0), which are associated with resistance to either fluoroquinolones or SLIDs. The presence of mutations in these regions does not necessarily imply resistance to all the drugs within a particular group. Although specific mutations within these regions may be associated with different levels of resistance (i.e. different minimum inhibitory concentrations) to each drug within these groups, the extent of cross resistance is not completely understood.

Objectives, rationale and methods used to develop the guidance

This document provides a summary of the evidence and recommendations for the use of SL-LPA for the detection of mutations associated with resistance to fluoroquinolones and SLID **in patients with RR-TB and/or MDR-TB**.

The objectives of this policy guidance are:

- To assess and compare the diagnostic accuracy of SL-LPA for the detection of resistance to fluoroquinolones in sputum specimens (using direct testing) and culture isolates (using indirect testing) confirmed as *M. tuberculosis* complex.
- To assess and compare the diagnostic accuracy of MTBDR*sl* for the detection of resistance to SLIDs in sputum specimens (using direct testing) and culture isolates (using indirect testing) confirmed as *M. tuberculosis* complex.
- To guide the clinical use of the assay in initiation of an appropriate MDR-TB treatment regimen.

Since 2013, WHO has developed interim policy recommendations for the use of two new medicines – bedaquiline and delamanid – as part of MDR-TB regimens[1,2]. In addition, in 2012, WHO published advice about how the shorter standardised MDR-TB regimen[3] - may be introduced by national programmes under operational research conditions[4]. Based on the evidence gathered from observational studies of a shorter MDR-TB regimen in different settings, the *WHO guidelines for drug-resistant TB - 2016 update* recommends a wider use of this regimen by programmes preferably after resistance to both fluoroquinolones and SLID has been excluded by WHO-recommended laboratory tests. The availability of reliable and rapid tests for susceptibility or resistance to these two groups of second-line medicines would be a valuable tool to decide within a few days which patients would be eligible for shorter MDR-TB regimens. Moreover, among patients prescribed a conventional MDR-TB regimen, these tests would also help decide from the start who would benefit from adding one of the new drugs to strengthen the regimen[5].

In response to requests from end-users in the field for guidance on the appropriate use of SL-LPA given their potential to help guide the initiation of appropriate treatment for patients with MDR-TB, WHO commissioned a systematic review of the accuracy of SL-LPA to guide their clinical utility.

On 3 March 2016, the WHO Global TB Programme convened a GDG in Montreux, Switzerland to review the evidence for the use of SL-LPA. The meeting was chaired by an expert in evidence synthesis. Recommendations were developed based on consensus among the GDG members and were subsequently confirmed by an External Review Group. The evidence reviewed and this policy guidance applies to the use of the commercial MTBDR*sl* assay only. Other assays for the detection of resistance to second-line anti-TB agents were not evaluated. Any new or generic assay that detects the presence of mutations associated with drug resistance to second-line anti-TB agents should be subject to adequate evaluation and validation in the settings of intended use.

The WHO policy recommendations developed from the evidence synthesis process by the GDG are summarized below.

WHO's policy recommendations

These recommendations apply to testing of patients with confirmed rifampicin-resistant TB or MDR-TB using second-line lineprobe assays (SL-LPA):

For patients with confirmed rifampicin-resistant TB or MDR-TB, SL-LPA may be used as the initial test, instead of phenotypic culture-based DST, to detect resistance to fluoroquinolones

(Conditional recommendation; moderate certainty in the evidence for test accuracy for direct testing of sputum specimens; low certainty in the evidence for test accuracy for indirect testing of *Mycobacterium tuberculosis* cultures).

---

1 The use of bedaquiline in the treatment of multidrug-resistant tuberculosis. Interim policy guidance (WHO/HTM/TB/2013.6) [Internet]. Geneva, World Health Organization. 2013.
Available from: http://apps.who.int/iris/bitstream/10665/84879/1/9789241505482_eng.pdf

2 The use of delamanid in the treatment of multidrug-resistant tuberculosis. Interim policy guidance (WHO/HTM/TB/2014.23) [Internet]. Geneva, World Health Organization. 2014.
Available from: http://apps.who.int/iris/bitstream/10665/137334/1/WHO_HTM_TB_2014.23_eng.pdf

3 Van Deun A, Maug AKJ, Salim MAH, Das PK, Sarker MR, Daru P, et al. Short, highly effective, and inexpensive standardized treatment of multidrug-resistant tuberculosis. Am J Respir Crit Care Med. 2010 Sep 1;182(5):684–92.

4 The use of short regimens for treatment of multidrug-resistant tuberculosis [Internet]. [cited 2012 Nov 9].
Available from: http://www.who.int/tb/challenges/mdr/short_regimen_use/en/

5 Companion handbook to the WHO guidelines for the programmatic management of drug-resistant tuberculosis. (WHO/HTM/TB/2014.11) [Internet]. Geneva, World Health Organization. 2014. Available from: http://apps.who.int/iris/bitstream/10665/130918/1/9789241548809_eng.pdf

For patients with confirmed rifampicin-resistant TB or MDR-TB, SL-LPA may be used as the initial test, instead of phenotypic culture-based DST, to detect resistance to the second-line injectable drugs

(Conditional recommendation; low certainty in the evidence for test accuracy for direct testing of sputum specimens; very low certainty in the evidence for test accuracy for indirect testing of *Mycobacterium tuberculosis* cultures).

Remarks

a. These recommendations apply to the use of SL-LPA for testing sputum specimens (direct testing) and cultured isolates of *M. tuberculosis* complex (indirect testing) from both pulmonary and extrapulmonary sites. Direct testing on sputum specimens allows for the earlier initiation of appropriate treatment;

b. These recommendations apply to the direct testing of sputum specimens from rifampicin-resistant TB or MDR-TB, irrespective of the smear status, while acknowledging that the indeterminate rate is higher when testing smear-negative sputum specimens compared with smear-positive sputum specimens;

c. These recommendations apply to the diagnosis of XDR-TB while acknowledging that the accuracy for detecting resistance to the fluoroquinolones and to the SLIDs differs and hence the accuracy of a diagnosis of XDR-TB overall is reduced ;

d. These recommendations do not eliminate the need for conventional phenotypic DST capacity which will be necessary to confirm resistance to other drugs and to monitor the emergence of additional drug resistance;

e. Conventional phenotypic DST can still be used in the evaluation of patients with a negative SL-LPA result, particularly in populations with a high pre-test probability for resistance to fluoroquinolones and/or SLID;

f. These recommendations apply to the use of SL-LPA in children with confirmed rifampicin-resistant TB or MDR-TB based on the generalisation of data from adults;

g. Resistance conferring mutations detected by SL-LPA are highly correlated with phenotypic resistance to ofloxacin and levofloxacin. However, the correlation of these mutations with phenotypic resistance to moxifloxacin and gatifloxacin is unclear and the inclusion of moxifloxacin or gatifloxacin in a MDR-TB regimen is best guided by phenotypic DST results;

h. Resistance conferring mutations detected by SL-LPA are highly correlated with phenotypic resistance to SLID and are an indication to use a MDR-TB regimen which is appropriately strengthened;

i. Given high specificity for detecting resistance to fluoroquinolones and SLID the positive results of SL-LPA could be used to guide the implementation of appropriate infection control precautions.

1. Background

Genotypic (molecular) methods have considerable advantages for scaling up programmatic management and surveillance of drug-resistant TB, offering speed of diagnosis, standardised testing, potential for high through-put, and fewer requirements for laboratory biosafety. Molecular tests for detecting drug resistance such as the Genotype® MTBDR*sl* assay (henceforth called MTBDR*sl*) have shown promise for the diagnosis of drug-resistant tuberculosis (TB). These tests are rapid (results are available within a single working day) and detect the presence of mutations associated with drug resistance. MTBDR*sl* belongs to a category of molecular genetic tests called second-line line probe assays (SL-LPA). MTBDR*sl* (version 1.0) was the first commercial line probe assay for detection of resistance to second-line TB drugs. In 2015, the manufacturer developed and made commercially available version 2.0 of the MTBDR*sl* assay. Version 2.0 detects the mutations associated with fluoroquinolone and second-line injectable drug (SLID) resistance detected by version 1.0, as well as additional mutations (described below). Once a diagnosis of rifampicin-resistant TB or multidrug-resistant TB (MDR-TB) has been established, SL-LPA could be used to detect second-line drug resistance.

In September 2013, a World Health Organization (WHO) Guideline Development Group reviewed the evidence (11 published and 7 unpublished studies) and recommended that SL-LPA not be used as a replacement test for phenotypic drug susceptibility testing (DST) and noted, in addition, that the assay did not allow for detection of specific resistance to individual drugs within the SLID of fluoroquinolone groups (WHO 2013). Subsequently, at least two systematic reviews have been published, and a new version of the test (version 2.0) has become commercially available. A Cochrane review on the diagnostic accuracy of version 1.0 of MTBDR*sl* (Theron 2014) included 21 studies published up to 30 January 2014. Several additional studies have since been performed, including studies that have evaluated MTBDR*sl* version 2.0, necessitating an updated systematic review.

Since 2013, WHO has developed interim policy recommendations for the use of two new medicines – bedaquiline and delamanid – as part of MDR-TB regimens[6,7]. In addition, in 2012, WHO published advice about how the shorter standardised MDR-TB regimen[8] - may be introduced by national programmes[9]. In the *WHO guidelines for drug-resistant TB - 2016 update* it is recommended that, preferably, susceptibility to fluoroquinolones and SLIDs be determined by WHO-recommended DST methods before a shorter standardised MDR-TB regimen is started. The availability of reliable and rapid tests for susceptibility/resistance to these two groups of second-line drugs would be a valuable tool to decide which patients would be eligible for shorter MDR-TB regimens. Moreover, among patients prescribed a conventional MDR-TB regimen, such tests would also help decide from the start who would benefit from new medicines to strengthen the regimen[10].

---

[6] The use of bedaquiline in the treatment of multidrug-resistant tuberculosis. Interim policy guidance (WHO/HTM/TB/2013.6) [Internet]. Geneva, World Health Organization. 2013.
Available from: http://apps.who.int/iris/bitstream/10665/84879/1/9789241505482_eng.pdf

[7] The use of delamanid in the treatment of multidrug-resistant tuberculosis. Interim policy guidance (WHO/HTM/TB/2014.23) [Internet]. Geneva, World Health Organization. 2014.
Available from: http://apps.who.int/iris/bitstream/10665/137334/1/WHO_HTM_TB_2014.23_eng.pdf

[8] Van Deun A, Maug AKJ, Salim MAH, Das PK, Sarker MR, Daru P, et al. Short, highly effective, and inexpensive standardized treatment of multidrug-resistant tuberculosis. Am J Respir Crit Care Med. 2010 Sep 1;182(5):684–92.

[9] The use of short regimens for treatment of multidrug-resistant tuberculosis [Internet]. [cited 2012 Nov 9].
Available from: http://www.who.int/tb/challenges/mdr/short_regimen_use/en/

[10] Companion handbook to the WHO guidelines for the programmatic management of drug-resistant tuberculosis. (WHO/HTM/TB/2014.11) [Internet]. Geneva, World Health Organization. 2014.
Available from: http://apps.who.int/iris/bitstream/10665/130918/1/9789241548809_eng.pdf

In accordance with WHO standards for assessing evidence when formulating policy recommendations, the GRADE approach (the Grading of Recommendations Assessment, Development and Evaluation (http://www.gradeworkinggroup.org/) was used. GRADE provides a structured framework for evaluating the accuracy of diagnostic tests, and assessing their impact on patient and public health. The systematic review assessed the accuracy of SL-LPA in the direct testing of sputum specimens and in the indirect testing of *M. tuberculosis* culture isolates from patients with confirmed RR-TB or MDR-TB.

The assay reviewed is the commercially available test Genotype MTBDR*sl* assay (versions 1.0 and 2.0) Hain Lifesciences, Nehren, Germany. The evidence reviewed and this policy guidance applies to the use of this commercial SL-LPA only. Other assays for the detection of resistance to second-line anti-TB agents were not evaluated. Any new or generic assay for the detection of the presence of mutations associated with drug resistance to second-line anti-TB agents should be subject to adequate evaluation and validation in the settings of intended use as per WHO policy[11].

Index test

The index test is MTBDR*sl* and the different characteristics of versions 1.0 and 2.0 are presented in Table 1. SL-LPA detect specific mutations associated with resistance to the class of fluoroquinolones (including ofloxacin, levofloxacin, moxifloxacin and gatifloxacin) and SLIDs (including kanamycin, amikacin, and capreomycin) in the *Mycobacterium tuberculosis* complex. Version 1.0 detects mutations in the *gyrA* quinolone resistance-determining region (codons 85-97) and *rrs* (codons 1401, 1402, and 1484). Version 2.0 additionally detects mutations in the *gyrB* quinolone resistance-determining region (codons 536-541) and the *eis* promoter region (codons -10 to -14)[12]. As mutations in these regions may cause additional resistance to the fluoroquinolones or SLIDs respectively, version 2.0 is expected to have improved sensitivity for resistance to these drug classes. Mutations in some regions (e.g., the *eis* promoter region) may be responsible for causing resistance to one drug in a class more than other drugs within that class. For example, the *eis* C14T mutation is associated with kanamycin resistance in strains from Eastern Europe[13]. Version 1.0 also detects mutations in *embB* that may encode for resistance to ethambutol. As ethambutol is a first-line drug and was omitted from version 2.0, this review did not determine the accuracy for ethambutol resistance.

Table 1. Characteristics of Genotype MTBDRsl versions 1.0 and 2.0 as per manufacturer

Detection	**Version 1.0** *M. tuberculosis* complex and resistance to fluoroquinolones, SLIDs and ethambutol	**Version 2.0** *M. tuberculosis* complex and resistance to fluoroquinolones and SLIDs
Samples	Smear-positive specimens and culture isolates	Smear-positive and smear-negative specimens and culture isolates
Fluoroquinolone resistance	Mutations in resistance-determining region of the *gyrA* gene	Mutations in resistance-determining regions of the *gyrA* and *gyrB* genes

[11] WHO 2015. Implementing tuberculosis diagnostics. Policy Framework. WHO/HTM/TB/2015.11 http://apps.who.int/iris/bitstream/10665/162712/1/9789241508612_eng.pdf?ua=1

[12] Hain Lifescience 2015. Rapid Diagnosis of Tuberculosis Brochure. http://www.hain-lifescience.de/uploadfiles/file/produkte/mikrobiologie/mykobakterien/tb_eng.pdf (accessed 23 December 2015).

[13] Gikalo MB, Nosova EY, Krylova LY, Moroz AM. The role of eis mutations in the development of kanamycin resistance in Mycobacterium tuberculosis isolates from the Moscow region. J Antimicrob Chemother. 2012 Sep;67(9):2107-9.

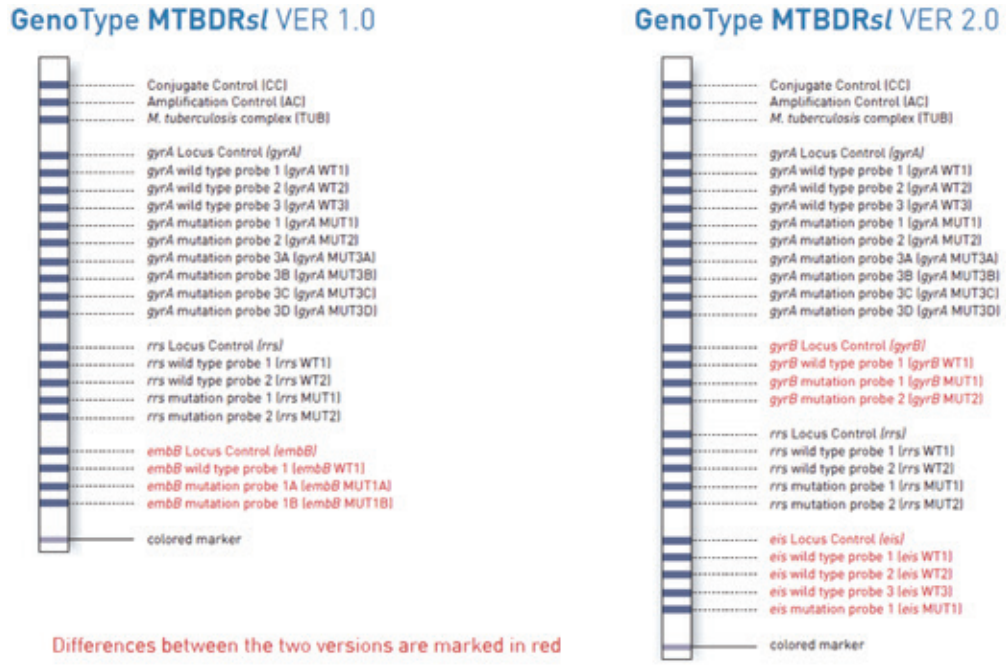
SLID resistance	Mutations in resistance determining region of the *rrs* gene	Mutations in resistance determining region *rrs* gene and the *eis* promoter region
Ethambutol resistance	Mutations in the *embB* gene	Not included

The MTBDR*sl* assay incorporates probes to detect mutations within genes (*gyr*A and *rrs* for version 1.0 and, in addition, *gyr*B and the eis promoter for version 2.0), which are associated with resistance to the class of fluoroquinolones or the SLIDs. The presence of mutations in these regions does not necessarily imply resistance to all the drugs within that class. Although specific mutations within these regions may be associated with different levels of resistance (i.e. different minimum inhibitory concentrations) to each drug within these classes, the extent of cross resistance is not completely understood. More data is needed to better understand the correlation of the presence of certain fluoroquinolone resistance-conferring mutations with phenotypic DST resistance and with patient outcomes.

The assay procedure can be performed **directly** using a processed sputum sample (see steps 1-4 below) or **indirectly** using DNA isolated and amplified from a culture of *M. tuberculosis* (see steps 2-4 below). **Direct testing** involves the following steps: 1) decontamination (e.g. with sodium hydroxide) and concentration of a sputum specimen by centrifugation; 2) isolation and amplification of DNA; 3) detection of the amplification products by reverse hybridization; and 4) visualization using a streptavidin-conjugated alkaline phosphatase colour reaction. The observed bands, each corresponding to a wild type or resistance genotype probe, can be used to determine the drug susceptibility profile of the analysed specimen. The assay can be performed and completed with in a single working day.

Figure 1 shows an example of MTBDR*sl* results for version 1.0 and 2.0. A band for the detection of the *M. tuberculosis* complex (the "TUB" band) is included, as well as two internal controls (conjugate and amplification controls) and a control for each gene locus (version 2.0: *gyr*A, *gyr*B, *rrs, eis*). The two internal controls plus each gene locus control should be positive; otherwise the assay cannot be evaluated for that particular drug. A result can be indeterminate for one locus but valid for another (on the basis of a gene-specific locus control failing).

Figure 1. Examples of different GenoType® MTBDRsl strip readouts



A template is supplied by the manufacturer to help read the strips where the banding patterns are scored by eye, transcribed, and reported. In high-volume settings, the GenoScan®, an automated reader, can be incorporated to interpret the banding patterns automatically and give a suggested interpretation. If the operator agrees with the interpretation, the results are automatically uploaded, thereby reducing possible transcription errors.

2. Methods

2.1. Evidence synthesis

In March 2016, a Guideline Development Group was convened to assess available data on the use of SL-LPA. WHO commissioned a systematic review on the accuracy of SL-LPA for the detection of mutations associated with resistance to fluoroquinolones and second-line injectable drugs (SLIDs) in patients with rifampicin resistant and/or multi-drug resistant tuberculosis.

The evaluation used the GRADE system to determine the quality of the evidence and provide information on the strength of the recommendations using PICO questions agreed by the GDG. PICO refers to the following four elements that should be included in questions that govern a systematic search of the evidence: the **P**opulation targeted by the action or intervention (in the case of systematic reviews of diagnostic test accuracy, P is the population of interest); the **I**ntervention (I is the index test; the **C**omparator (C is the comparator test(s); and the **O**utcome (O is usually sensitivity and specificity). The PICO questions for the review are given below.

Overarching question

Should SL-LPA results be used to guide clinical decisions to use second line anti-TB drugs (fluoroquinolones, SLIDs) in patients diagnosed with rifampicin-resistant TB or MDR-TB as compared to phenotypic DST (based on diagnostic accuracy, prevalence of the resistance to particular anti-TB drugs, cross resistance between anti-TB drugs)?

PICO questions addressed by the GDG

1. Should SL-LPA be used to guide clinical decisions to use fluoroquinolones in the direct testing of patient specimens with confirmed RR-TB or MDR-TB?

2. Should SL-LPA be used to guide clinical decisions to use SLID in the direct testing of patient specimens with confirmed RR-TB or MDR-TB?

3. Should SL-LPA be used to guide clinical decisions to use fluoroquinolones in the indirect testing of culture isolates from patients with confirmed rifampicin-resistant TB or MDR-TB?

4. Should SL-LPA be used to guide clinical decisions to use SLID in the indirect testing of culture isolates from patients with confirmed RR-TB or MDR-TB?

The systematic reviews were conducted according to the standards outlined by Cochrane in the Cochrane Handbook for Diagnostic Test Accuracy Reviews.[14] A comprehensive search of the following databases was performed up to 21 September 2015, without language restrictions: Cochrane Infectious Diseases Group Specialized Register; MEDLINE (PubMed); EMBASE OVID; ISI Web of Knowledge (Science Citation Index - Expanded, Conference Proceedings Citation Index- Science (CPCI-S), BIOSIS Previews, LILACS (http://lilacs.bvsalud.org/en/ and SCOPUS. A search of the metaRegister of Controlled Trials (mRCT; http://www.controlled-trials.com/) and the search portal of the World Health Organization (WHO) International Clinical Trials Registry Platform (www.who.int/trialsearch) were performed to identify ongoing trials, and ProQuest Dissertations & Theses A&I to identify relevant dissertations.

[14] Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and Presenting Results. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* Version 1.0. The Cochrane Collaboration, 2010. Available from: http://methods.cochrane.org/sdt/handbook-dta-reviews

All studies that determined the diagnostic accuracy of the index test in comparison with a defined reference standard, including case-control designs, were included. Included studies were those from which data could be extracted for true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) to calculate sensitivity and specificity estimates and 95% confidence intervals (CI) for individual studies. The individual study results were graphed by plotting the estimates of sensitivity and specificity (and their 95% CIs) in forest plots. Patients of any age who had rifampicin-resistant or MDR-TB or may have had resistance to any of the second-line TB drugs, irrespective of background burden and patient population were included. Unpublished studies reported only in abstracts and conference proceedings were excluded.

The following reference standards were used:

1. Phenotypic culture-based DST using either solid or liquid culture incorporating the drug of interest;

2. Sequencing of the *gyr*A or *rrs* genes (for evaluation of version 1.0) or additionally the *gyr*B and *eis* promoter regions (for evaluation of version 2.0);

3. A composite reference standard that included both phenotypic culture-based DST and sequencing of the same specimens. If a specimen was resistant according to culture-based DST or had a resistance conferring mutation associated with a particular drug, the specimen was classified as being resistant to a particular drug. If both culture-based DST and sequencing indicated susceptibility, the specimen was classified as being susceptible to a particular drug.

Phenotypic culture-based DST is the current accepted reference standard method. However, the results are only reliable when an appropriate concentration of the drug of interest is used to determine the threshold to differentiate probable susceptible from probable resistant strains. Sequencing is considered a reliable method to detect mutations known to be associated with phenotypic drug-resistance, however, not all resistance-determining mechanisms for the fluoroquinolones and SLIDs are completely known. As a consequence, targeted sequencing may not detect all strains with phenotypic resistance due to the presence of mutations that confer resistance occurring outside the area of a particular gene targeted for sequencing.

When phenotypic culture-based DST was performed using more than one drug from the class of fluoroqunolones (ofloxacin, levofloxacin, or moxifloxacin) or group of SLIDs (amikacin, kanamycin or capreomycin), data were extracted for each drug and for each class or group overall. If the reference standard indicated resistance to at least one drug in that class or group, the sample was classified as resistant to that class or group of drugs. Accuracy estimates for individual drugs were reported within the drug class or group when that drug was used as part of the phenotypic culture-based DST reference standard. For determining resistance to the drug class or group, the following approach was used:

For a culture-based DST reference standard, some studies might have used detection of ofloxacin resistance and other studies might have used detection of moxifloxacin resistance to confirm a SL-LPA fluoroquinolone-resistant result. In such a scenario, if phenotypic culture-based DST was positive for resistance to one of the drugs in the drug class and the SL-LPA result was concordant, the index test result was classified as a true positive for resistance to the fluoroquinolones.

For sequencing as a reference standard, if the index test reported resistance to fluoroquinolones and the presence of mutations known to be associated with fluoroquinolone resistance was confirmed in the same regions of the genome targeted by SL-LPA, then the test result was recorded as concordant and the index test was classified as a true positive for resistance to the fluoroquinolones. The same approach was adopted for the SLIDs.

Using the GRADE framework, calculations of test sensitivity and specificity were used as proxy measures for patient outcomes; these outcomes were based on the relative importance or impact of false-positive and false-negative results: Poor sensitivity would result in *false-negative* results so that

patients with RR-TB or MDR-TB would not be placed on an effective treatment regimen, which would have negative consequences in terms of delayed initiation of an effective regimen, the development of additional drug resistance, morbidity, mortality and further transmission of disease. Poor specificity would result in *false-positive* results so that patients without additional drug resistance would be prescribed unnecessary treatment which may have increased adverse effects.

Rates for true positives, true negatives, false positives and false negatives were calculated using agreed pretest probabilities of resistance to second-line anti-TB drugs. Prevalences of 5%, 10%, and 15% were used to cover the lower and upper levels of resistance associated with resistance to fluoroquinolones or SLIDs among patients with RR-TB or MDR-TB. Prevalences of 1%, 5%, and 10% were used to cover both the lower and upper levels of prevalence XDR-TB in different epidemiological settings. These thresholds were chosen based on the findings from global surveillance of second-line resistance among patients with RR-TB or MDR-TB from 75 countries[15].

The evaluation of the impact on patients was based on a balance among the following values:

- *true positives* – the benefit to patients from rapid diagnosis and treatment;
- *true negatives* – the benefit to patients who would be spared unnecessary treatment; the benefit of reassurance and alternative diagnosis;
- *false positives* – the likelihood of anxiety and morbidity caused by additional testing, unnecessary treatment and possible adverse effects; possible stigma associated with a diagnosis of drug-resistant TB; and the chance that a false positive may halt further diagnostic evaluation;
- *false negatives* – the increased risk of morbidity and mortality, delayed treatment initiation and the continued risk of transmission of drug-resistant TB.

2.2 Guideline Development Group meeting

The WHO Steering Group (Annex 3) was responsible for scoping the guideline, drafting the PICO questions and overseeing the evidence retrieval and analyses. The Steering Group was also responsible for selecting members for the GDG (Annex 1) and External Review Group (Annex 2), for managing declarations of interest, and for organising the GDG meeting. Brief biographies of the GDG members were made available for public scrutiny on the WHO Global TB Programme website (http://www.who.int/tb/laboratory/policy_statements/en/) two weeks prior to the GDG meeting.

PICO questions were drafted by the WHO Steering Group and were presented to the GDG for discussion and modification. The Steering Group also prepared an initial list of relevant outcomes, including desirable effects and undesirable effects, and requested the GDG to identify any other important outcomes.

A webinar was conducted with members of the GDG prior to the meeting to review the preliminary findings from the systematic reviews, refine and finalize the proposed patient outcomes and to rate their relative importance. The following outcomes for each PICO question were determined, and the ratings of their importance were unanimously agreed:

- Critical outcomes – diagnostic accuracy as reflected by true-positive, true-negative, false-positive and false-negative results;
- Important outcomes – Impact on the time-to-diagnosis, ease of use and acceptability of the test, and cost.

[15] World Health Organization 2015. Global Tuberculosis Report 2015.
Available at: http://apps.who.int/iris/bitstream/10665/191102/1/9789241565059_eng.pdf

The format for the Evidence to Recommendations tables was discussed and agreed upon by the Guideline Development Group members during the webinar. The format included the following sections: description of the problem; diagnostic test accuracy; patient values and preferences; certainty of the evidence for test accuracy; benefits and harms of the test's use; resources required; equity; acceptability; feasibility and neutral language to formulate the draft recommendations.

Evidence to Recommendations tables were developed for each of the PICO questions in order to guide the process of development of the recommendations.

The meeting was chaired by a guideline methodologist with expertise in guideline development processes and methods. The methodologist participated in the initial planning, scoping and development of the key questions for the GDG meeting. During the meeting, the methodologist helped the GDG formulate recommendations based on the evidence presented. Decisions were based on consensus, ie. unanimous agreement among all GDG members.

The full set of evidence to recommendations tables are included in Annex 6.

2.3 External Review Group

The findings and recommendations from the GDG meeting were sent to an External Review Group of international experts in the field of TB laboratory diagnostics, which included representatives from the WHO TB Supranational Reference Laboratory Network, the WHO TB Strategic and Technical Advisory Group (STAG-TB) and the Core Group members of the Global Laboratory Initiative Working Group of the Stop TB Partnership. The External Review Group agreed with the GDG's recommendations and with the subsequent WHO policy guidance.

3. Scope

This document provides a pragmatic summary of the evidence and recommendations on the use of SL-LPA for the detection of mutations associated with resistance to fluoroquinolones and SLID in patients with RR-TB and/or MDR-TB. It should be read in conjunction with the 2015 WHO Framework for implementing TB diagnostics, which provides guidance on implementing the diagnostic tools and methods approved by WHO within the context of a country's infrastructure, resources, epidemiology or drug resistant TB and HIV. These documents are available at

http://www.who.int/tb/areas-of-work/laboratory/policy_statements/en/.

3.1 Target audience

This policy guidance is intended to be used by clinicians treating patients with RR-TB or MDR-TB, managers and laboratory directors working in TB programmes in coordination with external laboratory consultants, donor agencies, technical advisers, laboratory technicians, procurement officers for laboratory equipment, service providers in the private sector, relevant government sectors, and implementation partners that are involved in country-level strengthening of MDR-TB diagnostic and treatment services. Individuals responsible for programme planning, budgeting, mobilizing resources and implementing training activities for the programmatic management of drug resistant TB may also benefit from this document.

Date of review: 2020 or earlier should significant additional evidence become available.

4. Evidence base for policy formulation

Twenty nine unique studies were identified. All studies but two (Fan 2011, in Chinese, and Chikamatsu 2012, in Japanese) were written in English. 26 studies evaluated the MTBDR*s*/ version 1.0 assay which included 21 studies from the original Cochrane review by Theron et al, 2014 (Ajbani 2012; Barnard 2012; Brossier 2010; Chikamatsu 2012; Fan 2011; Ferro 2013; Hillemann 2009; Huang 2011; Ignatyeva 2012; Jin 2013; Kiet 2010; Kontsevaya 2011; Kontsevaya 2013; Lacoma 2012; Lopez-Roa 2012; Miotto 2012; Said 2012; Surcouf 2011; Tukvadze 2014; van Ingen 2010) and five new studies (Catanzaro 2015; Kambli 2015a; Kambli 2015b; Simons 2015; Tomasicchio 2016). Three studies (one published study (Tagliani 2015) and two unpublished evaluation studies (FIND 2016; NICD 2015) evaluated version 2.0.

Data for version 1.0 and version 2.0 of the MTBDR*s*/ assay were analysed separately. A phenotypic culture-based DST reference standard was used for the primary analyses. These analyses were stratified first by susceptibility or resistance to a particular drug and second by type of SL-LPA testing (indirect testing or direct testing). Within each stratum estimates of the studies observed sensitivities and specificities were plotted in forest plots with 95% confidence intervals (CI). Where adequate data were available, data were combined for meta-analysis by fitting the bivariate random-effects model[16,17] (Macaskill 2010; Reitsma 2005) using Stata version 14 with the metandi and xtmelogit commands. In situations with few studies, a meta-analysis was performed where appropriate by reducing the bivariate model to either univariate random effects logistic regression models or, when there was little observed or no heterogeneity, to fixed effects models[18].

Results from studies of direct testing were compared with results from studies of indirect testing by adding a covariate for the type of testing to the model. The significance of the differences in sensitivity and specificity estimates between studies was assessed in which SL-LPA were performed by direct testing or indirect testing. Where data were sufficient, comparative analyses were performed including only those studies that made direct comparisons between test evaluations with the same participants. Otherwise, all studies with available data were included. Comparative studies are preferred to non-comparative studies when deriving evidence of diagnostic test accuracy (Takwoingi 2013). The pooled estimates for the diagnostic accuracy of MTBDR*s*/ (version 1.0) for resistance to fluoroquinolones, SLIDs and detection of XDR-TB by direct and indirect testing, compared with a phenotypic culture-based DST reference standard are given in Table 2.

Of the total 29 studies (26 studies evaluated version 1.0; three studies evaluated version 2.0), 20 (69%) studies evaluated patients from low-income or middle-income countries. The median sample size (interquartile range) was 138 (59,256).

[16] Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and presenting results In: Deeks JJ, Bossuyt PM, Gatsonis C (editors). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 2010. Available from: http://srdta.cochrane.org/

[17] 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.

[18] Takwoingi Y, Guo B, Riley RD, Deeks JJ. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. Stat Methods Med Res 2015;pii: 0962280215592269.

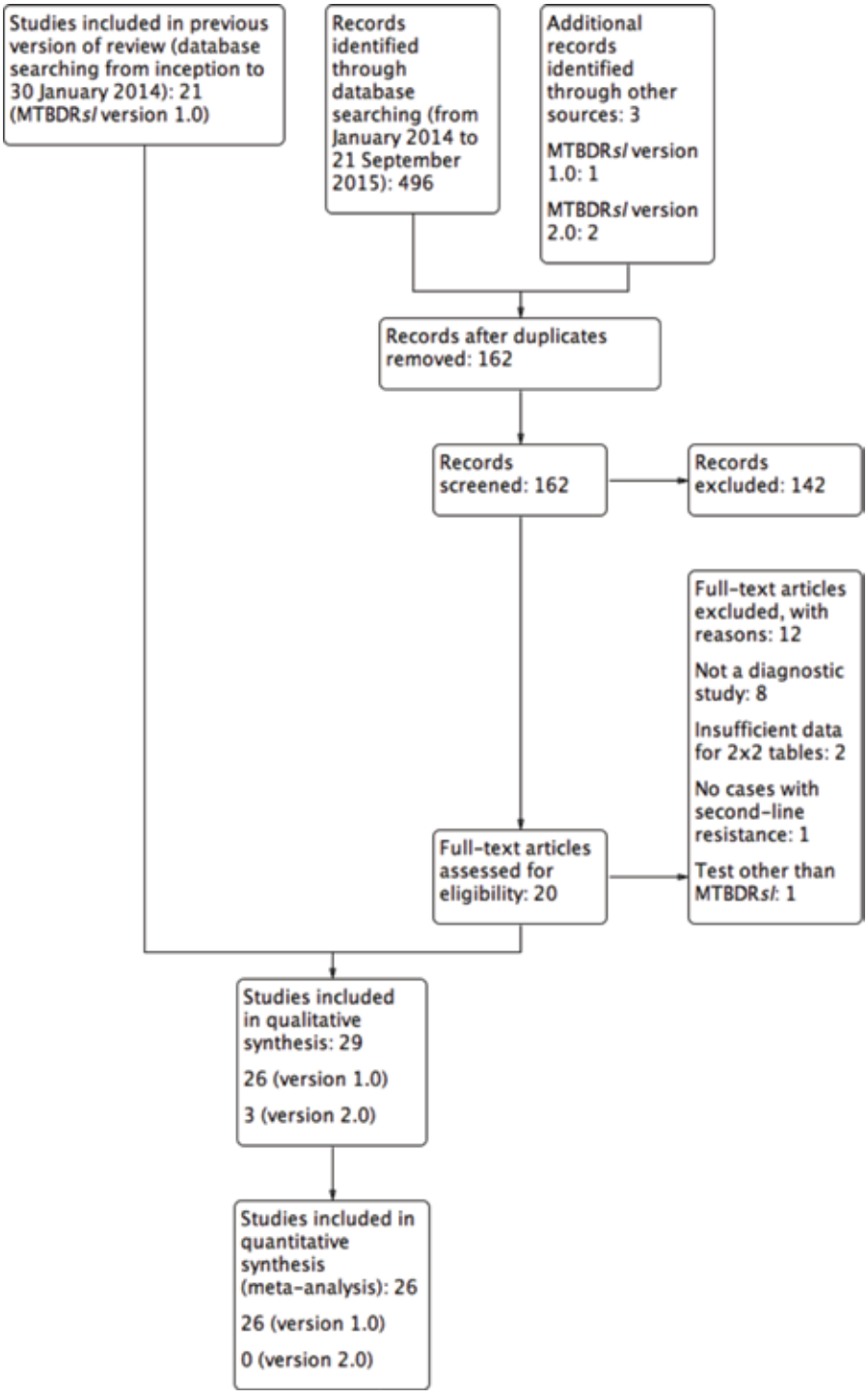
Table 2. Accuracy of MTBDRsl (version 1.0) for fluoroquinolone and second-line injectable drug resistance and XDR-TB, indirect and direct testing (smear-positive specimens), phenotypic culture-based DST reference standard

Pooled sensitivity (95% CI)	**Pooled specificity (95% CI)**	**Pooled sensitivity (95% CI)**	**Pooled specificity (95% CI)**	**Pooled sensitivity P value[1]**	**Pooled specificity P value[1]**
Fluoroquinolones, indirect testing (19 studies, 2 223 participants)		**Fluoroquinolones, direct testing** (9 studies, 1 771 participants)			
85.6% (79.2 to 90.4)	98.5% (95.7 to 99.5)	86.2% (74.6 to 93.0)	98.6% (96.9 to 99.4)	0.932	0.333
Second-line injectable drugs, indirect testing (16 studies, 1 921 participants)		**Second-line injectable drugs, direct testing** (8 studies, 1 639 participants)			
76.5% (63.3 to 86.0)	99.1% (97.3 to 99.7)	87.0% (38.1 to 98.6)	99.5% (93.6 to 100.0)	0.547	0.664
XDR-TB, indirect testing (8 studies, 880 participants)		**XDR-TB, direct testing** (6 studies, 1 420 participants)			
70.9% (42.9 to 88.8)	98.8% (96.1 to 99.6)	69.4% (38.8 to 89.0)	99.4% (95.0 to 99.3)	0.888	0.855

[1] Likelihood ratio test for evidence of a significant difference between accuracy estimates.

Figure 2. shows the flow of studies in the review. Included and excluded studies and the reasons for their exclusion are given in online Annex 8.
Available at: http://www.who.int/tb/areas-of-work/laboratory/policy_statements/en/

Figure 2. Selection of studies evaluating the accuracy of the MTBDRsl assay for the detection of resistance to the fluoroquinolones and SLID



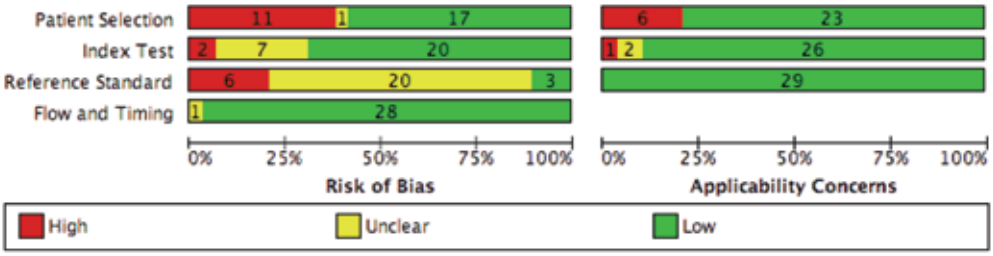
4.1 Assessment of methodological quality

The quality of the included studies was appraised with the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool[19]. QUADAS-2 consists of four domains: patient selection, index test, reference standard, and flow and timing. All domains were assessed for the potential for risk of bias and the first three domains for concerns regarding applicability. Signalling questions in each domain were used to form judgments about the risk of bias. One review author (GT) piloted the tool with two included studies and finalized the tool based on experience gained from the pilot testing. Three review authors (GT, JP, and KRS) then independently assessed methodological quality of included studies with the finalized tool and finalized judgments by discussion.

More than 50% of studies were considered to have low risk of bias for the patient selection, index test, and flow and timing domains. For the reference standard domain, only three (10%) studies were considered to have low risk of bias because these studies used the WHO-recommended critical concentrations for every drug they included in the culture-based DST reference standard whereas the other studies did not. Regarding applicability, there was low concern for all QUADAS-2 domains. There was low concern about the applicability of the included studies despite the evaluations having predominately been conducted with the MTBDR*s*l (version 1.0) assay. It was reasoned that with the addition of new probes targeting more resistance-conferring mutations, the diagnostic accuracy of MTBDR*s*l (version 2.0) would be the same or higher than that of MTBDR*s*l version 1.0 and the specificity to be the same or slightly lower than that of version 1.0 (because of the small likelihood of at least one of the probes mispriming and giving false-positive results). The findings were therefore considered applicable to the test. The included studies assessed sensitivity and specificity in research settings which may have overestimated the accuracy compared with routine practice settings.

Figures 3 shows the risk of bias and applicability concerns for the 29 included studies.

Figure 3. Risk of bias and applicability concerns graph: review authors' judgments about each domain presented as percentages across included studies



Assessment of the certainty of the evidence

The certainty of the evidence (also called quality of evidence or confidence in effect estimates) was assessed using the GRADE approach[20,21] and GRADEpro Guideline Development Tool

[19] 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.
[20] Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE,Falck-Ytter Y, Meerpohl J, Norris S, Guyatt GH. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64((4)):401-6
[21] Schünemann H, Brożek J, Guyatt G, Oxman A, editors. The GRADE Working Group. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. Available from www.guidelinedevelopment.org/handbook.

software[22] (GRADEpro 2015). In the context of a systematic review, the ratings of the certainty of the evidence reflect the extent of confidence that the estimates of the effect (including test accuracy and associations) are correct. The certainty of the evidence was rated as high (no points subtracted), moderate (one point subtracted), low (two points subtracted), or very low (greater than two points subtracted) based on five domains: risk of bias, indirectness, inconsistency, imprecision, and publication bias. One point was subtracted when there was a serious issue identified or two points when there was a very serious issue identified in any of the domains used to judge the certainty of the evidence.

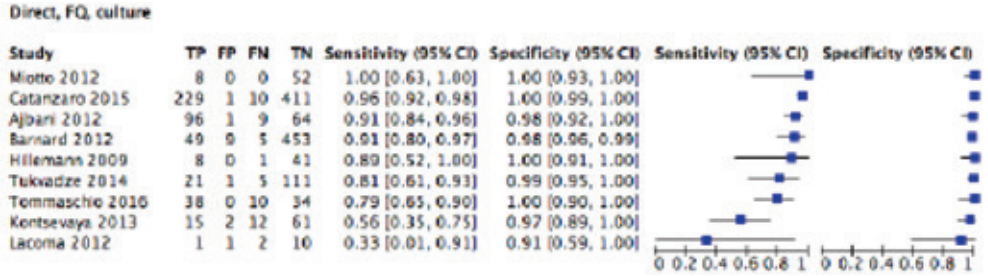
4.2 Accuracy for the detection of fluoroquinolone resistance

4.2.1 Direct testing MTBDRsl version 1.0

Nine studies evaluated MTBDRs*l* in 1771 participants, including 519 (29.3%) confirmed cases of TB with resistance to fluoroquinolones by direct testing on sputum specimens compared with a phenotypic culture-based DST reference standard (Figure 4). For individual studies, sensitivity estimates ranged from 33% to 100% and specificity estimates ranged from 91% to 100%. The pooled sensitivity and specificity (95% CI) were 86.2% (74.6% to 93.0%) and 98.6% (96.9% to 99.4%) respectively.

Nine studies reported indeterminate results for the MTBDRs*l* assay by direct testing of smear-positive sputum specimens compared with phenotypic culture-confirmed resistance to fluoroquinolones. Of 2059 results, 147 (7.1%) were indeterminate. Of the indeterminate results, 68 were culture-based DST resistant, 73 culture-based DST susceptible and six did not have a culture-based DST result. The indeterminate rates for direct testing for each smear-grade (smear-negative, scanty, 1+, 2+, 3+) were 61/190 (32.1%), 28/133 (21.1%), 35/272 (12.9%), 19/211 (9.0%), and 44/388 (11%), respectively.

Figure 4. Forest plots of MTBDRsl sensitivity and specificity for fluoroquinolone resistance by direct testing in comparison with a culture-based DST reference standard



TP true positive; FP false positive; FN false negative; TN true negative; CI confidence interval

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). The individual studies are ordered by decreasing sensitivity. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

Sensitivity and specificity estimates for MTBDRs*l* by direct testing of individual drugs within the class of fluoroquinolones is given in Figure 5. Seven studies involving 1667 participants evaluated the MTBDRs*l* assay for the detection of ofloxacin resistance by direct testing. Sensitivity estimates ranged from 79% to 100% and specificity estimates ranged from 98% to 100%. The pooled

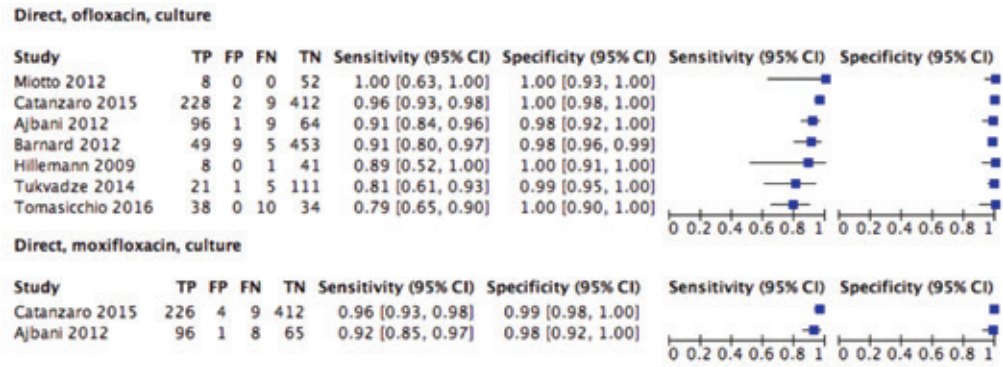
[22] GRADEpro GDT: GRADEpro Guideline Development Tool [Software]. McMaster University, 2015 (developed by Evidence Prime, Inc.). Available from gradepro.org

sensitivity and specificity (95% CI) were 90.9% (84.7% to 94.7%) and 98.9% (97.8% to 99.4%) respectively. Based on all data, there was no evidence of a statistically significant difference in accuracy between indirect and direct testing for ofloxacin resistance (P values for differences in sensitivity and specificity of 0.180 and 0.161, respectively).

Two studies involving 821 participants evaluated the MTBDR*sl* assay for detection of moxifloxacin resistance by direct testing (Figure 5). Catanzaro 2015 reported a sensitivity of 96% and specificity of 99% and Ajbani 2012 reported a sensitivity of 92% and specificity of 98%. The pooled sensitivity and specificity (95% CI) were 95.0% (92.1% to 96.9%) and 99.0% (97.5% to 99.6%) respectively. Based on all data, there was no evidence of a statistically significant difference in accuracy between indirect and direct testing for moxifloxacin resistance (P values for differences in sensitivity and specificity of 0.820 and 0.365, respectively).

No studies were identified that performed direct testing for detection of levofloxacin or gatifloxacin resistance.

Figure 5. Forest plots of MTBDRsl sensitivity and specificity for ofloxacin and moxifloxacin resistance, test performed directly, phenotypic culture-based DST reference standard



TP true positive; FP false positive; FN false negative; TN true negative; CI confidence interval

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). The individual studies are ordered by decreasing sensitivity. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

4.2.2 Direct testing MTBDRsl version 2.0

For MTBDR*sl* version 2.0 the data were either too sparse or too heterogeneous to combine in a meta-analysis or to compare indirect and direct testing.

Two studies evaluated MTBDR*sl* version 2.0 in 187 individuals, including 37 (19.8%) confirmed cases of TB with resistance to fluoroquinolones by direct testing on a processed sputum sample compared with a phenotypic culture-based DST reference standard. Sensitivity and specificity estimates (95% CI) were 100% (59% to 100%) and 100% (86% to 100%) for FIND 2016 and 97% (83% to 100%) and 98% (93% to 100%) for Tagliani 2015.

Two studies evaluated the MTBDR*sl* version 2.0 assay in 22 individuals, including eight confirmed cases of TB and resistance to fluoroquinolones by direct testing on sputum smear-negative specimens. Sensitivity and specificity estimates (95% CI) were 100% (29% to 100%) and 90% (55% to 100%) for FIND 2016 and 80% (28% to 99%) and 100% (40% to 100%) for Tagliani 2015. Of two studies that evaluated indeterminate MTBDR*sl* results for direct testing of smear-positive

specimens, no indeterminate results were reported. The indeterminate rate for direct testing for smear-negative specimens was 8/30 (26.7%) (four culture-DST resistant and four culture-DST susceptible).

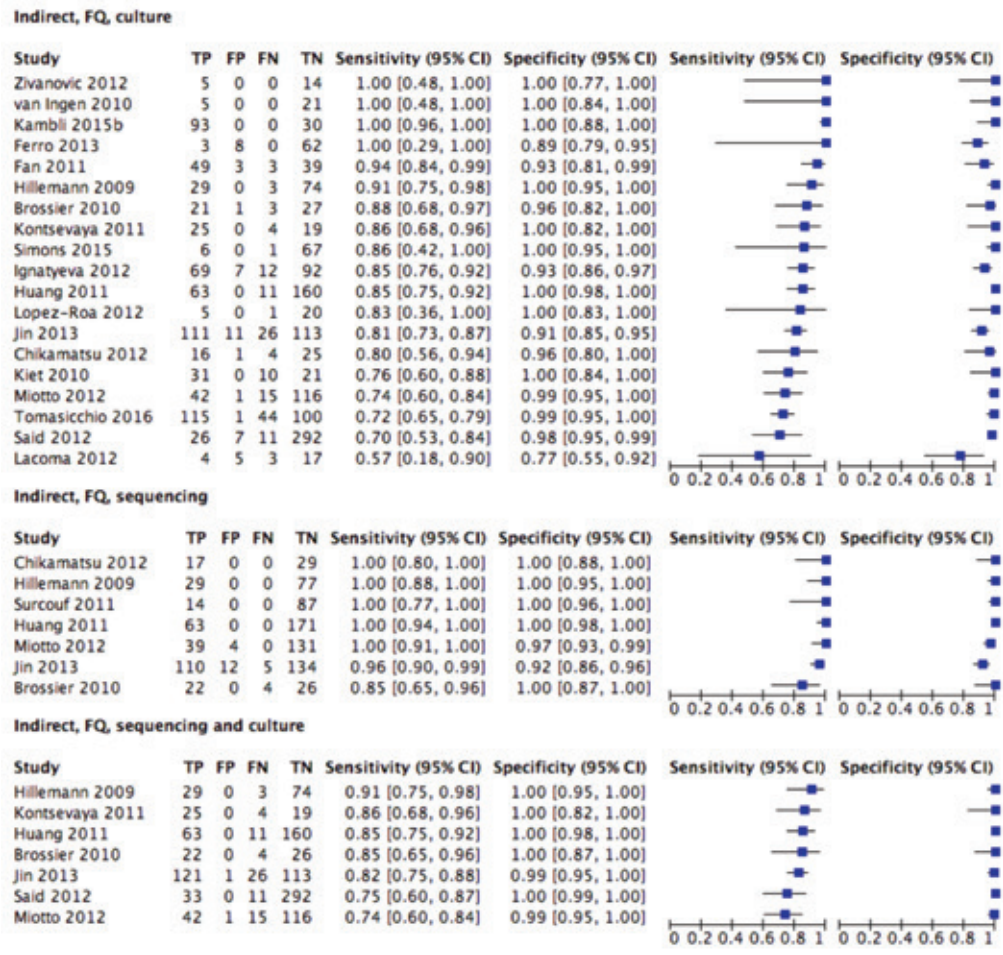
4.2.3 Indirect testing MTBDRsl version 1.0

Nineteen studies evaluated MTBDR*sl* in 2223 individuals, including 869 (39.1%) confirmed cases of TB with resistance to fluoroquinolones by indirect testing on a culture of *M. tuberculosis* compared with a phenotypic culture-based DST reference standard (Figure 6). For individual studies, sensitivity estimates ranged from 57% to 100% and specificity estimates ranged from 77% to 100%. The pooled sensitivity and specificity (95% CI) were 85.6% (79.2% to 90.4%) and 98.5% (95.7% to 99.5%) respectively. Fourteen studies reported indeterminate results for the MTBDR*sl* assay by indirect testing of a culture of *M.tuberculosis* with phenotypic culture-confirmed resistance to fluoroquinolones. Of 2065 results, eight (0.4%) results were indeterminate (seven culture-based DST resistant and one culture-based DST susceptible).

Seven studies evaluated MTBDR*sl* in 974 individuals compared with a sequencing reference standard (Figure 6). Sensitivity estimates ranged from 85% to 100% and specificity estimates ranged from 92% to 100%. The pooled sensitivity and specificity (95% CI) were 99.3% (81.2% to 100.0%) and 99.3% (90.8% to 100.0%). There was evidence of a statistically significant higher sensitivity using sequencing as the reference standard compared with phenotypic culture-based DST (P value < 0.001), but not specificity (P value of 0.735).

Seven studies evaluated MTBDR*sl* in 1211 individuals compared with a reference standard of phenotypic culture-based DST and sequencing combined (i.e. both investigations performed in all isolates) (Figure 6). Sensitivity estimates ranged from 74% to 91% and specificity estimates ranged from 99% to 100%. The pooled sensitivity and specificity (95% CI) were 82.0% (77.7% to 85.6%) and 99.8% (98.5% to 100.0%) respectively. There was no evidence of a statistically significant difference in accuracy using both phenotypic culture-based DST and sequencing as the reference standard compared with culture-based DST (P values for differences in sensitivity and specificity of 0.664 and 0.070, respectively).

Figure 6. Forest plots of MTBDRsl sensitivity and specificity for fluoroquinolone resistance, test performed indirectly, different reference standards



TP true positive; FP false positive; FN false negative; TN true negative; CI confidence interval

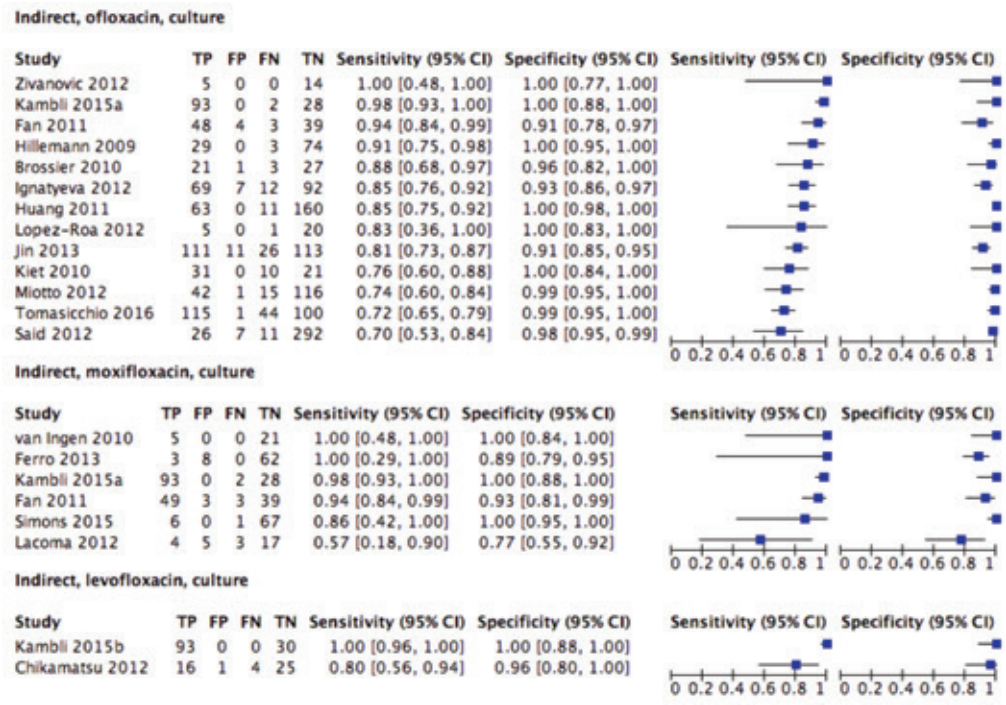
The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). The individual studies are ordered by decreasing sensitivity. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

Indirect testing of the individual drugs within the class of fluoroquinolones is given in Figure 7. Thirteen studies involving 1927 participants evaluated the MTBDRs*l* assay for the detection of ofloxacin resistance by indirect testing. Sensitivity estimates ranged from 70% to 100% and specificity estimates ranged from 91% to 100%. The pooled sensitivity and specificity (95% CI) were 85.2% (78.5% to 90.1%) and 98.5% (95.6% to 99.5%) respectively.

Six studies involving 419 participants evaluated the MTBDRs*l* assay for detection of moxifloxacin resistance by indirect testing (Figure 7). Sensitivity estimates ranged from 57% to 100% and specificity estimates from 77% to 100%. The pooled sensitivity and specificity (95% CI) were 94.0% (82.2% to 98.1%) and 96.6% (85.2% to 99.3%) respectively.

Two studies were identified that use the MTBDRs*l* assay for detection of levofloxacin resistance by indirect testing. Sensitivity and specificity estimates (95% CI) were 80% (56% to 94%) and 96% (80% to 100%) for Chikamatsu 2012 and 100% (96% to 100%) and 100% (88% to 100%) for Kambli 2015b. A meta-analysis for detection of levofloxacin resistance was not performed.

Figure 7. Forest plots of MTBDRsl sensitivity and specificity for ofloxacin, moxifloxacin, and levofloxacin resistance, the test performed indirectly, phenotypic culture-based DST reference standard[23]



TP true positive; FP false positive; FN false negative; TN true negative; CI confidence interval

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). The individual studies are ordered by decreasing sensitivity. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

Based on analysis of all data, there was no evidence of a statistically significant difference in MTBDRs*l* accuracy for fluoroquinolone resistance between indirect and direct testing (smear-positive specimens) when using phenotypic culture-based DST as a reference standard (P values for differences in sensitivity and specificity of 0.932 and 0.333, respectively). Within-study comparisons were not possible because no studies performed MTBDRs*l* testing on specimens and isolates from the same patients.

Annex 7[24]. Table 1 shows the drug concentrations for ofloxacin, levofloxacin, and moxifloxacin used in culture-based DST in relation to the WHO-recommended critical concentrations.

[23] Ferro 2013 performed indirect agar proportion method on Middlebrook 7H10 agar with moxifloxacin tested at 2.0 ug/ml. Ferro noted that only one strain was a true false positive; the other seven strains are part of 15 strains that showed resistance to ciprofloxacin due to common *gyrA* gene-conferred resistance.

[24] Annex 7: Dug concentrations used in phenotypic culture-based DST for each included study. Available at: http://www.who.int/tb/areas-of-work/laboratory/policy_statements/en/

4.2.4 Indirect testing MTBDRsl version 2.0

For MTBDR*sl* version 2.0 the data were either too sparse or too heterogeneous to combine in a meta-analysis or to compare indirect and direct testing.

Three studies evaluated the MTBDR*sl* version 2.0 in 562 individuals, including 111 confirmed cases of TB with fluoroquinolone resistance by indirect testing on a culture of *M. tuberculosis* compared with a phenotypic culture-based DST reference standard. Sensitivity estimates ranged from 84% to 100% and specificity estimates ranged from 99% to 100% respectively.

Three studies reported indeterminate results for the MTBDR*sl* assay by indirect testing of a culture of *M. tuberculosis* with phenotypic culture-confirmed resistance to fluoroquinolones. Of 570 results, eight (1.4%) were indeterminate (six culture-based DST resistant and two culture-based DST susceptible).

4.3 Accuracy for the detection of SLID resistance

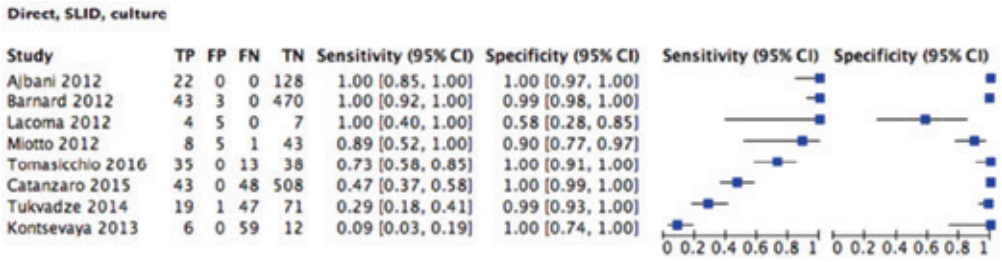
4.3.1 Direct testing MTBDRsl version 1.0

Eight studies evaluated the MTBDR*sl* assay in 1639 individuals, including 348 (21.2%) confirmed cases of SLID-resistant TB, by direct testing for detection of SLID resistance (Figure 8). For individual studies, sensitivity estimates ranged from 9% to 100% and specificity estimates ranged from 58% to 100%. The variability was explained in part by the use of different drugs, critical concentrations, and types of culture media in the reference standard and likely presence of *eis* resistance-conferring mutations in patients in Eastern European countries. The pooled sensitivity and specificity (95% CI) were 87.0% (38.1% to 98.6%) and 99.5% (93.6% to 100.0%) respectively.

Four studies reported indeterminate results for the MTBDR*sl* assay by direct testing of smear-positive sputum specimens compared with phenotypic culture-confirmed resistance to SLID. Of 1627 results, 219 (13.5%) were indeterminate (34 culture-based DST resistant and 165 culture-based DST susceptible and 20 did not have a culture-based DST result).

The indeterminate rates for direct testing for each smear-grade (smear-negative, scanty, 1+, 2+, 3+) were 76/180 (42.2%), 35/91 (38.5%), 47/213 (22.1%), 29/200 (14.5%), and 70/364 (19.2%), respectively.

Figure 8. Forest plots of MTBDRsl sensitivity and specificity for SLID resistance, test performed directly, phenotypic culture-based DST reference standard



TP true positive; FP false positive; FN false negative; TN true negative; CI confidence interval

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). The individual studies are ordered by decreasing sensitivity. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

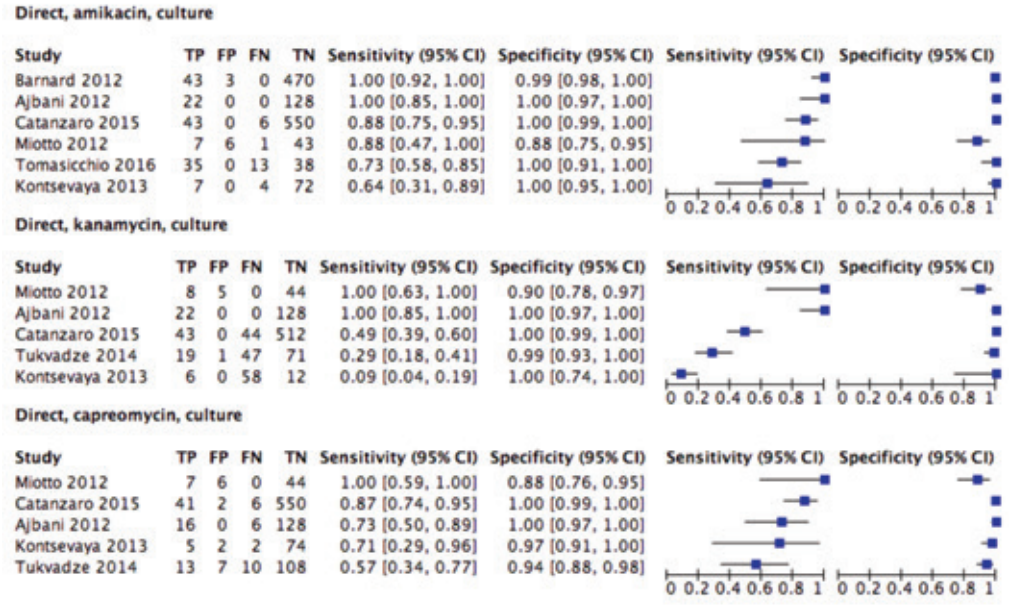
Sensitivity and specificity estimates for the detection of resistance to the individual SLID are given in Figure 9.

Six studies involving 1491 individuals evaluated the MTBDRs*l* assay for the detection of amikacin resistance by direct testing. Sensitivity estimates ranged from 64% to 100% and specificity estimates ranged from 88% to 100%. The pooled sensitivity and specificity (95% CI) were 91.9% (71.5% to 98.1%) and 99.9% (95.2% to 100.0%) respectively. Based on all data, there was no evidence of a statistically significant difference in accuracy between indirect and direct testing for amikacin resistance (P values for differences in sensitivity of specificity 0.338 and 0.213, respectively).

Five studies involving 1020 individuals evaluated the MTBDRs*l* assay for the detection of kanamycin resistance by direct testing. Sensitivity estimates ranged from 9% to 100% and specificity estimates ranged from 90% to 100%. The pooled sensitivity and specificity (95% CI) were 78.7% (11.9% to 99.0%) and 99.7% (93.8% to 100.0%) respectively. Based on all data, there was no evidence of a statistically significant difference in accuracy between indirect and direct testing for kanamycin resistance (P values for differences in sensitivity of specificity 0.836 and 0.445, respectively).

Five studies involving 1027 individuals evaluated the MTBDRs*l* assay for the detection of capreomycin resistance by direct testing. Sensitivity estimates ranged from 57% to 100% and specificity estimates ranged from 88% to 100%. The pooled sensitivity and specificity (95% CI) were 76.6% (61.1% to 87.3%) and 98.2% (92.5% to 99.6%) respectively. Based on all data, there was no evidence of a statistically significant difference in accuracy between indirect and direct testing for capreomycin resistance (P values for differences in sensitivity and specificity of 0.841 and 0.353, respectively).

Figure 9. Forest plots of MTBDRsl sensitivity and specificity for resistance to amikacin, kanamycin, and capreomycin, test performed directly, phenotypic culture-based DST reference standard



TP true positive; FP false positive; FN false negative; TN true negative; CI confidence interval

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). The individual studies are ordered by decreasing sensitivity. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

4.3.2 Direct testing MTBDRsl version 2.0

For MTBDR*sl* version 2.0 the data were either too sparse or too heterogeneous to combine in a meta-analysis or to compare indirect and direct testing.

Two studies evaluated MTBDR*sl* version 2.0 in 200 individuals, including 41 confirmed cases of TB with resistance to SLIDs by direct testing on smear-positive sputum specimens compared with a phenotypic culture-based DST reference standard. Sensitivity and specificity estimates (95% CI) were 62% (36% to 86%) and 91% (72% to 99%) for FIND 2016 and 89% (72% to 98%) and 90% (84% to 95%) for Tagliani 2015. The indeterminate rate for direct testing for smear-positive specimens was 1/201 results (0.5%) (one culture-DST susceptible) on a scanty smear-positive specimen.

Two studies evaluated MTBDR*sl* version 2.0 in 24 individuals, including 11 confirmed cases of TB with resistance to SLID by direct testing on smear-negative sputum specimens compared with a phenotypic culture-based DST reference standard. Sensitivity and specificity estimates (95% CI) were 83% (36% to 100%) and 78% (40% to 97%) for FIND 2016 and 80% (28% to 99%) and 100% (40% to 100%) for Tagliani 2015.

The indeterminate rate for direct testing for smear-negative sputum specimens was 6/30 results (20.0%) (three culture-DST resistant and three culture-DST susceptible).

4.3.3 Indirect testing MTBDRsl version 1.0

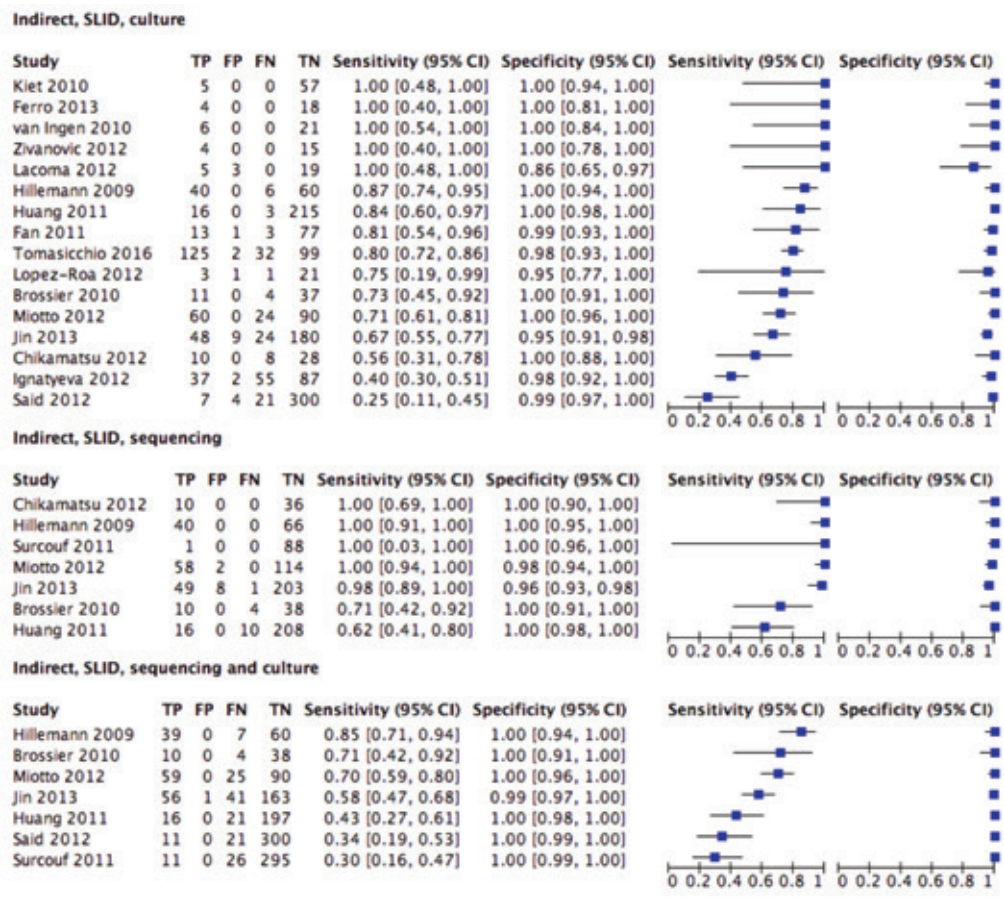
Sixteen studies evaluated MTBDR*sl* in 1921 individuals, including 575 (29.9%) confirmed cases of SLID-resistant TB evaluated MTBDR*sl* by indirect testing on a culture of *M. tuberculosis* compared with a phenotypic culture-based DST reference standard (Figure 10). For individual studies, sensitivity estimates ranged from 25% to 100% and specificity estimates ranged from 86% to 100%. The pooled sensitivity and specificity (95% CI) were 76.5% (63.3% to 86.0%) and 99.1% (97.3% to 99.7%) respectively.

Ten studies reported indeterminate results for the MTBDR*sl* assay by indirect testing of a culture of *M. tuberculosis* with phenotypic culture-confirmed resistance to SLIDs. Of 1316 results, seven (0.5%) were indeterminate (two culture-based DST resistant and five culture-based DST susceptible).

Seven studies evaluated the MTBDR*sl* assay in 962 individuals compared with a sequencing reference standard (Figure 10). Sensitivity estimates ranged from 62% to 100% and specificity estimates ranged from 96% to 100%. The pooled sensitivity and specificity (95% CI) were 97.0% (77.0% to 99.7%) and 99.5% (94.5% to 100.0%) respectively. There was evidence of a statistically significant higher sensitivity using sequencing as the reference standard compared with phenotypic culture-based DST (P value of 0.034), but not for specificity (P value of 0.456).

Seven studies evaluated the MTBDR*sl* assay in 1491 individuals compared with a composite reference standard of sequencing and phenotypic culture-based DST combined (Figure 10). Sensitivity estimates ranged from 30% to 85% and specificity estimates ranged from 99% to 100%. The pooled sensitivity and specificity (95% CI) were 61.3% (45.8% to 74.8%) and 99.9% (99.0% to 100.0%) respectively. There was no evidence of a statistically significant difference in MTBDR*sl* accuracy using culture and sequencing as the reference standard compared with phenotypic culture-based DST (P values for differences in sensitivity and specificity of 0.458 and 0.203, respectively).

Figure 10. Forest plots of MTBDRsl sensitivity and specificity for SLID resistance, test performed indirectly, different reference standards



TP true positive; FP false positive; FN false negative; TN true negative; CI confidence interval

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). The individual studies are ordered by decreasing sensitivity. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

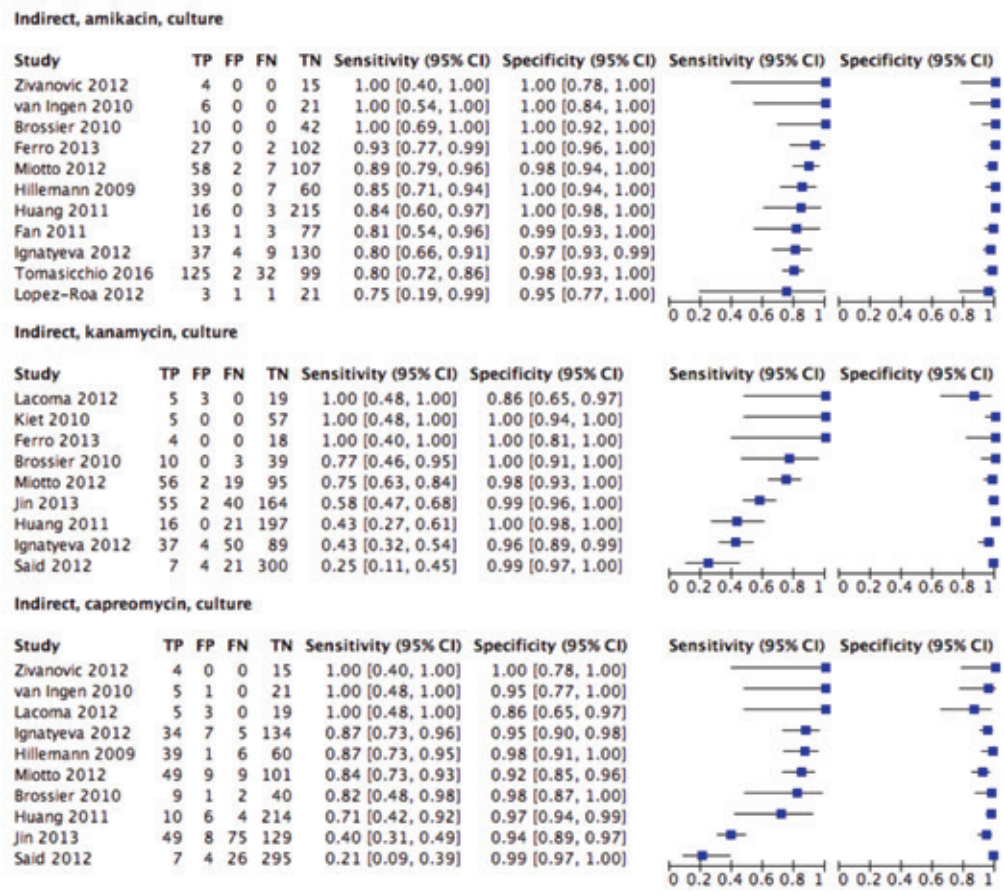
Sensitivity and specificity estimates for MTBDRsl by indirect testing of the individual SLIDs is given in Figure 11.

Eleven studies involving 1301 individuals evaluated the MTBDRsl assay for the detection of amikacin resistance by indirect testing. Sensitivity estimates ranged from 75% to 100% and specificity estimates ranged from 95% to 100%. The pooled sensitivity and specificity (95% CI) were 84.9% (79.2% to 89.1%) and 99.1% (CI 97.6% to 99.6%) respectively.

Nine studies involving 1342 individuals evaluated the MTBDRsl assay for the detection of kanamycin resistance by indirect testing. Sensitivity estimates ranged from 25% to 100% and specificity estimates ranged from 86% to 100%, (nine studies, 1342 participants). The pooled sensitivity and specificity (95% CI) were 66.9% (44.1% to 83.8%) and 98.6% (96.1% to 99.5%) respectively.

Ten studies involving 1406 individuals evaluated the MTBDRs*l* assay for the detection of capreomycin resistance by indirect testing. Sensitivity estimates ranged from 21% to 100% and specificity estimates from 86% to 100%. The pooled sensitivity and specificity (95% CI) were 79.5% (58.3% to 91.4%) and 95.8% (93.4% to 97.3%) respectively.

Figure 11. Forest plots of MTBDRsl sensitivity and specificity for resistance to amikacin, kanamycin, and capreomycin, test performed indirectly, phenotypic culture-based DST reference standard



TP true positive; FP false positive; FN false negative; TN true negative; CI confidence interval

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). The individual studies are ordered by decreasing sensitivity. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

Based on analysis of all data, there was no evidence of a statistically significant difference in MTBDRs*l* accuracy for SLID resistance between indirect and direct testing when using phenotypic culture-based DST as a reference standard (P values for differences in sensitivity or specificity of 0.547 and 0.664, respectively). Within-study comparisons were not possible because no studies performed MTBDRs*l* testing on specimens and isolates from the same patients.

Annex 7[25]. Table 2 shows amikacin, kanamycin, and capreomycin drug concentrations used in culture-based DST in relation to the WHO-recommended critical concentrations.

4.3.4 Indirect testing MTBDRsl version 2.0

Three studies evaluated the MTBDR*sl* version 2.0 in 620 individuals, including 237 confirmed cases of TB with SLID resistance by indirect testing on a culture of *M. tuberculosis* compared with a phenotypic culture-based DST reference standard. Sensitivity estimates ranged from 72% to 89% and specificity estimates ranged from 90% to 99%.

Three studies reported indeterminate results for the MTBDR*sl* assay version 2.0 by indirect testing of a culture of *M. tuberculosis* with phenotypic culture-confirmed resistance to SLIDs. Of 631 results, eleven (1.7%) were indeterminate (ten culture-based DST resistant and one culture-basedDST susceptible).

4.4 Accuracy for the detection of XDR-TB

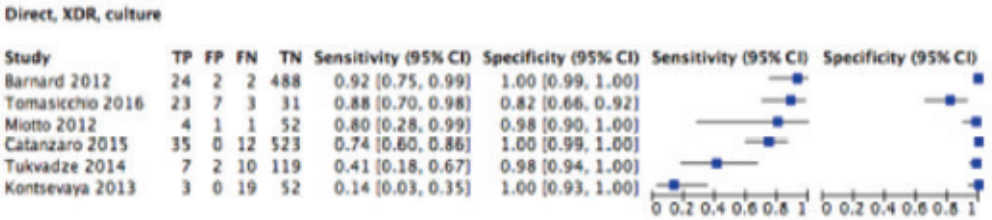
4.4.1 Direct testing MTBDRsl version 1.0

Six studies evaluated the MTBDR*sl* assay in 1420 individuals, including 143 (10.1%) confirmed cases of XDR-TB, by direct testing (Figure 12). For individual studies, sensitivity estimates ranged from 14% to 92% and specificity estimates ranged from 82% to 100%. The variability was partly explained by the use of different drugs, critical concentrations, and types of culture media in the reference standard and likely presence of *eis* mutation in patients in Eastern European countries. The pooled sensitivity and specificity (95% CI) were 69.4% (38.8% to 89.0%) and 99.4% (95.0% to 99.3%) respectively.

Seven studies reported indeterminate results for the MTBDR*sl* assay by direct testing of smear-positive sputum specimens compared with phenotypic culture-confirmed XDR-TB. Of 1665 results, 224 (13.5%) were indeterminate (37 culture-based DST resistant, 186 culture-based DST susceptible and 11 did not have a culture-based DST result).

The indeterminate rates for direct testing for each smear-grade (smear-negative, scanty, 1+, 2+, 3+) were 81/183 (44.2%), 39/186 (21.0%), 53/225 (23.5%), 33/301 (11.0%), and 82/177 (46.3%), respectively.

Figure 12 Forest plots of MTBDRsl sensitivity and specificity for detection of XDR-TB, the test performed directly, culture-based DST reference standard.



TP true positive; FP false positive; FN false negative; TN true negative; CI confidence interval

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). The individual studies are ordered by decreasing sensitivity. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

[25] Annex 7: Dug concentrations used in phenotypic culture-based DST for each included study. Available at: http://www.who.int/tb/areas-of-work/laboratory/policy_statements/en/

4.4.2 Direct testing MTBDRsl version 2.0

For MTBDR*sl* version 2.0 the data were either too sparse or too heterogeneous to combine in a meta-analysis or to compare indirect and direct testing.

Two studies evaluated MTBDR*sl* version 2.0 in 199 individuals, including 20 confirmed cases of XDR-TB by direct testing on smear-positive sputum specimens compared with a phenotypic culture-based DST reference standard. Sensitivity and specificity estimates (95% CI) were 100% (54% to 100%) and 100% (88% to 100%) for FIND 2016 and 79% (49% to 95%) and 97% (93% to 99%) for Tagliani 2015. There were no indeterminate results reported for direct testing for smear-positive specimens.

Two studies evaluated MTBDR*sl* version 2.0 in 22 individuals, including 5 confirmed cases of XDR-TB by direct testing on smear-negative sputum specimens compared with a phenotypic culture-based DST reference standard. Sensitivity and specificity estimates (95% CI) were 100% (29% to 100%) and 90% (55% to 100%) for FIND 2016 and 50% (1% to 99%) and 100% (59% to 100%) for Tagliani 2015. The indeterminate rate for direct testing for smear-negative specimens was 9/31 results (29.0%) (seven culture-based DST resistant and two culture-based DST susceptible).

4.4.3 Indirect testing MTBDRsl version 1.0

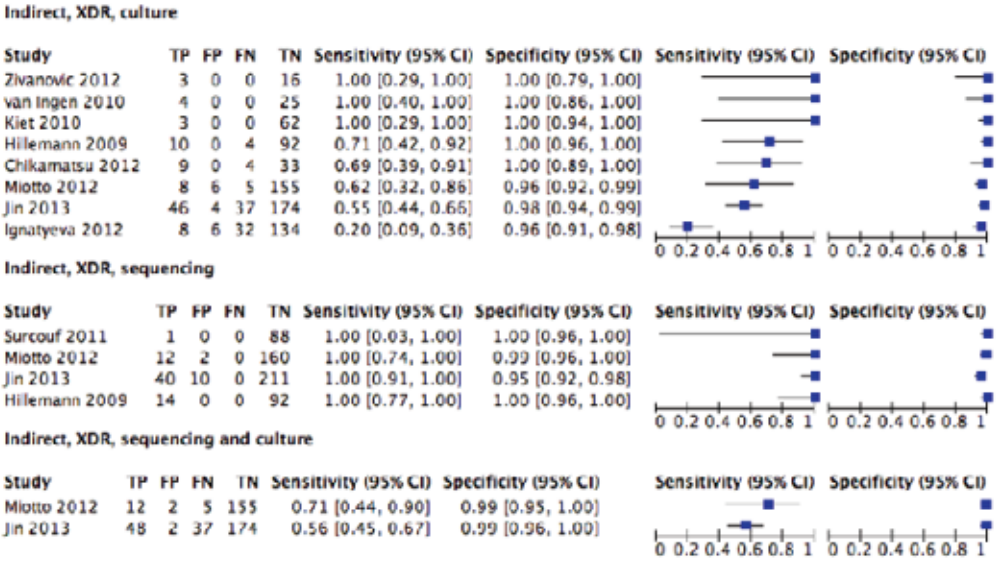
Eight studies evaluated MTBDR*sl* in 880 individuals, including 173 (19.7%) confirmed cases of XDR-TB evaluated MTBDR*sl* by indirect testing on a culture of *M. tuberculosis* compared with a phenotypic culture-based DST reference standard (Figure 13). For individual studies, sensitivity estimates ranged from 20% to 100% and specificity estimates ranged from 96% to 100%. The pooled sensitivity and specificity (95% CI) were 70.9% (42.9% to 88.8%) and 98.8% (96.1% to 99.6%) respectively.

Six studies that reported indeterminate results for the MTBDR*sl* results by indirect testing of a culture of *M. tuberculosis* with phenotypic culture-based DST confirmed XDR-TB. Of 554 results, only one (0.2%) was indeterminate.

Four studies evaluated the MTBDR*sl* assay in 630 individuals compared with a sequencing reference standard (Figure 13). Sensitivity estimates were 100% in all four studies and specificity estimates ranged from 95% to 100%. The pooled sensitivity and specificity (95% CI) were 100.0% (94.6% to 100.0%) and 97.9% (96.3% to 98.8%) respectively. The statistical difference between sensitivity and specificity estimates comparing sequencing and phenotypic culture-based DST reference standards could not be determined as the estimates were derived using different models.

Two studies evaluated the MTBDR*sl* assay in 435 individuals compared with a composite reference standard of sequencing and phenotypic culture-based DST combined (Figure 13). Sensitivity and specificity estimates (95% CI) were 56% (45% to 67%) and 99% (96% to 100%) for Jin 2013 and 71% (44% to 90%) and 99% (95% to 100%) for Miotto 2012. A meta-analysis could not be performed.

Figure 13 Forest plots of MTBDRsl sensitivity and specificity for detection of XDR-TB, test performed indirectly, different reference standards



TP true positive; FP false positive; FN false negative; TN true negative; CI confidence interval

The figure shows the estimated sensitivity and specificity of each study (square) and its 95% CI (horizontal line). The individual studies are ordered by decreasing sensitivity. Values for test results are the number of each type of result (true positive, false positive, false negative, true negative).

4.4.4 Indirect testing MTBDRsl version 2.0

Three studies evaluated the MTBDRs*l* version 2.0 in 621 individuals, including 88 confirmed cases of XDR-TB by indirect testing on a culture of *M. tuberculosis* compared with a phenotypic culture-based DST reference standard. Sensitivity estimates ranged from 75% to 80% and specificity estimates ranged from 91% to 100%.

Three studies reported indeterminate results for the MTBDRs*l* assay version 2.0 by indirect testing of a culture of *M. tuberculosis* with phenotypic culture-based DST confirmed XDR-TB. Of 621 results, eleven (1.6%) were indeterminate.

5. Summary of evidence to recommendations

In patients with RR-TB or MDR-TB, a positive SL-LPA result for fluoroquinolone resistance (as a class) or SLID resistance (as a group) can be treated with confidence. The diagnostic accuracy of SL-LPA are similar when performed directly on sputum specimens or indirectly on cultured isolates of *M. tuberculosis*.

Phenotypic resistance to ofloxacin and levofloxacin is highly correlated with resistance conferring mutations detected by SL-LPA. However, uncertainty remains about the susceptibility to moxifloxacin and gatifloxacin for such strains with specific mutations in the *gyr*A and *gyr*B genes.

SL-LPA have high specificity for the detection of resistance conferring mutations in the *rrs* gene to all of the SLIDs (including kanamycin, amikacin and capreomycin). However, mutations associated with the *eis* promoter region correlate with phenotypic resistance to kanamycin only.

Given the confidence in a positive result and the ability of the test to provide rapid results, the GDG felt that SL-LPA may be considered for use as an initial test for resistance to the fluoroquinolones and SLIDs. However, when the test shows a negative result, phenotypic culture-based DST may be necessary, especially in settings with a high pre-test probability for resistance to either fluoroquinolones and or SLIDs.

The use of SL-LPA in routine care should improve the time to the diagnosis of fluoroquinolone and SLIDs especially when used for the direct testing of a sputum specimens of patients with confirmed rifampicin-resistant TB or MDR-TB. Early detection of drug resistance should allow for the earlier initiation of appropriate patient therapy and improved patient health outcomes.

Overall, the test performs well in the direct testing of sputum specimens from patients with confirmed rifampicin-resistant or MDR-TB although the indeterminate rate is higher when testing smear-negative sputum specimens compared with smear-positive sputum specimens. When the MTBDR*sl* assay is used in the direct testing of smear-negative sputum specimens from a population of patients with confirmed drug-resistant TB, there may be up to 44% indeterminate results detected (less with version 2.0, although very limited data) that would require repeat or additional testing. However, if the same test were to be applied to the testing of smear-negative sputum specimens from patients without confirmed TB or drug-resistant TB (i.e. patients suspected of having drug-resistant TB), the indeterminate rate for the test would be significantly higher.

Given the test's sensitivity and specificity when SL-LPA are done directly on sputum, the GDG felt that SL-LPA can be used for the testing of all sputum specimens from patients with confirmed rifampicin-resistant TB or MDR-TB, irrespective of whether the microscopy result is positive or negative.

Online Annex 5[26] contains the GRADE summary of findings tables 1 to 6 which summarize the review findings for direct and indirect testing of the MTBDR*sl* (version 1.0) assay by applying the results to a hypothetical cohort of 1000 individuals with RR-TB or MDR-TB thought to have additional resistance to fluoroquinolones or SLIDs, or to both fluoroquinolones and SLIDs.

5.1 SL-LPA for the detection of resistance to fluoroquinolones

By direct testing, SL-LPA will detect 86% of patients with fluoroquinolone resistance and will rarely give a positive result for those without resistance. In a population of 1000 patients, where 150 (15% pre-test probability) have fluoroquinolone resistance, SL-LPA will correctly identify 129 people with fluoroquinolone resistance and miss 21 patients. In this same population, where 850 patients do not have fluoroquinolone resistance, the test will correctly classify 838 patients and misclassify 12.

[26] Annex 5: GRADE summary of findings tables.
Available at: http://www.who.int/tb/areas-of-work/laboratory/policy_statements/en/

The GDG felt that both false-positive and false-negative fluoroquinolone resistance results may contribute harms to the patient. The consequences for the 12 patients wrongly diagnosed (FP) with fluoroquinolone resistance were likely patient anxiety, possible delays in further diagnostic evaluation, and prolonged and unnecessary treatment with drugs that may have additional adverse effects. The consequences for the 21 false negative (FN) results for patients in the hypothetical scenario of 1000 patients were a potential increased risk of patient morbidity and mortality, and continued community transmission of drug-resistant TB. The harm for these patients with resistance to fluoroquinolones not detected by SL-LPA could be lessened as these patients could benefit from a MDR-TB regimen that may include a later generation fluoroquinolone such as moxifloxacin or gatifloxacin. Conventional phenotypic culture-basesd DST should be used in the follow-up evaluation of patients with a negative result especially in settings with a high pre-test probability for resistance to fluoroquinolones.

Given the speed for performing the test and considering the balance of harms and benefits the GDG made a conditional recommendation in favour of using SL-LPA as the initial test to detect fluoroquinolone resistance among patients with confirmed rifampicin-resistant TB or MDR-TB (Annex 6[27]).

5.2 SI-LPA for the detection of resistance to SLIDs

By direct testing, SL-LPA will detect 87% of patients with SLID resistance and rarely give a positive result for patients without resistance. In a population of 1000 patients, where 150 (15% pre-test probability) have SLID resistance, SL-LPA will correctly identify 131 patients with second-line injectable drug resistance and miss 19. In this same population, where 850 do not have second-line injectable drug resistance, the test will correctly classify 846 patients and misclassify four.

The GDG felt that both false-positive and false-negative SLID resistance results may contribute harms to the patient. The consequences for the 4 patients wrongly diagnosed (FP) with SLID resistance were likely patient anxiety, possible delays in further diagnostic evaluation, and prolonged and unnecessary treatment with drugs that may have additional adverse effects. The consequences for the 19 false negative (FN) results for patients in the hypothetical scenario of 1000 persons were a potential increased risk of patient morbidity and mortality, and continued community transmission of drug-resistant TB. The harm for these patients with resistance to SLID not detected by SL-LPA may be the initiation of an MDR-TB regimen which would include an injectable agent with doubtful efficacy. Conventional phenotypic DST should be used in the follow-up evaluation of patients with a negative result especially in settings with a high pre-test probability for resistance to SLIDs.

Given the speed for performing the test and considering the balance of harms and benefits the GDG made a conditional recommendation in favour of using SL-LPA as the initial test to detect SLID resistance among patients with confirmed rifampicin-resistant TB or MDR-TB (Annex 6[23]).

5.3 SL-LPA for the detection of XDR-TB

By direct testing, SL-LPA will 69% of patients with XDR-TB and rarely give a positive result for patients without resistance. In a population of 1000 patients, where 100 (10% pre-test probability) have XDR-TB, SL-LPA will correctly identify 69 patients with XDR-TB and miss 31. In this same population, where 900 do not have XDR-TB, the test will correctly classify 891 people and misclassify nine patients.

[27] Annex 6: GRADE evidence to recommendations tables.
Available at: http://www.who.int/tb/areas-of-work/laboratory/policy_statements/en/

The GDG felt that both false-positive and false-negative XDR-TB results may contribute harms to the patient. The consequences for the 9 patients wrongly diagnosed (FP) as XDR-TB were likely patient anxiety, possible delays in further diagnostic evaluation, and prolonged and unnecessary treatment with drugs that may have additional serious adverse effects. The consequences for the 31 false negative (FN) results in the hypothetical scenario of 1000 patients were a potential increased risk of patient morbidity and mortality, and continued community transmission of drug-resistant TB. The harm for these patients without XDR-TB detected by SL-LPA may be the initiation of a MDR-TB regimen with doubtful efficacy.

The SL-LPA have varying accuracy for the detection of resistance conferring mutations to the fluoroquinolones and to the SLIDs which combined lower the overall accuracy for the detection of XDR-TB. As a consequence, the GDG decided that the assay can be used for the diagnosis of XDR-TB while acknowledging that the diagnostic accuracy is sub-optimal. The GDG did feel that the assay could be used for surveillance of XDR-TB given statistical approaches to adjust for lower sensitivity and specificity during surveillance studies.

6. WHO Policy recommendations

Given the GRADE evidence assessment and considering the relative benefits and harms associated with the use of SL-LPA to test patients with confirmed rifampicin-resistant TB or MDR-TB, WHO recommends that:

For patients with confirmed rifampicin-resistant TB or MDR-TB, SL-LPA may be used as the initial test, instead of phenotypic culture-based DST, to detect resistance to fluoroquinolones

(Conditional recommendation; moderate certainty in the evidence for test accuracy for direct testing of sputum specimens; low certainty in the evidence for test accuracy for indirect testing of *Mycobacterium tuberculosis* cultures).

For patients with confirmed rifampicin-resistant TB or MDR-TB, SL-LPA may be used as the initial test, instead of phenotypic culture-based DST, to detect resistance to the second-line injectable drugs

(Conditional recommendation; low certainty in the evidence for test accuracy for direct testing of sputum specimens; very low certainty in the evidence for test accuracy for indirect testing of *Mycobacterium tuberculosis* cultures).

Remarks

a. These recommendations apply to the use of SL-LPA for testing sputum specimens (direct testing) and cultured isolates of *M. tuberculosis* (indirect testing) from both pulmonary and extrapulmonary sites. Direct testing on sputum specimens allows for the earlier initiation of appropriate treatment;

b. These recommendations apply to the direct testing of sputum specimens from rifampicin-resistant TB or MDR-TB, irrespective of the smear status, while acknowledging that the indeterminate rate is higher when testing smear-negative sputum specimens compared with smear-positive sputum specimens;

c. These recommendations apply to the diagnosis of XDR-TB while acknowledging that the accuracy for detecting resistance to the fluoroquinolones and to the SLIDs differs and hence the accuracy of a diagnosis of XDR-TB overall is reduced ;

d. These recommendations do not eliminate the need for conventional phenotypic DST capacity which will be necessary to confirm resistance to other drugs and to monitor the emergence of additional drug resistance;

e. Conventional phenotypic DST can still be used in the evaluation of patients with negative SL-LPA results, particularly in populations with a high pre-test probability for resistance to fluoroquinolones and/or SLID;

f. These recommendations apply to the use of SL-LPA in children with confirmed rifampicin-resistant TB or MDR-TB based on the generalisation of data from adults;

g. Resistance conferring mutations detected by SL-LPA are highly correlated with phenotypic resistance to ofloxacin and levofloxacin. However, the correlation of these mutations with phenotypic resistance to moxifloxacin and gatifloxacin is unclear and the inclusion of moxifloxacin or gatifloxacin in a MDR-TB regimen is best guided by phenotypic DST results;

h. Resistance conferring mutations detected by SL-LPA are highly correlated with phenotypic resistance to SLID and are an indication to use a MDR-TB regimen which is appropriately strengthened;

i. Given high specificity for detecting resistance to fluoroquinolones and SLID the positive results of SL-LPA could be used to guide the implementation of appropriate infection control precautions.

7. Implementation considerations

The SL-LPA should only be used to test specimens from patients with confirmed rifampicin-resistance TB and/or MDR-TB. Adoption of SL-LPA does not eliminate the need for conventional culture and DST capability. Despite good specificity of SL-LPA for the detection of resistance to fluoroquinolones and the SLIDs, culture and phenotypic DST is required to completely exclude resistance to these drug classes as well as to other second-line drugs. The following implementation considerations apply:

- SL-LPA cannot determine resistance to individual drugs in the class of fluoroquinolones. Resistance conferring mutations detected by SL-LPA are highly correlated with phenotypic resistance to ofloxacin and levofloxacin. However, the correlation of these mutations with phenotypic resistance to moxifloxacin and gatifloxacin is unclear and the inclusion of moxifloxacin or gatifloxacin in a MDR-TB regimen is best guided by phenotypic DST results;
- Mutations in some regions (e.g., the *eis* promoter region) may be responsible for causing resistance to one drug in a class (group) more than other drugs within that class (group). For example, the *eis* C14T mutation is associated with kanamycin resistance in strains from Eastern Europe;
- SL-LPA should be used in the direct testing of sputum specimens irrespective of whether samples are smear-negative or smear-positive;
- SL-LPA are designed to detect TB and resistance to fluroquinolones and second-line injectable drugs from sputum samples. Other respiratory samples (e.g. bronchoalveolar lavage and gastric aspirates) or extrapulmonary samples (tissue samples, CSF or other body fluids) have not been adequately evaluated;
- Culture and phenotypic DST plays a critical role in the monitoring of patients' response to treatment and for detecting additional resistance to second-line drugs during treatment;
- SL-LPA are suitable for use at the central or national reference laboratory level; they have the potential to be used at the regional level if the appropriate infrastructure can be ensured (three separate rooms are required);
- All patient identified by SL-LPA should have access to appropriate treatment and ancillary medications.

7.1 Plans for disseminating the WHO policy guidance on the use of SL-LPA

This WHO policy guidance will be published online (**http://www.who.int/tb/areas-of-work/laboratory/policy_statements/en/**) and disseminated through WHO/GTB Department listserves to all WHO Regional and Country Offices, Member States, the Global Laboratory Initiative, the TB/HIV Working Group and New Diagnostics Working Groups of Stop TB Partnership, donors, technical agencies and other stakeholders.

8. Research needs

Current recommendations on SL-LPA should not prevent or restrict further research on new rapid molecular DST tests, especially for assays that can be used as close as possible to where patients are initially diagnosed with RR-TB and MDR-TB and where treatment can be initiated. Further operational research on SL-LPA should focus on the following priorities:

- Develop improved understanding of the correlation between the detection of resistance-conferring mutations with phenotypic DST results and with patient outcomes;
- Develop improved knowledge of the presence of specific mutations detected with SL-LPA correlated with minimum inhibitory concentrations for individual drugs within the class of fluoroquinolones and group of SLIDs;
- Determine the limit of detection of SL-LPA for the detection of heteroresistance;
- Gather more evidence on the impact of MTBDR*sl* on appropriate MDR-TB treatment initiation and mortality;
- Strongly encourage that future studies follow the recommendations in the Standards for Reporting Diagnostic Accuracy (STARD)[28] statement to improve the quality of reporting;
- Perform country-specific cost-effectiveness and cost-benefit analyses of the use of SL-LPA in different programmatic settings.

[28] STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies BMJ 2015; 351 Available at: http://www.bmj.com/content/351/bmj.h5527

9. References to studies for the review of the diagnostic accuracy of SL-LPA

1. Ajbani K, Nikam C, Kazi M, Gray C, Boehme C, Balan K, et al. Evaluation of genotype MTBDR*s*/ assay to detect drug resistance associated with fluoroquinolones, aminoglycosides and ethambutol on clinical sediments. PLoS One 2012;7(11):e49433.

2. Barnard M, Warren R, Gey Van Pittius N, van Helden P, Bosman M, Streicher E, et al. Genotype MTBDR*s*/ line probe assay shortens time to diagnosis of extensively drug-resistant tuberculosis in a high-throughput diagnostic laboratory. American Journal of Respiratory and Critical Care Medicine 2012;186(12):1298-305.

3. Brossier F, Veziris N, Aubry A, Jarlier V, Sougakoff W. Detection by GenoType MTBDR*s*/ test of complex mechanisms of resistance to second-line drugs and ethambutol in multidrug-resistant Mycobacterium tuberculosis complex isolates. Journal of Clinical Microbiology 2010;48(5):1683–9.

4. Catanzaro A, Rodwell TC, Catanzaro DG, Garfein RS, Jackson RL, Seifert M, et al. Performance Comparison of Three Rapid Tests for the Diagnosis of Drug-Resistant Tuberculosis. PLoS One: 2015;10(8):e0136861.

5. Chikamatsu K, Aono A, Yamada H, Mitarai S. Evaluation of GenoType MTBDR*s*/ for testing resistance of Mycobacterium tuberculosis isolates to fluoroquinolone, aminoglycoside, and ethambutol. Kekkaku (Tuberculosis) 2012;87(10):641–7.

6. Fan QW, Guo J, Zhang HZ, Wu XY, Hu XN, Qian XQ, et al. The characteristics of drug resistant relevant genes in multidrug-resistant and extensively drug-resistant tuberculosis by fast molecular assay [Chinese]. Chinese Journal of Microbiology and Immunology 2011;31(12):1133-7.

7. Ferro B, García P, Nieto LM, van Soolingen D. Predictive value of molecular drug resistance testing of Mycobacterium tuberculosis isolates in Valle del Cauca, Colombia. Journal of Clinical Microbiology 2013;51(7):2220-4.

8. FIND. Multi-country evaluation of MTBDR*s*/ version 2.0. 2015. (unpublished)

9. Hillemann D, Rüsch-Gerdes S, Richter E. Feasibility of the GenoType MTBDR*s*/ assay for fluoroquinolone, amikacin-capreomycin, and ethambutol resistance testing of Mycobacterium tuberculosis strains and clinical specimens. Journal of Clinical Microbiology 2009;47(6):1767–72.

10. Huang WL, Chi TL, Wu MH, Jou R. Performance assessment of the GenoType MTBDR*s*/ test and DNA sequencing for detection of second-line and ethambutol drug resistance among patients infected with multidrug-resistant Mycobacterium tuberculosis. Journal of Clinical Microbiology 2011;49(7):2502–8.

11. Ignatyeva O, Kontsevaya I, Kovalyov A, Balabanova Y, Nikolayevskyy V, Toit K, et al. Detection of resistance to second-line antituberculosis drugs by use of the genotype MTBDR*s*/ assay: a multi-center evaluation and feasibility study. Journal of Clinical Microbiology 2012;50(5):1593-7.

12. Jin J, Shen Y, Fan X, Diao N, Wang F, Wang S, et al. Underestimation of the resistance of Mycobacterium tuberculosis to second-line drugs by the new GenoType MTBDR*s*/ test. Journal of Molecular Diagnostics 2013;15(1):44–50.

13. Kambli P, Ajbani K, Sadani M, Nikam C, Shetty A, Udwadia Z, et al. Correlating Minimum Inhibitory Concentrations of ofloxacin and moxifloxacin with gyrA mutations using the genotype MTBDR*s*/ assay. Tuberculosis (Edinb) 2015;95(2):137-41.

14. Kambli P, Ajbani K, Nikam C, Khillari A, Shetty A, Udwadia Z, Georghiou SB, Rodwell T, Catanzaro A, Rodrigues C. Determination of MICs of levofloxacin for Mycobacterium tuberculosis with gyrA mutations. The International Journal of Tuberculosis and Lung Disease 2015;19:1227-1229.

15. Kiet VS, Lan NT, An DD, Dung NH, Hoa DV, van Vinh Chau N, et al. Evaluation of the MTBDRs/ test for detection of second-line-drug resistance in Mycobacterium tuberculosis. Journal of Clinical Microbiology 2010;48(8):2934–9.

16. Kontsevaya I, Mironova S, Nikolayevskyy V, Balabanova Y, Mitchell S, Drobniewski F. Evaluation of two molecular assays for rapid detection of Mycobacterium tuberculosis resistance to fluoroquinolones in high-tuberculosis and -multidrug-resistance settings. Journal of Clinical Microbiology 2011;49(8):2832–7.

17. Kontsevaya I, Ignatyeva O, Nikolayevskyy V, Balabanova Y, Kovalyov A, Kritsky A, et al. Rapid diagnosis of extensively-drug resistant TB in HIV co-infected patients: diagnostic accuracy of the GenoType(R) MTBDRs/ assay. Journal of Clinical Microbiology 2013;51(1):243-8.

18. Lacoma A, García-Sierra N, Prat C, Maldonado J, Ruiz-Manzano J, Haba L, et al. GenoType MTBDRs/ for molecular detection of second-line-drug and ethambutol resistance in Mycobacterium tuberculosis strains and clinical samples. Journal of Clinical Microbiology 2012;50(1):30–6.

19. López-Roa P, Ruiz-Serrano MJ, Alcalá L, García-Escribano Ráez N, García de Viedma D, Bouza E. Susceptibility testing to second-line drugs and ethambutol by GenoType MTBDRs/ and Bactec MGIT 960 comparing with agar proportion method. Tuberculosis 2012;92(5):417–21.

20. Miotto P, Cabibbe AM, Mantegani P, Borroni E, Fattorini L, Tortoli E, et al. GenoType MTBDRs/ performance on clinical samples with diverse genetic background. European Respiratory Journal 2012;40(3):690-8.

21. National Institute Communicable Diseases, South Africa. Validation study MTBDRs/ version 2.0. 2015. (Unpublished)

22. Said HM, Kock MM, Ismail NA, Baba K, Omar SV, Osman AG, et al. Evaluation of the GenoType® MTBDRs/ assay for susceptibility testing of second-line anti-tuberculosis drugs. International Journal of Tuberculosis and Lung Disease 2012;16(1):104-9.

23. Simons SO, van der Laan T, de Zwaan R, Kamst M, van Ingen J, Dekhuijzen PN, et al. Molecular drug susceptibility testing in the Netherlands: performance of the MTBDRplus and MTBDRs/ assays. Int J Tuberc Lung Dis 2015;19(7):828-33.

24. Surcouf C, Heng S, Pierre-Audigier C, Cadet-Daniel V, Namouchi A, Murray A, et al. Molecular detection of fluoroquinolone-resistance in multi-drug resistant tuberculosis in Cambodia suggests low association with XDR phenotypes. BMC Infectious Diseases 2011;11:255-62.

25. Tagliani E, Cabibbe AM, Miotto P, Borroni E, Toro JC, Mansjö M, et al. Diagnostic Performance of the New Version (v2. 0) of GenoType MTBDRs/ Assay for Detection of Resistance to Fluoroquinolones and Second-Line Injectable Drugs: a Multicenter Study. Journal of clinical microbiology 2015;53(9):2961-9.

26. Tomasicchio M, Theron G, Pietersen E, Streicher E, Stanley-Josephs D, van Helden P, et al. The diagnostic accuracy of the MTBDRplus and MTBDRs/ assays for drug-resistant TB detection when performed on sputum and culture isolates. Scientific Reports 2016;6:doi:10.1038/srep17850.

27. Tukvadze N, Bablishvili N, Apsindzelashvili R, Blumberg HM, Kempker RR. Performance of the MTBDRs/ assay in Georgia. International Journal of Tuberculosis and Lung Disease 2014;18(2):233-9.

28.van Ingen J, Simons S, de Zwaan R, van der Laan T, Kamst-van Agterveld M, Boeree MJ, et al. Comparative study on genotypic and phenotypic second-line drug resistance testing of Mycobacterium tuberculosis complex isolates. Journal of Clinical Microbiology 2010;48(8):2749–53.

29. Živanović I, Vuković D, Dakić I, Stefanović G, Savić B. Detection of drug-resistant Mycobacterium tuberculosis strains isolated in Serbia by the GenoType MTBDRs/ assay. Archives of Biological Sciences 2012;64(4):1311-8.

10. Annexes

Annex 1. Meeting participants

Guideline Development Group (GDG) Members

Holger Schunemann – Chair of the GDG
McMaster University
Canada
schuneh@mcmaster.ca

Gavin Churchyard
The Aurum Institute NPC
South Africa
gchurchyard@auruminstitute.org

Daniela Maria Cirillo
HSR San Raffaele Scientific Institute (SRL Milan)
Italy
cirillo.daniela@hsr.it

Chris Coulter
Queensland Mycobacterium Reference Laboratory
Australia
Chris_Coulter@health.qld.gov.au

Greg Fox
University of Sydney
Australia
greg.fox@sydney.edu.au

Moses Joloba
National Reference Laboratory of the National TB and
Leprosy Programme
Uganda
mlj10@case.edu

James Posey
Centers for Disease Control
USA
hzp9@cdc.gov

Michael Rich
Partners in Health
United States of America
mrich@pih.org

Leen Rigouts
Prince Leopold Institute of Tropical Medicine
Belgium
LRigouts@itg.be

Thomas M Shinnick
Former Chair, Global Laboratory Initiative
United States of America
tms1@bellsouth.net

Rebecca Tadokera (unable to attend**)**
Human Sciences Research Council (HSRC)
South Africa
tadokerar@gmail.com

Marie Alice Telles
Tuberculosis Laboratory Independent Consultant for PAHO
Brazil
atelles.msh@gmail.com

Francis Varaine
Médecins Sans Frontières
France
Francis.VARAINE@paris.msf.org

Observers

Sevim Ahmedov
USAID
USA
sahmedov@usaid.gov

Emmanuelle Cambau
Groupe Hospitalier Lariboisière-Fernand Widal
France
emmanuelle.cambau@aphp.fr

Miranda Langendam
University of Amsterdam
The Netherlands
m.w.langendam@amc.uva.nl

Thomas Schön
Kalmar County Hospital and Linköping University
Sweden
tschon@hotmail.com

Belay Tessema
FIND
Switzerland Belay.Tessema@finddx.org

Annex 2. Members of the External Review Group

Heather Alexander
US Centers for Disease Control and Prevention
United States of America
drz5@cdc.gov

Martina Casenghi
Médecins Sans Frontières-Access Campaign
Martina.CASENGHI@geneva.msf.org

Kathleen England
KNCV tuberculosis foundation, Netherlands
kathleen.england@kncvtbc.org

Rumina Hassan
SRL Pakistan, Aga Khan University
Karachi, Pakistan
rumina.hasan@aku.edu

Nazir Ahmed Ismail
Candidate SRL, National Institute of Communicable Diseases, South Africa
naziri@ncid.ac.za

Richard Lumb
Adelaide SRL,
Adelaide, Australia
Richard.Lumb2@health.sa.gov.au

Beatrice Mutayoba
Ministry of Health and Social Welfare
Tanzania
beatricemutayoba@yahoo.com

Alaine Nyaruhirira
Management Sciences for Health,
Pretoria, South Africa
alainenyaruhirira@hotmail.com

Rebecca Tadokera
Human Sciences Research Council (HSRC)
South Africa
tadokerar@gmail.com

Annex 3. WHO Steering Group

Christopher Gilpin
WHO HQ, Global TB Programme
gilpinc@who.int

Alexei Korobitsyn
WHO HQ, Global TB Programme
korobitsyna@who.int

Karin Weyer
WHO HQ, Global TB Programme
weyerk@who.int

Fuad Mirzayev
WHO HQ, Global TB Programme
mirzayevf@who.int

Dennis Falzon
WHO HQ, Global TB Programme
falzond@who.int

Matteo Zignol
WHO HQ, Global TB Programme
zignolm@who.int

Annex 4. Declarations of Interests

None declared
Holger Schünemann (Chair)
Moses Joloba
James Posey
Francis Varraine

Declared, insignificant
Chris Coulter: Short-term consultancies for WHO (self, Total AUD<5,000). Research grant – TB transmission in Australia, whole genome sequencing study (NHMRC to Research collaboration, AUD<18,000). Laboratory support services to Papua New Guinea (Australian government to Brisbane, SRL, AUD 240,000).

Gregory Fox: Otsuka / Union Young Innovators' Award at the 2015 World Lung Health Conference: (USD 10,000 in airfare, accommodation and per diem for the IUATLD meeting).

Michael Rich: Employed in Partners in Health to work on clinical care guidelines and in the programmatic management of MDR-TB. Consulting on behalf of WHO. Conducting research on Bedaquiline and Delamanid as recipient of the UNITAID's endTB grant.

Thomas Shinnick: An employee of the United States Centres for Disease Control and Prevention (CDC). CDC supports travel and research related to his work on the laboratory services needed for tuberculosis control; represented CDC's positions on laboratory services needed for tuberculosis diagnosis, treatment, and control. Served on the Data and Safety Monitoring Board (DSMB) organized by Otsuka for the clinical trial of delamanid – no remuneration received.

Daniela Cirillo: Research on new diagnostics, including for Xpert MTB/RIF (FIND, EUR 17,000).

Maria Alice Telles: Consultant providing training on Xpert MTB/RIF (FIND, USD 4,000). Participation in a meeting on MGIT (BD, funding travel and per diem)

Gavin Churchyard: Research grant receivedto evaluate the national roll out of GeneXpert in South Afirca (Bill and Melinda Gates Foundation provided USD 11 million to the Aurum Institute, South Africa).

Leen Rigouts: Laboratory support to clinical trial C208-C209, phase II (Janssen, no personal renumeration, funding for Mycobacteriology Unit, Institute of Tropical Medicine, Antwerp). Involved in evaluation of the NIPRO/LPAs (PZA and second/line) (NIPRO, LPA kits provided by the company).

Declared, significant
None

Technical resource persons
Karen Steingart
Grant Theron

Observers
Belay Tessema: none declared

David Dolinger: Employment with commercial entity (USD 190,000 per year). Working with FIND on assessment of new diagnostics

Miranda Langendam: none declared

Emmanuelle Cambau: Reimbursement from ESCMID for participation in European Congress on Clinical Microbiology and Infectious Diseases

Thomas Schön: None declared

Sevim Ahmedov: Participation in the current meeting is covered by USAID

Online annexes available at http://www.who.int/tb/areas-of-work/laboratory/policy_statements/en/

Annex 5. GRADE summary of findings tables

Annex 6. Evidence to recommendations tables

Annex 7. Drug concentrations used in phenotypic culture-based DST for each included study

Annex 8. References to studies included and excluded from the review

THE
END TB
STRATEGY

World Health Organization

For further information, contact:
Global TB Programme
World Health Organization
Avenue Appia 20, CH-1211
Geneva-27, Switzerland
Information Resource Centre HTM/GTB:
Email: tbdocs@who.int
Website: www.who.int/tb

9 789241 510561