Meeting Report

High-priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting

28–29 April 2014

Geneva, Switzerland

World Health Organization

Meeting Report

High-priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting

28–29 April 2014
Geneva, Switzerland

The meeting was convened by the Global TB Programme of the World Health Organization on behalf of the Global Laboratory Initiative and the New Diagnostics Working Groups of the Stop TB Partnership.

World Health Organization

Designed by GPS Publishing

WHO/HTM/TB/2014.18

Contents

Acknowledgements

This document was prepared by Claudia Denkinger (the Foundation for Innovative New Diagnostics, FIND) with Sandra Kik (McGill University) and Martina Casenghi (Médecins Sans Frontières) on the basis of consensus achieved at a meeting on high-priority target product profiles for new tuberculosis diagnostics, which was convened by the Global TB Programme of the World Health Organization on behalf of the Global Laboratory Initiative and the New Diagnostics Working Groups of the Stop TB Partnership.

This document was finalized following consideration of all comments and suggestion made by participants of the meeting.

The contributions of the following individuals are gratefully acknowledged: the chairs of the meeting (Catharina Boehme, Martina Casenghi, Daniela Cirillo, Tom Shinnick and Karin Weyer), the rapporteurs of the meeting (Martina Casenghi, Claudia Denkinger, Chris Gilpin and Sandra Kik) and the presenters at the meeting who provided important background or additional material to inform the discussions (David Alland, Daniela Cirillo, Frank Cobelens, David Dolinger, Carl Mendel, Madhukar Pai, Marco Schito and Matteo Zignol).

Madhukar Pai (McGill University) and his group as well as FIND are acknowledged for their considerable work in developing the target product profiles before the meeting. A number of stakeholders, including the Critical Path to TB Drug Regimens DST assay working group, have given valuable feedback on several earlier versions of the target product profiles.

Note to the reader

Because of the richness of the discussion and in an attempt to keep this report simple and readable, comments have not been attributed unless their content rendered attribution necessary. This report attempts to convey the themes addressed in each session, rather than attempting to provide a chronological summary of the dialogue.

Abbreviations

AG	aminoglycoside
CAP	capreomycin
DST	drug-susceptibility testing
EMB	ethambutol
FIND	Foundation for Innovative New Diagnostics
FQ	fluoroquinolone
HIV	human immunodeficiency virus
HRZE	isoniazid , rifampicin, pyrazinamide, ethambutol
INH	isoniazid
LAM	lipoarabinomannan
LVX	levofloxacin
MDR-TB	multidrug-resistant tuberculosis
MOX	moxifloxacin
NAAT	nucleic acid amplification test
NPV	negative predictive value
PaMZ	Pa824, moxifloxacin, pyrazinamide
PCR	polymerase chain reaction
PPV	positive predictive value
PZA	pyrazinamide
RIF	rifampicin
REMox	rifampicin, moxifloxacin, pyrazinamide, ethambutol
SNP	single nucleotide polymorphism
TB	tuberculosis
TPP	target product profile
WHO	World Health Organization
XDR-TB	extensively drug-resistant tuberculosis

Executive summary

Globally, one third of all tuberculosis (TB) cases are not notified, and many patients' samples do not undergo drug-susceptibility testing (DST). To achieve the targets for TB prevention, care and control that have been agreed for after 2015, new health-system strategies and diagnostic tools are critically important (1). The development of target product profiles (TPPs) helps to align the needs of end-users with the targets and specifications that product developers should meet for the performance and operational characteristics of a test. The meeting convened by the World Health Organization in April 2014 aimed to build consensus around four TPPs that were identified by stakeholders to be of high priority:

- a point-of-care non-sputum-based test capable of detecting all forms of TB by identifying characteristic biomarkers or biosignatures (known as the biomarker test);
- a point-of-care triage test, which should be a simple, low-cost test that can be used by first-contact health-care providers to identify those who need further testing (the triage test);
- a point-of-care sputum-based test to replace smear microscopy for detecting pulmonary TB (the smear-replacement test);
- a rapid drug-susceptibility test that can be used at the microscopy-centre level of the health-care system to select first-line regimen-based therapy (the rapid DST test).

Stakeholders were surveyed before the meeting to identify high-priority TPPs (2), and a Delphi-like process was used to facilitate consensus building around the TPPs. Shortened TPPs (including only key characteristics) were sent to invited participants (excluding industry representatives); participants were requested to provide a statement reflecting their level of agreement with each of the proposed characteristics for each of the TPPs. In total, 47 individuals were asked to participate in this process, of which 39 responded (response rate 83%). Prespecified agreement levels were achieved (that is, more than 50% of respondents gave a score of at least 4 (that is, mostly or fully agree)) for all characteristics. Characteristics for which less than 75% of the respondents agreed or with which a distinct subgroup disagreed were discussed at the meeting.

Further discussions at the meeting led to agreement on all key characteristics of the first three TPPs. Extensive discussions on the fourth TPP (for the rapid DST test) led to the consensus that the following anti-TB agents should be included in resistance testing at the microscopy level (they are presented in order of decreasing priority):

- rifampicin
- fluoroquinolones (including moxifloxacin)
- isoniazid and pyrazinamide, and
- second-line injectable agents (aminoglycosides and capreomycin).

Optimally, all anti-TB agents would be included, but as a minimum at least rifampicin should be included. Discussions about the highest acceptable price of a DST test did not result in an agreement.

Revised TPPs with suggested improvements are included in this report.

Background

Globally, one third of all tuberculosis (TB) cases are not notified, and many patients' samples do not undergo drug-susceptibility testing (DST) (3). The global strategy for TB prevention, care and control after 2015 aims at a 95% reduction in TB deaths and a 90% reduction in TB incidence by 2035. To achieve these goals the implementation of new health-system strategies and diagnostic tools are critically important (1).

The TB community has expressed the need for several additional TB diagnostic tests (4). The tests needed include those that can be used for triage and screening (5), tests for patients whose disease can be difficult to diagnose (such as children, HIV-positive patients and patients with extrapulmonary TB) (6), a simple point-of-care-test for active pulmonary TB (7), a molecular test to replace smear microscopy (8) at the microscopy-centre level of the health-care system, DST (9), predictive biomarker tests for latent TB infection (10) and tests to monitor treatment (11).

An increasing number of companies have shown interest in developing TB diagnostics, and several tests are in the pipeline. Partly this interest has been triggered by the success of the Xpert MTB/RIF assay (Cepheid, Sunnyvale, CA, United States) (12, 13). For test developers it is important to understand which of these products are perceived to be the highest priority, the necessary performance and operational characteristics of the different tests and what the potential market size would be for these new tests.

A priority-setting exercise was conducted by McGill University and its partners to identify which of the diagnostic needs should be the highest priority for further TPP development, and investment in research and development. Predefined expert groups were asked to rank nine diagnostic needs (Figure 1) based on the following criteria: their prioritization according to key stakeholders; the potential impact of the test on TB transmission, morbidity and mortality; the market potential of the test; and how easy the test would be to implement and scale-up.

Figure 1. Diagnostic tests needed, by main indication

Triage test to rule out disease and for systematic screening
Triage test for those seeking care
Test that could be used in an HIV/ART clinic to rule out active TB
Systematic screening test for active case-finding
Rapid TB diagnosis (with optional drug-susceptibility testing)
Rapid, sputum-based cartridge-based molecular test for use in microscopy centres (with the option of an add-on DST cartridge-based test)
Rapid biomarker-based instrument-free test for non-sputum samples that could also detect childhood TB and extrapulmonary TB
Multiplexed test for TB and other infectious diseases
Next-generation drug-susceptibility test
Centralized, high-throughput drug-susceptibility test incorporating testing for new anti-TB agents to support the roll-out of new TB treatment regimens after 2014
Treatment monitoring test
Test to monitor treatment (to test for cure)
Predictive test for latent TB infection
Predictive test for latent TB infection in patients at high risk of active TB

HIV, human immunodeficiency virus; ART, antiretroviral therapy.

The three diagnostic priorities identified were (2):

- a point-of-care biomarker-based non-sputum-based test to detect TB;
- a point-of-care test that could be used for triage;
- a point-of-care sputum-based test that could be used as a replacement for smear microscopy with the capacity to do DST (at the microscopy-centre level).

1.1. Developing target product profiles

Manufacturers need TPPs at an early stage in the diagnostic development process to inform the targets and specifications for the performance and operational characteristics of a test that will also meet the needs of end users. At a minimum, the TPPs for diagnostic tests should specify the clinical purpose of the test, the goal to be met (for example, to initiate treatment), the target population that will be tested, the level of implementation in the health-care system, and the intended end-users (4). In addition, TPPs should outline the most important performance and operational characteristics (with the term "minimal" used to refer to the lowest acceptable output for a characteristic, and "optimal" used to refer to the ideal target for a characteristic). The optimal and minimal characteristics define a range. Products should meet at least all of the minimal characteristics, but preferably they would meet as many of the optimal characteristics as possible.

For the three key diagnostic needs identified by the prioritization exercise, the following TPPs were developed by McGill University and the Foundation for Innovative New Diagnostics (FIND) with input from several stakeholders and a technical advisory group (funding was provided by the Bill and Melinda Gates Foundation):

- a TPP for a point-of-care non-sputum-based test capable of detecting all forms of TB by identifying characteristic biomarkers or biosignatures (referred to as the biomarker test);
- a TPP for a point-of-care triage test, which should be a simple, low-cost test that can be used by first-contact health-care providers to rule-out TB (referred to as a triage test);
- a TPP for a point-of-care sputum-based test to be used as a replacement for smear microscopy (the smear-replacement test); and
- a TPP for rapid drug-susceptibility tests that can be used at microscopy centres (referred to as a rapid DST test).

The initial TPPs were detailed and tried to incorporate information about as many characteristics as possible (Annex A). The timeline envisioned in the TPPs for the development of the tests is 5 years. For several of the characteristics, only limited evidence was available and expert advice was sought from at least 20 stakeholders and technical experts for each of the TPPs (Annex A).

In addition to the consultative process that was carried out during the preparation of the TPPs, it was considered important to bring the TPPs to a larger stakeholder audience, including clinicians, implementers and representatives of countries and national TB programmes, before they were finalized.

To meet this aim, a consensus-gathering meeting was organized in April 2014 by the Global TB Programme at WHO on behalf on the Global Laboratory Initiative and New Diagnostic Working Groups of the Stop TB Partnership. For the purpose of the meeting on high-priority TPPs, key characteristics for each of the TPPs were identified in order to shorten the TPPs and to facilitate the process of finding consensus on the most important characteristics.

1.2. Delphi process

Before the meeting, a Delphi-like process was used to facilitate consensus building. The shortened TPPs were sent to all invited participants (excluding individuals working in the industry in order to avoid possible bias). Participants were requested to provide a statement on how much they agreed with each of the proposed characteristics for each of the TPPs. Agreement was scored on a Likert scale ranging from 1 to 5 (1-disagree, 2-mostly disagree, 3-don't agree or disagree, 4-mostly agree, 5-fully agree). Individuals were asked to provide comments when they did not agree with a statement (that is, when they scored a characteristic at 3 or lower).

To reach consensus it was prespecified that more than 50% of respondents should have provided a score of at least 4 on the Likert scale (indicating that they mostly or fully agreed with a characteristic). In total, 47 individuals were asked to participate in this process, and 39 responded (response rate, 83%).

Nearly half of respondents were from academia or product-development partnerships. About 25% of respondents worked for national tuberculosis programmes or in TB Supranational Reference Laboratories; 13% were implementing partners of diagnostic technologies and clinicians. Six respondents (15%) worked for donors, funders or an international body such as WHO.

Before the meeting, it had been anticipated that two rounds of the Delphi survey would be needed, but since predefined consensus for all characteristics was reached after the first round, no second round was initiated. Characteristics on which less than 75% of the respondents agreed, or on which a distinct subgroup disagreed, were discussed during the meeting with the comments that had been provided.

1.3. Consensus meeting

The meeting focused on building further consensus around the four high priority TPPs, the intended use of the diagnostics, and their performance, operational characteristics and pricing.

Participants represented stakeholders from technical and funding agencies, researchers, implementers, representatives from countries and civil-society organizations, and representatives from companies working on the development of TB diagnostics (see Annex B and Annex C).

2. A rapid biomarker-based non-sputum-based test for detecting TB

Presenter: Sandra V. Kik, McGill University, International TB Centre, Montreal, Canada

2.1. Background information

The majority of cases of pulmonary TB are diagnosed by sputum-smear microscopy. However, smear microscopy has suboptimal sensitivity, and children and individuals infected with the human immunodeficiency virus (HIV) often have difficulty providing good quality sputum samples (14–16). From 15% to 25% of all TB cases are extrapulmonary TB. The diagnosis of extrapulmonary TB is often missed by sputum evaluation, and requires invasive sample collection (3). While the Xpert MTB/RIF assay is more sensitive than sputum smears, it cannot be used in the majority of microscopy centres owing to the need for a continuous and stable power supply to operate the GeneXpert instrument. In addition, the type of samples required for the Xpert MTB/RIF assay (a sputum specimen or extrapulmonary specimens, such as cerebrospinal fluid, or samples of lymph nodes or other tissues) may be difficult to obtain in peripheral health-care settings and from individuals who are suspected of having extrapulmonary TB or paucibacillary TB (17).

According to the analysis of stakeholders' requirements performed by McGill University, the most urgent need is for a rapid biomarker-based non-sputum-based diagnostic test that uses an easily accessible sample and is able to accurately diagnose pulmonary TB (and ideally also extrapulmonary TB) (2).

Ideally, a biomarker test should be easy to perform and implement at health posts, where health-care workers with limited training would be able to conduct and interpret the test, and where standard TB treatment regimens can be initiated. Such a biomarker test would conceivably increase the number of individuals suspected of having TB who are tested as well as the number of patients diagnosed with TB. Additionally, if the sensitivity of the biomarker test were greater than that of smear microscopy for pulmonary TB and extrapulmonary TB, this would further add to the value of the test. However, when the aim of a biomarker test is to replace smear microscopy, at a minimum it should be able to accurately diagnose most smear-positive cases since these are known to contribute most to transmission (18). Also, the test would ideally be able to diagnose adults and children, and pulmonary TB and extrapulmonary TB alike. However, if biomarkers or biosignatures for adult TB are suboptimal for childhood TB, or different for pulmonary TB and extrapulmonary TB, then the need for separate products should be considered.

To be successfully implemented at microscopy centres, health posts and primary-care clinics, the test should use an easily accessible sample (such as urine, blood, or breath condensate). Non-sputum-based samples are preferred, both for biosafety reasons and because of the stigma associated with TB.

The simpler, more portable and more durable that the test is, the more likely it is to be implemented in peripheral settings. Due to conditions that prevail in peripheral settings in countries with a high burden of TB, it is necessary that the test uses simple sample preparation, has minimal operational and maintenance requirements, and provides results that are easy to interpret (8) (Table 1).

Table 1. Scope for a biomarker-based diagnostic test for TB as defined prior to Delphi survey

Characteristic	Optimal requirements	Minimal requirements
Scope		
Goal	To develop a rapid biomarker-based test that can diagnose **pulmonary TB** and, ideally, also **extrapulmonary TB** using non-sputum samples (such as, urine, blood, oral mucosal transudates, saliva, exhaled air) for the purpose of initiating TB treatment during the same clinical encounter or on the same day	
Target population	Target groups are adults and children including those who are HIV-positive who are suspected of having active TB; this includes both pulmonary TB and extrapulmonary TB in countries with a medium prevalence to a high prevalence of TB as defined by WHO[a]	
Target user of test	Health-care workers with a minimum of training	Trained microscopy technicians
Setting (level of the healthcare system)	Health posts without attached laboratories (a level lower than microscopy centres)	Primary health clinics that have laboratories; peripheral microscopy centres

HIV, human immunodeficiency virus.
a High-prevalence countries are those with > 40 cases per 100 000 population; medium-prevalence countries are those with 20–40 cases per 100 000 population; and low-prevalence countries are those with < 20 cases per 100 000 population (19).

2.2. Discussion

Table 2 lists the seven key performance and operational characteristics of the biomarker test as phrased for the Delphi process. The proportions of respondents that agreed with the optimal and minimal key characteristics as proposed are provided together with the most common areas of disagreement raised in the Delphi survey. For all seven key characteristics the majority of respondents (>50%) agreed with both the optimal specifications and the minimal specifications .

During the meeting the discussion focused on three characteristics for which there was less than 75% agreement.

Table 2. Delphi survey results for seven proposed key characteristics of a rapid biomarker-based non-sputum-based test for detecting TB

Characteristic	Optimal requirements	% (range) agreeing with the optimal requirements for the characteristic[a]	Minimal requirements	% (range) agreeing with the minimal requirements for the characteristic[a]	Comments
1 Diagnostic sensitivity for pulmonary TB in adults	Sensitivity should be ≥ 98% for smear-positive culture-positive pulmonary TB, and ≥ 68% for smear-negative, culture-positive pulmonary TB in adults (that is, similar to the sensitivity of the Xpert MTB/RIF assay); overall pooled sensitivity should be ≥ 80% in adults with HIV infection	90% (75–91%)	Overall sensitivity should be > 65%, but should be > 98% among patients with smear-positive culture-positive pulmonary TB (that is, it should be similar to smear microscopy)	95% (79–96%)	Optimal: sensitivity should be higher (better than the Xpert MTB/RIF assay) Minimal: if the test had sensitivity < 65% it could still be extremely useful when combined with smear microscopy
2 Diagnostic sensitivity for extrapulmonary TB in adults	Sensitivity should be ≥ 85% in lymph node aspirates or tissue; sensitivity should be ≥ 80% in cerebrospinal fluid for microbiologically confirmed TB (that is, sensitivity should be equal to that of the Xpert MTB/RIF assay)	85% (70–87%)	No lower range of sensitivity was defined	69% (57–74%)	Optimal: participants commented that sensitivity should not be defined since it will depend on the type of sample Minimal: a minimal sensitivity should be defined and should be at least that of Xpert MTB/RIF for the most common forms of extrapulmonary TB)
3 Diagnostic sensitivity in children	Sensitivity for childhood intrathoracic TB should be ≥ 66% for microbiologically confirmed TB (that is, equal to the sensitivity of the Xpert MTB/RIF assay)	85% (70–87%)	No lower range of sensitivity was defined	72% (60–77%)	Optimal: some participants thought that the sensitivity should be higher than that of Xpert MTB/RIF Minimal: some participants thought that a minimal sensitivity should be defined and should be at least similar to Xpert MTB/RIF

4 Diagnostic specificity	Should be at least as specific as the Xpert MTB/RIF assay for detecting pulmonary TB, extrapulmonary TB and childhood TB (that is, the test should have 98% specificity against a microbiological reference standard); the test should be able to distinguish between active TB and latent or past infection	92% (77–94%)	Specificity should be as described for the optimal value	87% (72–89%)	Minimal: there should be a difference between the optimal and minimal values. The minimal specificity may be lower, but it should be considered with the test's sensitivity
5 Time to result	< 20 minutes including time spent preparing the sample	97% (81–98%)	< 1 hour including time spent preparing the sample	85% (70–87%)	Minimal: comments about this value were diverse; both shorter and longer times until results were proposed
6 Maintenance and calibration	Disposable, no maintenance required	97% (81–98%)	Preventative maintenance should not be needed until after 1 year or > 1000 samples; only simple tools and minimal expertise should be required; an alert to indicate when maintenance is needed should be included; the instrument should be able to be calibrated remotely or no calibration should be needed	90% (75–91%)	Minimal: even minimal maintenance requirements may be challenging in many peripheral health-care settings
7 Price of individual test (cost of reagent and consumables only; after scale-up; ex-works [manufacturing costs only, excluding shipping])	<US$ 4.00	74% (62–79%)	<US$ 6.00	77% (64–81%)	Optimal: the price is too high Minimal: comments about pricing were diverse; the price was considered both too high and too low by respondents; the price of an individual test should be considered together with other test characteristics

HIV, human immunodeficiency virus.
a A range was calculated assuming that non-respondents would either all agree (high margin) or disagree (low margin) with the proposed characteristic.

2.3. Sensitivity for detecting extrapulmonary TB in adults

In the proposed TPP, no minimal threshold was defined for the sensitivity of diagnosing extrapulmonary TB in adults. The main reason for this omission was that data to support an evidence-based threshold for a minimal sensitivity were limited, and any added sensitivity would in fact be better than what is achieved by smear microscopy. During the discussion the point was made that the wording in the TPP for this characteristic has the risk of not sufficiently emphasizing the benefits to individual patients that would result from a test that is able to diagnose extrapulmonary TB. The importance of this was emphasized even though the population-level impact of detecting extrapulmonary TB is limited because it generally does not result in transmission. It was also mentioned that a test for extrapulmonary TB should be able to detect all forms of extrapulmonary TB, not only those included in WHO's policy update on the use of Xpert MTB/RIF (that is, it should be able to detect TB in samples from lymph nodes, other tissues, cerebrospinal fluid and pleural fluid), and that the reference standard should be clarified (17).

The TPP was revised to highlight the importance of detecting extrapulmonary TB, and a footnote was added to clarify the reference standard for determining the accuracy of the test.

2.4. Sensitivity for detecting childhood TB

In the TPP no minimal sensitivity was proposed for diagnosing TB in children. The main reason for this omission was because unpublished modelling data have shown that over a 10-year timeframe, an improved test for detecting pulmonary TB in adults would have a greater population-level impact on mortality from childhood TB than would a test that specifically targeted TB in children (C. Denkinger et al, unpublished data, 2013). Ideally, however, a test would improve the diagnosis of TB in both adults and in children because the individual-level impact on mortality of a test for childhood TB is more immediate. The revised wording in the TPP highlights the benefits to individual patients that would result from a test that is able to diagnose childhood TB. It was also noted that current tests for biomarkers (such as lipoarabinomannan, or LAM) generally perform better in adults than in children. This might make separate tests necessary for children and adults.

For optimal sensitivity it was suggested that it is important to specify the reference standard against which the performance of a test for childhood TB will be assessed. This information was included in a footnote.

2.5. Price of the test

In order to inform the discussion on pricing, the results of a continuing study on costs and affordability (A. Pantoja et al, unpublished data, 2014) were presented at the meeting. This analysis has shown that compared with 2014 spending on TB diagnosis in three countries (India, Kenya and the Russian Federation), a test at the minimal price target (US$ 4.00) would be affordable in those countries, assuming that no confirmatory testing would be needed. At the higher cost (US$ 6.00), the test would be affordable only in the Russian Federation. A test was considered affordable if it did not result in a higher proportion of a current TB budget being spent on TB diagnostics using the current diagnostic algorithm.

It was concluded that a test needed to be affordable (that is, within the limitations of a programme's current budget). However, the TPP is developed for a 5-year timeframe and, therefore, it might not be worthwhile to provide detailed specifications for a price range at this stage. Furthermore, the cost of the test would be driven by the types of biomarkers assessed and the number of markers included in the test.

One additional comment made was that the costs of the biomarker test should not be more than the current alternative (that is, smear

microscopy). However, while costs for reagents and consumables for smear microscopy might be cheaper, the overall costs, including the costs of quality assurance, may come close to the minimum cost suggested for the biomarker test. If the biomarker-based test were simple enough that it did not require substantial external quality assurance, then it would result in savings.

In conclusion, there was general agreement on the importance of having an affordable test and on the price range provided in the revised TPP (Table 3).

2.6. Other comments

Additional points made during the discussion are summarized here.

First, participants pointed out that there was no mention of a minimal sensitivity for diagnosing pulmonary TB in patients coinfected with HIV. Wording was added to improve the completeness of the description.

Second, a biomarker-based test might be able to diagnose earlier stages of disease than culture does. This will be captured only if a composite reference standard is used.

Third, requirements for higher operating temperatures than the maximum of 45 °C that was incorporated into the TPP should be considered (in Brazil, India and other areas, temperatures often reach 50 °C).

Fourth, if the test is introduced by a national TB programme at levels of the health-care system that are below microscopy centres, decentralized diagnosis may be possible but patients may still need to be referred to start treatment unless links to treatment can be established.

Additionally, the ability to perform multi-analyte testing for different diseases on one platform would be considered an important advantage.

Table 3. Revised target product profile (TPP) for a rapid biomarker-based non-sputum-based test to detect TB, using input from a Delphi survey and discussions at a consensus meeting, 2014

Characteristic	Optimal requirements	Minimal requirements
Scope		
Goal	To develop a rapid biomarker-based test that can diagnose **pulmonary TB** and optimally also **extrapulmonary TB** using non-sputum samples (for example, urine, blood, oral mucosal transudates, saliva, exhaled air) for the purpose of initiating TB treatment during the same clinical encounter or on the same day	
Target population	Target groups are adults and children including those who are HIV-positive and suspected of having active pulmonary TB or extrapulmonary TB in countries with a medium prevalence to a high prevalence of TB as defined by WHO[a]	
Target user of test[b]	Health-care workers with a minimum of training	Trained microscopy technicians
Setting (level of the healthcare system)	Health posts without attached laboratories (that is, levels below microscopy centres) or higher levels of the health-care system	Primary health-care clinics with attached laboratories; peripheral microscopy centres or higher levels of the health-care system

Performance characteristics		
Diagnostic sensitivity for pulmonary TB in adults[b]	Sensitivity should be ≥ 98% for smear-positive culture-positive pulmonary TB, and ≥ 68% for smear-negative culture-positive pulmonary TB in adults (that is, sensitivity should be similar to that of the Xpert MTB/RIF assay) Overall pooled sensitivity should be ≥ 80% in adults with HIV infection	Overall sensitivity should be ≥ 65% but should be > 98% among patients with smear-positive culture-positive pulmonary TB (that is, sensitivity should be similar to that of smear microscopy) Overall pooled sensitivity should be better than the sensitivity of smear microscopy in adults with HIV infection
Diagnostic sensitivity for extrapulmonary TB in adults	Ideally, sensitivity should be ≥ 80% for all forms of microbiologically confirmed extrapulmonary TB[c,d]	Diagnosis of extrapulmonary TB is an important need, and a test that can diagnose extrapulmonary TB in addition to pulmonary TB will have significant benefits for individual patients; additionally, it is likely to be better accepted in the community of care providers. No lower range of sensitivity was defined
Diagnostic sensitivity in children	Sensitivity for childhood intrathoracic TB should be ≥ 66% for microbiologically confirmed TB (that is, similar to the sensitivity of the Xpert MTB/RIF assay)[e]	Diagnosis of childhood TB is an important need, and a test that improves the diagnosis of TB in children will have significant benefits for individual patients; additionally, it is likely to be better accepted in the community of care providers. No lower range of sensitivity was defined
Diagnostic specificity[b]	At least as specific as the Xpert MTB/RIF assay for detecting pulmonary TB, extrapulmonary TB and childhood TB (that is, the test should have 98% specificity when compared against a microbiological reference standard); the test should distinguish between active TB and latent or past infection	
Operational characteristics		
Sample type	Not invasive or minimally invasive, non-sputum samples (such as, urine, blood, oral transudates, saliva, exhaled air)	
Manual preparation of samples (steps needed after obtaining sample)	Sample preparation should be integrated or manual preparation should not be required	A limited number of steps only; precise measuring should not be needed for any step (such as precise measuring of volumes or time)
Time to result[b]	< 20 minutes including time spent preparing the sample	< 1 hour including time spent preparing the sample

Instrument and power requirement	No instrument needed	Small, portable or hand-held instrument (weighing < 1 kg) that can operate on battery or solar power in places where power supplies may be interrupted
Maintenance and calibration[b]	Disposable, no maintenance required	Preventative maintenance should not be needed until after 1 year or > 1 000 samples; only simple tools and minimal expertise should be required; an alert to indicate when maintenance is needed should be included; the instrument should be able to be calibrated remotely or no calibration should be needed
Operating temperature and humidity level	Between +5 °C and +50 °C with 90% humidity	Between +5 °C and +40 °C with 70% humidity
Result capturing, documentation, data display	An instrument-free test with the ability to save results using a separate, attachable reader	The test menu must be simple to navigate; the instrument should have an integrated LCD screen, simple keypad or touch screen, and the ability to save results using either the instrument or a separate reader
Internal quality control	Internal controls should be included for processing the sample and detecting TB	Internal control included only for processing the sample
Pricing		
Price of individual test[b] (costs of reagents and consumables only; after scale-up; ex-works [manufacturing costs only, excluding shipping])	< US$ 4.00	< US$ 6.00

HIV, human immunodeficiency virus; LCD, liquid crystal display.
a High-prevalence countries are those with > 40 cases per 100 000 population; medium-prevalence countries are those with 20–40 cases per 100 000 population; and low-prevalence countries are those with < 20 cases per 100 000 population (19).
b These characteristics were considered to be the most important, and specific consensus was asked for and reached through a Delphi survey.
c The sensitivity for detecting extrapulmonary TB should also be tested against a composite reference standard that includes culture with or without a nucleic acid amplification test, histology, smear microscopy, biochemical testing, presenting signs, and response to treatment with anti-TB therapy, depending on site of infection. Xpert MTB/RIF testing has an estimated sensitivity for diagnosing TB of 84% for lymph node aspirates or other tissue samples, and 55% sensitivity for samples of cerebrospinal fluid, when compared with a composite reference standard, but Xpert MTB/RIF testing requires invasive samples (17).
d Xpert MTB/RIF has an estimated sensitivity for microbiologically confirmed TB of 85% for detecting TB in lymph node aspirates or other tissue samples, 80% for cerebrospinal fluid, and 44% for pleural fluid but testing requires invasive samples (from aspiration, biopsy, lumbar puncture or thoracentesis).
e The test's sensitivity in children should be evaluated against a composite reference standard as defined by an international panel of experts (6).

3. Community-based triage or referral test for identifying people suspected of having TB

Presenter: Claudia Denkinger, FIND, Geneva, and McGill University, Montreal, Canada

3.1. Background information

The Xpert MTB/RIF assay has greater accuracy in diagnosing active TB than sputum-smear microscopy, and it is being rolled out in more than 90 countries around the world (13). However, the placement of the GeneXpert instrument in peripheral settings is limited by a lack of sufficient infrastructure (such as a stable power supply) and by extreme environmental conditions (20).

In most countries with a high burden of TB, 2 weeks of cough is widely used to define individuals suspected of having active pulmonary TB who require diagnostic testing. However, the prevalence of TB among these patients is relatively low (21). Since most individuals suspected of having TB do not have TB, a triage test could help narrow the population that needs costly confirmatory testing.

A triage test needs to be a simple low-cost test that can be used by first-contact providers in the community (such as community health workers) to rule out TB (that is, to identify people who are triage-test negative) and direct individuals who require further evaluation (that is, those who are triage-test positive) to a confirmatory test (such as the Xpert MTB/RIF assay or similar molecular test) (Table 4). Triage testing could take place at the same level of care as confirmatory testing, especially in settings that see a large number of patients (such as outpatient clinics), but typically a triage test is conceived to be used at lower levels of care (for example, at the microscopy-centre level and below).

Ideally, the sensitivity of a triage test would be as good as that of the confirmatory test. However, if a triage test is done at lower levels of care, or if it is easier to use than a confirmatory test and more people suspected of having TB will be tested, then the test might increase the number of TB patients identified even if its sensitivity were lower than that of the confirmatory test. There are no data that can provide information on this trade-off between sensitivity and ease of use, and on the impact of lower sensitivity.

A triage strategy that reduces the use of expensive tests and potentially tests more patients could benefit patients and health systems. It may reduce costs for the programme and the workload of health-care staff, and may increase the number of patients who are identified as having TB (5).

The specificity required must consider the prevalence of TB in the population tested. The prevalence of TB among individuals who seek care and testing in the community is likely to be lower (estimated to be less than 5%) than that among individuals who seek care and testing in clinics or health posts (estimated to be less than 10%). Therefore, when testing in the community a test with a higher specificity is necessary in order to achieve a sufficiently high post-test probability that justifies referral (for example, when there is a 5% prevalence among people seeking care, if a test with 95% sensitivity and 80% specificity is used then the post-test probability after a positive triage test is 24%). A test with low specificity (resulting in a high number of false-positive results) may jeopardize the motivation of patients and health-care workers to act on a positive result.

To be successfully implemented at the community level, a test would ideally use an easily accessible sample (for example, urine or blood from a finger-stick). The simpler, more portable and more durable the test is, the more likely it will be implemented in peripheral settings. Ideally, the test would not require a device or the device

would be battery operated or not require power (8, 20). The ideal time to result (including sample preparation and processing time) has not been studied, and might vary significantly between countries and between settings where patients are tested; however, a rapid test will be more likely to be integrated into the workflow and result in a decision being made during the same visit.

Furthermore, it will be essential that only simple tools and minimal expertise are necessary for both operation and maintenance since the expertise of operators might be limited. An internal quality control needs to be included in the test (22). Given the expected environmental conditions in endemic countries, tolerance of high temperatures and humidity is also required (8, 20).

Table 4. Scope for a triage test for detecting TB as defined prior to Delphi survey

Characteristic	Optimal requirements	Minimal requirements
Scope		
Goal	To develop a test that can be used at the entry point to the health-care system to identify patients with signs and symptoms of **active TB** including those who may be coinfected with HIV, those who do not have TB and those in need of referral for further confirmatory testing	To develop a test that can be used at the entry point to the health-care system to identify patients with signs and symptoms of active **pulmonary TB**, including those who may be coinfected with HIV, those who do not have TB and those in need of referral for further confirmatory testing
Target population	Target groups are adults and children with signs and symptoms of **active TB at any site** in countries with a medium prevalence to a high prevalence of TB as defined by WHO[a]	Target groups are adults and children with signs and symptoms of **active pulmonary TB** in countries with a medium prevalence to a high prevalence of TB as defined by WHO[a]
Target user of test	Community health workers who have had a minimum of training, and informal providers	Health workers trained to the level of auxiliary nurses

HIV, human immunodeficiency virus.
a High-prevalence countries are those with >40 cases per 100 000 population; medium-prevalence countries are those with 20–40 cases per 100 000 population; and low-prevalence countries are those with <20 cases per 100 000 population (19).

3.2. Discussion

Table 5 lists the six key performance and operational characteristics of the triage test as they were phrased for the Delphi process. The proportions of respondents that agreed with the optimal and minimal key characteristics as proposed are provided together with the most common areas of disagreement raised in the Delphi survey. For all six key characteristics, the majority of respondents agreed with both the optimal specifications as well as the minimal specifications.

The discussion focused on the two key operational characteristics for which more than 75% agreement was reached: the minimal number of manual steps required to perform the test and its price.

Table 5. Delphi survey results for six proposed key characteristics of a community-based triage or referral test for identifying people suspected of having TB

Characteristic	Optimal requirements	% (range) agreeing with the optimal requirements for the characteristic[a]	Minimal requirements	% (range) agreeing with the minimal requirements for the characteristic[a]	Comments
1 Diagnostic sensitivity	Overall sensitivity should be >95% compared with the confirmatory test for pulmonary TB; no lower range of sensitivity was defined for extrapulmonary TB	84% (66–87%)	Overall sensitivity should be >90% compared with the confirmatory test for pulmonary TB	81% (65–84%)	Several participants commented that considerations of sensitivity, specificity and price are interrelated. One suggestion was that a triage test should have higher sensitivity or sensitivity equal to that of the confirmatory test. On the other hand, it was commented that sensitivity targets might be overambitious and that it would be useful to model the effect of a screening test that had lower levels of sensitivity. Another comment was that if a test could do multiplexed testing for other diseases (for example, for HIV or malaria as well) that may make a lower sensitivity more acceptable.
2 Diagnostic specificitya	Specificity should be > 80% when compared with the confirmatory test	81% (64–85%)	Specificity should be > 70% when compared with the confirmatory test	81% (64–85%)	One respondent commented that reduced specificity would be acceptable only if sensitivity is > 95%. Several suggested that the test could be used for triage (that is, in patients with symptoms) only in areas where the prevalence is typically at least 10 %; for low-prevalence settings, the positive predictive value would be too low as a false-positive rate of 20–30% is too high.
3 Manual preparation of samples (steps needed after obtaining sample)	Sample preparation should be either integrated or manual preparation should not be required (excluding waste disposal); precise timing and measuring should not be required	78% (62–83%)	Should require only 2 steps (excluding waste disposal); precise timing and measuring should not be required	73% (57–79%)	Minimal: if implemented at the community level, 2 steps might be prohibitive and also would require additional quality control; Optimally, the same sample would be able to be used for a confirmatory test

4 Maintenance and calibration	Device should be disposable, and should not require maintenance or calibration	86% (68–89%)	Preventative maintenance should not be needed until after 1 year or 1 000 samples; only simple tools and minimal expertise should be required; an alert to indicate when maintenance is needed should be included; the device should be able to be calibrated remotely, should calibrate itself, or no calibration should be required	81% (64–85%)	Minimal: requiring any maintenance or calibration at lower levels of the health-care system is challenging; for quality assurance purposes, any maintenance would need trained personnel
5 Time to result	< 5 minutes	89% (70–91%)	< 30 minutes	81% (60–85%)	Minimal: 3 respondents suggested that 30 minutes is too long for a triage test to be feasible in crowded settings; 2 respondents suggested that the minimal time requirement should be extended to 1 hour Optimal: 5 respondents considered 5 minutes to be unrealistic and suggested <15 minutes
6 Price of individual test (costs of reagents and consumables only; after scale-up; ex-works [manufacturing costs only, excluding shipping])	< US$ 1.00	70% (55–77%)	< US$ 2.00	62% (49–70%)	Minimal: equal numbers of respondents argued that price should be higher (highest price suggested was US$ 5.00) or lower (lowest price suggested was <US$ 1.00). The differences in price between the minimal and optimal requirements were considered too small

a A range was calculated assuming that non-respondents would either all agree (high margin) or disagree (low margin) with the proposed characteristic.

3.3. Manual preparation of samples

Manual preparation of samples that takes two steps (as described by the minimal scenario) was considered to possibly be prohibitive at the lower levels of the health-care system where this test might be used, and a two-step procedure would also require additional quality control. Additionally, the confirmatory test would ideally use the same sample as the triage test had used (that is, confirmatory testing would not require an additional visit by the patient or another specimen to be collected). However, participants at the meeting agreed with the specification as defined under 'minimal' for the sample preparation, and no changes were made to the specifications.

3.4. Price of the test

In order to inform the discussion on pricing, the results of an analysis of costs and affordability were presented at the meeting (A. Pantoja et al, unpublished data, 2014). This analysis has shown that a triage test followed by Xpert MTB/RIF as a confirmatory test would result in increased spending on diagnosing TB in the three countries studied (India, Kenya and the Russian Federation), even when the triage test would cost only US$ 1.00. The results of this affordability analysis contrast with data published in 2013 that showed that if triage testing were used at the same level of the health-care system as the Xpert MTB/RIF assay (that is, at the district level), it would be cost effective even when the triage test costs more than US$ 1.00 (5). The differences between the two analyses relate to the populations tested: the cost–effectiveness analysis assumed that the same patient population receives the triage test and the confirmatory test, while the affordability analysis assumed that 30% more patients would be reached (the additional patients would be reached at the health-post level) and there is testing for extrapulmonary TB and paucibacillary TB.

As shown in the Delphi survey (Table 5), equal numbers of respondents suggested a more expensive price for the test as suggested a cheaper price (the maximum was US$ 5.00 versus a minimum of less than US$ 1.00). However, in general it was felt that the triage test should be the cheapest of all the tests described in the TPPs, and this is reflected in the proposed target price in this TPP. There was general agreement with the importance of having an affordable test and with the price range provided in the revised TPP (Table 6).

3.5. Other comments

Other general points that were raised related in part to the characteristics specified in the TPP that were not part of the Delphi survey. First, it was noted that even though it is not specifically outlined in the TPP, the triage test should ideally maintain its sensitivity in HIV-positive patients. Second, in the Delphi survey one respondent voiced concern over the ease of interpreting a visual display of results and the likelihood of variability in interpreting the results among readers, noting that both should be considered during test development.

It was suggested that when describing the goal of the test, a clarification was needed as to which individuals should be tested. For instance, if symptom screening (for example, 2 weeks of cough) is used as a requirement before testing, then this might limit the number of patients tested and might still miss people with TB. Therefore, several participants suggested specifying that individuals with any symptoms suggestive of TB or risk factors for TB in the absence of symptoms (for example, people with HIV or diabetes) should be eligible for testing. The text in the TPP was adjusted accordingly.

A triage test should not be confused with a test that can be used for screening (active case-finding). The characteristics of a screening test may be considerably different, in particular with regard to its specificity. Therefore, it is necessary to develop a separate TPP for a screening test.

Another important point was that earlier stages of disease and extrapulmonary TB might be diagnosed with a biomarker-based test, and it might be challenging to differentiate between false-positive and true-positive results using a microbiological reference standard or

confirmatory test. Thus, it is important to consider using a composite reference standard that includes both clinical and microbiological results.

Lastly, it was mentioned that a prerequisite for a triage test should be that patients will have access to a good confirmatory test as well (that is, it is not sufficient to use smear microscopy as a confirmatory test).

Table 6. Revised target product profile (TPP) for a community-based triage or referral test to identify people suspected of having TB, using input from a Delphi survey and discussions at a consensus meeting, 2014

Characteristic	Optimal requirements	Minimal requirements
Scope		
Goal	To develop a test that can be used during a patient's first encounter with the health-care system to identify patients with **any symptoms of or risk factors for active TB**, including patients coinfected with HIV, those who do not have TB and those who need referral for further confirmatory testing	To develop a test that can be used during a patient's first encounter with the health-care system to identify patients with **any symptoms of or risk factors for active pulmonary TB**, including patients coinfected with HIV, those who do not have TB and those who need referral for further confirmatory testing
Target population	Adults and children with signs and symptoms of **active TB at any site** in countries with a medium prevalence to a high prevalence of TB as defined by WHO[a]	Adults and children with signs and symptoms of **active pulmonary TB** in countries with a medium prevalence to a high prevalence of TB as defined by WHO[a]
Target user of the test[b]	Community health workers and informal providers who have had a minimum of training	Health workers trained to the level of auxiliary nurses
Setting (level of the health-care system)	Community level or village level or higher levels of the health-care system	Health posts and primary-care clinics or higher levels of the health-care system
Performance characteristics		
Diagnostic sensitivity[b]	Overall sensitivity should be > 95% when compared with the confirmatory test for pulmonary TB[c] no lower range of sensitivity was defined for extrapulmonary TB[d]	Overall sensitivity should be > 90% compared with the confirmatory test for pulmonary TB[c]
Diagnostic specificity[b]	Specificity should be > 80% compared with the confirmatory test	Specificity should be > 70% compared with the confirmatory test
Operational characteristics		
Sample type	Non-sputum samples (such as urine, oral mucosal transudates, saliva, exhaled air or blood from a finger-stick)	Sputum; non-sputum samples are preferred (such as urine, oral mucosal transudates, saliva, exhaled air, or blood from a finger-stick; imaging technology

Manual preparation of samples (steps needed after obtaining sample)	Sample preparation should be integrated or manual preparation should not be required (excluding waste disposal); precise timing and measuring should not be required	2 steps (excluding waste disposal); precise timing and measuring should not be required
Time to result[b]	< 5 minutes	< 30 minutes
Instrument and power requirement	None	Small, portable or hand-held device (weighing < 1 kg); should have an option for battery power or solar power
Maintenance and calibration[b]	Disposable, no maintenance required	Preventative maintenance should not be needed until after 1 year or 1 000 samples; only simple tools and minimal expertise should be required; an alert to indicate when maintenance is needed should be included; the device should be able to be calibrated remotely, should calibrate itself, or no calibration should be required
Operating temperature and humidity level	Between +5 °C and +50 °C with 90% humidity	Between +5 °C and +40 °C with 70% humidity
Result capturing, documentation and data display	An instrument-free test with visual readout and with the ability to save results using a separate, attachable reader	The test menu must be simple to navigate; the instrument should have an integrated LCD screen, a simple keypad or touch screen, and the ability to save results using either the instrument or a separate reader
Internal quality control	Internal controls should be included for processing the sample and detecting TB	Internal control included only for processing the sample
Pricing		
Price of individual test[b] (costs of reagents and consumables only; after scale-up; ex-works [manufacturing costs only, excluding shipping])	< US$ 1.00	< US$ 2.00

HIV, human immunodeficiency virus; LCD, liquid crystal display.
a High-prevalence countries are those with > 40 cases per 100 000 population; medium-prevalence countries are those with 20–40 cases per 100 000 population; and low-prevalence countries are those with < 20 cases per 100 000 population (19).
b These characteristics were considered to be the most important, and specific consensus was asked for and reached through a Delphi survey.
c The performance characteristics of the triage test need to match those of the confirmatory test that will be used.
d The sensitivity of the triage test should be compared with the sensitivity of a composite reference standard (that includes culture with or without a nucleic acid amplification test, histology, smear microscopy, biochemical testing, presenting signs and response to treatment with anti-TB therapy, depending on site of infection) to account for the fact that the test may detect cases of early TB or extrapulmonary TB in cases in which a standard microbiological reference standard might not perform well.

4. Rapid sputum-based test for detecting TB at the microscopy-centre level of the health-care system

Presenter: Claudia Denkinger, FIND, Geneva, and McGill University, Montreal, Canada

4.1. Background information

Although progress has been made in increasing TB cure rates, WHO's latest global report indicates that 33% of all TB cases are still not being notified (3). Additionally, the incidence of TB decreased by only 2% in 2013, and a substantial amount of transmission continues to occur (3).

A rapid, sensitive, easy to perform sputum-based test for detecting TB that could replace sputum-smear microscopy in decentralized settings, such as microscopy centres, has the potential to improve programmes and patient care in two ways. First, it would increase the number of patients with TB who are diagnosed and treated and thus reduce transmission; and, second, it would reduce morbidity because patients would be diagnosed and treated earlier (23). A test that is available at the level of a microscopy centre would leverage the infrastructure that is used for smear microscopy. It is conceivable that nucleic acid amplification tests (NAATs) or other technologies could replace smear microscopy. A test that can replace smear microscopy for both initial diagnosis and the monitoring of treatment (for example, by detecting viable bacteria) would be likely to be adopted more widely compared with a test that would facilitate only the initial diagnosis of TB.

Due to conditions that prevail in microscopy centres in countries with a high burden of TB, it will be necessary for the test to be robust, require simple sample preparation and have minimal operational requirements (Table 7). The scale at which such a test is adopted will likely depend on how well a new product meets the specified operational characteristics (8, 20).

Table 7. Scope for a test to replace smear microscopy for diagnosing TB as defined prior to Delphi survey

Characteristic	Optimal requirements
Scope	
Goal	To develop a sputum-based test to detect **pulmonary TB** at the microscopy-centre level of the health-care system to support initiation of TB therapy during the same clinical encounter or on the same day
Target population	Target groups are all patients suspected of having pulmonary TB who are able to produce sputum in countries with a medium prevalence to a high prevalence of TB as defined by WHO[a]
Target user of test	Health-care workers with a minimum of training
Setting (level of the healthcare system)	Microscopy-centre level (that is, primary health-care centres with attached peripheral laboratories)

a High-prevalence countries are those with > 40 cases per 100 000 population; medium-prevalence countries are those with 20–40 cases per 100000 population; and low-prevalence countries are those with < 20 cases per 100000 population (19).

4.2. Discussion

Table 8 lists the seven key performance and operational characteristics of the rapid sputum-based test as they were phrased for the Delphi process. The proportions of respondents that agreed with the optimal and minimal key characteristics as proposed are provided together with the main areas of disagreement. For all key characteristics, the majority of respondents agreed with both the optimal specifications as well as the minimal specifications.

The discussion focused on the price of the test since this was the only characteristic for which less than 75% agreement was reached. In addition, the need to be able to monitor treatment was discussed because one subgroup of respondents disagreed with the necessity of being able to use the test to monitor treatment. Specificity was also discussed.

Table 8. Delphi survey results for seven proposed key characteristics of a rapid sputum-based test for detecting TB at the microscopy-centre level

Characteristic	Optimal requirements	% (range) agreeing with the optimal requirements for the characteristic[a]	Minimal requirements	% (range) agreeing with the minimal requirements for the characteristic[a]	Comments
1 Diagnostic sensitivity	Sensitivity should be > 95% for a single test when compared with culture (for smear-negative cases it should be > 68%; for smear-positive it should be 99%)	95% (77–96%)	Sensitivity should be > 80% for a single test when compared with culture (for smear-negative cases it should be >60%; for smear-positive it should be 99%)	92% (74–94%)	Minimal: the usefulness of a test that is less sensitive was questioned; 2 respondents suggested a sensitivity of > 68% for smear-negative TB Optimal: a sensitivity of 95% compared with culture appears to be unrealistic
2 Diagnostic specificity	Specificity should be > 99% for a single test compared with culture	97% (79–98%)	Specificity should be > 99% for a single test compared with culture	89% (72–91%)	Several respondents thought that 99% specificity might be too high and suggested 98%
3 Possibility of using the test for treatment monitoring	Yes	97% (79–98%)	No	82% (67–85%)	4 respondents (mostly from high-burden countries) highlighted as an essential requirement the importance of having a test that also could be used to monitor treatment (it is currently listed as an optimal requirement); one respondent suggested that the capability for treatment monitoring needs to be evaluated in trials separately, and that further specifications should be outlined

4 Manual preparation of samples (steps needed after obtaining sample)	< 2 steps; precise volume control and precise timing should not be required	95% (77–96%)	≤ 3 steps; precise volume control and precise timing should not be required	79% (64–83%)	Optimal: no sample preparation steps should be needed (that is, the sample would go in, and the result would come out) Minimal: could allow for a maximum of 3 steps if no steps require precise volume control; In general, it is not so much the number but the types of steps to be performed
5 Time to result	< 20 minutes	92% (74–94%)	< 1 hour	92% (74–94%)	Minimal: 2 respondents suggested that 1 hour would be too long for tests performed in the field; 1 respondent suggested 2 hours would be acceptable; the simplicity of the test is the key to avoiding batching, or the time to result will not matter
6 Maintenance and calibration	Preventative maintenance or calibration should not be needed until after 2 years or 5 000 samples; only simple tools and minimal expertise should be required; an alert to indicate when maintenance is needed should be included; the instrument should be able to be calibrated remotely or should not require calibration	92% (74–94%)	Preventative maintenance should not be needed until after 1 year or 1 000 samples; only simple tools and minimal expertise should be required; an alert to indicate when maintenance is required should be included; the instrument should be able to be calibrated remotely, should calibrate itself, or no calibration should be required	92% (74–94%)	Optimal characteristics should be clarified (too many options are listed) Optimal: 1 respondent suggested that the instrument should not require maintenance

7 Price of individual test (cost of reagent only; after scale-up; ex-works [manufacturing costs only, excluding shipping])	< US$ 4.00 for TB detection	76% (62–80%)	< US$ 6.00 for TB detection	68% (55–74%)	Some respondents thought the price was too high, and others thought it was too low - Optimal: 5 respondents suggested that the price was too high and should be < US$ 3.00 (the lowest cost suggested is the same as for smear microscopy, which is <US$ 1.00); 4 respondents considered the price to be too low to be realistic and to stimulate manufacturers' interest One respondent suggested that "we should allow a minimal price that is higher if we get more for the money, but should push for an optimal price that is lower"

a A range was calculated assuming that non-respondents would either all agree (high margin) or disagree (low margin) with the proposed characteristic.

4.3. Specificity

Several respondents stated that a specificity of 99% might not be achievable because of difficulties with the reference standard. Therefore, the specificity target was adjusted to 98%.

4.4. Treatment monitoring

Despite a high percentage of agreement being reached in the Delphi survey on the treatment-monitoring characteristic, representatives from countries with a high burden of TB consistently argued that a smear replacement test should also be able to be used to monitor treatment in order to completely replace smear microscopy. Otherwise, two tests would be necessary. However, other participants indicated that the absence of a capability for monitoring treatment did not make a test useless if the test surpasses smear microscopy in other aspects. The decision was made to leave the characteristic as it is but to emphasize the importance of treatment monitoring as a competitive advantage.

4.5. Price of the test

In order to inform the discussion on pricing, again the results of the continuing study on costs and affordability were presented (A. Pantoja et al, unpublished data, 2014). This analysis has shown that in three countries (India, Kenya and the Russian Federation) the use of a replacement test for smear microscopy without further confirmatory testing would be affordable (that is, lead to expenditures from a TB budget on diagnostics that were comparable to those of the current algorithm) at the optimal cost of US$ 4.00 per test only in the Russian Federation. The main reasons for the higher cost for a replacement test for smear microscopy compared with current practice were the assumed capital costs for the instrument (optimal, less than US$ 4000; minimal, less than US$ 8000).

In order to decrease the initial cost and barriers to implementation, the cost of a new instrument would need to be low, particularly since a large number of them would be needed. Several participants perceived the capital costs as described in the revised TPP as too high (Table 9), especially if the cost is per module, given that many instruments will be necessary.

Considering that the price for a GeneXpert module is around US$ 4250 (calculated using the price of the four-module instrument) and that on average three to five times the number of instruments would be needed for a test to be implemented at a microscopy centre (compared with a district-level hospital where the GeneXpert instrument is placed preferentially), the minimal cost for a new instrument should be less than US$ 1400. The optimal cost would be less than US$ 500 per module or platform capable of performing one test.

During the meeting, general agreement was reached on the range of the price per test proposed in the TPP. The capital costs were revised to reflect the calculations in the preceding paragraph.

4.6. Other comments

During the discussion, a couple of additional points were made that related in part to characteristics that were specified in the TPP but were not scored in the Delphi survey.

Several participants commented on the number of manual steps needed to prepare the samples. Three steps were considered to be too many for implementation at the microscopy-centre level since each step would increase error rates and the need for quality assurance. The importance of not needing to monitor timing precisely and not needing to pipette was emphasized.

Several participants suggested that the feasibility of performing DST using the same platform would be beneficial and would avoid the need to have several different platforms for testing. However, an improved test for detection would be the most important to curb transmission. Additionally, the

value of the test would increase and capital costs might decrease if the test could also be used to identify non-tuberculous mycobacteria or could test multiple analytes.

It was suggested that the humidity level under which the instrument could operate should be increased in the optimal scenario. Up to 90% humidity is often reached in the Amazonas state of Brazil or in east India. Additionally, operating temperatures above 45 °C were also mentioned as being relatively common in some regions. Therefore, the operating temperature under the optimal scenario was increased to 50 °C.

Requirements for tests might differ substantially in urban and rural settings (for example, with regard to requirements for daily throughput). As outlined, the TPP has tried to ensure applicability to the majority of countries with a high burden of TB, but there may be specific differences between and within countries.

Table 9. Revised target product profile (TPP) for a test to replace smear microscopy for detecting TB, using input from a Delphi survey and discussions at a consensus meeting, 2014

Characteristic	Optimal requirements	Minimal requirements
Scope		
Goal	To develop a sputum-based test for detecting **pulmonary TB** at the microscopy-centre level of the health-care system to support the initiation of TB therapy during the same clinical encounter or the same day	
Target population	Target groups are all patients suspected of having pulmonary TB who are able to produce sputum, in countries with a medium prevalence to a high prevalence of TB as defined by WHO[a]	
Target user of the test[b]	Health-care workers with a minimum amount of training (that is, with skills that are similar to or less demanding than those needed for performing smear microscopy)	
Setting (level of the health-care system)	Microscopy-centre level (primary health-care centres with attached peripheral laboratories) or higher levels of the health-care system	
Performance characteristics		
Diagnostic sensitivity[b]	Sensitivity should be > 95% for a single test when compared with culture (for smear-negative cases it should be > 68%; for smear-positive it should be 99%)	Sensitivity should be > 80% for a single test when compared with culture (for smear-negative cases it should be > 60%; for smear-positive it should be 99%)
Diagnostic specificity[b]	> 98% specificity when compared with culture	
Possibility of using test for treatment monitoring	Yes: a test that is able to replace smear microscopy and also be used to monitor treatment is more likely to be adopted and more likely to completely replace smear microscopy	No
Operational characteristics		
Manual preparation of samples (steps needed after obtaining sample)	No steps or 1 step; precise volume control and precise timing should not be required	A maximum of 2 steps; precise volume control and precise timing should not be required

Reagent integration	All reagents should be contained in a single device	A maximum of 2 external reagents should be required; these should be part of test kit
Data export (connectivity and interoperability)	Integrated ability for all data to be exported (including data on use of the device, error rates and rates of invalid tests, and personalized, protected results) over a USB port and network	Integrated ability for all data to be exported (including data on use of the device, error rates and rates of invalid tests, and non-personalized results) over a USB port
Time to result[b]	< 20 minutes	< 2 hours
Power requirements	Battery operated with recharging capability and a circuit protector	
Maintenance and calibration[b]	Preventative maintenance and calibration should not be needed until after 2 years or 5 000 samples; only simple tools and minimal expertise should be required; an alert to indicate when maintenance is needed should be included; the device should be able to be calibrated remotely or no calibration should be required	Preventative maintenance should not be needed until after 1 year or 1 000 samples; only simple tools and minimal expertise should be required; an alert to indicate when maintenance is needed should be included; the device should be able to be calibrated remotely, should calibrate itself or no calibration should be required
Operating temperature and humidity level	Between +5 °C and +50 °C with 90% humidity	Between +5 °C and +40 °C with 70% humidity
Reagent kit – storage, stability, and stability during transport	2 years at 0 °C to +50 °C with 90% humidity; should be able to tolerate stress during transport (72 hours at +50 °C); no cold chain should be required	12 months at 0 °C to +40 °C with 70% humidity; should be able to tolerate stress during transport (72 hours at +50 °C); no cold chain should be required
Internal quality control	Full internal process controls are necessary, including controls for sample processing and amplification (for NAAT)	
Pricing		
Price of individual test[b] (costs of reagent only; after scale-up; ex-works [manufacturing costs only, excluding shipping])	< US$ 4.00 for detecting TB	< US$ 6.00 for detecting TB
Capital costs for instrument[b]	< US$ 500 per module	< US$ 1400 per module

NAAT, nucleic acid amplification test.
a High-prevalence countries are those with >40 cases per 100 000 population; medium-prevalence countries are those with 20–40 cases per 100 000 population; and low-prevalence countries are those with <20 cases per 100 000 population (19).
b These characteristics were considered to be the most important, and specific consensus was asked for and reached through a Delphi survey.

5. Next-generation drug-susceptibility testing at microscopy centres

Presenter: Claudia Denkinger, FIND, Geneva, and McGill University, Montreal

5.1. Background information

Although progress has been made in increasing TB cure rates, these gains are threatened by the global spread of drug-resistant TB (3). WHO's latest global report indicates that samples from only 5% of new bacteriologically confirmed TB cases and 9% of bacteriologically confirmed retreatment cases undergo DST (3). A test that detects both TB and resistance to first-line anti-TB agents, and that is used at lower levels of the health-care system, can aid care providers in choosing the most appropriate first-line regimen to use.

Currently, there is only one first-line regimen, which includes isoniazid (INH), rifampicin (RIF), pyrazinamide (PZA) and ethambutol (EMB) (the regimen is known as HRZE). However, two alternative regimens have been evaluated for first-line therapy in phase II clinical studies (24, 25): REMox – that is, RIF, moxifloxacin (MOX), PZA and EMB; and PaMZ – that is, Pa824, MOX and PZA. If phase III studies show benefits from either one or both of these regimens, it is likely that these regimens will be introduced and implemented, at least initially, in parallel with HRZE. Consequently, DST at the microscopy-centre level will be needed to aid the choice of first-line regimen, and this TPP was designed to be used in this setting (Table 10).

Novel diagnostic tests used at a microscopy centre should ideally test for resistance to RIF, INH, PZA and MOX to enable the most appropriate treatment regimen to be selected.

This prioritization of testing to these for drugs is based on an assessment of the importance of each anti-TB agent in the three different first-line regimens and on the prevalence of resistance.

- Testing for resistance to RIF:
  - RIF is a key component of HRZE and REMox;
  - RIF is also an indicator drug – if resistance to RIF is present, resistance against other anti-TB agents is more likely to be present; resistance to PZA and INH is highly associated with resistance to RIF, and resistance to fluoroquinolones (FQs) is moderately associated with resistance to RIF – that is, more than 80% of RIF-resistant strains are also resistant to PZA or INH, or both, and 10–30% are also resistant to FQs (26–30);
  - the prevalence of resistance to FQs and PZA is expected to be less than 3% in most countries across all patients presenting for testing. The specificity of DST for MOX and PZA would have to exceed 99.7% to ensure that 90% of positive results are true positives in settings where the prevalence of resistance to FQs and PZA is less than 3%. If DST for MOX and PZA is done only when a test is positive for resistance to RIF, this stringent specificity requirement can be relaxed because resistance to RIF increases the pretest probability for resistance to additional anti-TB agents. Therefore, it is important to test for resistance to RIF independently of whether RIF is included in the regimen.
- MOX is a key component of REMox, PaMZ, and therapy for multidrug-resistant TB (MDR-TB). Ideally, a test would detect resistance across all FQs; however, data on cross-resistance among FQs are limited, and only MOX is included in first-line regimens.
- PZA is included in all first-line regimens and is a key component of PaMZ.
- INH is a key component of HRZE. Although modelling data for India show that on a population level testing for resistance to

INH has minimal incremental value over testing for resistance to RIF alone to control MDR-TB and INH monoresistance (9), this might change as more INH preventative therapy is rolled out (10).

Performance characteristics of DST need to be optimized but compromises can be made that depend on the trade-off between a false-positive and a false-negative result.

- A compromise on specificity can be made if the trade-off for a false-positive result is not substantial because alternative regimens are available that are effective, safe and not too cumbersome. For instance, a true-negative resistance to RIF with a false-positive resistance to MOX will lead to a treatment decision to use HRZE rather than REMox. However, a true-positive resistance to RIF with a false-positive resistance to PZA would lead to initiation of treatment for MDR-TB rather than the use of PaMZ.
- A compromise on sensitivity can be made if the efficacy of a regimen will not be substantially reduced and the likelihood of developing additional resistance on a suboptimal regimen is limited.

When DST is performed at higher levels of care where additional tests for resistance are available, DST can aid providers in making decisions about second-line therapy. For second-line anti-TB agents, DST for aminoglycosides (AGs) and capreomycin (CAP) would be critical to ensure that patients are not treated for MDR-TB when they have pre-extensively drug-resistant (XDR) TB. Modelling data suggest that to attain cost effectiveness, a sensitivity of at least 88% should be achieved for detecting resistance to second-line anti-TB agents (4).

The operational requirements of the test will vary depending on the site of implementation. However, the most important considerations pertain to quality control, maintenance and calibration, and the ability to export data. This TPP aimed at implementing DST at peripheral levels of the health-care system (that is, microscopy centres); thus, more challenging minimal operational characteristics (such as ensuring an adequate power supply) need to be met to successfully implement DST at this level than at the centralized level (8, 20, 31).

Table 10. Scope for a next-generation drug-susceptibility test for TB to be implemented at the microscopy-centre level of a health-care system as defined prior to Delphi survey

Characteristic	Optimal requirements	Minimal requirements
Scope		
Goal	**Diagnosis** of TB disease **and** detection of **drug resistance** simultaneously using only a single reaction to inform decision-making **about optimal first-line therapy for a patient (HRZE, REMox or PaMZ), the presence of additional resistance to second-line anti-TB** agents and the need for further testing	**Diagnosis** of TB disease using only one reaction followed by a second reaction to detect **drug resistance** to inform decision-making **about optimal first-line therapy** (HRZE, REMox or PaMZ) **and the need for further testing**
Target population	Target groups are all patients suspected of having TB, with a special focus on those at high risk of morbidity and mortality from drug-resistant TB, such as people living with HIV, and those at high risk of having MDR-TB (for example, household contacts of patients diagnosed with MDR-TB, and persons with a history of TB, especially those for whom first-line therapy has failed) in countries with a medium prevalence to a high prevalence of TB as defined by WHO[a]	

Lowest setting of implementation (level of the health-care system)	Microscopy-centre level	Microscopy-centre level

HRZE, isoniazid , rifampicin, pyrazinamide, ethambutol; REMox, rifampicin, moxifloxacin, pyrazinamide, ethambutol; PaMZ, Pa824, moxifloxacin, pyrazinamide; MDR-TB, multidrug-resistant TB; HIV, human immunodeficiency virus.
a High-prevalence countries are those with > 40 cases per 100 000 population; medium-prevalence countries are those with 20–40 cases per 100 000 population; and low-prevalence countries are those with < 20 cases per 100 000 population (19).

5.2. Discussion

Table 11 lists the key performance and operational characteristics of next-generation DST capable of being implemented at the microscopy-centre level of the health-care system as they were phrased for the Delphi process. The proportions of respondents that agreed with the optimal and minimal key characteristics as proposed are provided together with the most common areas of disagreement raised in the Delphi survey. For all key characteristics the majority of respondents agreed with both the optimal specifications as well as the minimal specifications.

As a result of concerns and disagreements expressed through the Delphi process, the following topics were prioritized for discussion:

- the priority of the anti-TB agents to be tested under the optimal and minimal scenarios; more than 75% of respondents agreed with the proposals but several concerns were raised during the Delphi process; thus, it was considered important to discuss this critical aspect of the TPP;
- the diagnostic sensitivity (less than 75% of respondents agreed on this characteristic);
- the price per test (less than 75% of respondents agreed on the minimal requirement).

Table 11. Delphi survey results for proposed key characteristics of a next-generation drug-susceptibility test for TB to be used at peripheral microscopy centres

	Characteristic	Optimal requirements	% (range) agreeing with the optimal requirements for the characteristic[a]	Minimal requirements	% (range) agreeing with the minimal requirements for the characteristic[a]	Comments
1	**Priority of anti-TB agents for testing**	In order of decreasing importance: 1. RIF 2. FQs 3. PZA 4. INH 5. AG and CAP	87% (72–89%)	In order of decreasing importance: 1. RIF 2. MOX	85% (70–87%)	Minimal: 4 respondents suggested the order (1) RIF, (2) MOX, (3) INH; 3 respondents favoured including PZA Optimal: One respondent suggested that considering timelines for regimen roll-out and the 5-year timeframe for the TPP, a more objective priority list for DST assays intended to be used at microscopy centres is (1) RIF, (2) MOX, (3) INH, (4) AG and CAP, (5) PZA. One respondent questioned whether it was realistic to include PZA under optimal. One respondent stated that priority should be given to bactericidal agents or sterilizing agents over static drugs. One respondent suggested that the test should detect resistance to all FQs and not be restricted only to MOX if REMox turns out to be unsuccessful (LVX is more widely used than MOX). One participant commented that the focus should be more on status quo and not on hypothetical regimens that might get rolled out. One respondent suggested that new drugs should be included such as bedaquiline and delaminid.
2	**Diagnostic sensitivity of DST**	Should have >95% sensitivity for detecting resistance to RIF, FQ, PZA, INH and AGs when compared with phenotypic liquid culture DST	79% (66–82%)	Should have >95% sensitivity for detecting resistance to RIF and >90% sensitivity for detecting resistance to MOX when compared with phenotypic liquid culture DST	72% (60–77%)	Several respondents suggested that the optimal requirements are unrealistic, especially for PZA and INH. The requirement for >95% sensitivity for AG and CAP will be hard to meet without affecting the specificity according to one respondent. In relation to the reference standard, the new test should at least reach the sensitivity for clinically relevant resistance – that is, the higher values for minimum inhibitory concentrations. Independent of the sensitivity specifications here, several respondents reiterated their disagreement with the priority list for testing anti-TB agents.

3a	**Diagnostic specificity of DST compared against phenotypic DST**	Should be ≥98% for any anti-TB agent for which the test is able to identify resistance irrespective of the prevalence of underlying resistance; the estimates of specificity for molecular tests when compared with phenotypic testing as a reference standard might be falsely low since the reference standard has limited sensitivity	92% (77–94%)	Should be ≥98% for any anti-TB agent for which the test is able to identify resistance irrespective of the prevalence of underlying resistance; the estimates of specificity for molecular tests when compared with phenotypic testing as a reference standard might be falsely low since the reference standard has limited sensitivity	90% (74–91%)	Optimal: 98% specificity requirement might be hard to meet for PZA and probably also for the second-line injectable agents (AGs and CAP). The requirements should not make statements about imperfect reference standards but should instead specify what the reference standard is. A respondent noted that the profile could opt for lower specificity in higher prevalence settings; a respondent asked whether there could be a profile for two different prevalence settings. One respondent stated for AGs and CAP there is a critical trade-off with sensitivity. Another respondent suggested that the ability to rule out (that is, have high sensitivity) seems important given the high levels of toxicity associated with treatment with second-line anti-TB agents, but clinicians may think otherwise
3b	**Diagnostic specificity for DST against genetic sequencing**	Should be ≥99.7% for any anti-TB agent for which the test is able to identify resistance if the prevalence of underlying resistance is <3%; if the prevalence of resistance is ≥20% (as determined using resistance to RIF as an indicator or by testing only high-risk patients), specificity of >97% is sufficient to achieve a positive predictive value of 90%	87% (72–89%)	Should be ≥99.7% for any anti-TB agent for which the test is able to identify resistance if the prevalence of underlying resistance is <3%; if the prevalence of resistance is ≥20% (as determined using resistance to RIF as an indicator or by testing only high-risk patients), specificity of >97% is sufficient to achieve a positive predictive value of 90%	84% (70–87%)	Participants commented that we should understand the likely sensitivity and specificity of the reference standard for genetic sequencing relative to phenotypic testing and then require the new test to have a certain accuracy relative to the reference standard for genetic sequencing

4	**Operating temperature and humidity level**	+5 °C to +45 °C with 70% humidity	95% (79–96%)	+5 °C to +40 °C with 70% humidity	90% (74–91%)	Minimal: several experts suggested that the temperature should be +45 °C; 1 respondent mentioned that a lower temperature (+30 °C) would be acceptable for microscopy centres in most endemic settings
5	**Time to result**	< 30 minutes	85% (70–87%)	< 2 hours	79% (66–83%)	Participants commented that if DST is done only in cases with confirmed TB (a 2-step approach), they would be willing to wait longer than 30 minutes Minimal: could be a range between 2 hours and 24 hours Optimal: might be unachievable
6	**Data analysis**	Integrated data analysis (no requirement for a PC); exported data should be capable of being analysed on a separate or networked PC	97% (81–98%)	Integrated data analysis (no requirement for a PC); exported data should be capable of being analysed on a separate or networked PC	87% (72–89%)	Participants commented that they must be able to see raw data, although this may not be a critical aspect because it raises the need for an electronic device Minimal: might be reasonable to require a PC for integrated data analysis; some asked whether we want to force DST to be better than Xpert MTB/RIF in its minimal characteristics
7	**Data export**	Integrated ability for all data to be exported (including data on use of the device, error rates and rates of invalid tests, and personalized, protected results) over a USB port and over a mobile phone network	94% (79–95%)	Integrated ability for all data to be exported (including data on use of the device, error rates and rates of invalid tests, and non-personalized results) over a USB port	86% (72–88%)	Minimal: it would be acceptable to tolerate print-outs from or uploads to other computers; requirement should include transmissibility of data; should include exporting personalized data to a web link or to third-party devices Optimal: should include Internet connectivity, which is important for quality assurance

8	**Maintenance and calibration**	Preventative maintenance should not be needed until after 2 years or > 5000 samples; an alert to indicate when maintenance is needed should be included; the instrument should be able to be calibrated remotely or should not require calibration	100% (84–100%)	Preventative maintenance should not be needed until after 1 year or 1000 samples; an alert to indicate when maintenance is needed should be included; the instrument should be able to be calibrated remotely or should not require calibration	95% (79–96%)	No comments
9	**Internal quality control**	Full process controls are necessary, including controls for sample processing, amplification and detection of TB	92% (77–94%)	Full process controls are necessary, including controls for sample processing, amplification and detection of TB	87% (72–89%)	No comments
10	**Price of individual test**	<US$ 10.00 for TB detection and DST	90% (74–91%)	<US$ 45.00 for DST only	64% (53–70%)	Minimal: respondents said that the price needed to be lower; respondents also commented that testing for resistance to FQs should be considered as a reflex test after testing for resistance to RIF to reduce cost; US$ 45.00 is not a sustainable price Optimal: 1 respondent said that the optimal price was too high; 1 respondent asked what the industry could achieve and what countries could afford? A participant noted that the technical and operational specifications are directly proportional to costs so trade-offs will need to be made
	Capital costs for instrument	US$ 5000		US$ 14000		

DST, drug-susceptibility testing; RIF, rifampicin; FQ, fluoroquinolone; PZA, pyrazinamide; INH, isoniazid; AG, aminoglycoside; CAP, capreomycin; LVX, levofloxacin; PC, personal computer.
a A range was calculated assuming that non-respondents would either all agree (high margin) or disagree (low margin) with the proposed characteristic.

5.3. Prioritizing anti-TB agents for testing

Although the prioritization of anti-TB agents as presented for the Delphi survey resulted in consensus prior to the meeting for both the initial criteria (more than 50% of respondents agreed with the proposed criteria) and the more stringent criteria (more than 75% agreement among respondents), there was a perceived need for additional discussion.

The TPP assumes that new first-line regimens will be available. This was perceived as problematic by some participants because uncertainty about the availability of new regimens in the near future makes it difficult to decide which anti-TB agents should be included in the TPP. While the need for rapid DST for regimens that are currently available is clear, it is difficult to predict if and how new regimens will be rolled out, and thus it is difficult to predict future needs. However, other participants argued that some anti-TB agents (for example, PZA and FQs) are likely to be important components of many future regimens. Therefore, the recommendation emphasizing PZA and FQs is independent of the regimens that are currently being evaluated in late-phase trials (such as REMox and PaMZ).

There was extensive debate about which anti-TB agents should be tested for resistance by next-generation DST; the primary objective was to identify an acceptable compromise between current and future needs for rapid DST.

Testing for resistance to PZA

- The inclusion of PZA in the optimal requirements was not questioned, although concerns were raised over the feasibility of testing for resistance to PZA at the specified performance characteristics.
- Some participants also advocated for the inclusion of PZA in the minimal requirements, given the importance of PZA in novel regimens.
- Knowledge about the molecular basis of resistance to PZA is limited, and given the large number of single-nucleotide polymorphism (SNPs) associated with resistance, the tests would likely have to be performed in a separate, additional step.
- It was considered whether resistance to RIF could be used as a proxy for resistance to PZA, given the good negative predictive value of susceptibility to RIF (that is, if a strain is susceptible to RIF then it is likely to be susceptible to PZA). However, the positive predictive value was considered to be too variable to be useful, and it would limit the use of PaMZ because depending on the epidemiological setting, up to 70% of patients could be treated with PaMZ rather than with MDR-TB therapy.
- If detection of resistance to PZA is not included initially with other resistance tests but only as a reflex test – for example, if it is used when resistance to RIF is detected – then the market volume would be small (essentially, it would be only patients with MDR-TB), which would limit the industry's interest in such a test.

Testing for resistance to INH

- The overall prevalence of resistance to INH is high (on average, it is three times higher than resistance to RIF). Some settings (such as countries in the former Soviet Union) have a very high prevalence of INH-resistant strains. Therefore, detecting INH resistance might be more important in some settings than in others.
- Individual patients will benefit if resistance to INH is recognized. This needs to be considered even if the population-level impact of testing might not be substantial (data from a model representative of the South-East Asian epidemic) (32).
- As the use of INH preventive therapy increases, resistance to INH is likely to increase.

- The perceived need for tests to detect resistance to INH was high among clinicians in a survey done by the Stop TB Partnership's New Diagnostics Working Group and presented to the meeting; therefore, the acceptance of a test that includes testing for resistance to INH will likely be high.
- Novel regimens do not include INH but INH will continue to be used in individualized therapy.

Testing for resistance to FQs or only MOX

- FQs are likely to be important components of many novel regimens, and they are a critical component of therapy for MDR-TB.
- While new first-line regimens currently under evaluation include only MOX, levofloxacin (LVX) is commonly used to treat MDR-TB. The ability of a test to differentiate between resistance to different FQs is necessary for individualized therapy. Although the need for tests that can guide clinical decision-making about individualized therapy might be more relevant at higher levels of care in some settings, it was emphasized that in practise the decentralized use of DST assays described in the TPP will likely also be used to detect MDR-TB and guide treatment for MDR-TB.
- Differentiating resistance among FQs is more a function of interpreting mutations (that is, evaluating the hierarchical structure of mutations) rather than detecting different mutations. This can be achieved with bioinformatics, and is not a function of targets included in the test.

Testing for resistance to AGs

- Detection of incomplete cross-resistance is limited; but
- With revised lower critical concentrations for capreomycin resistance, cross-resistance is almost uniform (J. Posey et al, unpublished data, 2014).

General considerations around the inclusion of testing for anti-TB-drugs

- The importance of achieving high sensitivity in all tests that detect resistance to anti-TB agents was emphasized. Because resistance to FQs and PZA correlate well with resistance to RIF, RIF resistance can be used as a triage test to decide which samples should be tested for resistance against other agents. If regimens are available despite observed resistance, then treatment can occur at the microscopy-centre level. If it is not possible to treat patients at the microscopy-centre level, then the patient must be referred to a higher level of care, where confirmatory testing and additional DST can be performed.
- Survey data have shown that a positive predictive value (PPV) of more than 90% for detecting resistance would be necessary to result in health-care providers excluding an anti-TB agent from a regimen. This supports the need for high specificity in settings with a low prevalence of resistance or for testing only high-risk patients in areas with a high prevalence of resistance. However, providers found that a test would be useful if it had a high negative predictive value (NPV) even if the PPV were lower, which supports the idea of developing a triage test for DST.

Testing anti-TB agents not mentioned in the initial TPP

- One participant suggested, that testing for resistance to delamanid and bedaquiline should be included because they will be available before new regimens are. The counter-argument was that delamanid and bedaquiline are used in individualized therapy, which is unlikely to be provided at the microscopy-centre level given the type of training given to staff at this level and the limited resources available there. However, the availability of rapid DST assays targeting these anti-TB agents would

be important at higher levels of care. These assays would be valuable for surveillance purposes in the short term, and for guiding individualized therapy in the medium term or long term.

- The STREAM regimen will be the first short-course regimen available to treat MDR disease (http://www.controlled-trials.com/ISRCTN78372190/STREAM). Decentralized, rapid DST for FQs, and AGs and CAP could be useful for selecting patients at lower levels of the health-care system and deciding who should be enrolled into short-course treatment for MDR-TB. Although STREAM therapy is unlikely to be initiated at the microscopy-centre level, initially triaging patients at microscopy centres might be useful to ensure that patients are referred for the most appropriate treatment.

Additional points in the debates were raised on the sequence of testing

- In low-prevalence settings, high specificity has to be achieved to ensure high PPVs.
- Up-front combined testing for TB detection and drug susceptibility is preferred because it is likely to result in DST being performed more quickly, although this depends on the testing strategy. Loss to follow-up with two-step testing might not be substantially higher if patients are informed quickly of their TB diagnosis.
- There will be trade-offs with costs. From the standpoint of a national TB programme, testing for TB and DST at the same visit might be more costly than testing for detection and following that with a second test for DST.

Comments from representatives of industry

- The industry needs to know for which anti-TB agents resistance testing will be necessary for at least the next 3 years to ensure that it can plan appropriately.
- Assay designs cannot be flexible owing to stringent regulatory processes.
- Decisions about testing and the sequence of testing will affect the size of the market, and this will affect what companies will do and want to do.
- The need to prove a test's ability to detect resistance to every anti-TB agent requested is costly.
- The industry's commitment is likely to be low if the possible market volume is low (for example, when DST would be used only after a test to detect TB and resistance to RIF).

5.4. Conclusions

- Ultimately, consensus was reached on prioritizing anti-TB agents for testing in both optimal and minimal requirements in the following order: (1) RIF, (2) FQs (including MOX), (3) INH and PZA (equally important), (4) AGs/CAP. This acknowledges that detecting resistance to INH is a priority for current regimens, while detecting resistance to PZA is an equally important priority considering future regimens. The TPP was adjusted accordingly (Table 12).
- With regard to the sequence of testing (that is, detecting TB and drug susceptibility simultaneously or as part of a two-step test), there are insufficient data to express preferences for tests that simultaneously detect TB and drug susceptibility or tests that detect TB and then use DST or those that detect TB and resistance to RIF simultaneously and follow this with additional DST.
- There was consensus that independent of implementing DST at the microscopy-centre level, centralized DST needs to be improved to facilitate individualized therapy, which may include additional anti-TB agents such as delamanid and bedaquiline.

5.5. Diagnostic sensitivity of anti-TB agents

Seventy two percent of those who responded to the Delphi survey agreed with the minimal requirements for sensitivity that were outlined for the different anti-TB agents. However, the majority of respondents commented that they disagreed with how the anti-TB agents had been prioritized, but did not provide any information about why they disagreed with the proposed sensitivity. During the meeting no additional comments were made about this characteristic. No changes were made to the TPP.

5.6. Price of the test

In order to inform the discussion on pricing, again the results of a continuing study on costs and affordability were presented for three representative countries (India, Kenya and the Russian Federation) (A. Pantoja et al, unpublished data, 2014). For the affordability study it was assumed that DST would be used without further confirmatory testing. Compared with current expenditures in TB budgets for diagnostics, a test that initially detected TB and was then followed by DST (when TB is detected) would be more costly than current practices used to diagnose MDR-TB (that is, DST for all first-line agents) even if a target of US$ 15.00 per test and a capital cost of US$ 14 000 were considered. Therefore, such a test might not be affordable for many highly endemic countries. A test that combines DST and detection and costs US$ 10.00 per test with a capital cost of US$ 5 000 would be affordable only in the Russian Federation.

The discussion highlighted, that similar to the considerations of capital costs related to the rapid sputum-based test, the cost of instruments would need to be low to decrease the initial costs and barriers to implementation, particularly since the number of instruments needed will be sizeable (one instrument per microscopy centre). The participants perceived that the capital cost in the TPP was too high, especially because the cost was indicated per module.

Considering that the price for a GeneXpert module is about US$ 4250 (calculated from the cost of the four-module instrument) and that on average three to five times the number of instruments would be needed for a test to be implemented at a microscopy centre (compared with a district-level hospital where the GeneXpert instrument is placed primarily), the minimal cost should be less than US$ 1400.

Based on the discussion, the proposed capital costs were lowered.

Time did not allow discussion on the price per test proposed in the TPP. In general, agreement was reached about the importance of having an affordable test and on the optimal target price per test (less than US$ 10.00) but further discussion on what would be an acceptable price range could not be pursued.

Table 12. **Revised target product profile (TPP) for a next-generation drug-susceptibility test to be implemented at peripheral levels of the health-care system, using input from a Delphi survey and discussions at a consensus meeting, 2014**

Characteristic	Optimal requirements	Minimal requirements
Scope		
Goal	To develop a test to diagnose TB disease and detect drug resistance to inform decision-making about optimal current and likely future first-line therapy (HRZE, REMox or PaMZ), and possibly to detect resistance to additional second-line anti-TB agents and determine whether there is a need for further testing	

Target population	Target groups are all patients suspected of having TB, with a special focus on those at high risk of morbidity and mortality from drug-resistant TB, such as people living with HIV, and those at high risk of having MDR-TB (for example, household contacts of patients diagnosed with MDR-TB, and persons with a history of TB, especially those for whom first-line therapy has failed) in countries with a medium prevalence to a high prevalence of TB as defined by WHO[a]	
Priority of anti-TB agents for testing[b]	1. RIF 2. FQs (including MOX) 3. INH and PZA (equally important) 4. AGs and CAP Ordered by preference. Optimally, all anti-TB agents would be included, but as a minimum at least RIF should be included	
Lowest setting of implementation[b] (level of the health-care system)	Microscopy-centre level or higher levels of the health-care system	
Performance characteristics		
Diagnostic sensitivity compared against genetic sequencing as the reference standard[b]	Sensitivity should be > 98% for detecting targeted SNPs for resistance to RIF, FQs, PZA, INH, and AGs and CAP when compared with genetic sequencing	Sensitivity should be > 98% for detecting targeted SNPs for resistance to RIF, and 95% for detecting SNPs for resistance to FQs, PZA, INH, and AGs and CAP when compared with genetic sequencing
Diagnostic sensitivity compared against phenotypic DST as the reference standard	Sensitivity should be > 95% for detecting resistance to RIF, FQs, PZA, INH, and AGs and CAP when compared with the recommended reference phenotypic culture DST for a specific anti-TB agent	Sensitivity should be > 95% for detecting resistance to RIF and > 90% for detecting resistance to FQs, PZA, INH and AGs when compared with the recommended reference phenotypic culture DST for a specific anti-TB agent
Diagnostic specificity compared against genetic sequencing as the reference standard[b]	Specificity should be ≥ 98% for any anti-TB agent for which the test is able to identify resistance when compared against genetic sequencing as the reference standard The positive predictive value (PPV) of a resistance call will depend on the underlying prevalence of resistance in the population tested: if the prevalence is < 3% then very high specificity (that is, ≥ 99.7%) must be achieved to obtain a PPV of > 90%; if the prevalence of resistance is ≥ 20% (for example, by testing only high-risk patients), a specificity of > 97% is sufficient to achieve a PPV of 90%	
Diagnostic specificity compared against phenotypic DST as the reference standard	Targeted sequencing for the mutations included in the test that indicate resistance against any anti-TB agent should exceed 98% when compared against the phenotypic reference standard recommended for each anti-TB agent.	

Operational characteristics		
Maintenance and calibration[b]	Preventative maintenance should not be needed until after 2 years or > 5000 samples; an alert to indicate when maintenance is needed should be included; the instrument should be able to be calibrated remotely or should not require any calibration	Preventative maintenance should not be needed until after 1 year or 1000 samples; an alert to indicate when maintenance is needed should be included; the instrument should be able to be calibrated remotely or should not require any calibration
Time to result[b]	< 30 minutes	< 2 hours
Data analysis	Data analysis should be integrated (there should be no requirement for a PC); exported data should be capable of being analysed on a separate or networked PC	
Data export (connectivity and interoperability)	Integrated ability for all data to be exported (including data on use of the device, error rates and rates of invalid tests, and personalized, protected results) over a USB port and network	Integrated ability for all data to be exported (including data on use of the device, error rates and rates of invalid tests, and non-personalized results) over a USB port
Operating temperature and humidity level	Between +5 °C and +50 °C with 90% humidity	Between +5 °C and +40 °C with 70% humidity
Reagent kit – storage, stability, and stability during transport	2 years at +5 °C to +50 °C with 90% humidity; no cold chain should be required; should be able to tolerate stress during transport for a minimum of 72 hours at 0 °C to +50 °C	18 months at +5 °C to +40 °C with 70% humidity; no cold chain should be required; should be able to tolerate stress during transport for a minimum of 72 hours at 0 °C to +40 °C
Internal quality control	Full controls are necessary, including controls for sample processing, amplification and detection of TB	
Price of individual test[b, c] (reagent and consumables only; after scale-up; ex-works [manufacturing costs only, excluding shipping])	Optimal price: < US$ 10.00 Meeting participants emphasized the critical need for the price to be kept within an affordable range. A price higher than currently available technologies (for example, molecular testing for simultaneous detection of TB and resistance to RIF or for resistance to RIF and INH both cost approximately US$ 10.00 per test) would be justified only if the new tests add substantial value in terms of improved performance, greater suitability for decentralization, and the number of anti-TB agents for which resistance can be detected. The market size for reflex testing will also need to be taken into account. Consensus was not reached on the minimal requirement (that is, the highest acceptable price point)	
Capital costs for instrument[c]	US$ 1400 per module (for a test combining detection and DST)	US$ 1400 per module for DST only

HRZE, isoniazid, rifampicin, pyrazinamide, ethambutol; REMox, rifampicin, moxifloxacin, pyrazinamide, ethambutol; PaMZ, Pa824, moxifloxacin, pyrazinamide; MDR-TB, multidrug-resistant TB; RIF, rifampicin; FQ, fluoroquinolone; MOX, moxifloxacin; PZA, pyrazinamide; INH, isoniazid; AG, aminoglycoside; CAP, capreomycin; SNP, single nucleotide polymorphism; PPV, positive predictive value; PC, personal computer.

a High-prevalence countries are those with > 40 cases per 100000 population; medium-prevalence countries are those with 20–40 cases per 100000 population; and low-prevalence countries are those with < 20 cases per 100000 population (19).

b These characteristics were considered to be the most important, and specific consensus was asked for and reached through a Delphi survey.

c Italics indicate characteristics for which no consensus was reached.

6. Closing remarks

After 2 days of informative discussion, consensus was reached for all of the characteristics of the first three TPPs and a large part of the TPP for DST. While many of the characteristics were defined based on expert opinion, further work to shape the TPPS is continuing. Also, a team composed of representatives from McGill University, FIND, UNITAID, the Bill and Melinda Gates Foundation, and country-level partners is working to develop estimates of market size for the priority TPPs. This work is expected to be completed in the autumn of 2014. The TPPs represent the needs perceived by the TB community to be the highest priority, and they will hopefully translate into concrete tests that can help to curb the TB epidemic.

References

1. *Global strategy and targets for tuberculosis prevention, care and control after 2015*. Geneva, World Health Organization, 2014 (WHA 67.1) (available at http://apps.who.int/gb/ebwha/pdf_files/WHA67/A67_R1-en.pdf).

2. Kik SV, Denkinger CM, Casenghi M et al. Tuberculosis diagnostics: which target product profiles should be prioritised? *European Respiratory Journal*, 2014, doi:10.1183/09031936.00027714.

3. *Global tuberculosis report 2013*. Geneva, World Health Organization, 2013 (WHO/HTM/TB/2013.11) (available at http://www.who.int/tb/publications/global_report/en/).

4. Houben RM, Dowdy DW, Vassall A et al. How can mathematical models advance tuberculosis control in high HIV prevalence settings? *International Journal of Tuberculosis and Lung Disease*, 2014, 18:509–514.

5. Van't Hoog A, Cobelens FG, Vassall A et al. Optimal triage test characteristics to improve the cost-effectiveness of the Xpert MTB/RIF assay for TB diagnosis: a decision analysis. *PLoS One*, 2013, 8(12):e82786.

6. Graham SM, Ahmed T, Amanullah F et al. Evaluation of tuberculosis diagnostics in children: 1. Proposed clinical case definitions for classification of intrathoracic tuberculosis disease. Consensus from an expert panel. *Journal of Infectious Diseases*, 2012, 205(Suppl. 2):S199–S208.

7. Batz H-G, Cooke GS, Reid SD. *Towards lab-free tuberculosis diagnosis*. Geneva, Médecins Sans Frontières, 2011.

8. Denkinger CM, Kik SV, Pai M. Robust, reliable and resilient: designing molecular tuberculosis tests for microscopy centers in developing countries. *Expert Review of Molecular Diagnostics*, 2013, 13:763–767.

9. Wells WA, Boehme CC, Cobelens FG et al. Alignment of new tuberculosis drug regimens and drug susceptibility testing: a framework for action. *Lancet Infectious Diseases*, 2013, 13:449–458.

10. Pai M, Denkinger CM, Kik SV et al. Gamma interferon release assays for detection of Mycobacterium tuberculosis infection. *Clinical Microbiology Reviews*, 2014,, 27:3–20.

11. Wallis RS, Kim P, Cole S et al. Tuberculosis biomarkers discovery: developments, needs, and challenges. *Lancet Infectious Diseases*, 2013, 13:362–372.

12. *Tuberculosis: diagnostics technology and market landscape*, 2nd ed. Geneva, UNITAID, 2013.

13. *TB diagnostics and laboratory strengthening: WHO monitoring of Xpert MTB/RIF roll-out*. Geneva, World Health Organization, 2013 (http://who.int/tb/laboratory/mtbrifrollout/en/, accessed 7 August 2014).

14. Davis JL, Cattamanchi A, Cuevas LE et al. Diagnostic accuracy of same-day microscopy versus standard microscopy for pulmonary tuberculosis: a systematic review and meta-analysis. *Lancet Infectious Diseases*, 2013, 13:147–154.

15. Zar HJ, Connell TG, Nicol M. Diagnosis of pulmonary tuberculosis in children: new advances. *Expert Review of Anti-Infective Therapy*, 2010, 8:277–288.

16. Peter JG, Theron G, Pooran A et al. Comparison of two methods for acquisition of sputum samples for diagnosis of suspected tuberculosis in smear-negative or sputum-scarce people: a randomised controlled trial. *Lancet Respiratory Medicine*, 2013, 1:471–478.

17. *Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF assay for the diagnosis of pulmonary and extrapulmonary TB in adults and children. Policy update*. Geneva, World Health Organization, 2013 (WHO/HTM/TB/2013.16) (available at http://apps.who.int/iris/bitstream/10665/112472/1/9789241506335_eng.pdf).

18. Behr MA, Warren SA, Salamon H et al. Transmission of Mycobacterium tuberculosis from patients smear-negative for acid-fast bacilli. *Lancet* 1999, 353:444–449.

19. *Global tuberculosis control: WHO report 2011*. Geneva, World Health Organization, 2011 (WHO/HTM/TB/2011.16) (available at http://www.who.int/tb/publications/global_report/2011/gtbr11_full.pdf).

20. Denkinger CM, Nicolau I, Ramsay A et al. Are peripheral microscopy centres ready for next generation molecular tuberculosis diagnostics? *European Respiratory Journal*, 2013, 42:544–547.

21. *A systematic review of the sensitivity and specificity of symptom- and chest-radiography screening for active pulmonary tuberculosis in HIV-negative persons and persons with unknown HIV status*. Geneva, World Health Organization, 2013 (http://www.who.int/tb/Review2Accuracyofscreeningtests.pdf, accessed 7 August 2014).

22. Parsons LM, Somoskovi A, Gutierrez C et al. Laboratory diagnosis of tuberculosis in resource-poor countries: challenges and opportunities. *Clinical Microbiology Reviews*, 2011, 24:314–350.

23. Sun AY, Pai M, Salje H et al. Modeling the impact of alternative strategies for rapid molecular diagnosis of tuberculosis in Southeast Asia. *American Journal of Epidemiology*, 2013, 178:1740–1749.

24. Diacon AH, Dawson R, von Groote-Bidlingmaier F et al. 14-day bactericidal activity of PA-824, bedaquiline, pyrazinamide, and moxifloxacin combinations: a randomised trial. *Lancet*, 2012, 380:986–993.

25. Dorman SE, Johnson JL, Goldberg S et al. Substitution of moxifloxacin for isoniazid during intensive phase treatment of pulmonary tuberculosis. *American Journal of Respiratory and Critical Care Medicine*, 2009, 180:273–280.

26. Ando H, Mitarai S, Kondo Y et al. Pyrazinamide resistance in multidrug-resistant Mycobacterium tuberculosis isolates in Japan. *Clinical Microbiology and Infection*, 2010, 16:1164–1168.

27. Jonmalung J, Prammananan T, Leechawengwongs M et al. Surveillance of pyrazinamide susceptibility among multidrug-resistant Mycobacterium tuberculosis isolates from Siriraj Hospital, Thailand. *BMC Microbiology*, 2010, 10:223.

28. Nguyen D, Brassard P, Westley J et al. Widespread pyrazinamide-resistant Mycobacterium tuberculosis family in a low-incidence setting. *Journal of Clinical Microbiology*, 2003, 41:2878–2883.

29. Falzon D, Gandhi N, Migliori GB et al. Resistance to fluoroquinolones and second-line injectable drugs: impact on MDR-TB outcomes. *European Respiratory Journal*, 2013, 42:156–168.

30. Smith SE, Kurbatova EV, Cavanaugh JS et al. Global isoniazid resistance patterns in rifampin-resistant and rifampin-susceptible tuberculosis. *International Journal of Tuberculosis and Lung Disease*, 2012, 16:203–205.

31. Banoo S, Bell D, Bossuyt P et al. Evaluation of diagnostic tests for infectious diseases: general principles. *Nature Reviews. Microbiology*, 2010, 8(Suppl.):S17–S29.

32. Denkinger CM, Pai M, Dowdy DW. Do we need to detect isoniazid resistance in addition to rifampicin resistance in diagnostic tests for tuberculosis? *PLoS One*, 2014, 9:e84197.

33. Denkinger CM, Kik SV, Pai M. Robust, reliable and resilient: designing molecular tuberculosis tests for microscopy centers in developing countries. *Expert Review of Molecular Diagnostics*, 2013, 13:763–767.

34. Dowdy D, Houben RM, Cohen T et al. Impact and cost-effectiveness of current and future TB diagnostics: the contribution of mathematical modelling. *International Journal of Tuberculosis and Lung Disease*, 2014 (In press).

35. Pai NP, Vadnais C, Denkinger C et al. Point-of-care testing for infectious diseases: diversity, complexity, and barriers in low- and middle-income countries. *PLoS Medicine*, 2012, 9:e1001306.

36. *Diagnostics for tuberculosis: global demand and market potential*. Geneva: TDR, FIND, 2006 (available at http://www.who.int/tdr/publications/documents/tbdi.pdf).

37. Kik SV, Denkinger CM, Chedore P et al. Replacing smear microscopy for the diagnosis of tuberculosis: what is the market potential? *European Respiratory Journal*, 2014, doi:10.1183/09031936.00217313.

38. Sun AY, Pai M, Salje H et al. Modeling the Impact of alternative strategies for rapid molecular diagnosis of tuberculosis in Southeast Asia. *American Journal of Epidemiology*, 2013 178:1740–1749.

39. Tortoli E, Russo C, Piersimoni C et al. Clinical validation of Xpert MTB/RIF for the diagnosis of extrapulmonary tuberculosis. *European Respiratory Journal*, 2012, 40:442–447.

40. Vadwai V, Boehme C, Nabeta P et al. Xpert MTB/RIF: a new pillar in diagnosis of extrapulmonary tuberculosis? *Journal of Clinical Microbiology*, 2011, 49:2540–2545.

41. Denkinger CM, Schumacher SG, Boehme CC et al. Xpert MTB/RIF assay for the diagnosis of extrapulmonary tuberculosis: a systematic review and meta-analysis. *European Respiratory Journal, 2014,* doi:10.1183/09031936.00007814.

42. Steingart K, Sohn H, Schiller I et al. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database of Systematic Reviews*, 2013, (1):CD009593.

43. Nkonki L, Cliff J, Sanders D. Lay health worker attrition: important but often ignored. *Bulletin of the World Health Organization*, 2011, 89:919–923.

44. Cobelens F, van den Hof S, Pai M et al. Which new diagnostics for tuberculosis, and when? *Journal of Infectious Diseases*, 2012, 205 (Suppl. 2): S191–S198.

45. Keeler E, Perkins MD, Small P et al. Reducing the global burden of tuberculosis: the contribution of improved diagnostics. *Nature*, 2006, 444 (Suppl. 1):49–57.

46. Lemaire JF, Casenghi M. New diagnostics for tuberculosis: fulfilling patient needs first. *Journal of the International AIDS Society*, 2010, 13:40.

47. Boehme CC, Nabeta P, Hillemann D et al. Rapid molecular detection of tuberculosis and rifampin resistance. *New England Journal of Medicine*, 2010, 363:1005–1015.

48. Boehme CC, Nicol MP, Nabeta P et al. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *Lancet*, 2011, 377:1495–1505.

49. Helb D, Jones M, Story E et al. Rapid detection of Mycobacterium tuberculosis and rifampin resistance by use of on-demand, near-patient technology. *Journal of Clinical Microbiology*, 2010, 48:229–237.

50. Claassens MM, du Toit E, Dunbar R et al. Tuberculosis patients in primary care do not start treatment. What role do health system delays play? International *Journal of Tuberculosis and Lung Disease*, 2013, 17:603–607.

51. Sreeramareddy CT, Kishore PV, Menten J et al. Time delays in diagnosis of pulmonary tuberculosis: a systematic review of literature. *BMC Infectious Diseases*, 2009, 9:91.

52. *Tuberculosis laboratory biosafety manual*. Geneva, World Health Organization, 2012 (WHO/HTM/TB/2012.11) (available at http://www.who.int/tb/publications/2012/tb_biosafety/en/).

53. Denkinger CM, Grenier J, Stratis AK et al. Mobile health to improve tuberculosis care and control: a call worth making. *International Journal of Tuberculosis and Lung Disease*, 2013, 17:719–727.

54. Grosset JH, Singer TG, Bishai WR. New drugs for the treatment of tuberculosis: hope and reality. *International Journal of Tuberculosis and Lung Disease*, 2012, 16:1005–1014.

55. Mills HL, Cohen T, Colijn C. Community-wide isoniazid preventive therapy drives drug-resistant tuberculosis: a model-based analysis. *Science Translational Medicine*, 2013, 5:180ra49.

56. Clouse K, Page-Shipp L, Dansey H et al. Implementation of Xpert MTB/RIF for routine point-of-care diagnosis of tuberculosis at the primary care level. *South African Medical Journal*, 2012, 102:805–807.

57. Louw GE, Warren RM, Donald PR et al. Frequency and implications of pyrazinamide resistance in managing previously treated tuberculosis patients. *International Journal of Tuberculosis and Lung Disease*, 2006, 10(7): 802-7.

58. Dowdy DW, Hoog AV, Shah M, Cobelens F. Cost-effectiveness of rapid susceptibility testing against second-line drugs for tuberculosis. *International Journal of Tuberculosis and Lung Disease*, 2014, 18:647–654.

59. Blakemore R, Story E, Helb D et al. Evaluation of the analytical performance of the Xpert MTB/RIF assay. *Journal of Clinical Microbiology*, 2010, 48:2495–2501.

60. Barnard M, Warren R, Gey Van Pittius N et al. Genotype MTBDRsl 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, 186:1298–1305.

61. Ling DI, Zwerling AA, Pai M. GenoType MTBDR assays for the diagnosis of multidrug-resistant tuberculosis: a meta-analysis. *European Respiratory Journal*, 2008, 32:1165–1174.

62. Feng Y, Liu S, Wang Q et al. Rapid diagnosis of drug resistance to fluoroquinolones, amikacin, capreomycin, kanamycin and ethambutol using genotype MTBDRsl assay: a meta-analysis. *PLoS One*, 2013, 8:e55292.

63. Hillemann D, Rusch-Gerdes S, Richter E. Feasibility of the GenoType MTBDRsl assay for fluoroquinolone, amikacin-capreomycin, and ethambutol resistance testing of Mycobacterium tuberculosis strains and clinical specimens. *Journal of Clinical Microbiology*, 2009, 47:1767–1772.

64. Said HM, Kock MM, Ismail NA et al. Evaluation of the GenoType(R) MTBDRsl assay for susceptibility testing of second-line anti-tuberculosis drugs. *International Journal of Tuberculosis and Lung Disease*, 2012, 16:104–109.

65. Dalton T, Cegielski P, Akksilp S et al. Prevalence of and risk factors for resistance to second-line drugs in people with multidrug-resistant tuberculosis in eight countries: a prospective cohort study. *Lancet*, 2012, 380:1406–1417.

66. Horne DJ, Pinto LM, Arentz M et al. Diagnostic accuracy and reproducibility of WHO-endorsed phenotypic drug susceptibility testing methods for first-line and second-line antituberculosis drugs. *Journal of Clinical Microbiology*, 2013, 51:393–401.

67. *Xpert MTB/RIF implementation manual. Technical and operational 'how-to': practical considerations*. Geneva, World Health Organization, 2014 (WHO/HTM/TB/2014.1) (available at http://apps.who.int/iris/bitstream/10665/112469/1/9789241506700_eng.pdf).

68. Zhang Y, Mitchison D. The curious characteristics of pyrazinamide: a review. *International Journal of Tuberculosis and Lung Disease*, 2003, 7:6–21.

Annex A. Detailed target product profiles

Detailed target product profiles (TPPs) were prepared before the consensus meeting by Claudia Denkinger (FIND), Sandra Kik (McGill University) and Madhukar Pai (McGill University), with input from numerous stakeholders; funding was provided by the Bill and Melinda Gates Foundation. These detailed TPPs formed the basis of the short TPPs that were discussed at the consensus meeting and used during the Delphi process. The detailed TPPs are provided here for background information for readers.

Definitions

Characteristic: This is a specific requirement or specification that is measurable; every characteristic stands by itself (for example, an optimal requirement would be that no instrument is needed to perform a test; however, if an instrument were to be used, then the optimal requirement would be that data export functions should be integrated into the instrument).

Minimal requirement: For a specific characteristic the minimal requirement is the lowest acceptable output for that characteristic. For clarification, solutions must meet the minimal requirements for a characteristic in order to be acceptable (caveat: a test may still be acceptable if targets are missed only marginally or shortcomings pertain only to desirable but not completely necessary characteristics).

Optimal requirement: For a specific characteristic the optimal requirement provides information about the ideal output for that characteristic. Characteristics that meet the optimal requirements provide the greatest impact for end-users, clinicians and patients. Ideally, developers would design and develop their solutions to meet the optimal requirements for all characteristics.

Table A1. Detailed target product profile (TPP) for a rapid, biomarker-based non-sputum-based test for detecting TB

Characteristic	Optimal (ideal) requirements	Minimal requirements	Explanations	References
Scope				
Goal and potential market	A rapid, biomarker-based test that can diagnose **pulmonary TB** and ideally also **extrapulmonary TB** using non-sputum samples (for example, urine, blood, oral mucosal transudates, saliva, exhaled air) for the purpose initiating TB treatment during the same clinical encounter or on the same day		The goal is to have a simple test to diagnose any form of active TB that uses easily accessible samples (such as, urine, blood, oral mucosal transudates, saliva, exhaled air) and that can be performed at peripheral-level microscopy centres, primary-care clinics and even health posts that do not have attached laboratories. If it is possible to identify a biomarker or biosignature that has acceptable performance for both pulmonary TB and extrapulmonary TB in both adults and children, then having one test for both disease presentations in both populations would be preferred. However, if the biomarkers or biosignatures for adult TB are suboptimal for detecting childhood TB, or if the biomarkers or biosignatures for pulmonary TB are suboptimal for detecting extrapulmonary TB, then separate products need to be considered. If the sensitivity for diagnosing pulmonary TB is sufficiently high, this test is likely to replace sputum-smear microscopy tests or sputum-based rapid molecular tests. **Potential market:** In 2006, the potential market size for such a test was predicted by FIND and TDR to be about 50 million tests globally in 2020. More recently, in the public sector alone it was estimated that 62 million sputum smears are performed for the initial diagnosis of pulmonary TB each year (for approximately 31 million individuals suspected of having pulmonary TB) in 22 high-burden countries (37). The potential market should increase further when the diagnosis of extrapulmonary TB and childhood TB are included, and the use of non-sputum samples and testing at lower levels of the health-care system, as well as testing in the private sector. Nevertheless, even if a biomarker test does not meet all required infrastructure needs detailed in this TPP, there can still be a considerable potential market for the test in selected countries or a subset of the facilities that are better equipped, or both. This may be the case for countries with emerging economies such as the BRICS (Brazil, Russian Federation, India, China and South Africa).	(12, 33–37)
Drug-susceptibility testing	None	None	A biomarker-based test will most likely not be able to detect drug resistance. Testing for drug resistance will need to be done as a reflex test using a different sample and technology, and perhaps at a higher level in the health-care system, where second-line treatment can be offered.	

Target population	Target groups are adults and children including those who are HIV-positive who are suspected of having active pulmonary TB or extrapulmonary TB in countries with a medium prevalence to a high prevalence of TB as defined by WHO		Since a biomarker-based test uses non-sputum samples, its target population is broad and diverse. Using easily accessible, non-sputum samples will enable diagnosis in patients who are difficult to diagnose with conventional sputum-based diagnostics (for example, children, people who are HIV-positive, patients with extrapulmonary TB) as well as patients with pulmonary TB. **WHO's categories:** High-prevalence countries are those with >40 cases per 100 000 population; medium-prevalence countries are those with 20–40 cases per 100 000 population; and low-prevalence countries are those with <20 cases per 100 000 population (19).	(19)
Target user of the testa	Health-care workers with a minimum of training	Trained microscopy technicians		
Setting (level of the health-care system)	Health posts without attached laboratories (that is, levels below microscopy centres) or higher levels of the health-care system	Primary health-care clinics (with attached laboratories), peripheral-level microscopy centres or higher levels of the health-care system	Conditions that prevail in microscopy centres in high-burden countries have been described, and the characteristics required for novel tests at this level of the health-care system have also been proposed. Ideally, a biomarker test with the proper specifications (that is, one that does not require an instrument, can be used in a setting without a laboratory and can be operated by staff with a minimum of training) should be able to be used at a level of the health-care system that is lower than peripheral-level microscopy centres; this would enhance access to TB diagnostics. Depending on how well a new product meets the characteristics needed for these settings, its potential for wider use and scale-up may increase.	(20, 33)
Pricing				
Price of individual test (costs of reagent only; after scale-up; ex-works [manufacturing costs only, excluding shipping])a, b	< US$ 4.00	< US$ 6.00	The final price may depend on whether volume-based pricing models are used. If the sensitivity for detecting pulmonary TB is at least as good as the sensitivity of smear microscopy then this test is expected to replace smear microscopy or alternative tests that are used for initial diagnosis. In addition, the test will be used for the initial testing of people suspected of having extrapulmonary TB or childhood TB. Since millions of smear microscopy tests are performed every year, the expected market for this product is likely to be substantial (37) and, therefore, volume-based pricing agreements may be possible. A more expensive test is less likely to penetrate the market. If the sensitivity of the biomarker-based test for diagnosing pulmonary TB is lower than the sensitivity of alternatives tests, it may not completely replace available diagnostics for initial testing for pulmonary TB but it could still be adopted in specific settings (for example, if no electricity is required for the test then it may be adopted at lower levels of the health-care system, such as primary-care clinics) or be used for specific populations (for example, people with HIV, people suspected of having extrapulmonary TB, children). In this case, the overall volume of tests will be lower, although still sizeable, while the target price likely would need to remain the same.	(37)

Capital costs for the instrument	No instrument needed	< US$ 500	If possible, an instrument should not be necessary in order to ensure that the test is suitable for a diverse range of settings, including lower levels of the health-care system. When the test does require an instrument, the lower the capital costs of the instrument are, the lower the initial cost would be, and thus the barrier to implementation would also be lower, particularly since the volume of instruments that would be distributed would be sizeable. The cost of the instrument should include warranties, service contracts and technical support.	
Performance				
Diagnostic sensitivity[a,b]	**For pulmonary TB in adults:** Sensitivity should be ≥ 98% for smear-positive culture-positive pulmonary TB, and ≥68% for smear-negative culture-positive pulmonary TB (that is, it should be similar to the sensitivity of the Xpert MTB/RIF assay) Overall pooled sensitivity should be ≥ 80% in adults with HIV infection **For extrapulmonary TB in adults:** Ideally, sensitivity should be ≥ 80% for all forms of microbiologically confirmed extrapulmonary TB[c, d]	**For pulmonary TB in adults:** Overall sensitivity should be ≥ 65%, but should be > 98% among patients with smear-positive culture-positive pulmonary TB (that is, it should be similar to smear microscopy) Overall pooled sensitivity should be better than the sensitivity of smear microscopy in adults with HIV infection **For extrapulmonary TB in adults:** The need to diagnose extrapulmonary TB is important and a test that can diagnose it will have significant benefits for individual patients; additionally, it is likely to be accepted in the community of care providers; no lower range of sensitivity was defined	Ideally, this test would be at least as sensitive and specific as the current standards for diagnosing pulmonary TB and extrapulmonary TB. The performance of a biomarker-based test might not differ in detecting smear-positive culture-confirmed TB and smear-negative culture-confirmed TB. However, when a biomarker-based test aims to replace smear microscopy its performance should be evaluated separately for smear-positive culture-positive TB and smear-negative culture-positive TB. Moreover, it should be determined whether the test can identify those patients who are most infectious and whether biomarker levels correlate with bacterial load in sputum (in pulmonary TB cases). This may depend on the selected biomarker or biomarkers. The test's performance should be studied separately for each specific target population (groups should be stratified by the different case definitions proposed in 6), including children, patients with extrapulmonary TB at different sites, HIV-positive patients and HIV-negative patients. It may be necessary to evaluate the potential confounding effects of previous vaccination with BCG, latent TB infection or a previous episode of active TB. Optimally, this test would be at least as sensitive as the Xpert MTB/RIF assay for diagnosing pulmonary TB, extrapulmonary TB and childhood TB, but it would be easier to implement at lower levels of the health-care system. Systematic reviews that evaluated the performance of the Xpert MTB/RIF assay using non-sputum-based samples reported a high degree of heterogeneity in sensitivity for diagnosing extrapulmonary TB in adults and children, which depended on the type of sample used (biopsy, pleural fluid, gastric lavage aspirate, lymph fluid, urine, pus or cerebrospinal fluid) (17). Nevertheless, WHO recommends that the Xpert MTB/RIF assay should be used to diagnose extrapulmonary TB and childhood TB. Systematic reviews have suggested that the sensitivity of Xpert MTB/RIF is around 44% when the sample used is pleural fluid, 85% for lymph node tissue, 80% for cerebrospinal fluid, and 84% for other types of samples (17). Xpert MTB/RIF has a sensitivity of around 66% for diagnosing pulmonary TB in children using expectorated sputum (95% for induced sputum) (17).	(6, 17–42)

	For children: Sensitivity for childhood intrathoracic TB should be ≥ 66% for microbiologically confirmed TB (similar to the sensitivity of the Xpert MTB/RIF assay)[e]	**For children:** The need to diagnose childhood TB is important, and a test that can diagnose TB in children will have significant benefits for individual patients; additionally, it is likely to be accepted in the community of care providers; no lower range of sensitivity was defined	Currently, patients with extrapulmonary TB and/or childhood TB are treated either empirically or referred to the district level for further invasive diagnosis (such as biopsy or gastric lavage). Therefore, no minimum sensitivity has been set for this test, since any confirmatory diagnosis of extrapulmonary TB and/or childhood TB at the microscopy-centre level or lower levels of the health-care system would be an improvement over the current situation. A modelling study suggested that increasing the proportion of TB patients initiated on treatment is more important than an equivalent decrease in sensitivity. This model showed that in WHO's South-East Asia Region if a biomarker test with a sensitivity of 40% for smear-negative TB were used at the point of care in the most peripheral health-care setting it would result in a reduction in TB incidence similar to using a test with 70% sensitivity for smear-negative TB at the district level (for example, Xpert MTB/RIF) (38). However, the exact trade-off between a lower sensitivity for smear-negative TB and an increase in access to testing will depend on the setting, and modelling studies might be needed to identify the most suitable settings to implement the tests. It is difficult to diagnose TB in children because it is difficult to obtain sputum specimens, and the disease is often paucibacillary. Moreover, cultures may not be positive in all children suspected of having TB. Therefore, children are considered one of the main target populations for a non-invasive biomarker-based test. Xpert MTB/RIF has been shown to have a sensitivity of around 66% for pulmonary TB in children using expectorated sputum (95% for induced sputum) (17); therefore, an ideal biomarker-based test would be as least as good.	
Diagnostic specificity[a]	The test should be at least as specific as the Xpert MTB/RIF assay for detecting pulmonary TB, extrapulmonary TB and childhood TB (that is, the test should have 98% specificity when compared against a microbiological reference standard); the test should distinguish between active TB and latent or past infection		The specificity of the test should be high to avoid the need for additional confirmatory testing at this level of the health-care system.	(17, 42)
Quantitation	Qualitative	Qualitative	The test should give an objective qualitative result (such as, yes/no/invalid). However, if this test were to replace smear microscopy to monitor treatment or for follow-up, a semiquantitative result would be required that correlated with the bacterial load (infectiousness) as well as with treatment success.	
Operational characteristics				
Sample type	Non-sputum-based samples (such as urine, blood, oral mucosal transudates, saliva, exhaled air)		If blood samples are needed, finger-stick samples would be preferred to venous blood samples, especially for samples from children.	

Sample volume	Urine: <5 ml Finger-stick blood: < 10 mcl Saliva or transudates: < 0.1 ml	Urine: < 10 ml Finger-stick blood: < 25 mcl Saliva or transudates: < 0.2 ml	The lowest volume possible for all types of samples would be preferred, especially since children might have difficulty providing a larger volume for a sample; however, this requirement for the lowest volume possible should not come at the expense of decreased sensitivity. The test should also be able to accommodate larger volumes when they are available if they would potentially increase the sensitivity.	
Sample preparation and assay processing (steps needed after obtaining sample)	Sample preparation should either be integrated or preparation should not be required	Only a limited number of steps should be required; precise measuring should not be needed for any step (for example, for either volumes or time)	The Xpert MTB/RIF assay has 2 steps. A new test should be as simple to perform as Xpert MTB/RIF. Devices such as a centrifuge or heat block are available only infrequently at the level of microscopy centres; therefore, they should not be required for novel assays. The expertise needed to operate a micropipette is also often lacking.	(20)
Reagent integration	All reagents should be contained in a consumable device	Only 1 external reagent (provided by the manufacturer)		
Time to result[a, b]	< 20 minutes including time spent preparing the sample	< 1 hour including time spent preparing the sample	Most rapid, lateral flow assays produce results within 20 minutes. However, the ideal time to result has not been studied, and might vary among countries and settings where the patient is tested. But in order to be usable as a test at the point of care, the result should be available during the same visit	
Daily throughput	> 25 tests per 6-hour working day; ability to test individual samples so that batching is not required	> 10 tests per 6-hour working day; no need to batch tests	The need for a short turnaround time, the possibility of testing individual samples without batching (asynchronous testing) and testing multiple samples at the same time are interrelated. The time to result and the possibility of asynchronous testing are probably the most important parameters for implementing the test at the point of care and limiting the wait time for patients to prevent loss to follow-up. The number of suspected cases that need to be tested per day at an average peripheral microscopy centre, health post or primary-care clinic will vary, but it will be smaller than the number at more centralized laboratories or clinics.	
Sample capacity	Multiple samples should be able to be tested at the same time	One sample at a time		
Biosafety	No need for a biosafety cabinet; consumables should be able to be disposed of as general laboratory waste	No need for a biosafety cabinet; consumables should be able to be disposed of as general laboratory waste	A biosafety cabinet is not commonly available at the level of a microscopy centre. If a novel test has increased biosafety (if it is safer to use) then its acceptability will be enhanced.	

Waste disposal – solid	Simple trash; recyclable or compostable plastics and consumables	Simple trash; recyclable or compostable plastics and consumables	Reducing the steps required to dispose of waste will facilitate the test's use in the most peripheral settings.	
Waste disposal – infectious	No infectious waste should be generated	Incineration of infectious materials		
Multiuse platform	None	None	A multiuse platform is not required, but multiplexed testing for other diseases might increase acceptability of a test.	
Instrument	No instrument needed	Preferably the test would not require an instrument; if an instrument is required, it should be small, portable or hand-held and weigh < 1 kg; it should be able to operate on battery or solar power in places where the power supply may be interrupted	The simpler, more portable and more durable the test is, the more likely it is to be implemented in peripheral settings. Many peripheral microscopy centres in countries with a high burden of TB do not have uninterrupted power supplies, and some have no access to clean water.	(20)
Power requirements	None	Optional battery or solar-powered operation	The fewer infrastructure requirements that an instrument has (such as for power, water and the skill levels of the staff), the more likely it is that the test can be adopted at lower levels of the health-care system, such as at health posts. Continuous power is not always available at the microscopy-centre level, and it is even less likely to be available at primary-care clinics; therefore, a battery-operated device that can be recharged using solar power would be ideal and would allow the test to be used in many different settings.	(20)
Maintenance[a]	Disposable, no maintenance required	Preventative maintenance should not be needed until after 1 year or > 1 000 samples; only simple tools and minimal expertise should be required; an alert should be included to indicate when maintenance is needed; the mean time to failure should be ≥ 12 months	If a device is anticipated to have a longer lifespan, then a maintenance alert is essential to ensure that it functions properly in settings where it is unlikely that the device will be always used by the same person, and that records will be kept about the duration of use. It is essential that only simple tools and minimal expertise are necessary to perform maintenance, given the quantity of devices that are likely to be used.	

Calibration	None required	The instrument should be able to be calibrated remotely, to calibrate itself, or no calibration should be needed
Result capturing, documentation, data display	Ideally, the test would not require an instrument; there should be a visual readout with the ability to save results using a separate, attachable reader	The test menu must be simple to navigate; the instrument should have an integrated LCD screen, a simple keypad or touch screen, and the ability to save results using either the instrument or a separate reader
Regulatory requirements	Manufacturing of the assay and system should comply with ISO EN 13485 or higher standards or regulations, and comply with ISO IEC 62304 Medical Device Data Systems; the manufacturing facility should be certified and authorized for use by a regulatory authority that is a member of the International Medical Device Regulators Forum, formerly known as the Global Harmonization Task Force; the assay must be registered for in vitro diagnostic use	Manufacturing of the assay and system should comply with ISO EN 13485 or higher standards or regulations, and comply with ISO IEC 62304 Medical Device Data Systems; the manufacturing facility should be certified and authorized for use by a regulatory authority that is a member of the International Medical Device Regulators Forum, formerly known as the Global Harmonization Task Force; the assay must be registered for in vitro diagnostic use

Data export (connectivity and interoperability)	Ideally, the test would not require an instrument but data should be able to be exported, for instance using a reader with a mobile phone All data should be able to be exported (including data on use of the device, error rates and rates of invalid tests, and personalized, protected results) over a USB port and network; network connectivity should be available through a GSM/UMTS mobile broadband modem; results should be encoded using a documented standard (such as HL7) and be formatted as JSON text; JSON data should be able to be transmitted through HTTP(S) to a local or remote server as results are generated; results should be stored locally, and queued during network interruptions to be sent as a batch when connectivity is restored	All data should be able to be exported (including data on use of the device, error rates and rates of invalid tests, and personalized, protected results) over a USB port and network; network connectivity should be available through Ethernet, Wi-Fi, or GSM/UMTS mobile broadband modem, or a combination of these; results should be encoded using a documented standard (such as HL7) and be formatted as JSON text; JSON data should be transmitted through HTTP(S) to a local or remote server as results are generated; results should be stored locally, and queued during network interruptions to be sent as a batch when connectivity is restored	The ability to export data will enhance surveillance and allow for better supply-chain management.
Electronics and software	None	Integrated	
Data analysis	Integrated	Integrated	Integrating the capacity for data analysis into the test will reduce interpretation errors.

Operating temperature and humidity level	Between +5 °C and +50 °C with 90% humidity	Between + 5 °C and +40 °C with 70% humidity	High environmental temperatures and high humidity are often present in countries where TB is endemic. (20)
Reagent kit – transport	All reagents and consumables should be included in the device or kit (including water, alcohol swabs and lancets if needed); no cold chain should be required	All reagents and consumables should be included in the device or kit (including water, alcohol swabs and lancets if needed); no cold chain should be required	
Reagent kit – storage and stability	2 years at +5 °C to +40 °C with 90% humidity; should be able to tolerate stress during transport (72 hours at +50 °C); no cold chain should be required	12 months at +5 °C to +35 °C with 70% humidity; should be able to tolerate stress during transport (72 hours at +50 °C); no cold chain should be required	
Reagent calibration	No reagent should be necessary	No calibration should be necessary	
Additional supplies (not included in the kit)	None	None	
Internal quality control	Internal controls for sample processing and detection of TB should be included	Internal controls for sample processing should be included	Internal quality controls should be used in addition to external quality assurance.
Training and education	< 1 day for a primary health-care worker	< 3 days for a health-care worker	The need for training and education in how to use the test should be low, given the high turnover of health-care workers. The test should be able to be conducted by non-laboratory personnel after a short training session.

FIND, Foundation for Innovative New Diagnostics; TDR, Special Programme for Research and Training in Tropical Diseases; BCG, bacille Calmette–Guérin; LCD, liquid crystal display; GSM, Global System for Mobile Communications; UMTS, Universal Mobile Telecommunications System.

a These characteristics were considered to be the most important. b These characteristics were associated with the most uncertainty.

c The sensitivity for detecting extrapulmonary TB should also be tested against a composite reference standard that includes culture with or without a nucleic acid amplification test, histology, smear microscopy, biochemical testing, presenting signs, and response to treatment with anti-TB therapy, depending on site of infection. Xpert MTB/RIF testing has an estimated sensitivity for diagnosing TB of 84% in lymph node aspirates or other tissue samples, and 55% sensitivity for samples of cerebrospinal fluid, when compared with a composite reference standard, but Xpert MTB/RIF testing requires invasive samples (17).

d Xpert MTB/RIF has an estimated sensitivity for microbiologically confirmed TB of 85% for detecting TB in lymph node aspirates or other tissue samples, 80% for cerebrospinal fluid and 44% for pleural fluid, but testing requires invasive samples (aspirate, biopsy, lumbar puncture or thoracentesis).

e The test's sensitivity in children should be evaluated against a composite reference standard as defined by an international panel of experts (6).

Table A2. Detailed target product profile (TPP) for a community-based triage or referral test to identify people suspected of having TB

Characteristic	Optimal (ideal) requirements	Minimal requirements	Explanations	References
Scope				
Goal and potential market	A test used during a patient's first encounter with the health-care system to identify patients with **any symptoms or risk factors for active TB**, including patients coinfected with HIV , those who do not have TB and those in need of referral for further confirmatory testing	A test used during a patient's first encounter with the health-care system to identify patients with **any symptoms or risk factors for active pulmonary TB**, including patients coinfected with HIV, those who do not have TB and those in need of referral for further confirmatory testing	Since most individuals suspected of having TB do not have TB, a triage test can help narrow the population that needs confirmatory testing. A triage test needs to be a simple low-cost test that can be used by first-contact providers in the community (for example, community health workers or informal providers) to test persons with any symptoms or risk factors suggestive of TB who are seeking care (that is, it would not be used for active case-finding) to rule out TB (identify those who are triage-test negative) and direct individuals who require further evaluation (those who are triage-test positive) to a confirmatory test. The triage test should have higher accuracy than currently available symptom screening (84% sensitivity and 74% specificity for any symptom in settings with a high prevalence of HIV, and 70% sensitivity and 60% specificity for any symptom in settings with a low prevalence of HIV). Furthermore, triage testing could ensure that patients are rapidly isolated if necessary during their first presentation to a clinical setting, where crowding could result in increased transmission.	(4, 12, 21, 35, 36,)
Drug-susceptibility testing	None	None	Drug-susceptibility testing will be done higher levels of care (for example, at microscopy centres or the district level).	
Target population	Target groups are adults and children suspected of having active TB at any site in countries with a medium prevalence to a high prevalence of TB as defined by WHO	Target groups are adults and children suspected of having active, **pulmonary TB** in countries with a medium prevalence to a high prevalence of TB as defined by WHO	**WHO's categories**: High-prevalence countries are those with > 40 cases per 100 000 population; medium-prevalence countries are those with 20–40 cases per 100 000 population; and low-prevalence countries are those with < 20 cases per 100 000 population (19).	(19)
Target user of the testa	Community health workers and informal providers who have had a minimum of training	Staff trained to the level of auxiliary nurses		

Setting (level of the health-care system)	Community level, village level or higher levels of the health-care system	Health posts and primary-care clinics or higher levels of the health-care system	The lowest level of implementation for a triage test is at a lower level than a microscopy centre, where TB testing and anti-TB agents are available. Implementing a triage test at the peripheral level may increase the number of patients screened for TB, and potentially shorten the time to diagnosis. In addition, a triage test could be useful in all settings where there is likely to be a large influx of patients at different levels of the health-care system (such as in crowded outpatient clinics or refugee camps). In these settings it can serve to enhance infection control by allowing early identification of patients who may have TB and early implementation of infection control measures. A triage test can also be implemented at higher levels of care (that is, at the level of the confirmatory test) if it is at least as sensitive as the confirmatory test and is cheaper; it this case, it would reduce the cost of the confirmatory test.	(23)
Pricing				
Price of individual test (reagent costs only; after scale-up; ex-works [manufacturing costs only, excluding shipping])[a, b]	< US$ 1.00	< US$ 2.00	The acceptable cost of the triage test depends largely on the cost of the confirmatory test that is available, and on the specificity of the triage test. The lower the specificity, the more likely it is that confirmatory testing will be needed and, thus, the lower the cost of the triage test will have to be for the overall testing strategy to be cost effective. It is expected that an affordable, molecular test will be available at the microscopy-centre level within the next 2–5 years, and it will likely cost US$ 4.00 per test. Thus, the cost of the triage test should be lower than that of the confirmatory test. Given the expected large potential market for a triage test, volume-based pricing agreements may be available and may increase access.	(5)
Capital costs for the instrument	No instrument needed	< US$ 50.00	The cost of the device has been estimated for a 1-year lifespan. Alternatively, a more expensive device could be considered but maintenance and calibration would need to be optimized to function remotely and to require simple tools and minimal expertise (or not to be necessary at all), considering that a large number of devices will be required for use in health posts and clinics.	

Performance				
Diagnostic sensitivity (compared with confirmatory test)[a, b]	Overall sensitivity should be > 95% when compared with the confirmatory test for pulmonary TB[c] no lower range of sensitivity was defined for extrapulmonary TB	Overall sensitivity should be > 90% when compared with the confirmatory test for pulmonary TB[c]	Performance characteristics need to be similar to those of the confirmatory test that will be used. Ideally, the test would be at least as sensitive as the available confirmatory test (for example, Xpert MTB/RIF has a diagnostic sensitivity > 98% for culture-positive smear-positive TB and > 67% for culture-positive smear-negative TB). If the sensitivity of the triage test is not similar to that of the confirmatory test then the willingness to implement triage testing would likely depend on how much the sensitivity decreases in comparison with that of the confirmatory test and the standard used at the microscopy centre, as well as the expected increase in the number of patients tested for TB using the triage strategy (which would thus presumably increase the number of TB cases identified). A lower sensitivity would be acceptable only if there were a higher specificity (that is, if the sensitivity were 85% and the specificity were 70% at a TB prevalence of 5%, then the post-test probability of the confirmatory test being positive is only 13%, but if the specificity were 80%, the post-test probability would be 18%). If the instrument can perform multiplex testing for other diseases (for example, for HIV or malaria), a lower sensitivity may be more acceptable.	(5)
Diagnostic specificity[a, b]	Overall specificity should be > 80% when compared with the confirmatory test	Overall specificity should be > 70% when compared with the confirmatory test	The prevalence of TB in the population being tested should be considered when planning the specificity requirement. The prevalence of TB in the community is likely to be lower than at clinics or health posts (prevalence in the community estimated at < 5%; prevalence at clinics or health posts estimated at < 10%). Thus, a higher specificity is necessary for testing in the community to achieve a sufficient post-test probability that justifies referral (for example, at a 5% prevalence among people seeking care, using a test with 95% sensitivity and 80% specificity gives a post-test probability of 24%). The trade-off between cost and specificity, as well as the burden on the health system when a large number of false-positive patients are tested, need to be considered. Furthermore, there might be a psychological and social component to triage testing if a large number of patients are referred but then are identified as not having TB. These psychological and social burdens might be less if the sample – not the patient – is sent for a confirmatory test.	(5)
Quantitation	Qualitative			

Operational characteristics				
Sample type	Non-sputum samples (such as urine, oral mucosal transudates, saliva, exhaled air, or blood from a finger-stick)	Sputum; a non-sputum sample type would be preferred (such as urine, oral mucosal transudates, saliva, exhaled air, or blood from a finger-stick); imaging technology	For successful implementation at the community level, a test would use an easily accessible sample (such as urine, blood, breath condensate). Non-sputum-based samples are preferred, both for biosafety and because of the stigma associated with TB.	
Sample volume	Urine: < 5 ml Finger-stick blood: < 10 mcl Saliva or transudates: < 0.1 ml	Urine: < 10 ml Finger-stick blood: < 25 mcl Saliva or transudates: < 0.2 ml Sputum: < 0.1 ml	The lowest volume possible for all types of samples would be preferred, especially since children might have difficulty providing a larger volume for a sample; however, this requirement for the lowest volume possible should not come at the expense of decreased sensitivity. The test should also be able to accommodate larger volumes when they are available if they would potentially increase the sensitivity.	
Manual preparation of samples (steps needed after obtaining sample)	Sample preparation should be integrated or not required (excluding waste disposal); precise timing and measuring should not be required	2 steps (excluding waste disposal); precise timing and measuring should not be required	The simpler, more portable and more durable the test is, the more likely it is to be implemented in peripheral settings.	(8, 20)
Reagent integration	All reagents should be contained in a consumable device	Only 1 external reagent should be required and should be part of the test kit		
Time to result[a, b]	< 5 minutes including time spent preparing the sample	< 30 minutes including time spent preparing the sample	The time to result includes the time needed to prepare the sample and processing time. The ideal time to result has not been studied and might vary significantly among countries and among settings where patients are tested (such as a busy clinic versus in the community).	
Daily throughput	> 30 tests per 8-hour day; ability to test individual samples so that batching is not required	> 20 tests per 8-hour day; ability to test individual samples so that batching is not required	The need for a short turnaround time, the possibility of batching or testing individual samples without batching (asynchronous testing), and the ability to test multiple samples at the same time are interrelated. The time to result is probably the most important parameter since extending the wait time for patients may result in loss to follow-up.	(4)

Sample capacity	Multiple samples should be able to be tested at the same time	One sample at a time	If the time to result outlined above is achieved, then the capacity to test multiple samples simultaneously might not be necessary.
Biosafety	No need for a biosafety cabinet; consumables should be able to be disposed of as general laboratory waste; a sharps container should be provided if a lancet is necessary	No need for a biosafety cabinet; standard procedures should be used if sputum is the sample of choice; a sharps container should be provided if a lancet is necessary	Reducing concerns about biohazards will increase the acceptability of the test.
Waste disposal – solid	Simple trash, or recyclable or compostable plastics and consumables should be used	Simple trash, or recyclable or compostable plastics and consumables should be used	Reducing the steps required to dispose of waste will facilitate the test's use in the most peripheral settings.
Waste disposal – infectious	No infectious waste should be generated	Incineration of infectious materials	Simplifying waste disposal by using only non-infectious samples will facilitate the test's use in the most peripheral settings.
Multiuse platform	No platform	None	Providing multiplexed testing for other diseases may increase the acceptability of the test.
Instrument	None	A small, portable or hand-held device weighing < 1 kg	
Power requirements[a]	None	Optional battery or solar-powered operation	If no power is required or if an independent power source can be used, the test is more likely to be useful in peripheral settings.
Maintenance[a]	Disposable, no maintenance required	Preventative maintenance should not be needed until after 1 year or 1 000 samples; only simple tools and minimal expertise should be required; an alert should be included to indicate when maintenance is needed	If a device is anticipated to have a longer lifespan, then a maintenance alert is essential to ensure proper functioning in settings where it is unlikely that the same person will always handle the device, and that records will be kept about the duration of use. Furthermore, it will be essential that only simple tools and minimal expertise are necessary to maintain the device, given the quantity of devices that is likely to be in use.

Calibration	None required	The instrument should be able to be calibrated remotely, to calibrate itself, or no calibration should be needed	
Result capturing, documentation, data display	Ideally, the test would not require an instrument; there should be a visual readout with the ability to save results using a separate, attachable reader	The test menu must be simple to navigate; the instrument should have an integrated LCD screen, a simple keypad or touch screen, and the ability to save results using either the instrument or a separate reader	Output from the device should not need interpretation (for example, colour bands may be used to indicate results), and results should be clear (for example, positive/negative).
Regulatory requirements	The assay and system manufacturing should comply with ISO EN 13485 or higher standards or regulations, and comply with ISO IEC 62304 Medical Device Data Systems; the manufacturing facility should be certified and authorized for use by a regulatory authority that is a member of the of the International Medical Device Regulators Forum, formerly known as the Global Harmonization Task Force; the assay must be registered for in vitro diagnostic use		

Data export (connectivity and interoperability)	All data should be able to be exported (including data on use of the device, error rates and rates of invalid tests, and personalized, protected results) over a USB port and network; network connectivity should be available over an Ethernet, Wi-Fi or GSM/UMTS mobile broadband modem, or a combination of these; results should be encoded using a documented standard (such as HL7) and be formatted as JSON text; JSON data should be able to be transmitted through HTTP(S) to a local or remote server as results are generated; results should be stored locally, and queued during network interruptions to be sent as a batch when connectivity is restored	None required	The ability to export data may facilitate reporting to patients and providers, connectivity to an electronic record, improved surveillance, and allow for better supply-chain management.	(8, 20)
Electronics and software	None	Integrated		
Data analysis	Integrated		Integrating the capacity for data analysis into the test will reduce interpretation errors.	
Operating temperature and humidity level[a]	Between +5 °C and +50 °C with 90% humidity	Between +5 °C and +40 °C with 70% humidity		

Reagent kit – transport	A cold chain should not be required; there should be no special requirements		A cold chain and other special requirements should not be necessary since refrigerators are not likely to be present and continuous power is likely to be limited in the settings where the test will be used.	(8, 20)
Reagent kit – storage and stability	2 years at +5 °C to +40 °C with 90% humidity; should be able to tolerate stress during transport (72 hours at +50 °C); no cold chain should be required	12 months at +5 °C to +35 °C with 70% humidity; should be able to tolerate stress during transport (72 hours at +50 °C); no cold chain should be required		
Reagent calibration	No reagent should be necessary	No reagent should be necessary		
Additional supplies (not included in kit)	None			
Internal quality control	Internal controls for sample processing and detection of TB should be included	Internal controls for sample processing should be included	Internal quality controls should be included in addition to external quality assurance.	
Training and education	< 1 hour for a health-care worker who has had a minimum of training	< 1/2 day (4 work hours) for a health-care worker	The need for training and education about how to use the test should be low, given the high turnover of health-care workers.	(43)

LCD, liquid crystal display; GSM, Global System for Mobile Communications; UMTS, Universal Mobile Telecommunications System.
a These characteristics were considered to be the most important.
b These characteristics were associated with the most uncertainty.
c The performance characteristics of the triage test need to be matched with those of the confirmatory test that will be used.

Table A3. Detailed target product profile (TPP) for a rapid sputum-based test for detecting TB at the microscopy-centre level of the health-care system

Characteristic	Optimal (ideal) requirements	Minimal requirements	Explanations	References
Scope				
Goal and potential market	A sputum-based test to detect **pulmonary TB** to be used at the level of a microscopy centre to support initiation of TB therapy during the same clinical encounter or on the same day	A sputum-based test to detect **pulmonary TB** to be used at the level of a microscopy centre to support initiation of TB therapy during the same clinical encounter or on the same day	The test should be optimized for TB detection. It is conceivable that NAAT or another technology with simple sample preparation could be developed. The test would be used at peripheral-level microscopy centres as a replacement for sputum-smear microscopy. If the test has greater sensitivity than smear microscopy and is able to be implemented more widely than the Xpert MTB/RIF assay or culture, it has the potential both to increase the number of patients who are treated and to ensure that they are treated earlier in the course of their disease.	(37, 44–46)
Drug-susceptibility testing	None		Drug-susceptibility testing could be performed during the second step of the test (see the TPP for DST, Table 4). Ideally, this would be performed using the same platform.	
Target population	Target groups are all patients suspected of having pulmonary TB and able to produce sputum, in countries with a medium prevalence to a high prevalence of TB as defined by WHO		The target population described under the optimal and minimal requirements corresponds to WHO's 2013 recommendations for using the Xpert MTB/RIF assay, although the resource implications need to be considered. **WHO's categories:** High-prevalence countries are those with > 40 cases per 100 000 population; medium-prevalence countries are those with 20–40 cases per 100 000 population; and low-prevalence countries are those with < 20 cases per 100 000 population (19).	(17, 19)
Target user of test[a]	Health-care workers who have had a minimum of training (that is, trained to a level that is similar to that necessary for performing smear microscopy)			
Setting (level of the health-care system)	Microscopy-centre level (that is, primary health centres with attached peripheral laboratories)		Conditions that prevail in microscopy centres in high-burden countries have been described, and the characteristics required for novel tests at this level of the health-care system have also been proposed. The potential of the new product and the scale at which it is adopted will depend on whether the product has the necessary characteristics. A test that can be used at the level of a microscopy centre would leverage the infrastructure used for smear microscopy and could potentially reach large numbers of individuals suspected of having TB.	(8, 20)

Pricing				
Price of individual test (cost of reagent only; after scale-up; ex-works [manufacturing costs only, excluding shipping])[a, b]	<US$ 4.00 for TB detection	<US$ 6.00 for TB detection	The final price may depend on whether volume-based pricing models are used. Since millions of smear microscopy tests are done every year, the expected market for this product is likely to be substantial (17) and, therefore, volume-based pricing agreements may be possible. A more expensive test is less likely to penetrate the market.	(17)
Capital costs for the instrument	<US$ 500 per module	<US$ 1400 per module	The lower the capital cost of the instrument is, the lower the initial cost would be and, thus, the barrier to implementation would also be lower, particularly since the volume of instruments that would be distributed to microscopy centres is sizeable. The cost of the instrument should include warranties, service contracts and technical support.	
Performance				
Analytical and diagnostic sensitivity for detecting TB[a]	Sensitivity for TB case-detection should be better than Xpert MTB/RIF – that is, < 4.5 genome equivalents/reaction and < 10e2 CFU/assay using 1 sample; ideally, the sensitivity would be as good as liquid culture (resulting in a diagnostic sensitivity of > 95% for a single test when compared with culture); sensitivity should be > 99% for culture-positive smear-positive TB and > 68% for culture-positive smear-negative TB using 1 sample	Sensitivity for TB case-detection should be between that of smear microscopy and Xpert MTB/RIF – that is, between 10e2 CFU/assay and 10e5 CFU/assay using 1 sample, resulting in sensitivity ≥ 80% for a single test when compared with culture; sensitivity for smear-negative TB should be > 60%, and for smear-positive it should be 99%	To identify patients early and reduce transmission it is important to aim for a test with a better sensitivity than that of using 1 Xpert MTB/RIF assay to detect TB. However, if the new test were used at a microscopy centre with good links to treatment, then a test with performance characteristics that are better than smear microscopy and less good than Xpert MTB/RIF would reduce transmission more than using Xpert MTB/RIF only at the district level.	(23,42, 47–49)

Analytical and diagnostic specificity for detecting TB[a]	Specificity for TB case-detection should be at least as good as Xpert MTB/RIF when compared with culture (≥ 98% specificity); there should be no cross-reactivity with other organisms, including non-tuberculous mycobacteria	Specificity for TB case-detection should be at least as good as Xpert MTB/RIF when compared with culture (≥ 98% specificity); there should be no cross-reactivity with other organisms, including non-tuberculous mycobacteria		(42, 47, 48)
Treatment monitoring possible	Yes	No	A test that is able to replace smear microscopy for treatment monitoring (for example, by detecting viable bacteria) is more likely to be adopted and to completely replace smear microscopy.	
Operational characteristics				
Sample type	Unprocessed sputum	Unprocessed sputum		
Sample volume	0.1—10 ml	0.5—5 ml	The lowest volume possible for all types of samples should be 0.1 ml, especially since children and HIV-positive patients may have difficulty providing a sample; however, this ideal requirement should not come at the expense of decreased sensitivity. If a higher volume is available, the test should be able to use it if doing so would increase sensitivity.	
Manual preparation of samples (steps needed after obtaining sample)[a]	No steps or 1 step; precise volume control and precise timing should not be required	Maximum of 2 steps; precise volume control and precise timing should not be required	The Xpert MTB/RIF assay has 2 steps. Devices such as a centrifuge or heat block are available only infrequently at the level of microscopy centres; therefore, these should not be required for novel assays. Also, the expertise needed to operate a micropipette is often lacking; if pipetting is needed, this should be able to be done using a Pasteur pipette.	(8, 20)
Reagent integration	All reagents should be contained in a single device		A maximum of 2 external reagents should be needed and these should be included as part of the test kit	
Time to result[a, b]	< 20 minutes including time spent preparing the sample	< 2 hours including time spent preparing the sample	These limits include the time needed for sample preparation and processing. Two hours is the current time to result for the Xpert MTB/RIF assay.	(50, 51)
Daily throughput	15–20 tests	10 tests	The daily throughput for the GeneXpert IV system is about 10 tests. The daily throughput needed in most microscopy centres is ≤10 tests (C. Denkinger, personal communication, 2013).	

Sample capacity	Single or multiple samples should be able to be tested at the same time	None	The need for a short turnaround time, the possibility of batching or testing individual samples without batching (asynchronous testing), and the ability to test multiple samples at the same time are interrelated. The time to result is probably the most important parameter since extending the wait time for patients may result in loss to follow-up. If the optimal time to result is met, then in most settings the capacity for performing multiple tests at the same time will not be needed.	
Walk-away operation, random access, STAT sampling	These features must be integrated into the test	None of these features	Ideally, 1 sample should not occupy the instrument without it still being able to process other samples (that is, random access or parallel analyses should be possible). If the platform is multiplexed, then running different assays at the same time should be feasible. If the optimal time to result is met, then in most settings random access or STAT sampling will not be needed.	
Biosafety	Should have the same requirements as Xpert MTB/RIF	Same as smear microscopy	A biosafety cabinet is not commonly available at the level of a microscopy centre; therefore, a sample-collection device that can be sealed immediately or can be rapidly decontaminated, or both, is important for subsequent processing without a cabinet. If a novel test has increased biosafety (if it is safer to use) then its acceptability to providers will be enhanced. Further information is provided in WHO's Tuberculosis Laboratory Biosafety Manual.	(8, 20, 52)
Waste disposal – solid	Should require no more than smear microscopy	Should require no more than Xpert MTB/RIF	Further information is provided in WHO's Tuberculosis Laboratory Biosafety Manual and the policy update on using the Xpert MTB/RIF assay.	(17, 52)
Waste disposal – infectious	Should require no more than smear microscopy	Should require no more than Xpert MTB/RIF	Further information is provided in WHO's Tuberculosis Laboratory Biosafety Manual and the policy update on using the Xpert MTB/RIF assay.	(17, 52)
Multiuse platform	Yes (to test for HIV and other diseases, depending on the local epidemiology)	None	If the platform can perform multiplex testing or can be used for different tests, then the acceptability of the new test will likely be increased, especially in the private sector.	
Instrument	Ideally, a single device that is an integrated system	Up to 2 instruments that are independent of each other	Ideally, a single device is preferred but up to 2 instruments would be acceptable (for example, for separate sample processing and detection).	
Power requirements[a]	Battery operated with recharging capability and a circuit protector		Continuous power is not always available at the level of a microscopy centre; therefore, a battery-operated device that can be recharged using solar power would be ideal and would allow the test to be used in microscopy centres in many different settings.	(8, 20)

Maintenance and calibration[a]	Preventative maintenance and calibration should not be needed until after 2 years or 5000 samples; only simple tools and minimal expertise should be required; an alert should be included to indicate when maintenance is needed	Preventative maintenance should not be needed until after 1 year or 1000 samples; only simple tools and minimal expertise should be required; an alert should be included to indicate when maintenance is needed	A maintenance alert is necessary to ensure proper functioning in settings where it is unlikely that the same person will always handle the device and that records will be kept about the duration of use. Furthermore, it will be essential that only simple tools and minimal expertise are necessary to perform maintenance, given the quantity of devices that are likely to be used; additionally, service visits are unlikely to be feasible outside of urban settings.	(31)
Calibration	The instrument should be able to be calibrated remotely or no calibration should be needed	The instrument should be able to be calibrated remotely, to calibrate itself, or no calibration should be needed		
Result capturing, documentation, data display	An integrated results screen and the ability to save and print results should be included; external printing would be acceptable	Integrated LCD screen and ability to save results		
Regulatory requirements	Manufacturing of the assay and system should comply with ISO EN 13485 or higher standards or regulations, and comply with ISO IEC 62304 Medical Device Data Systems; the manufacturing facility should be certified and authorized for use by a regulatory authority that is a member of the International Medical Device Regulators Forum, formerly known as the Global Harmonization Task Force; the assay must be registered for in vitro diagnostic use			

Data export (connectivity and interoperability)	All data should be able to be exported (including data on use of the device, error rates and rates of invalid tests, and personalized, protected results) over a USB port and network; network connectivity should be available through an Ethernet, Wi-Fi, or GSM/UMTS mobile broadband modem, or a combination of these; results should be encoded using a documented standard (such as HL7) and be formatted as JSON text; JSON data should be able to be transmitted through HTTP(S) to a local or remote server as results are generated; results should be stored locally, and queued during network interruptions to be sent as a batch when connectivity is restored	Integrated ability for all data to be exported from the device in a user-friendly format (including data on use of the device, error rates and rates of invalid tests, and non-personalized results) over a USB port	A data-export feature could be leveraged for data export, quality control, supply-chain management and surveillance.	(8, 20, 53)
Electronics and software	Should be integrated into the instrument		If a separate PC is needed, it will likely limit the ability to update software since staff with the skills needed to operate a PC are not present in all microscopy centres. Furthermore, security will be an issue, and separate PCs might be stolen.	(8, 20)
Data analysis	Integrated			
Operating temperature and humidity level[a]	Between +5 °C and +50 °C with 90% humidity	Between +5 °C and +40 °C with 70% humidity	High environmental temperatures and high humidity are often present in countries where TB is endemic. Dust also is a problem in these settings, and the need to adequately protect optics should be considered.	(20, 31)

Reagent kit – transport	No cold chain should be required	No cold chain should be required	Refrigeration is often not available at customs or during transportation.	(8, 20)
Reagent kit – storage and stability	2 years at +5 °C to +40 °C with 90% humidity; should be able to tolerate stress during transport (72 hours at +50 °C); no cold chain should be required	12 months at +5 °C to +35 °C with 70% humidity; should be able to tolerate stress during transport (72 hours at +50 °C); no cold chain should be required	High environmental temperatures and high humidity are often present in countries where TB is endemic.	(8, 20, 31)
Reagent calibration	The kit should have a bar code for calibration, or calibration should not be required	External controls provided by the manufacturer should be tested once a week		(31)
Additional supplies (not included in the kit)	None	Items that would be present in a laboratory that is capable of testing blood chemistry		(8, 20)
External quality control	Both positive and negative controls for *Mycobacterium tuberculosis* detection should be included as well as controls for drug-susceptibility testing	Both positive and negative controls for *M. tuberculosis* detection and drug-susceptibility testing should be included	External controls should be used periodically. The need for external quality control should not exceed what has been established for smear microscopy.	
Internal quality control	Controls for sample processing and amplification should be included		The need for external quality control should not exceed what has been established for smear microscopy.	(31)
Training and education	6 work hours for staff at the level of a microscopy technician	3 days (or 24 work hours) for staff at the level of a laboratory technician	Training to use the Xpert MTB/RIF assay takes 1–3 days.	(43)

NAAT, nucleic acid amplification test; CFU, colony forming units; LCD, liquid crystal display; GSM, Global System for Mobile Communications; UMTS, Universal Mobile Telecommunications System; PC, personal computer.
a These characteristics were considered to be the most important.
b These characteristics were associated with the most uncertainty.

Table A4. Detailed target product profile (TPP) for a next-generation drug-susceptibility test to be implemented at peripheral levels of the health-care system to inform decisions about first-line treatment regimens

Characteristic	Optimal (ideal) requirements	Minimal requirements	Explanations	References
Scope				
Key assumptions	The development **timeline is < 5 years**; this approach would use **1 platform** for TB detection and DST; this TPP has taken the developers' perspective by assuming that new regimens will be implemented and available in parallel with current standard–of-care regimens, at least initially			
Rationale	To provide support for **effective first-line anti-TB therapy** in the context of the roll-out of new regimens after 2014 (that is, HRZE, REMox or PaMZ); to provide the characteristics and qualities of a test that would have a sufficiently rapid turnaround time (that is, results can be provided during the same visit) for TB detection and would provide data about DST that can be used to inform treatment decisions			
Goal	Diagnosis of TB disease and detection of drug resistance to inform decision-making about the optimal first-line regimen (HRZE, REMox or PaMZ) for treatment, and possibly to detect the presence of additional resistance to second-line anti-TB agents and the need for further testing		The market for a test that includes DST and detection is all patients tested for TB, which is approximately 10 times the number of detected cases, or about 60 million patients. If DST were performed in a second step, the market would be all patients in whom TB had been detected (or about 9 million). The market for a test to detect PZA resistance is different because the current achievable performance characteristics of a molecular test for PZA resistance is a maximum of 95% for both sensitivity and specificity; therefore, a test for PZA resistance could be used as a follow-on test only if RIF resistance has been confirmed (a higher prevalence of resistance leads to a higher PPV for the detection of resistance to a particular anti-TB agent). This means that the market for testing for PZA resistance is only as large as the number of patients confirmed to have MDR-TB, which is about 450 000, although the number is likely to increase as testing for MDR-TB increases.	

Priority of anti-TB agents for testing[a]	In order of decreasing importance: 1.RIF 2. FQs (including MOX) 3. INH and PZA (equally important) 4. AGs and CAP Optimally all drugs would be included, but as a minimum at least RIF should be included	**Detecting resistance to RIF and FQs:** If there is only resistance to RIF, then HRZE or REMox are not an option. RIF is an indicator: if resistance to RIF is present then resistance against other anti-TB agents is more likely to be present. Resistance to PZA and INH is highly associated with resistance to RIF; resistance to FQs is moderately associated with resistance to RIF; that is, >80% of RIF-resistant strains are also resistant to PZA or INH, or both, and 10–30% are resistant to FQs (26–30). MOX is included in the REMox and PaMZ regimens. If there is monoresistance to MOX, then REMox is not an option. The importance of sensitivity depends on the trade-off with a false-negative result. Limited data are available about the impact of undiagnosed monoresistance to MOX on outcomes with REMox: it seems to reduce efficacy by about 15% relative to 2HRZE/4HR (2 months of HRZE followed by 4 months of HR). There are insufficient data on cross-resistance occurring across different FQs. The differentiation of resistance among FQs is more a function of interpreting mutations (that is, evaluating the hierarchical structure of mutations) rather than detecting differenæt mutations. No data are available on the transmissibility of MOX strains. Detecting resistance to FQs other than MOX is important at higher levels of care where other FQs are often used for individualized therapy (for example, LVX). **Detecting resistance to PZA:** PZA resistance is highly associated with RIF resistance. With the advent of the PaMZ regimen, confirmation of PZA sensitivity will be important to avoid treating patients using only two anti-TB agents and risking the development of resistance or referring patients directly for MDR-TB therapy. The importance of confirming PZA sensitivity if the HRZE/REMox regimen is used is less clear since the prevalence of PZA resistance occurring independently of resistance to RIF is low (about 3%). PZA is likely to be an important component of other future regimens independent of PaMZ. No data are available on the impact of undiagnosed PZA resistance on outcomes with PaMZ. No data are available on the transmissibility of monoresistant PZA strains.	(3, 30, 32, 54–56)

Detecting INH resistance: Modelling data have shown that testing for resistance to INH has only minimal incremental value over testing for resistance to RIF alone on the incidence of MDR-TB and monoresistance to INH, as well as on mortality at the population level (for example, in a setting like India). These findings might not apply to countries with high rates of coinfection with HIV or with very high prevalence of resistance to INH; the need for detecting INH resistance might increase as more preventive therapy using INH is rolled out. Furthermore, other data suggest that detecting resistance to INH is important at the individual level, and the acceptance of a test among providers might be enhanced if testing for INH resistance were incorporated (this was a stakeholder's opinion).

Detecting AG resistance: Testing for resistance to AGs might be beneficial for directing care and, at higher levels of care, avoiding treating patients with pre-XDR-TB with therapy for MDR-TB.
There are insufficient data on cross-resistance across different AGs.
The GenoTypeMTBDRsl test (Hain Lifescience, Nehren, Germany) has a specificity of 98% for detecting resistance to AGs.

Sequence of detection and drug-resistance testing: No data are available about whether it is better to detect drug-resistance and TB simultaneously or to test initially only for TB and then use a second step for DST. The data necessary to inform this decision include information about how time to diagnosis relates to loss to follow-up. In other words, if initially testing for TB and drug-resistance comes at the expense of rapid TB detection and thus leads to a loss to follow up, then TB detection alone followed by a separate step for DST might be beneficial over simultaneous testing for TB and DST. This scenario might vary substantially among countries.
Also, initially testing for TB and DST might come at the expense of the sensitivity for TB detection, depending on the platform used and cost. However, a delay in DST might result in patients receiving inappropriate treatment until they return, assuming that the DST result will not be known in time to inform initial decisions about treatment. The acceptability of a longer wait time might vary among countries, and informing the patient of results on the same day if the result is not available during the first visit, might be associated with substantial costs.
Performing DST for resistance to RIF in conjunction with TB detection would allow RIF to serve as an indicator and reduce the specificity required of a test used in low-prevalence settings (see information on specificity below).

Assay design	The assay should be designed in such a manner that the addition of or removal of analytes does not require extensive analytical and clinical re-verification and revalidation of the assay	The assay should be designed in such a manner that it is capable of being updated as needed, with minimal redevelopment required.	
Target population	Target groups are all patients suspected of having TB, with a special focus on those at high risk of morbidity and mortality from drug-resistant TB, such as people living with HIV and those at high risk of having MDR-TB (for example, household contacts of patients diagnosed with MDR-TB, and persons with a history of TB, especially those for whom first-line therapy has failed) in countries with a medium prevalence to a high prevalence of TB as defined by WHO	The optimal target population corresponds to WHO's 2013 recommendations for using the Xpert MTB/RIF assay, although the resource implications need to be considered. Children aged <11 years have limited ability to produce sputum for testing. Therefore, initial validation studies should focus on adults. **WHO's categories:** High-prevalence countries are those with > 40 cases per 100 000 population; medium-prevalence countries are those with 20–40 cases per 100 000 population; and low-prevalence countries are those with < 20 cases per 100 000 population (19).	(3, 19)
Target user of test[a]	Health-care workers with training necessary for performing smear microscopy		
Setting (level of the health-care system)	Microscopy-centre level or higher levels of the health-care system	Implementation at the microscopy-centre level should be feasible using the specifications as outlined. This would embed the test in an infrastructure that is based around smear microscopy. However, the test could be implemented at higher levels of care as well. Testing for resistance to the anti-TB agents included in second-line therapy could be incorporated into separate reactions, but ideally it would be feasible to test the same specimen.	(8, 20, 26, 27, 44, 45, 57)
Pricing			
Price of individual test (cost of reagent and consumables only; after scale-up; ex-works [manufacturing costs only, excluding shipping])[a]	< US$ 10.00	Meeting participants emphasized the critical need for the price to be kept within an affordable range. A price that is higher than available technologies (for example, molecular testing for simultaneous detection of *Mycobacterium tuberculosis* and RIF resistance or testing for resistance to RIF and INH, both of which cost approximately US$ 10.00 per test) would be justified only if the new test brings substantial added value in terms of vastly improved performance, greater suitability for decentralization, and the number of anti-TB agents for which resistance can be detected. The market size for reflex testing will also need to be taken into account. *Consensus was not reached on the minimal requirement (that is, the highest acceptable price).*	(58)

Capital costs for the instrument	US$ 1400 per module for a test combining TB detection and DST	US$ 1400 per module for DST only	These cost estimates are based on the likely volume of tests. The lower the capital costs of the instrument are, the lower the initial cost would be, and thus the barrier to implementation would also be lower, particularly since the volume of instruments that would be distributed to microscopy centres is sizeable. The cost of the instrument should also include warranties, service contracts and technical support.	
Performance				
Diagnostic sensitivity for TB detection[a]	Sensitivity for detecting TB should be > 95% for a single test when compared with 2 liquid cultures; for smear-negative TB it should be > 68%; for smear-positive TB it should be 99%	Sensitivity should be > 80% for a single test when compared with 2 liquid cultures; for smear-negative TB it should be > 60%; for smear-positive TB it should be 99%	The optimal sensitivity specified has been raisedto that of the Xpert MTB/RIF assay; the minimal sensitivity is substantially better than that of smear microscopy. Sun et al found that a test implemented at the microscopy level with only 50% sensitivity for smear-negative TB will lead to a substantial improvement in TB detection over a test like the Xpert MTB/RIF assay that is implemented at the district level (23). There are no data to inform discussions about the sensitivity needed by a test to gain the confidence of providers and curb empirical treatment in people who test negative.	(23)
Diagnostic specificity for TB detection[a]	Specificity should be > 98% for a single test when compared with culture	Specificity should be > 98% for a single test when compared with culture		(17, 48, 59)
Diagnostic sensitivity for DST compared against genetic sequencing as the reference standard[a]	Sensitivity should be > 98% for detecting targeted SNPs for resistance to RIF, FQs, PZA, INH, and AG/CAP when compared with genetic sequencing	Sensitivity should be > 98% for detecting targeted SNPs for resistance to RIF, and 95% sensitivity for detecting targeted SNPs for resistance to FQs, PZA, INH, and AG/CAP when compared with genetic sequencing		
Diagnostic sensitivity for DST compared against phenotypic DST as a reference standard[a]	> 95% sensitivity for detecting RIF, FQ, PZA, INH and AG/CAP resistance in comparison to recommended phenotypic culture reference DST for specific anti-TB agent	> 95% sensitivity for detecting RIF resistance; > 90% for detection of FQ, PZA, INH and AG resistance in comparison to recommended phenotypic culture reference DST for specific anti-TB agent	Modelling data suggest that for rapid DST to be more cost effective than culture, on a GeneXpert platform it must attain an aggregated sensitivity of 88% for all clinically relevant mutations. A lower sensitivity could be tolerated for a test with high specificity, particularly if the prevalence, and thus the pretest probability, are high. The sensitivity achieved against a phenotypic reference standard will be only as good as the mutations that are targeted (that is, even if all known mutations conferring INH resistance are detected with 100% sensitivity when compared against a sequencing reference standard, 100% sensitivity cannot be achieved against a phenotypic reference standard because the knowledge of all molecular targets that confer resistance is not complete).	(3, 59–65)

Diagnostic specificity for DST compared against genetic sequencing as the reference standard [a]	Specificity should be ≥98% for any anti-TB agent for which the test is able to identify resistance when compared against genetic sequencing as the reference standard		If alternative regimens are available, effective, safe and not too cumbersome (for example, HRZE or REMox), then a lower PPV might be tolerated (for example, a false-positive MOX result may result only in 6 months of HRZE instead of 4 months of REMox). Because the pretest probability is low when all-comers without any additional risk factors are tested in settings with a low prevalence of resistance, the specificity has to be very high: if the prevalence of resistance is about 3% according to surveillance data, then a specificity of 99% results in a PPV of only 74%. A very high specificity (for example, ≥99.7%) is necessary in order to reach a PPV of >90%; if the prevalence of resistance is ≥20% (for example, when resistance to RIF is used as an indicator or when testing is only done in high-risk patients), a specificity of >97% is sufficient to achieve a PPV of 90%.	(65,66)
Diagnostic specificity for DST compared against phenotypic DST as a reference standard [a]	The specificity of targeted sequencing for the mutations included for any anti-TB agent for which the test is able to identify resistance should be ≥ 98% when compared against the phenotypic reference standard recommended for each anti-TB agent		The estimates of specificity for molecular tests in comparison with phenotypic testing as a reference-standard might be falsely low as the reference-standard has limited sensitivity. Therefore it is important to use the optimized phenotypic reference standard for a drug in comparison.	(3, 59–62, 65)
Limit of detection – TB detection after first reaction	Should be better than Xpert MTB/RIF for TB case-detection – that is, < 4.5 genome equivalents/reaction and < 10e2 CFU/ assay using 1 sample	Should be between smear microscopy and Xpert MTB/RIF for TB case-detection – that is, between 10e2 CFU/ assay and 10E5 CFU/ assay using 1 sample	Limit of detection testing should be performed as outlined in the United States Food and Drug Administration's draft guidance document dated 19 June 2013 (http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm357617.htm).	(17, 47, 48)
Limit of detection – TB detection after second reaction for DST	Should be no worse than Xpert MTB/RIF for TB case-detection – that is, ≥4.5 genome equivalents/reaction and 131 CFU/mL of sputum	Should be between smear microscopy and Xpert MTB/RIF for TB case-detection – that is, between 10e2 CFU/ assay and 10e5 CFU/ assay using 1 sample	A slightly decreased analytical sensitivity for TB detection in the second reaction for resistance testing (in comparison with the first reaction) both for the optimal and minimal requirements will avoid resistance calls (for example, no M. tuberculosis but resistance present) but will come at the expense of a slightly lower sensitivity for DST.	(49)
Analytical specificity for TB detection	No cross-reactivity with other organisms including non-tuberculous mycobacteria	No cross-reactivity with other organisms including non-tuberculous mycobacteria		

Indeterminate results during detection[a]	< 2%	< 5%	Indeterminate results may be caused by a lower sensitivity for detecting TB during the second reaction.	
Reproducibility	Interassay coefficients of variance should be ≤ 10.0% at the high and low extremes of the assay		This applies if the quantitative outcomes of a test are measurable (for example, for the limit of detection, and cycle threshold values).	
Interfering substances	No interference should be caused by those substances known to occur in the human respiratory and pulmonary tracts, including blood that could potentially inhibit a PCR reaction, and substances used to treat or alleviate respiratory disease or symptoms			
Treatment monitoring capability	Yes	No	A test that is able to replace smear microscopy for treatment monitoring (for example, by detecting viable bacteria) is more likely to be adopted and to completely replace smear microscopy; thus, it would have a larger market as well.	
Operational characteristics				
Sample type	Unprocessed sputum	Unprocessed sputum		
Sample volume	0.1–10 ml	< 0.5–2 ml	The lowest volume possible for all types of samples should be 0.1 ml, especially since HIV-positive patients may have difficulty providing a sample; however, this ideal requirement should not come at the expense of decreased sensitivity. If a higher volume is available, the test should be able to use it if doing so would increase sensitivity. Additionally, the ideal test would need only 1 sample even if requires 2 steps or reactions.	
Manual preparation of samples (steps needed after obtaining sample)[a]	No steps or 1 step; precise volume control and precise timing should not be required	Maximum of 2 steps; precise volume control and precise timing should not be required	Devices such as a centrifuge or heat block are available only infrequently at the level of microscopy centres; therefore, these should not be required for novel assays. Also, the expertise needed to operate a micropipette is often lacking.	(8, 20)
Reagent integration	All reagents should be contained in a single device	A maximum of 2 external reagents should be needed and if required, should be included in the test kit		
Time to result[a]	< 30 minutes for detection and DST	< 2 hours for DST alone	The need for rapid turnaround, the possibility of batching or using random access for testing, and the ability to test multiple samples at the same time are interrelated. The time to result is probably the most important parameter since extending the wait time for patients may result in loss to follow-up.	(50, 51)

Daily throughput	> 25 tests	> 5 tests	The daily throughput for the GeneXpert IV system is about 10 tests. The daily throughput needed in most microscopy centres is <10 tests per day. If the modules were cheap and throughput could be expanded if needed by using additional modules, then the daily throughput per module could be reduced to the minimal requirement (≥5) (C. Denkinger, personal communication, 2013).	
Sample capacity and throughput	Multiple samples should be able to be tested at the same time; random access should be possible	Batching should be possible	Ideally, 1 sample should not occupy the instrument without it still being able to process other samples (that is, random access or parallel analyses should be possible). If the platform is multiplexed, then running different assays at the same time should be feasible.	
Walk-away operation	These features are required; there should not be a need for operator intervention once the sample has been placed into or on the instrument	No more than 1 step of operator intervention should be needed once the sample has been placed into or on the system	Once the sample has been loaded into an instrument, then further operator intervention should not be required until detection has occurred. This characteristic is related to the characteristics for sample preparation and assay processing (that is, the steps needing to be completed after a sample has been obtained).	
Biosafety	Should have the same requirements as the Xpert MTB/RIF assay	Should have the same requirements as the Xpert MTB/RIF assay	A biosafety cabinet is not commonly available at the level of a microscopy centre; further information is provided in WHO's *Tuberculosis Laboratory Biosafety Manual*.	(8, 20, 52)
Waste disposal – solid	Should require no more than smear microscopy; should have the possibility of recycling some waste	Should require no more than Xpert MTB/RIF	Further information is provided in WHO's Tuberculosis Laboratory Biosafety Manual and the policy update on using the Xpert MTB/RIF assay.	(52, 67)
Waste disposal – infectious	Should require no more than Xpert MTB/RIF	Should require no more than Xpert MTB/RIF	Further information is provided in WHO's Tuberculosis Laboratory Biosafety Manual and the policy update on using the Xpert MTB/RIF assay.	(52, 67)
Multiuse platform	Yes	None required	Multiplex testing or the ability to use a platform to perform different tests will likely increase the acceptability of the new test, especially in the private sector.	
Instrument	Ideally, would be a single integrated system that is modular to allow throughput to be increased if needed	Up to 2 instruments within the system that are independent of each other	Ideally, a single device is preferred but up to 2 instruments are acceptable (for example, for separate sample processing and detection).	

Power requirements[a]	Battery operated with the ability to run for 1 day on the battery, and with recharging capability (which could be solar powered) and a circuit protector	Capable of running on standard electricity plus an uninterrupted power supply unit to enable a cycle to be completed in case of a power outage; a circuit protector should be included; the uninterrupted power supply and circuit protector must be integrated within the system	Continuous power is not always available at the level of a microscopy centre; therefore, a battery-operated device that can be recharged, possibly using solar power, would be ideal.	(8, 20)
Maintenance and calibration[a]	Preventative maintenance should not be needed until after 2 years or > 5000 samples; an alert should be included to indicate when maintenance is needed; should be able to be calibrated remotely, or no calibration should be needed	Preventative maintenance should not be needed until after 1 year or 1000 samples; an alert should be included to indicate when maintenance is needed; should be able to be calibrated remotely, or no calibration should be needed	A maintenance alert is necessary to ensure proper functioning in settings where it is unlikely that the same person will always handle the device and that records will be kept about the duration of use. Furthermore, it will be essential that only simple tools and minimal expertise are necessary to perform maintenance, given the quantity of devices that are likely to be used; additionally, service visits are unlikely to be feasible outside of urban settings.	(22, 31, 68)
Data analysis	Data analysis should be integrated into the device; a PC should not be required; exported data should be capable of being analysed on a separate or networked PC			
Result documentation, data display	An integrated results screen and the ability to save and print results should be included; the device should have a USB port	An integrated results screen and the ability to save results should be included; the device should have a USB port	Results should be simple to interpret (for example, positive versus negative for TB detection, or present versus absent for drug resistance). Information that would allow a more detailed interpretation of results should be available (for example, information on the mutations detected) for surveillance purposes or more differentiated clinical decision-making; however, it should be possible to hide this information if necessary.	

Regulatory requirements	Manufacturing of the assay and system should comply with ISO EN 13485 or higher standards or regulations, and comply with ISO IEC 62304 Medical Device Data Systems; the manufacturing facility should be certified and authorized for use by a regulatory authority that is a member of the International Medical Device Regulators Forum, formerly known as Global Harmonization Task Force; the assay must be registered for in vitro diagnostic use			
Data export (connectivity and interoperability)	All data should be able to be exported (including data on use of the device, error rates and rates of invalid tests, and personalized, protected results) over a USB port and network; network connectivity should be available through an Ethernet, Wi-Fi, or GSM/UMTS mobile broadband modem, or a combination of these; results should be encoded using a documented standard (such as HL7) and be formatted as JSON text; JSON data should be transmitted through HTTP(S) to a local or remote server as results are generated; results should be stored locally and queued during network interruptions to be sent as a batch when connectivity is restored	Integrated ability for all data to be exported from the device in a user-friendly format (including data on use of the device, error rates or rates of invalid tests, and non-personalized results) over a USB port	Mobile phone capacity is frequently available even at the level of microscopy centres. This could be leveraged for data export, quality control, supply-chain management and surveillance.	(8, 20, 53)

Electronics and software	Should be integrated into the instrument	Should be integrated into the instrument	If a separate PC is needed, it will likely limit the ability to update software, since staff with the skills needed to operate a PC are not present in all microscopy centres. Furthermore, security will be an issue, and separate PCs might be stolen.	
Operating temperature and humidity level[a]	Between +5 °C to +50 °C with 90% humidity	Between +5 °C and +40 °C with 70% humidity	High environmental temperatures and high humidity are often present in countries where TB is endemic. Dust also is a problem in these settings, and the need to adequately protect optics should be considered.	(20, 31)
Reagent kit – transport	No cold chain should be required; should be able to tolerate stress during transport for a minimum of 72 hours at -15 °C to +50 °C	No cold chain required; should be able to tolerate stress during transport for a minimum of 72 hours at -15 °C to +40 °C	Refrigerated transport is costly and often cannot be guaranteed for the entire transportation process. Frequent delays in transport are commonplace.	(8, 20)
Reagent kit – storage and stability	2 years at +5 °C to +40 °C with 90% humidity; should be able to tolerate stress during transport for a minimum of 72 hours at +50 °C; no cold chain should be required	12 months at +5 °C to +35 °C with 70% humidity; should be able to tolerate stress during transport for a minimum of 72 hours at +50 °C; no cold chain should be required	High environmental temperatures and high humidity are often present in countries where TB is endemic; they are especially problematic during the transport of reagents and systems.	(8, 20, 31)
Additional supplies (not included in kit)	None	None		
Internal quality control	Full controls for sample processing, amplification and detection of TB should be included			(22, 31, 68)
Training and education	6 work hours for staff at the level of a microscopy technician	3 days (or 24 work hours) for staff at the level of a laboratory technician		

DST, drug-susceptibility testing; HRZE, isoniazid , rifampicin, pyrazinamide, ethambutol; REMox, rifampicin, moxifloxacin, pyrazinamide, ethambutol; PaMZ, Pa824, moxifloxacin, pyrazinamide; PZA, pyrazinamide; RIF, rifampicin; PPV, positive predictive value; MDR-TB, multidrug-resistant TB; FQ, fluoroquinolone; MOX, moxifloxacin; INH, isoniazid; AG, aminoglycoside; CAP, capreomycin; LVX, levofloxacin; SNP, single nucleotide polymorphism; PCR, polymerase chain reaction; PC, personal computer; GSM, Global System for Mobile Communications; UMTS, Universal Mobile Telecommunications System.

a These characteristics were considered to be the most important.

Annex B. Participants

Participants from technical agencies, and researchers

Tope Adepoyibi
Senior Technical Officer, PATH
TB Team – TB/HIV Global Program
Washington, DC
USA
E-mail: tadepoyibi@path.org

Riccardo Alagna
Fondazione Centro San Raffaele
Milan
Italy
E-mail: alagna.riccardo@hsr.it

Heather Alexander
TB/OI Unit Lead
International Laboratory Branch
Division of Global HIV/AIDS
US Centers for Disease Control and Prevention (CDC)
Atlanta, GA
USA
E-mail: HAlexander1@cdc.gov

David Alland
Professor and Chief
Division of Infectious Disease
Rutgers New Jersey Medical School
Rutgers University
Newark, NJ
USA
E-mail: David.alland@rutgers.edu

Catharina Boehme
Foundation for Innovative New Diagnostics (FIND)
Geneva
Switzerland
E-mail: Catharina.boehme@finddiagnostics.org

Daniela Cirillo
New Diagnostics Working Group
Fondazione Centro San Raffaele
Milan
Italy
E-mail: cirillo.daniela@hsr.it

Frank Cobelens
Scientific Director
KNCV Tuberculosis Foundation
The Hague
Netherlands
E-mail: Frank.cobelens@kncvtbc.org

Sarah Cook-Scalise
Program Specialist, Market Access
TB Alliance
New York, NY
USA
E-mail: sarah.cook-scalise@tballiance.org

Claudia Denkinger
Foundation for Innovative New Diagnostics (FIND)
Geneva
Switzerland
E-mail: Claudia.denkinger@finddiagnostics.org

Gregory Dolganov
Senior Research Scientist
Division of Infectious Diseases
Stanford University Medical Center
Stanford, CA
USA
E-mail: gregoryd@stanford.edu

David Dolinger
Business and Technology Development Officer
Foundation for Innovative New Diagnostics (FIND)
Geneva
Switzerland
E-mail: david.dolinger@finddiagnostics.org

Kathleen England
TB Care I
PMU Laboratory Technical Advisor
KNCV Tuberculosis Foundation
The Hague
Netherlands
E-mail: kathleen.england@kncvtbc.org

Debra Hanna
Executive Director
CPTR Critical Path Institute
Tucson, AZ
USA
E-mail: dhanna@c-path.org

Sandra Kik
McGill International TB Centre and Department of Epidemiology and Biostatistics
Montreal Chest Institute
Montreal, QC
Canada
E-mail: Sandra.kik@mail.mcgill.ca

Ann Kolokathis
CPTR Critical Path Institute
Tuscon, AZ
USA
E-mail: kolokathis@gmail.com

Carl Mendel
Senior Vice President, R&D
TB Alliance
New York, NY
USA
E-mail: carl.mendel@tballiance.org

Madhukar Pai
Associate Professor
McGill University
Associate Director
McGill International TB Centre
Department of Epidemiology and Biostatistics
Montreal
Canada
E-mail: madhukar.pai@mcgill.ca

James E Posey
Lead Applied Research Team
Centers for Disease Control and Prevention (CDC)
Atlanta, GA
USA
E-mail: jposey@cdc.gov

Julien Reboud
Division of Biomedical Engineering
School of Engineering
University of Glasgow
Glasgow
Scotland
E-mail: Julien.Reboud@glasgow.ac.uk

Marco Schito
Senior Scientific Officer
NIH/CPTR Rapid TB-DST Consortium
Bethesda, MD
USA
E-mail: schitom@niaid.nih.gov

Gary K Schoolnik
Professor of Medicine, Microbiology and Immunology
Associate Director
Stanford Institute for Immunology, Transplantation and Infection
Stanford Medical School
Stanford, CA
USA
E-mail: gary.schoolnik@stanford.edu

Tom Shinnick
Associate Director of DTBE for Global Laboratory Activities
Centers for Disease Control and Prevention (CDC)
Atlanta, GA
USA
E-mail: tms1@cdc.gov

Alessandra Varga
Secretariat
New Diagnostics Working Group
Foundation for Innovative New Diagnostics (FIND)
Geneva
Switzerland
E-mail: Alessandra.varga@finddiagnostics.org

Dalene von Delft
TB PROOF
Cape Town
South Africa
E-mail: daleneduples@gmail.com

Participants from funding organizations

Jim Gallarda
Senior Program Officer
Bill and Melinda Gates Foundation Seattle, WA
USA
E-mail: Jim.gallarda@gatesfoundation.org

Janet Ginnard
Technical Officer
Market Dynamics
UNITAID
Geneva
Switzerland
E-mail: ginnardj@unitaid.who.int

Michael Kimerling
Senior Program Officer, Tuberculosis
Bill and Melinda Gates Foundation
Seattle
USA
E-mail: Michael.Kimerling@gatesfoundation.org

Regina B Osih
TB/HIV Adviser
Clinical Support Team
Clinton Health Access Initiative (CHAI)
South Africa
E-mail: Rosih@clintonhealthaccess.org

Amy Piatek
Senior TB Technical Advisor
USAID
Washington, DC
USA
E-mail: apiatek@usaid.gov

Namita Singh
Manager
Scientific New Diagnostics Project
Clinton Health Access Initiative (CHAI)
Haryana
India
E-mail: nsingh@clintonhealthaccess.org

Mark Ware
Manager
Diagnostic Services
Clinton Health Access Initiative (CHAI)
Edinburgh
Scotland
E-mail: mware@clintonhealthaccess.org

Participants from TB Supranational Reference Laboratories

Rumina Hasan
Department of Pathology and Microbiology
Aga Khan University
Karachi 74800
Pakistan
E-mail: rumina.hasan@aku.edu

Sven Hoffner
Director of SRL Stockholm
Public Health Agency of Sweden
Solna
Sweden
E-mail: sven.hoffner@folkhalsomyndigheten.se

Nazir Ismail
NHLS
Johannesburg
South Africa
E-mail: naziri@nicd.ac.za

Leen Rigouts
Laboratory Director
Mycobacteriology Unit
Institute of Tropical Medicine
Antwerp
Belgium
E-mail: lrigouts@itg.be

Sabine Rüsch-Gerdes
Head of the NRL for Mycobacteria
Borstel
Germany
E-mail: srueschg@fz-borstel.de

Implementers and clinicians

Martina Casenghi
Médecins Sans Frontières
Geneva
Switzerland
E-mail: martina.casenghi@geneva.msf.org

Charles Daley
Chief
Division of Mycobacterial and Respiratory Infections
National Jewish Health
Englewood, CO
USA
E-mail: daleyc@njhealth.org

Carole Jefferson
Consultant
New Freedom, PA
USA
E-mail: Cjefferson2520@gmail.com

Camilla Rodrigues
Consultant Microbiologist
Department of Microbiology
Hinduja Hospital
Mumbai
India
E-mail: dr_crodrigues@hindujahospital.com

Francis Varaine
Coordinator of MQSF TB WG
Médecins Sans Frontières
Medical Department
Paris
France
E-mail: fvaraine@msf.org

Participants from national TB programmes and countries with high burdens of TB and MDR-TB

Radhay S Gupta
Deputy Director General
Head, Central TB Division
Project Director RNTCP
Dte.GHS / MOHFW
Government of India
New Delhi
India
Email: ddgtb@rntcp.org

Lindiwe Mvusi
TB Control and Management
National Department of Health
Pretoria
South Africa
Email: MvusiL@health.gov.za

Elena Romancenco
Chief
National Reference Laboratory in TB
Institute of Phtihsiopneumology
Chisinau
Republic of Moldova
Email: eromancenco@yahoo.com

Alena Skrahina
Deputy Director
Republican Research and Practical
Centre for Pulmonology and TB
Minsk
Belarus
Email: alena_skrahina@tut.by

Ruy de Souza Junior
LAB Focal Point
National Tuberculosis Program
Planaltina - DF
Brazil
Email: ruy.souza@saude.gov.br

Sabira Tahseen
Technical Advisor
National TB Control Programme
National TB Reference Laboratory
Federal General Hospital
Islamabad
Pakistan
Email: Sabira.tahseen@gmail.com

Representatives from industry

Pia Azarschab
International Business Manager
Hain Lifescience GmbH
Nehren
Germany
E-mail: pia.azarschab@hain-lifescience.de

Martin Colla
Programme Manager – Asia
High Burden and Developing Countries
Cepheid
Maurens Scopont
France
E-mail: martin.colla@cepheidhbdc.com

Jean-François de Lavison
President and Founder
Ahimsa Partners SAS
Lyon
France
E-mail: jfdelavison@ahimsa-partners.com

David Hain
Chief Executive Officer
Hain Lifescience GmbH
Nehren
Germany
E-mail: david.hain@hain-lifescience.de

Glenn Johns
Director
Global Health Diagnostics
Ionian Technologies
San Diego, CA
USA
E-mail: gjohns@ionian-tech.com

Koné Kaniga
Senior Principal Scientist
Janssen Research and Development
Titusville, NJ
USA
E-mail: kkaniga@its.jnj.com

Shinichi Kojiya
General Manager
Business Development and Licensing
Molecular Diagnostics, Research
and Development Division
Eiken Chemical Co., Ltd.
Tochigi
Japan
E-mail: Shinichi_Kojiya@eiken.co.jp

Davide Manissero
Director Medical and Scientific Affairs
QuantiFERON, EMEA
QIAGEN Manchester Ltd.
Manchester
England
E-mail: Davide.manissero@qiagen.com

Yasuyoshi Mori
Manager
Section I, Department II, Biochemical Research
Laboratory II
Research and Development Division
Eiken Chemical Co., Ltd.
Tochigi
Japan
E-mail: Yasuyoshi_Mori@eiken.co.jp

Anita Suresh
Worldwide Marketing Manager
BD Diagnostics
Sparks, MA
USA
E-mail: Anita_Suresh@bd.com

Lawrence J Wangh
Professor
Department of Biology
Brandeis University
Waltham, MA
USA
E-mail: wangh@brandeis.edu

Advocate for patients

Colleen Daniels
Director TB/HIV
Treatment Action Group (TAG)
New York, NY
USA
E-mail: Colleen.daniels@treatmentactiongroup.org

WHO Headquarters staff

Anna Dean
TB Monitoring and Evaluation
WHO Global TB Programme
E-mail: deanan@who.int

Chris Gilpin
Laboratories, Diagnostics and Drug Resistance
WHO Global TB Programme
E-mail: gilpinc@who.int

Fuad Mirzayev
Laboratories, Diagnostics and Drug Resistance
WHO Global TB Programme
E-mail: mirzayevf@who.int

Francis Moussy
Public Health, Innovation and Intellectual Property
WHO Essential Medicines and Health Products
E-mail: moussyf@who.int

Karin Weyer
Laboratories, Diagnostics and Drug Resistance
WHO Global TB Programme
E-mail: weyerk@who.int

Diego Zallocco
Laboratories, Diagnostics and Drug Resistance
WHO Global TB Programme
E-mail: zalloccod@who.int

Matteo Zignol
TB Monitoring and Evaluation
WHO Global TB Programme
E-mail: zignolm@who.int

Annex C. Meeting agenda

Monday, 28th April 2014

Chairs: Karin Weyer (WHO) and Catharina Boehme (FIND); WHO Secretariat: C Gilpin

Session 1		
08:30 – 08:45	Registration	
08:45 – 09:00	Welcome, meeting scope and objectives	Karin Weyer/ Catharina Boehme
09:00 – 09:20	Overview: Need for TPPs, landscape of different TPPs, prioritization	Madhukar Pai
09:20 – 09:50	TPP biomarker detection test – including Delphi decisions/ outstanding questions	Sandra Kik
09:50 – 10:40	Discussion Moderator: Catharina Boehme Rapporteur: Martina Casenghi	
10:40 – 11:00	COFFEE BREAK	
11:00 – 11:30	TPP triage test – incl. Delphi decisions/outstanding questions	Claudia Denkinger
11:30 – 12:30	Discussions/Consensus building Moderator: Karin Weyer Rapporteur: Sandra Kik	
12:30 – 13:30	LUNCH BREAK	
13:30 – 14:00	TPP for a smear replacement test – including Delphi decisions/outstanding questions	Claudia Denkinger
14:00 – 14:45	Discussion Moderator: Martina Casenghi Rapporteur: Chris Gilpin	
14:45 – 15:00	Summary and next steps	Karin Weyer/ Catharina Boehme
15:00 – 15:20	COFFEE BREAK	

Session 2		
15:20 – 15:30	Introduction to the session and objectives in the context of The CPTR framework	Marco Schito
15:30 – 15:50	New regimens: the need for rapid Mox and PZA testing	Carl Mendel
15:50 – 16:10	TPP-relevant results from WHO-coordinated surveillance project on moxifloxacin and PZA resistance	Matteo Zignol
16:10 – 16:30	Review of results from the cost–effectiveness modelling studies	Frank Cobelens
16:30 – 16:50	Molecular detection of drug resistance: what can be achieved by current and pipeline technologies?	David Dolinger
16:50 – 17:30	Discussion	

Tuesday, 29th April 2014

Chairs: Daniela Cirillo (Co-Chair NDWG) and Tom Shinnick (Chair)

Session 3		
9:00 – 9:40	TPP for rapid DST - including Delphi decisions	Claudia Denkinger
09:40 – 9:55	Results from assay developer and laboratory user survey	David Dolinger
9:55 – 10:15	Priority medical needs for rapid DST: results from clinician survey	Daniela Cirillo
10:15 – 10:30	Cost – from affordability to feasibility	Martina Casenghi
10:30 – 10:50	COFFEE BREAK	
10:50 – 11:10	Genotypic versus phenotypic DST: state of knowledge	David Alland
11:10 – 11:30	Outstanding questions around TPP for DST	Martina Casenghi
11:30 – 12:30	Discussion Moderators: Daniela Cirillo and Tom Shinnick	
12:30 – 13:30	LUNCH BREAK	
13:30 – 15:00	Discussion continued Moderators: Daniela Cirillo and Tom Shinnick Rapporteur: Claudia Denkinger	
15:00 – 15:20	COFFEE BREAK	
15:20 – 16:00	Discussion regarding peripheral versus centralized molecular DST Moderators: Daniela Cirillo and Tom Shinnick	
16:00 – 17:00	Summary of consensus outcomes and next steps	Karin Weyer
	End of meeting	

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

New Diagnostics Working Group