

pre-publication copy



**World Health
Organization**



**For research on
diseases of poverty**

UNICEF • UNDP • World Bank • WHO

**NON-COMMERCIAL CULTURE METHODS AND
MYCOBACTERIOPHAGE-BASED ASSAYS FOR RAPID SCREENING
OF PATIENTS AT RISK OF DRUG-RESISTANT TUBERCULOSIS**

EXPERT GROUP MEETING REPORT

This report contains the collective views of an international group of experts, and does not necessarily represent the decisions or the stated policy of the World Health Organization. Endorsement of a technology does not imply endorsement of any specific commercial product.

Geneva: 3 November 2009

TABLE OF CONTENTS

1.	Background	1
2.	Evidence base	3
2.1	Process	3
2.2	Systematic reviews and meta-analyses	3
2.3	Evaluation of the strength of the evidence base	4
3.	Results	5
3.1	General considerations	5
3.2	Microscopic observation of drug susceptibility (MODS) assay	8
3.2.1	Quality of evidence	8
3.2.2	Balance of desirable and undesirable effects	9
3.2.3	Values and preferences of patients	9
3.2.4	Costs and requirements	9
3.2.5	Added value and potential for scale-up	10
3.2.6	Results for direct application	11
3.2.7	Main research gaps identified	11
	FINAL RECOMMENDATION	11
3.3	Thin layer agar (TLA) assay	12
3.3.1	Quality of evidence	12
3.3.2	Balance of desirable and undesirable effects	13
3.3.3	Values and preferences of patients	13
3.3.4	Costs and requirements	13
3.3.5	Added value and potential for scale-up	13
3.3.6	Results for direct application	14
3.3.7	Main research gaps identified	14
	FINAL RECOMMENDATION	14
3.4	Colorimetric redox indicator (CRI) methods	15
3.4.1	Quality of evidence	16
3.4.3	Values and preferences of patients	16
3.4.4	Costs and requirements	16
3.4.5	Added value and potential for scale-up	17
3.4.6	Results for direct application only	17
3.4.7	Main research gaps identified	17
	FINAL RECOMMENDATION	17
3.5	Nitrate reductase assay (NRA)	18
3.5.1	Quality of evidence	19
3.5.2	Balance of desirable and undesirable effects	19
3.5.3	Values and preferences of patients	19
3.5.4	Costs and laboratory requirements	19
3.5.5	Added value and potential for scale-up	20
3.5.6	Results for direct application	20
3.5.7	Main research gaps identified	20
	FINAL RECOMMENDATION	20
3.6	Phage-based assays	21
3.6.1	Quality of evidence	23
3.6.2	Balance of desirable and undesirable effects	24
3.6.3	Values and preferences of patients	24
3.6.4	Costs and requirements	24
3.6.5	Added value and potential for scale-up	24
3.6.6	Results for direct application	25
3.6.7	Main research gaps identified	25
	FINAL RECOMMENDATION	25
4.	Implementation considerations	26
4.1	Specimen collection, storage and transport	26
4.2	Biosafety	26
4.3	Equipment, supplies and reagents	26
4.4	Human resources and training	27
4.6	Quality assurance	27
5.	Policy recommendations	28

NON-COMMERCIAL CULTURE METHODS AND MYCOBACTERIOPHAGE-BASED ASSAYS FOR RAPID SCREENING OF PATIENTS AT RISK OF DRUG-RESISTANT TUBERCULOSIS

1. Background

Early detection of drug resistance in tuberculosis (TB) allows the use of appropriate treatment regimens for patients, which has an important impact for improved TB control. The development of rapid methods for drug susceptibility testing (DST) is crucial due to increasing rates of multidrug-resistant tuberculosis (MDR-TB) worldwide and the emergence of extensively drug-resistant tuberculosis (XDR-TB), with very high reported HIV-associated mortality. Spread of drug resistant strains of *Mycobacterium tuberculosis* and the management of patients diagnosed with drug resistant disease is one of the most formidable obstacles faced by national tuberculosis control programmes, compounded by a critical lack of appropriate diagnostic tools and vastly inadequate laboratory capacity.

Conventional culture and DST methods require prolonged periods to confirm mycobacterial growth and detect drug resistance, during which time patients may be inappropriately treated, drug resistant strains may continue to spread, and amplification of resistance may occur. Rapid diagnosis of TB and drug resistance will therefore have obvious patient- as well as public health benefits, including better prognosis, increased survival, prevention of acquisition of further drug resistance, and reduced spread of drug resistant strains to vulnerable populations.

No single test currently satisfies all the demands of 'quick', 'cheap', and 'easy'. Commercially available liquid culture systems and molecular line probe assays for rapid detection of MDR-TB have been endorsed by WHO; however, due to their complexity and cost, as well as the need for sophisticated laboratory infrastructure, uptake has been limited in many resource-constrained settings. Several non-commercial culture and DST methods have been developed at the same time, aimed at use in laboratories that lack access to more sophisticated infrastructure and techniques. Among these methods, microscopic observation of drug susceptibility (MODS), thin layer agar (TLA), colorimetric redox indicator (CRI) methods, the nitrate reductase assay (NRA) and mycobacteriophage-based assays are most advanced and have shown initial promise as rapid, inexpensive methods. They were therefore the focus of this Expert Group meeting.

Microscopic observation of drug susceptibility (MODS) assay

The MODS assay is a microcolony method, based on direct inoculation of patient specimens to drug-free and drug-containing liquid media followed by microscopic examination of early culture growth. Growth of *M. tuberculosis* is identified by typical cord formation under an inverted light microscope. Concurrent growth on drug-free and drug-containing media facilitates direct DST of rifampicin and isoniazid: Growth in drug-free media indicates a positive culture; growth in both drug-free and drug-containing media indicates resistance.

Thin layer agar (TLA) assay

The TLA assay is a microcolony direct method on solid culture media using a standard light microscope to simultaneously detect *M. tuberculosis* complex and indicate isoniazid and rifampicin resistance. Plates with a thin layer of agar medium are incubated and examined microscopically on alternate days for the first two weeks and less frequently thereafter. Growth on drug-free media indicates a positive culture; growth on both drug-free and drug-containing media indicates resistance.

Colorimetric redox indicator (CRI) methods

CRI methods are indirect methods based on the reduction of a colored indicator added to liquid culture medium in a microtiter plate after *M. tuberculosis* has been exposed *in vitro* to different antibiotics and different drug concentrations. Resistance is detected by a change in color of the indicator, which is proportional to the number of viable mycobacteria in the medium. Among the different growth indicators used are the tetrazolium salts XTT (2,3-bis(2-methoxy-4-nitro-5-sulfophenyl)-2H-tetrazolium-5-carboxanilide and MTT(3(4,5-dimethylethyl-thiazol-2-yl)-2,5-diphenyltetrazoliumbromide), and the redox-indicators Alamar blue and resazurin.

Nitrate reductase assay (NRA)

The NRA is a solid culture technique based on the capacity of *M. tuberculosis* to reduce nitrate to nitrite, which is detected by adding a specific reagent (Griess reagent) to conventional Löwenstein-Jensen (LJ) medium into which 1 mg/ml of potassium nitrate (KNO₃) has been incorporated. The reduction of nitrate is detected by a coloured reaction. Resistance testing is done by inoculating drug-free and drug-containing media. Detection the coloured reaction on the drug-free medium alone indicates a positive culture and drug susceptibility; growth in both drug-free and drug-containing media indicates resistance. The NRA test can be used as a direct or indirect test.

Phage-based assays

Phage-based assays utilize bacteriophages to infect and detect the presence of viable *M. tuberculosis* in clinical specimens and culture isolates. Two main approaches have been developed: 1) amplification of phages after their infection of *M. tuberculosis*, followed by detection of progeny phages using sensor bacteria and measuring plaque formation, and 2) detection of light produced by luciferase reporter phages (LRP) after their infection of live *M. tuberculosis*. When these assays detect *M. tuberculosis* in drug-free specimens, but fail to detect *M. tuberculosis* in drug-containing specimens, the strains are classified as drug susceptible.

Several investigators have evaluated in-house amplification assays using D29 phages and their results have been included in the systematic reviews. Currently, there is one commercial phage-based test on the market, ie. the *FASTPlaque*[™] assay (Biotec Laboratories Limited, Ipswich, Suffolk, UK). This assay is based on infection of *M. tuberculosis* present in the specimen by specific mycobacteriophages that, after amplification and release, are allowed to infect a lawn of non-pathogenic organisms within an agar plate, resulting in plaque formation on the surface of the agar. Appearance of plaques is indicative of *M. tuberculosis* growth. The first generation test for detection of rifampin resistance, the *FASTPlaque* – RIF[™] or *FASTPlaque* – MDRi[™], was used only with *M. tuberculosis* isolates from cultures, ie. in indirect testing. This has now been replaced by the *FASTPlaque* – Response[™] which can be used as a direct test on patient specimens as well as an indirect test on *M. tuberculosis* isolates.

2. Evidence base

2.1 Process

The systematic, structured, evidence-based process for policy generation as developed recently by WHO was followed: The first step constituted a systematic review and meta-analysis of available data (published and unpublished), using standard methods appropriate for diagnostic accuracy studies. The second step involved the convening of an Expert Group to a) evaluate the strength of the evidence base; b) recommend operational and logistical considerations for mainstreaming such tools/approaches into national TB control programmes; and c) identify gaps to be addressed in future research. Based on the Expert Group findings, the third and final step involves WHO policy guidance on the use of these tools/approaches, presented to the WHO Strategic and Technical Advisory Group for TB (STAG-TB) for endorsement, and subsequent dissemination to WHO member states for implementation.

The Expert Group (Annex 1) consisted of researchers, clinicians, epidemiologists, end-users (programme and laboratory representatives), and evidence synthesis experts. In order to comply with current standards for evidence assessment in formulation of policy recommendations the GRADE system (www.gradeworkinggroup.org), recently adopted by WHO for all policy and guidelines development, was used. The GRADE approach, assessing both the quality of evidence and strength of recommendations, aims to provide a comprehensive and transparent approach for developing policy guidance. Started about 10 years ago to assess treatment interventions, the GRADE approach has recently been refined for diagnostics;¹ however, while the latter process shares the fundamental logic of recommendations for other interventions (notably treatment), it also presents unique challenges, most often due to study limitations related to a lack of data on patient-important outcomes and impact.

Randomised controlled trials of alternative diagnostic approaches represent the ideal study design for informing eventual policy decisions; however, no such studies are available for the methods assessed - a common problem of diagnostic interventions in general. Recognising that test results are surrogates for patient-important outcomes, the Expert Group evaluated test accuracy while also drawing inferences on the likely impact of these methods on patient outcomes, as reflected by false-negatives (ie. cases missed) or false-positives. In addition, the Expert Group considered the implications of each test or method for programmatic implementation, including laboratory infrastructure, human resources, interface between patients and laboratory services, diagnosis and initiation of treatment, costs to the health system and to patients, and research gaps.

2.2 Systematic reviews and meta-analyses

Systematic reviews and meta-analyses were commissioned by WHO for each method under evaluation. One of the standardised objectives was, for each assay/method, to perform a systematic review of available literature (published and unpublished), followed by a meta-analysis (where appropriate), on data examining the diagnostic accuracy and performance characteristics of each assay for the detection of drug resistance in *M. tuberculosis*. All systematic reviews and meta-analyses followed standard protocols, using predetermined

eligibility criteria for the primary analyses. Detailed methodology is described in individual systematic review reports available at www.who.int/tb/dots/laboratory/policy. Studies using both direct testing on patient specimens and indirect testing on culture isolates were included.

Study limitations were assessed by QUADAS (Quality Assessment of Diagnostic Accuracy Studies) criteria, a validated tool based on a checklist of 14 essential items.² Among the QUADAS criteria the following quality characteristics were considered especially important for the systematic reviews: (i) blinded interpretation of the test result with reference standard results and vice-versa, (ii) complete verification of test results with the reference standard, (iii) recruitment of patients/specimens either consecutively or randomly, and (iv) study design.

In addition to pooled estimates of the sensitivity and specificity of the assay/method against the reference standard, the following additional outcomes were extracted from included studies: turnaround time (defined as the time from specimen receipt or processing in the laboratory to availability of results within the laboratory) and contamination rates (defined as the proportion of specimens contaminated on first inoculation). Pooled contamination rate estimates were obtained by also including studies that evaluated the respective methods for diagnosis of TB in addition to drug-resistant TB.

Primary meta-analyses were limited to, and provided separately for, rifampin and isoniazid resistance detection. Subgroup analyses were performed for the type of specimen inoculated, ie. direct testing on patient specimens vs. indirect testing on culture isolates grown from patient specimens, where sufficient studies were identified.

Results from the systematic review reports are summarized below.

2.3 Evaluation of the strength of the evidence base

Evaluation followed the GRADE system for grading quality of evidence and strength of recommendations for diagnostic tests and strategies. The quality of evidence was graded by six criteria:

- *Study design*
 - a. Cross-sectional: Random or consecutive selection of patients/specimens at risk
 - b. Case-control: Selection of patients/specimens according to reference standard
- *Risk of bias (as reflected by the QUADAS tool)*
 - Representativeness of the sample to the at-risk target population
- *Directness*
 - Presence of direct evidence of impact on patient-important outcomes
- *Inconsistency*
 - Unexplained inconsistency in sensitivity or specificity estimates
- *Imprecision*
 - Wide confidence intervals for pooled sensitivity or specificity estimates

- *Publication/reporting bias*

Publications of research based on their nature and outcome, eg. studies showing poor performance not being published, language bias, etc.

As called for by GRADE, the Expert Group also considered for each method/assay the strength of the recommendation (strong or weak), based on a balance of effects (advantages weighed against disadvantages), assumed patient values and preferences, and costs (human resources, laboratory infrastructure requirements (including biosafety), equipment and consumables). In the absence of relevant studies, patient values and preferences were considered to be primarily reflected in the turnaround time (TAT) of the respective assays, assuming that the quicker the result is available, the higher the test would be valued by patients.

In addition to the GRADE system, the Expert Group also weighed the added value (if any) of the non-commercial culture and DST methods in settings which already have conventional WHO-endorsed technologies/methods available, the potential for scale-up/decentralization of the new methods beyond the national reference laboratory, the potential to replace liquid culture and DST, and the programmatic conditions required for implementation. For each of the methods/assays under evaluation a final recommendation was then arrived at by consensus, using the following categories: (i) recommended for policy development; (ii) recommended for policy development with restrictions; (iii) not recommended for policy development but with suggestions for further research; (iii) not recommended for policy development or further research.

3. Results

3.1 General considerations

The following considerations guided and premised the final recommendations:

1. Rapid drug susceptibility testing (DST) is primarily important for identifying patients at risk of multidrug-resistant TB (MDR-TB) as the first priority in a screening strategy. Once a patient has been identified as having MDR-TB, additional first- and second-line drug susceptibility results should be obtained in a second step, based on current WHO recommendations³ and available laboratory capacity;
2. Rifampicin resistance is a valid and reliable indicator/proxy of MDR-TB.³ Therefore, the most important performance indicators of the methods/assays under evaluation are those relating to rifampicin. All evaluations included, in addition to rifampin, estimates of accuracy for detecting resistance to isoniazid; however, little or no evidence exists for effectiveness of the assays/methods under evaluation for detecting resistance to other first- or second-line drugs.
3. With the exception of one phage-based assay, none of the assays and methods involves design-locked, standardized and quality-assured tests. This has consequences for

comparability between studies, and potentially for reproducibility when used outside the settings in which they were originally developed and evaluated.

4. While phenotypic methods may be improved upon for rapid identification of MDR-TB, genotypic methods have considerable advantages for scaling-up programmatic management of drug-resistant TB, in particular with regard to speed, standardized testing, potential for high throughput, and biosafety. The ultimate aim should be to implement molecular assays such as the line-probe (or other well-validated and WHO-endorsed molecular platforms in the future) for rapid first-step identification of MDR-TB in the programmatic context. Rapid phenotypic DST methods therefore present an interim solution, especially in resource-constrained settings, while capacity for genotypic testing is being developed.
5. Large-scale demonstration studies that show the generalisability to programmatic settings of the results of validation/evaluation studies have not been performed for any of the methods/assays under consideration. While the Expert Group recognizes that such demonstration studies should ideally be done before a method or assay can be recommended for large-scale use, it saw reasons to relax this requirement: First, being (in all but one case) modifications of standard conventional DST methods, there has been little attention and support from industry or funding agencies for research and development of these methods/assays. Large-scale demonstration projects, which are generally expensive, would thus require extensive funding from other sources (eg. TB control programmes) which is highly unlikely to become available. Second, since large-scale demonstration projects may take several years, investing in such projects may not be an optimal use of scarce resources, as recommendations limited to demonstration studies without eventual scale-up are likely to result in the loss of affordable assays with potential to significantly reduce morbidity and mortality in resource-limited high MDR-TB prevalence settings.
6. In the GRADE system, quality of evidence is assessed both according to test accuracy and impact on patient outcomes. The quality of evidence from diagnostic studies, without direct evidence of benefit to patients, is by default considered to be low in the GRADE system. The Expert Group therefore agreed to use test accuracy as a proxy measure of patient-important outcomes, given the absence of direct data from current diagnostic studies.

GRADE recommends the rating of patient importance using a 9-point scale (1-3: not important; 4-6: important; 7-9: critical). The Expert Group therefore agreed to use the following approach to reach consensus on the balance of effects:

- True positives (Importance=9): Patients benefiting from rapid detection of (multi)drug resistance resulting in earlier treatment with appropriate second-line regimens;
- True negatives (Importance=7): Patients spared unnecessary treatment with lengthy, toxic and expensive second-line treatment regimens; benefit of reassurance and alternative diagnosis (including TB and treatment with more effective first-line drugs);
- False positives (Importance=8): Potential harm from unnecessary treatment with lengthy, toxic and expensive second-line treatment regimens;

- False negatives (Importance=9): Increased risk of patient morbidity and mortality continued risk of community transmission of (multi)drug-resistant TB and amplification of drug resistance.
7. Primary meta-analyses evaluated the accuracy of each method/assay in combined direct (ie. on specimens) and indirect testing (ie. on culture isolates grown from patient specimens). Secondary analyses assessed accuracy separately for direct and indirect testing where enough studies/data were available. **The Expert Group considered the use of rapid DST methods *directly* on sputum specimens to be the most important for policy recommendation, as the clinically relevant gain in turn-around time would have the highest patient benefit.** While good indirect test performance can be considered as proof-of-principle of the method/assay, direct performance is known to be affected by the initial mycobacterial inoculum size that is difficult to control without prior culture. Therefore, the results of indirect test performance cannot be simply extrapolated to direct application, and performance estimates for direct application are required. However, for most methods/assays under consideration, data on test performance for direct application were much less precise and generally much sparser (in terms of settings where studies had been performed).
 8. Data on direct application of the assays/methods have been available almost exclusively on smear-positive sputum specimens. Therefore, any recommendations on direct application are limited to sputum smear-positive specimens.
 9. With direct testing, bacterial and fungal contamination can be problematic. However, since the assays/methods under evaluation provide results rapidly, the possibility exists to subculture or to repeat culture from the original clinical specimen. Therefore, contamination is less of a problem than with conventional indirect DST methods.
 10. Since non-tuberculous mycobacteria may display innate resistance to rifampicin, the Expert Group agreed that the methods under consideration required mycobacterial speciation, at minimum whenever rifampicin resistance is observed.
 11. None of the currently available tests and methods for culture and DST are mutually exclusive. Given the evidence base for non-commercial methods as outlined below, none of these methods should replace liquid culture or line probe assays as the current reference standards.
 12. The policy formulation process for liquid culture and DST (2007), as well as for line probe assays (2008) preceded the development of GRADE for diagnostics and implementation by WHO of the GRADE process; nevertheless, policy formulation for these assays was subject to rigorous analysis by Expert Groups of the accuracy of these tests compared to conventional methods, and included extensive data from large-scale field demonstration studies which included cost-effectiveness data. In addition, the strength of the evidence base using the GRADE process was assessed for these two technologies by two of the groups responsible for the current systematic reviews. The quality of data for both liquid culture and line probe assays was found to be moderate, acknowledging limited data on patient-important

outcomes (although substantially more data than for the non-commercial methods). The strength of the recommendation (strong) for the use of liquid culture and line probe assays remained unaffected.

3.2 Microscopic observation of drug susceptibility (MODS) assay

Table 1 presents a summary of the findings from meta-analyses to evaluate assay performance results against conventional DST methods, as well as the outcomes of the GRADE process, estimating the anticipated impact at three hypothetical levels of rifampicin resistance prevalence in order to allow an assessment of patient impact (balance of effect) as outlined before.

3.2.1 Quality of evidence

Accuracy data showed that MODS is highly sensitive (pooled estimate 98.0%, 95% CI 94.5-99.3) and specific (pooled estimate 99.4%, 95% CI 95.7-99.9) for the detection of rifampicin resistance, and slightly less sensitive for isoniazid (pooled sensitivity 91.4%, 95% CI 86.7-94.6; pooled specificity 97.7%, 95% CI 93.6-99.2). Testing for other anti-tuberculosis drugs suffered from variable and suboptimal sensitivity (data not shown).

Table 1. Summary of findings on MODS

# Studies	9 (2 excluded, see below)			
# Participants	1,474			
<i>Pooled accuracy estimates from meta-analyses</i>				
Rifampicin sensitivity	98.0%			
Rifampicin specificity	99.4%			
Assumed prevalence of rifampicin resistance (pre-test probability)	Expected # patients			
	TP	TN	FP	FN
5%	49	944	6	1
10%	98	895	5	2
20%	196	795	5	4
<i>Other reported outcomes</i>				
Application for direct testing	High			
TAT (direct application)	11.6 days			
TAT (indirect application)	6.5 days post-isolation			
Contamination rate				
DST studies alone	7.4%			
DST and TB detection studies	6.3%			
<i>Quality indicators</i>				
Study design	6 cross-sectional, 3 unclear			
Risk of bias	Low			
Directness	Limited (-1 for generalisability)			
Inconsistency	Low			
Imprecision	Low			

Publication bias	Possible
QUALITY OF EVIDENCE	MODERATE
<i>Resource requirements*</i> Training Infrastructure Equipment (cost) Consumables	Extensive BSL2 (not fully resolved - see below) Medium Medium
STRENGTH OF RECOMMENDATION	STRONG

*Adapted from: 'New Laboratory Diagnostic Tools for Tuberculosis Control' (2008), published by: Stop TB Partnership Retooling Task Force and New Diagnostics Working Group.

Notes:

- Of the nine studies included in the meta-analysis, two studies were considered not to have used an appropriate reference standard. Excluding these from the meta-analyses did not affect the accuracy estimates;
- Concerns by the Expert Group were expressed about limited generalisability of existing data beyond research or reference laboratory settings;
- Sensitivity for detecting isoniazid resistance was higher at drug concentration 0.1mcg/ml (97.7%, 95% CI 94.4-99.1) than at 0.4 mcg/ml (90.0%, 95% CI 84.5-93.7), with slightly lower specificity (95.8% versus 98.6%).

3.2.2 Balance of desirable and undesirable effects

The balance of effects was considered favourable using rifampin resistance as indicator/proxy of MDR and the relative importance of false-positive and false-negative results as outlined above.

3.2.3 Values and preferences of patients

No studies available, see above section for use of turn-around time of test results as indirect measure. The average turnaround time (9.9 days; 95% CI 4.1-15.8), was similar to that of conventional liquid culture, much quicker than the conventional indirect proportion DST method on solid Löwenstein-Jensen (LJ) culture medium, and therefore considered favourable.

3.2.4 Costs and requirements

Compared to the standard proportion DST method on LJ solid medium, MODS requires additional staff skills, an additional inverted microscope, and additional consumables that may be more difficult to obtain; compared to the conventional indirect proportion DST method in liquid culture medium, MODS requires additional staff skills, and consumables that may be readily available.

Limited and robust cost comparisons indicate that MODS is as expensive as CRI, less expensive than commercial liquid culture, and more expensive than TLA or the NRA.

The MODS assay has been fully standardized, with testing protocols and support available through a dedicated website. A system of accreditation of laboratories to perform the MODS assay is being implemented by the Peruvian national mycobacterial reference laboratory with support by the developer.

The Expert Group could not reach consensus on the biosafety implications or the speciation discriminatory ability of the MODS assay. The technique involves liquid culture medium, but small volumes are used in microtitre plates which are sealed in transparent bags directly after inoculation of the sputum specimen. Therefore, selected Expert Group members felt that the assay can be considered to have a biosafety risk similar to that of conventional culture on solid medium (biosafety level 2), as manipulation of highly concentrated suspensions containing *M. tuberculosis* (such as in indirect DST methods) is not required.

The MODS assay involves speciation by microscopic determination based on observing typical cord formation patterns by *M. tuberculosis*. Expert Group members were divided on their opinion that cord-formation is sufficiently sensitive and specific in comparison to standard biochemical or genotypic speciation, and no meta-analysis data were available of studies that had performed such comparisons. Several Expert Group members insisted that biochemical or genotypic speciation was necessary, at least for MODS cultures showing rifampicin resistance.

Discussions are ongoing with the developer to include a well containing p-nitrobenzoic acid (PNB) in the microtitre plate, which would allow *M. tuberculosis* complex to be differentiated from non-tuberculous mycobacteria (*M. tuberculosis* complex failing to grow in the presence of PNB), as a further confirmation of speciation based on cord-forming (and therefore in line with current WHO recommendations for DST on conventional solid LJ media). Adding a PNB well to the microtitre plate also obviates the need to re-open the plate, thereby further reducing the biosafety risk.

(Note: Subsequent to the Expert Group meeting, the developer has started experiments to standardize the PNB concentration, and the revised MODS platform will then be implemented as routine in two government laboratories in an operational research exercise).

3.2.5 Added value and potential for scale-up

The potential of MODS to replace conventional liquid culture and DST was considered unknown due to lack of demonstration data. Lower cost, speed and not having to manipulate cultures to perform DST were regarded as comparative advantages. Laboratories in Peru had greater experience implementing MODS; nevertheless, MODS appears to perform equally well in reports from selected national/reference laboratories outside of Peru.

Given the need for additional staff skills and biosafety measures, the potential for scale-up/decentralization of MODS beyond the reference laboratory level was considered to be low.

3.2.6 Results for direct application

Six studies reported direct application of MODS. Sub-analyses of diagnostic accuracy did not differ significantly from those for direct and indirect testing combined, although the TAT was slightly longer (11.6 days; 95% CI 1.5-21.7 vs 6.5 days post-isolation when used in indirect testing).

Preliminary but very limited data suggested good sensitivity on specimens obtained during treatment.

3.2.7 Main research gaps identified

- Accuracy of microscopic speciation;
- Accuracy of MODS used directly on specimens from patients under treatment
- Accuracy of MODS used directly on clinical specimens from patients with smear-negative pulmonary and extrapulmonary TB;
- Applicability of MODS for detecting resistance to drugs other than isoniazid and rifampicin;
- External quality assurance of MODS.

FINAL RECOMMENDATION

The Expert Group agrees that there is sufficient evidence to recommend the use of MODS for rapid screening of patients suspected of having MDR-TB, under clearly defined programmatic and operational conditions, once speciation concerns have been adequately addressed and without compromising biosafety.

3.3 Thin layer agar (TLA) assay

Table 2 presents a summary of the findings from meta-analyses to evaluate assay performance results against conventional DST methods, as well as the outcomes of the GRADE process, estimating the anticipated impact at three hypothetical levels of rifampicin resistance prevalence in order to allow an assessment of patient impact (balance of effect) as outlined before.

3.3.1 Quality of evidence

Accuracy data showed that TLA is highly sensitive (100%) and specific (100%). However, the meta-analysis included only three studies (two on isoniazid), precluding reliable estimation of pooled accuracy estimates.

Table 2. Summary of findings on TLA

# Studies	3			
# Participants	439			
<i>Pooled accuracy estimates from meta-analyses</i>				
Rifampicin sensitivity	100%			
Rifampicin specificity	100%			
Assumed prevalence of rifampicin resistance (pre-test probability)	Expected # patients			
	TP	TN	FP	FN
5%	50	950	0	0
10%	100	900	0	0
20%	200	800	0	0
<i>Other reported outcomes</i>				
Application for direct testing	High			
TAT (direct application)	11.1 days			
TAT (indirect application)	No data			
Contamination rate				
DST studies alone	1.7%			
DST and TB detection studies	11.8%			
<i>Quality indicators</i>				
Study design	2 cross-sectional, 1 unclear			
Risk of bias	Low			
Directness	Limited (-1 for generalisability)			
Inconsistency	Low			
Imprecision	High (-2)			
Publication bias	Possible			
QUALITY OF EVIDENCE	VERY LOW			
<i>Resource requirements*</i>				
Training	Extensive			
Infrastructure	BSL 2			
Equipment (cost)	Medium			

Consumables	Medium
STRENGTH OF RECOMMENDATION	WEAK

*Adapted from: 'New Laboratory Diagnostic Tools for Tuberculosis Control' (2008), published by: Stop TB Partnership Retooling Task Force and New Diagnostics Working Group.

Notes:

- Serious concerns expressed about very limited data as well as generalisability of data beyond the research laboratory setting.

3.3.2 Balance of desirable and undesirable effects

The balance of effects was considered favourable using rifampicin resistance as indicator/proxy of MDR and the relative importance of false-positive and false-negative results as outlined above.

3.3.3 Values and preferences of patients

No studies available, see above section for use of turn-around time of test results as indirect measure. The mean turnaround time (11.1 days; 95% CI 10.1 - 12.0), was similar to that of conventional liquid DST culture, much quicker than the conventional indirect proportion DST method on solid Löwenstein-Jensen (LJ) culture medium, and therefore considered favourable.

3.3.4 Costs and requirements

Compared to the conventional indirect proportion DST method on LJ solid medium, TLA requires additional staff skills, similar equipment, and additional consumables that may be difficult to obtain; compared to the conventional indirect proportion DST method in liquid culture medium, TLA requires additional staff skills, less equipment, and consumables that may be readily available.

Limited and robust cost comparisons indicate that TLA is less expensive than MODS, commercial liquid culture and CRI methods, and slightly more expensive than the NRA.

Since TLA is based on solid medium and does not require manipulation of strains, the same biosafety requirements (biosafety level 2) as for solid culture apply.

3.3.5 Added value and potential for scale-up

The potential of TLA to replace conventional liquid culture and DST was considered unknown due to lack of demonstration data. Lower cost, speed and not having to manipulate cultures to perform DST were regarded as comparative advantages, especially if a simple colorimetric test could be used (such a test is currently under development by a group in Peru (Dr C Evans), with support from FIND).

Given the need for additional skills, the potential for scale-up/decentralization beyond the reference laboratory level was considered to be low.

3.3.6 Results for direct application

Two studies reported direct application of TLA. Although the diagnostic accuracy did not differ significantly from combined studies, numbers were very small and the quality of evidence of TLA application directly on sputum was therefore considered to be very low.

The pooled contamination rate (direct application only) was 1.7% (95% CI 0.5-4.4), but much higher when all studies on TLA reporting contamination rates were included (11.8%; 11.0-12.6).

3.3.7 Main research gaps identified

- Accuracy data in different settings, with a view to generalisability beyond the research laboratory setting, in particular for direct application;
- Accuracy of TLA used directly on specimens from patients under treatment;
- Accuracy of TLA used directly on clinical specimens from patients with smear-negative pulmonary and extrapulmonary TB;
- Applicability of TLA for detecting resistance to other drugs than isoniazid and rifampicin;
- Alternative decontamination methods that would obviate the need for centrifugation;
- Use of alternative, less expensive agar media;
- Cost-effectiveness for scale-up beyond the reference laboratory setting;

FINAL RECOMMENDATION

The Expert Group agrees that there is as yet insufficient evidence to recommend the use of TLA as a rapid test for rapid screening of patients suspected of having MDR-TB. TLA appears to be a promising diagnostic tool for rapid DST and further research is encouraged.

3.4 Colorimetric redox indicator (CRI) methods

Table 3 presents a summary of the findings from meta-analyses to evaluate assay performance results against conventional DST methods, as well as the outcomes of the GRADE process, estimating the anticipated impact at three hypothetical levels of rifampicin resistance prevalence in order to allow an assessment of patient impact (balance of effect) as outlined before.

Table 3. Summary of findings on CRI methods

# Studies	31			
# Participants	2 498			
<i>Pooled accuracy estimates from meta-analyses</i>				
Rifampicin sensitivity	98.0%			
Rifampicin specificity	99.0%			
Assumed prevalence of rifampicin resistance (pre-test probability)	Expected # patients			
	TP	TN	FP	FN
5%	49	941	9	1
10%	98	891	9	2
20%	196	792	8	4
<i>Other reported outcomes</i>				
Application for direct testing	No data			
TAT (direct application)	No data			
TAT (indirect application)	7 - 14 days post-isolation			
Contamination rate				
DST studies alone	5.0%			
DST and TB detection studies	No data			
<i>Quality indicators</i>				
Study design	13 cross-sectional, 17 case-control, 1 unclear			
Risk of bias	Low			
Directness	None (-1 for generalisability)			
Inconsistency	Low			
Imprecision	Low			
Publication bias	Possible			
QUALITY OF EVIDENCE	MODERATE			
<i>Resource requirements*</i>				
Training	Extensive			
Infrastructure	BSL 3			
Equipment (cost)	Medium			
Consumables	Medium			
STRENGTH OF RECOMMENDATION	MODERATE			

*Adapted from: 'New Laboratory Diagnostic Tools for Tuberculosis Control' (2008), published by: Stop TB Partnership Retooling Task Force and New Diagnostics Working Group.

Notes:

- Concerns expressed about limited generalisability of the data beyond the research or reference laboratory setting.

3.4.1 Quality of evidence

Accuracy data showed that CRI methods are highly sensitive and specific for the detection of rifampicin resistance (pooled sensitivity 98%, 95% CI 96-99; pooled specificity 99%, 95% CI 99-100) as well as for the detection of isoniazid resistance (pooled sensitivity 97%, 95% CI 96-98; pooled specificity 98%, 95% CI 97-99).

3.4.2 Balance of desirable and undesirable effects

The balance of effects was considered favourable using rifampicin resistance as indicator/proxy of MDR and the relative importance of false-positive and false-negative results as outlined before.

3.4.3 Values and preferences of patients

No studies available, see above section for use of turn-around time of test results as indirect measure. The average turnaround time of indirect testing was reported to be between 7 and 14 days, similar to that of conventional liquid DST culture, much quicker than the conventional indirect proportion DST method on solid Löwenstein-Jensen (LJ) culture medium, and therefore considered favourable.

3.4.4 Costs and requirements

Compared to the conventional indirect proportion DST method on LJ medium, CRI methods require additional staff skills, similar equipment and additional consumables that may be difficult to obtain; compared to the conventional indirect proportion DST method in liquid culture medium, CRI methods requires similar staff skills, less equipment, and consumables that may be readily available.

Limited and robust cost comparisons indicate that CRI methods are less expensive than conventional liquid culture but more expensive than MODS, TLA and the NRA.

Since CRI methods make use of manipulation of concentrated suspensions of mycobacteria, biosafety requirements are similar to those for liquid culture (biosafety level 3).

CRI methods have been fully standardized, with testing protocols available on www.tbevidence.org.

3.4.5 Added value and potential for scale-up

The potential of CRI methods to replace commercial liquid culture and DST was considered unknown due to lack of demonstration data, with lower cost being the only added comparative advantage.

Given the need for additional staff skills and biosafety measures, the potential for scale-up/decentralization beyond the reference laboratory level was considered to be low.

3.4.6 Results for direct application only

Only two studies reported direct application of CRI methods, showing a sensitivity of 83% and 100% respectively, and a specificity of 100% for detection of rifampin resistance. Quality of evidence of CRI directly on sputum was therefore considered to be very low. Data on contamination rates from direct testing were not available.

3.4.7 Main research gaps identified

- Accuracy data of CRI methods in different settings, with a view to generalizability (reproducibility, robustness) beyond the reference and research laboratory setting, in particular for direct application;
- Accuracy of CRI methods used directly on specimens;
- Accuracy of CRI methods used directly on clinical specimens from patients with smear-negative pulmonary and extrapulmonary TB;
- Risk of cross-contamination with CRI methods;
- Low-cost method(s) for speciation best used in combination with direct CRI methods.

FINAL RECOMMENDATION

The Expert Group agrees that there is sufficient evidence to recommend the use of CRI methods as an indirect test on culture isolates from patients suspected of having MDR-TB, under clearly defined programmatic and operational conditions, and acknowledging that time to detection of MDR-TB would not be faster (but less expensive) than conventional DST methods using commercial liquid culture.

3.5 Nitrate reductase assay (NRA)

Table 4 presents a summary of the findings from meta-analyses to evaluate assay performance results against conventional DST methods, as well as the outcomes of the GRADE process, estimating the anticipated impact at three hypothetical levels of rifampicin resistance prevalence in order to allow an assessment of patient impact (balance of effect) as outlined before.

Table 4. Summary of findings on NRA

# Studies	19			
# Participants	2 304			
<i>Pooled accuracy estimates from meta-analyses</i>	97.0%			
Rifampicin sensitivity	100%			
Rifampicin specificity				
Assumed prevalence of rifampicin resistance (pre-test probability)	Expected # patients			
	TP	TN	FP	FN
5%	49	950	0	1
10%	97	900	0	3
20%	194	800	0	6
<i>Other reported outcomes</i>				
Application for direct testing ¹	High			
TAT (direct application)	6 - 9 days			
TAT (indirect application)	7 - 21 days post-isolation			
Contamination rate				
DST studies alone	4.8.0%			
DST and TB detection studies	No data			
<i>Quality indicators</i>				
Study design	11 cross-sectional, 8 case-control			
Risk of bias	Low			
Directness	None (-1 for generalisability)			
Inconsistency	Low			
Imprecision	Low			
Publication bias	Possible			
QUALITY OF EVIDENCE	MODERATE			
<i>Resource requirements*</i>				
Training	Moderate			
Infrastructure	BSL 2			
Equipment (cost)	Medium			
Consumables	Medium			
STRENGTH OF RECOMMENDATION	MODERATE			

*Adapted from: 'New Laboratory Diagnostic Tools for Tuberculosis Control' (2008), published by: Stop TB Partnership Retooling Task Force and New Diagnostics Working Group.

¹ Subject to final verification of data by the Expert Group, following additional sub-analyses

Notes:

- Concerns about limited generalisability of the data beyond the research or reference laboratory setting;
- Some *M. tuberculosis* complex species do not produce nitrate reductase, ie. their growth would not be detected by NRA. No studies have purposefully included such strains.

3.5.1 Quality of evidence

Accuracy data show that NRA is highly sensitive and specific for the detection of rifampicin resistance (pooled sensitivity 97%, 95% CI 95-98; pooled specificity 100%, 95% CI 99-100) as well as for the detection of isoniazid resistance (pooled sensitivity 97%, 95% CI 95-98; pooled specificity 99%, 95% CI 99-100).

3.5.2 Balance of desirable and undesirable effects

The balance of effects was considered favourable using rifampicin resistance as indicator/proxy of MDR and the relative importance of false-positive and false-negative results as outlined above.

3.5.3 Values and preferences of patients

No studies available, see above section for use of turn-around time of test results as indirect measure. The average turnaround time for indirect testing was similar to that of liquid culture, quicker than the conventional indirect proportion DST method on solid LJ culture medium, and therefore considered favourable.

3.5.4 Costs and laboratory requirements

Compared to the conventional indirect proportion method DST on LJ solid medium, NRA requires similar staff skills, similar equipment, and no additional consumables; compared to the conventional indirect proportion DST method in liquid culture medium, NRA requires fewer staff skills, less equipment, and fewer consumables.

Reagents are non-proprietary and relatively inexpensive. Tubes can be re-used.

Limited and robust cost comparisons indicate that of NRA is the least expensive rapid culture-based DST method when compared to commercial methods, CRI methods, TLA and MODS.

Since NRA makes use of solid culture media, biosafety requirements are also similar (biosafety level 2), although culture tubes need to be opened repeatedly to add reagent. Addition of reagent does, however, pose a significant risk of aerosolisation which mandate this being done inside an appropriate biological safety cabinet.

NRA has been fully standardized, with testing protocols available on www.tbevidence.org.

3.5.5 Added value and potential for scale-up

The potential of NRA to replace liquid culture and DST was considered unknown due to lack of demonstration data. Lower cost and simplicity were regarded as comparative advantages.

Given the relatively limited need for additional skills and biosafety measures, the potential for scale-up/decentralization beyond the reference laboratory level was considered to be moderate.

3.5.6 Results for direct application

Four studies reported direct application of NRA. Sub-analyses of diagnostic accuracy did not differ significantly from those for direct and indirect testing combined, although individual studies showed greater variability in sensitivity (range 87% - 100%).² Contamination rates for direct testing were not available.

3.5.7 Main research gaps identified

- Accuracy data in different settings, with a view to generalizability (reproducibility, robustness) beyond the reference and research laboratory setting, in particular for direct application;
- Accuracy of NRA used directly on clinical specimens;
- Applicability of NRA for detecting resistance to drugs other than isoniazid and rifampin;
- Relative frequency and causes of invalid results, in particular for *M. africanum* type I;
- Low-cost method(s) for speciation to be used in combination with direct NRA;
- Performance of NRA on other solid medium, eg. Ogawa medium.

FINAL RECOMMENDATION

The Expert Group agrees that there is sufficient evidence to recommend the use of NRA for rapid screening of patients suspected of having MDR-TB, under clearly defined programmatic and operational conditions, and acknowledging that time to detection of MDR-TB in indirect application would not be faster (but less expensive) than conventional DST methods using commercial liquid culture.

----- 000 -----

² Subject to final verification of data by the Expert Group, following additional sub-analyses

3.6 Phage-based assays

Tables 4a to 4c present a summary of the findings from meta-analyses to evaluate assay performance results against conventional DST methods, as well as the outcomes of the GRADE process, estimating the anticipated impact at three hypothetical levels of rifampicin resistance prevalence in order to allow an assessment of patient impact (balance of effect) as outlined before.

Table 4a. Summary of findings on commercial phage-based assays

# Studies	12			
# Participants	2 945			
<i>Pooled accuracy estimates from meta-analyses</i>				
Rifampicin sensitivity	95.0%			
Rifampicin specificity	95.3%			
Assumed prevalence of rifampicin resistance (pre-test probability)	Expected # patients			
	TP ¹	TN ²	FP ³	FN ⁴
5%	43	905	45	2
10%	95	858	42	5
20%	190	762	38	10
<i>Other reported outcomes</i>				
TAT (direct application)	1 - 2 days			
TAT (indirect application)	1 - 2 days			
Uninterpretable results	20.4%			
Direct testing	5.8%			
Indirect testing				
<i>Quality indicators</i>				
Study design	3 cross-sectional, 9 case-control			
Risk of bias	Moderate (-1)			
Directness	None (-1 for generalisability)			
Inconsistency	Moderate - High (-1)			
Imprecision	Low			
Publication bias	Possible			
QUALITY OF EVIDENCE	VERY LOW			
<i>Resource requirements</i>				
Training	Not assessed due to quality of evidence			
Infrastructure	Not assessed due to quality of evidence			
Equipment (cost)	Not assessed due to quality of evidence			
Consumables	Not assessed due to quality of evidence			
STRENGTH OF RECOMMENDATION	WEAK			

Table 4b. Summary of findings on in-house phage-based assays

# Studies	11			
# Participants	1 037			
<i>Pooled accuracy estimates from meta-analyses</i>				
Rifampicin sensitivity	98.7%			
Rifampicin specificity	98.2%			
Assumed prevalence of rifampicin resistance (pre-test probability)	Expected # patients			
	TP ¹	TN ²	FP ³	FN ⁴
5%	49	933	17	1
10%	99	884	16	1
20%	197	786	14	3
<i>Other reported outcomes</i>				
Application for direct testing	Medium			
TAT (direct application)	1 – 2 days			
TAT (indirect application)				
Uninterpretable results	No data			
Direct testing	2.0%			
Indirect testing				
<i>Quality indicators</i>				
Study design	1 cross-sectional, 10 case-control			
Risk of bias	High (-2)			
Directness	None (-1 for generalisability)			
Inconsistency	Moderate - High (-1)			
Imprecision	Low			
Publication bias	Possible			
QUALITY OF EVIDENCE	VERY LOW			
<i>Resource requirements*</i>				
Training	Not assessed			
Infrastructure	Not assessed			
Equipment (cost)	Not assessed			
Consumables	Not assessed			
STRENGTH OF RECOMMENDATION	WEAK			

Table 4c. Summary of findings on luciferase reporter phage assays

# Studies	8
# Participants	664
<i>Pooled accuracy estimates from meta-analyses</i>	
Rifampicin sensitivity	99.6%

Rifampicin specificity	99.4%			
Assumed prevalence of rifampicin resistance (pre-test probability)	Expected # patients			
	TP ¹	TN ²	FP ³	FN ⁴
5%	50	944	6	0
10%	100	895	5	0
20%	199	795	5	1
<i>Other reported outcomes</i>				
TAT (direct application)	No data			
TAT (indirect application)	1 – 2 days			
Uninterpretable results				
Direct testing	No data			
Indirect testing	3.0%			
<i>Quality indicators</i>				
Study design	5 cross-sectional, 3 case-control			
Risk of bias	High (-2)			
Directness	None (-1 for generalizability)			
Inconsistency	Moderate - High (-1)			
Imprecision	High			
Publication bias	Possible			
QUALITY OF EVIDENCE	VERY LOW			
<i>Resource requirements*</i>				
Training	Not assessed due to quality of evidence			
Infrastructure	Not assessed due to quality of evidence			
Equipment (cost)	Not assessed due to quality of evidence			
Consumables	Not assessed due to quality of evidence			
STRENGTH OF RECOMMENDATION	WEAK			

3.6.1 Quality of evidence

Accuracy data showed that commercial phage-based assays have a wide range of sensitivities (81-100%) and specificities (73-100%). In-house phage-amplification assays showed similar variation (sensitivity range 88-100%; specificity range 84-100%). Luciferase reporter phage (LPR) assays yielded the most consistent estimates, with eight of nine studies reporting 100% sensitivity and five of nine studies reporting 100% specificity. Overall, LRP assays had the highest accuracy (sensitivity 99.6%, specificity 99.4%), followed by in-house phage amplification (sensitivity 98.7%, specificity 98.2%) and commercial assays (sensitivity 95.0%, specificity 95.3%).

Notes

- Concerns about design limitations, limited generalizability of the data beyond the research or reference laboratory setting, and inconsistency of results. For in-house

phage-amplification and LRP assays, there were additional concerns about publication bias;

- High rates of contamination or indeterminate results were reported (0-36%, mean 5.8%, 95% CI 2.6-9.0), in particular when used directly on sputum specimens. Highest rates of uninterpretable results were reported for commercial assays, although no studies had evaluated in-house phage-amplification assays and LRP assays directly on sputum. Unacceptably high uninterpretable results also caused FIND to interrupt their ongoing Demonstration Projects in 2007 using the *FASTPlaque-Response*[™] assay for detection of rifampin resistance.

3.6.2 Balance or desirable and undesirable effects

The balance of effects was considered favourable for in-house phage amplification and LRP assays. Serious concerns were raised about the commercial assays, in particular with regard to their relatively high false-negative and false-positive rates.

During the initial phase of the Demonstration Projects which FIND started at two trial sites in South Africa in 2007, the *FASTPlaque-Response*[™] test failed to meet required performance targets, with high rates of false-positive rifampin resistance results. As a consequence, FIND decided to discontinue activities with the *FASTPlaque*[™] assay until improvements or satisfactory alternatives are available. The manufacturer has made modifications in the protocol for the *FASTPlaque*[™] assays, but reported that while modifications increased the sensitivity in smear-negative specimens, this was still associated with high rates of false-positive results.

3.6.3 Values and preferences of patients

No studies available, see above section for use of turn-around time of test results as indirect measure. The average turnaround time much quicker than that of liquid culture and the conventional indirect proportion DST method on solid LJ culture medium, and therefore considered favourable.

3.6.4 Costs and requirements

Because of the very low quality of evidence in combination with concerns about high false-positivity rates for the commercial phage-based assays, and the absence of evidence on indirect application for the in-house phage-amplification LRP assays, the Expert Group did not further assess costs and laboratory requirements.

3.6.5 Added value and potential for scale-up

Because of the very low quality of evidence in combination with concerns about high false-positivity rates for the commercial phage-based assays, and the absence of evidence on indirect application for the in-house phage-amplification LRP assays, the Expert Group did not further consider added value and potential for scale-up.

3.6.6 Results for direct application

Only five studies of direct application of commercial assays were available. Quality of evidence on the use of these assays directly on sputum was considered very low. Studies of direct application to sputum specimens yielded lower sensitivity but higher specificity than those in which the assays were applied to culture isolates. Evidence on in-house phage-amplification assays and LRP assays directly on sputum was absent.

High rates of contamination or indeterminate results were reported when commercial assays were used directly on sputum specimens (range 3-36%, mean 20.4%, 95% CI 106-30.3). Addition of an antibiotic mixture (NOA) in three studies reduced the contamination rates from 19 to 1.2%, from 16 to 5%, and from 1.4 to 0.5%, respectively.

3.6.7 Main research gaps identified

Because of the very low quality of evidence in combination with concerns about high false-positivity rates for the commercial phage-based assays, and the absence of evidence on indirect application for the in-house phage-amplification LRP assays, the Expert Group did not further assess research gaps.

FINAL RECOMMENDATION

The Expert Group agrees that there is insufficient evidence to recommend the use of phage-based assays for rapid screening of patients suspected of having MDR-TB.

4. Implementation considerations

As with any new technology, method or approach, a range of implementation issues was identified, without which non-commercial culture and DST methods would not be useful. These include:

4.1 Specimen collection, storage and transport

The quality of sputum specimens submitted to the laboratory is critical in obtaining reliable results. A reliable specimen transport system will ensure that the full benefit is gained from use of any rapid assay, by reducing diagnostic delay times.

Current WHO recommendations call for MDR strains to be screened for XDR, both during surveys and in clinical settings where XDR-TB patients are suspected or confirmed. None of the non-commercial methods detect XDR-TB; therefore, additional specimens with refrigerated transport and rapid delivery systems are essential for conventional culture and DST procedures, requiring strict adherence to standard operating procedures for specimen collection, storage and transport.

4.2 Biosafety

Current WHO recommendations specify that specimen processing for mycobacterial culture be performed in a biological safety cabinet (BSC) under at least biosafety level 2 (BSL2) conditions. Procedures involving manipulation of *M. tuberculosis* cultures (identification, sub-culturing and indirect DST) must be performed in laboratories complying with BSL3 standards.

4.3 Equipment, supplies and reagents

In addition to the equipment required for initial digestion-decontamination of sputum specimens (such as BSCs and safety centrifuges), most non-commercial methods require specific equipment or supplies. CRI methods require microtitre plates and consumables for which a dedicated supply chain must be established, including proprietary medium (Middlebrook 7H9 broth) and a redox indicator. MODS require an inverted light microscope, Middlebrook 7H9 broth and a cold supply chain for OADC and PANTA. NRA requires several consumable products for which a dedicated supply chain is needed, including potassium nitrate and additional reagents depending on whether the test is performed with liquid reagents or a crystalline reagent.

Short expiration dates of reagents are a general concern for laboratories, especially in relatively inaccessible areas with complex customs clearance procedures. Management of inventory based on usage, shelf-life and lead time for deliver of orders is therefore needed.

4.4 Human resources and training

Successful implementation and interpretation of non-commercial culture and DST methods is highly dependent on the skill of laboratory staff performing the testing and the quality of supervision, the quality and training of staff and their adherence to strict working practices, including good laboratory practice (GLP) and good microbiological technique (GMP). Since skilled and highly trained personnel are required for performing these assays, the human resource requirements need to be carefully assessed prior to implementation.

4.5 Technical support

Coordination between commercial suppliers and customers with regard to ordering, shipping and customs clearance is critical to ensure smooth delivery of reagents and equipment and to avoid customs delays which may cause product deterioration due to inadequate storage conditions in transit.

A detailed plan for training, based on country-specific human resource needs, must be developed. In addition, ongoing technical support and continuous supply of consumables and reagents is essential, best provided for in a formal service contract between the supplier and customer. Such a contract should cover the following aspects:

- Maintenance of equipment and provision of a servicing contract, including the repair or replacement of faulty equipment at short notice;
- Supply of consumables and reagents with at least six months expiry after arrival at the laboratory;
- A detailed plan for provision of ongoing technical support and the channels through which this will be provided, eg. a local distributor, via helpline, or internet-based support;

4.6 Quality assurance

Standardized external quality assurance programmes for non-commercial methods are not yet available. Development of such systems is therefore an urgent priority, particularly since these assays are heavily operator-dependent.

4.7 Recording and reporting

In order to gain full benefit from implementation of rapid culture and DST methods, systems must be implemented to ensure that results are reported quickly to clinicians and patients to ensure that appropriate treatment is initiated. Furthermore, where conventional DST is used to confirm rapid assay results, the possibility of discrepant results must be considered, and a mechanism for explanation of implications of discrepancies to clinicians should be established.

5. Policy recommendations

In resource-constrained settings, the use of MODS and NRA are recommended by WHO as an interim solution for screening of patients suspected of MDR-TB.

In resource-constrained settings, the use of indirect CRI methods are recommended by WHO as an interim solution for screening of *M. tuberculosis* isolates from patients suspected of having MDR-TB, acknowledging that time to detection of MDR-TB would not be faster with indirect methods (but less expensive) than conventional DST methods using commercial liquid culture.

The following guiding principles apply to all methods recommended:

- 5.1 None of the non-commercial assays and methods constitutes design-locked, standardized and quality-assured tests that are produced under international diagnostic standards. Therefore, stringent laboratory protocols, standard operating procedures, and internal quality control mechanisms must be implemented and enforced.
- 5.2 Implementation must be preceded by appropriate training and validation at country level of the non-commercial methods against one of the international reference standards (conventional solid culture and DST, conventional liquid culture and DST, and/or line probe assay). Programmes for external quality assessment of laboratories involved in non-commercial culture and DST methods should be developed as a matter of priority in settings where implementation is considered.
- 5.3 Data on test performance for direct application were much less precise and generally much sparser (in terms of settings where studies had been performed) for most methods. In addition, data on direct application of the assays/methods have been available almost exclusively in smear-positive sputum specimens. Recommendations on direct application of the above non-commercial methods are therefore cautious and can only be extrapolated to smear-positive specimens.
- 5.4 Implementation of non-commercial culture and DST methods for detection of MDR-TB should be decided by Ministries of Health within the context of country plans for appropriate management of MDR-TB patients, including the development of country-specific screening algorithms and timely access to quality-assured second-line anti-tuberculosis drugs;
- 5.5 Implementation of any of the non-commercial DST methods does not eliminate the need for conventional culture and DST capability; culture remains necessary for primary isolation of *M. tuberculosis* strains in settings where CRI methods on their own are anticipated, and for definitive diagnosis of TB in smear-negative patients (including settings where NRA, CRI methods and MODS are anticipated). Conventional DST is required to diagnose XDR-TB, irrespective of the methods used to detect MDR-TB;

- 5.6 As current methods only detect resistance to rifampicin and/or isoniazid, countries with documented or suspected cases of XDR-TB should establish or expand conventional culture and DST capacity for quality-assured susceptibility testing of second-line drugs, based on current WHO policy guidance;
- 5.7 Implementation of non-commercial culture and DST methods for rapid detection of MDR-TB should be phased in, starting at national/central reference laboratories. Expansion should only be considered within the context of availability of suitable laboratory infrastructure and trained personnel, quality of specimen transport systems, and country capability to provide appropriate treatment and management of MDR-TB patients once diagnosed;
- 5.8 Adequate and appropriate laboratory infrastructure and equipment should be provided, ensuring that required precautions for biosafety and prevention of contamination are met:
 - 5.7.1 Specimen processing for culture must be performed in biological safety cabinets (BSCs) in at least BSL2 facilities;
 - 5.7.2 Procedures for manipulation of cultures (conventional identification, conventional DST) must be performed in BSL3 facilities;
 - 5.7.3 Successful establishment, staffing, and maintenance of BSL2 and BSL3 laboratories are demanding. Upgrading of facilities and establishment of the required infrastructure should be carefully planned and adequately financed;
- 5.9 Mechanisms for rapid reporting of laboratory results to clinicians must be established to provide patients with the benefit of an early diagnosis.