World Health Organization

Protocol for the performance evaluation of nucleic acid tests for the diagnosis of tuberculosis (with or without resistance detection) for WHO prequalification assessment

Protocol for the performance evaluation of nucleic acid tests for the diagnosis of tuberculosis (with or without resistance detection) for WHO prequalification assessment

ISBN 978-92-4-009105-4 (electronic version)
ISBN 978-92-4-009106-1 (print version)

Suggested citation. Protocol for the performance evaluation of nucleic acid tests for the diagnosis of tuberculosis (with or without resistance detection) for WHO prequalification assessment. Geneva: World Health Organization; 2024. Licence: CC BY-NC-SA 3.0 IGO.

Cataloguing-in-Publication (CIP) data. CIP data are available at https://iris.who.int/.

Contents

1. Introduction

1.1 Prequalification of In Vitro Diagnostics

World Health Organization (WHO) prequalification of in vitro diagnostics (IVDs) is coordinated through the department of Regulation and Prequalification. Focus is placed on IVDs for priority diseases and their suitability for use in resource-limited settings.

WHO prequalification of IVDs is a comprehensive quality assessment of individual IVDs through a standardized procedure aimed at determining whether the product meets WHO prequalification requirements (*1*). Two types of prequalification assessment can take place, depending on the regulatory version submitted and evidence from a previous stringent review by a Recognized Stringent Regulatory Authority.

The full prequalification assessment process consists of the following components:

- review of a full product dossier
- performance evaluation including operational characteristics
- inspection of the manufacturing site(s)
- labelling review.

The abridged prequalification assessment consists of the following components:

- review of an abridged product dossier
- performance evaluation including operational characteristics
- inspection of the manufacturing site(s)
- labelling review.

This document applies to the performance evaluation component, which is conducted as part of prequalification regardless of whether full or abridged assessment applies.

The performance evaluation will be conducted by a Performance Evaluation Laboratory (PEL) following a choice of two different mechanisms described here[1]. Performance evaluations conducted by a laboratory in List 1 will be coordinated and cost covered by WHO. Performance evaluations conducted by a laboratory in List 2 will be coordinated and cost incurred by the manufacturer.

1.2 Performance evaluation of tuberculosis molecular assays for prequalification assessment

This protocol describes the procedures required to perform an evaluation of molecular tests for the qualitative detection of *M. tuberculosis* complex (MTBC) with or without detection of resistance submitted for WHO prequalification assessment. **This protocol is not intended to replace validation studies that need to be conducted by the manufacturer in order to fulfil WHO prequalification product dossier requirements.**

This performance evaluation aims at independently verifying performance characteristics claimed by the manufacturer. These characteristics include clinical performance (clinical sensitivity, clinical specificity) determined on prospectively collected or archived specimens in a laboratory environment; and analytical performance (limit of detection, reproducibility, inclusivity, exclusivity, resistance detection and risk of

[1] https://extranet.who.int/prequal/vitro-diagnostics/performance-evaluation

cross-contamination). In addition, operational characteristics and ease of use are assessed to inform use in settings with limited infrastructure.

Given the variety of molecular assays available, this protocol remains generic in nature and some sections may be open to interpretation. Manufacturers are encouraged to contact WHO before the start of the evaluation to discuss applicability of the protocol to a specific product.

This protocol was developed in collaboration with the WHO Global TB Programme.

2. Intended audience

This document is intended to provide PELs and manufacturers with the WHO performance evaluation procedure for prequalification assessment.

This protocol is not intended to be a master protocol for the evaluation of molecular tests for the diagnosis of tuberculosis outside of prequalification assessment.

3. Objectives

3.1 Overall Objectives

The overall objective of the performance evaluation is to evaluate analytical and diagnostic performance of commercially available tuberculosis molecular assays undergoing prequalification assessment and independently verify a subset of performance claims from the manufacturer.

3.2 Specific Objectives

The specific objectives of the evaluation are:

- to assess sensitivity and specificity for the detection of MTBC using clinical specimens by comparing with a designated reference standard;
- to assess analytical performance
    - reproducibility
    - limit of detection
    - resistance detection using a standardized panel of strains
    - inclusivity/exclusivity using a standardized panel of strains
    - cross-contamination or carry-over;
- to describe and assess the operational characteristics and ease of use of the assay under evaluation and its suitability for use in countries with limited infrastructure and wide temperature and humidity ranges).

4. Implementation of the evaluation

4.1 Performance Evaluation Laboratories (PEL)

The performance evaluation will be exclusively conducted by a PEL. These laboratories have successfully undergone assessment for WHO listing, described here[2], which includes the submission of an Expression

[2] https://extranet.who.int/prequal/vitro-diagnostics/performance-evaluation-laboratories

of Interest (EoI) form, a stage 1 audit of the laboratory (desktop review of EoI and specific quality management system (QMS) documentation), and a stage 2 (on-site) audit to assess compliance with WHO requirements.

The list of PELs can be accessed here[3].

The laboratory shall hold the following certification for quality management within the laboratory: ISO15189 (Medical laboratories: Particular requirements for quality and competence), or ISO17025 (General requirements for the competence of testing and calibration laboratories), or equivalent.

The person(s) listed as the primary contact in the list of PELs will act as the Principal Investigator (PI) for the work performed by the PEL.

4.2 Training, performance evaluation and supervision

The following issues are key to minimizing error and maximizing the value of this evaluation.

- Only personnel having received specific training for this evaluation and showing successful competency testing will be employed in the evaluation.
- Accurate record keeping is crucial to the success of the evaluation and the PI will be responsible for ensuring that all paper and/or electronic data required for the evaluation are recorded as agreed.
- Worksheets should be prepared and tubes, test devices or plates labelled prior to commencement of any run / assay.
- Because objective, machine-generated, permanent results for some of the technologies available may not be feasible, it is essential that the PI emphasizes the need for accurate record keeping.
- To minimize the risk of error, results will be directly exported from the platform wherever possible. If this is not the case, results should be entered by one staff member and verified by another.

4.3 Safety

All types of specimens must be handled as potentially infectious. Appropriate precautions to minimize infectious hazards will be taken at all stages from the collection of specimens to the disposal of used materials from the laboratory. For specimens from patients with presumed or confirmed tuberculosis, airborne precautions are required, which include, based on site- and sample-specific risk assessments and in line with national regulations, N95 masks and other appropriate personal protective equipment, certified class II biosafety cabinets, and a functional BSL3 level laboratory with a pass-through autoclave for hazardous waste management. The WHO biosafety manuals (*2, 3*) and the evaluating site's guidelines on laboratory safety will be carefully followed by the laboratory staff.

4.4 Storage of assays

All reagents must be stored as indicated in the instructions for use. Calibrated thermometers are placed at each location where reagents and specimens are stored, i.e. ambient, refrigerator and freezer. Temperatures for spaces and equipment are recorded daily on temperature logs, or automatically in a central temperature recording system, and monitored by the PI or quality officer delegated by the PI. The lot numbers of the test kits received/used and their expiry dates are recorded on the individual run worksheets. New lot validation testing is performed according to accreditation quality standard and laboratory procedures and documented.

[3] https://extranet.who.int/prequal/vitro-diagnostics/prequalified/performance-evaluation-laboratories

Two separate production lots (with different lot numbers and different expiry dates) will be requested for the performance evaluation, according to the following definition of a lot[4]: "The amount of material that is uniform in its properties and has been produced in one process or series of processes. The material can be either starting material, intermediate material or finished product." Each lot shall comprise different production (or manufacturing, purification, etc.) runs of critical reagents. Furthermore, lots must be sourced from a representative production run and not produced especially for the purpose of this evaluation. Manufacturers will be requested to provide lots with a minimum of six months shelf-life at the time of delivery to the laboratory.

5. Specimens

5.1 Clinical performance specimen panel

As much as possible, specimens used in this evaluation will be representative of the intended use population.

The panel for the clinical evaluation will consist of a total of 300 specimens of the main specimen type, usually sputum (Table 1), which could be collected in more than one site. For at least half of them, the test under evaluation should be performed on an unprocessed sputum specimen, either freshly collected or left over (frozen) specimens for which reference results are available, possibly from a paired specimen (see Figure 1). The other half of specimens may consist of concentrated sediments.

In addition, if the test includes a claim for other specimen type(s) (extrapulmonary specimens), a convenience sample size of additional specimen types- including 30 TB positive and 50 TB negatives per specimen type will be included in the panel and analyzed separately. These additional specimen types may be prioritized based on clinical relevance and prevalence.

Furthermore, for tests that include a claim for drug resistance detection, additional stored left-over specimens (sputum or concentrated sediments)[5] will be tested to reach a total of 50 drug resistant specimens and 100 drug susceptible specimens for each of the drugs claimed.

The panel may include specimens collected purposefully for this evaluation and/or left-over specimens submitted for routine testing, if these were collected and stored in agreement with the recommendations provided in the instructions for use (IFU) of the assay under evaluation, and provided the specimens were homogenized prior to storage.

PELs will aim at storing left-over specimens, for which reference results are available, prior to the evaluations in order to reach the target sample size timeously.

Ethical considerations are described in section 10.

[4] ISO 18113-1:2009 In vitro diagnostic medical devices -Information supplied by the manufacturer (labelling) – Part 1: Terms, definitions and general requirements

[5] Other specimen types may be considered according to the IFU.

5.1.1 Criteria for eligible specimens

Specimens will be eligible for the evaluation if they meet the following criteria.

Inclusion criteria:

- Age: 18 years and above[6];
- Patients with signs and symptoms of TB;
- Not on TB treatment or less than one week of treatment;
- Written informed consent or, where applicable for left-over specimens, waiver from informed consent (see section 10 for ethical considerations).

For freshly collected specimens:

- Two specimens per patient of at least 2 mL each, collected either on the same day (a minimum of 90 minutes apart) or next visit in the same diagnostic work-up;
- Prospectively collected (no selection based on molecular test or microscopy results to avoid selection bias).

For stored specimens:

- At least 2 mL of specimen or 1.5 mL of sediment;
- Previously characterized in agreement with 5.1.4, on the same specimen or a paired specimen collected within the same diagnostic work-up;
- Randomly selected among culture positive and culture negative specimens based on known culture results.

Additional drug resistant specimens:

- At least 2 mL of stored specimen or 1.5 mL of sediment;
- Previously characterized in agreement with 5.1.4, on the same specimen or a paired specimen collected within the same diagnostic work-up;
- Randomly selected among relevant drug-resistant and drug-susceptible known MTBC-positive stored specimens.

Exclusion criteria:

- Specimens with > 1 freeze-thaw cycle (or according to IFU, if specified);
- Any exclusion criteria stated in the product IFU.

5.1.2 Sample size

The panel for the main specimen type, usually sputum, will include 100 TB-positive and 200 TB-negative specimens (Table 1). The sample size was calculated to ensure a width of the 95% confidence interval of about 14% around a sensitivity of 85% and about 5% around a specificity of 98%.

In addition, for tests that include a claim for drug resistance testing, additional specimens will be included to complement the main sample size in order to reach 50 drug resistant specimens for each claimed drug. Multidrug resistant specimens may be included in these panels and can account for resistant specimens for multiple drugs.

[6] Paediatric specimens will not be included in the evaluation as it is often challenging to collect the required specimen. Most of pediatric specimens are smear negative. Stratification of the sample size by smear result will ensure that smear-negative specimens are represented in the evaluation.

Finally, if the test includes a claim for other specimen type(s) (extrapulmonary specimens), a convenience sample size of 30 TB positive and 50 TB negative specimens will be used, including at least one third of fresh specimens. The extrapulmonary specimen types may be prioritized based on clinical relevance and prevalence.

Table 1. Number of specimens required for the clinical evaluation

	Number of specimens (main specimen type - sputum)	Number of specimens per other specimen type[b]
Positive for TB *Including at least*[a] *Smear negative* *Unprocessed specimen*	**100** *25-30* *50*	**30**
Negative for TB *Including at least* *Unprocessed specimen*	**200** *100*	**50**
Additional specimens for drug resistance claims	Additional specimens to reach a total of at least 50 drug resistant and 100 drug susceptible specimens per drug claimed	

[a] These categories are not mutually exclusive

[b] Other specimen types may be prioritized based on clinical relevance and prevalence

5.1.3 Specimen collection and storage

Collection and storage of specimens will be carried out in accordance with the current version of the instructions for use (IFU). The laboratory will ensure that:

- appropriate specimen type and preservative, if applicable, is used;
- processing of specimens happens in a timely manner respecting manufacturer's claims for specimen stability;
- when using archived specimens, these have been processed and stored respecting the previous two points.

Decontaminated sputum sediments and frozen sputum may be used in addition to fresh specimens. **If these specimen types are not acceptable to be used with the assay under evaluation, the manufacturer is requested to inform WHO before the evaluation commences.**

5.1.4 Characterization of the clinical specimens

The clinical specimens will be characterized using fluorescence smear microscopy, liquid culture (MGIT) and speciation, and phenotypic drug-susceptibility testing (DST) and sequencing if applicable (see below and Figure 1). Liquid culture followed by speciation (AFB stain followed by MPT64 antigen test and LPA for if the antigen test is negative) will be considered as the reference method for the detection of *M. tuberculosis* complex and the results of smear microscopy will be used for descriptive stratification of

results. Specimens with a positive microscopy result and negative culture result will be excluded from the panel.

Phenotypic DST for, at least, the drugs claimed by the manufacturer will be performed for all culture-positive specimens. For rifampicin (RIF) resistance testing, targeted sequencing of the *rpoB* gene will be performed for all culture-positive specimens in addition to phenotypic DST, and sequencing results will be considered as the reference results for RIF resistance. For other drugs, the results of phenotypic DST will be considered as the reference standard and sequencing will be done for the resolution of discrepant results (see section 6.2.1).

In addition, the specimens will be tested on a WHO recommended nucleic acid amplification test for the detection of TB, which will be considered as the comparator (benchmark) test. This is not necessary for the additional specimens for drug resistance testing.

All procedures used for characterization of the specimens will be done following international recommendations (*4, 5*). All commercial tests will be performed according to the IFU and used within expiry date.

The algorithm for testing prospectively collected sputum is described in Figure 1. For extrapulmonary specimens, the reference method and all tests (smear microscopy, liquid culture, comparator assay and index test) will be performed on the same specimen.

For stored specimens (frozen specimens or sediments), prior characterization data (fluorescence microscopy, MGIT results and resistance profile if applicable) will be used if available, or characterization data can be completed during the evaluation. If results of the comparator test are available, these will be used. If they are not available, part of the specimen will be used to perform both the index and comparator tests.

An alternative reference method may be selected by the PEL with prior agreement from WHO.

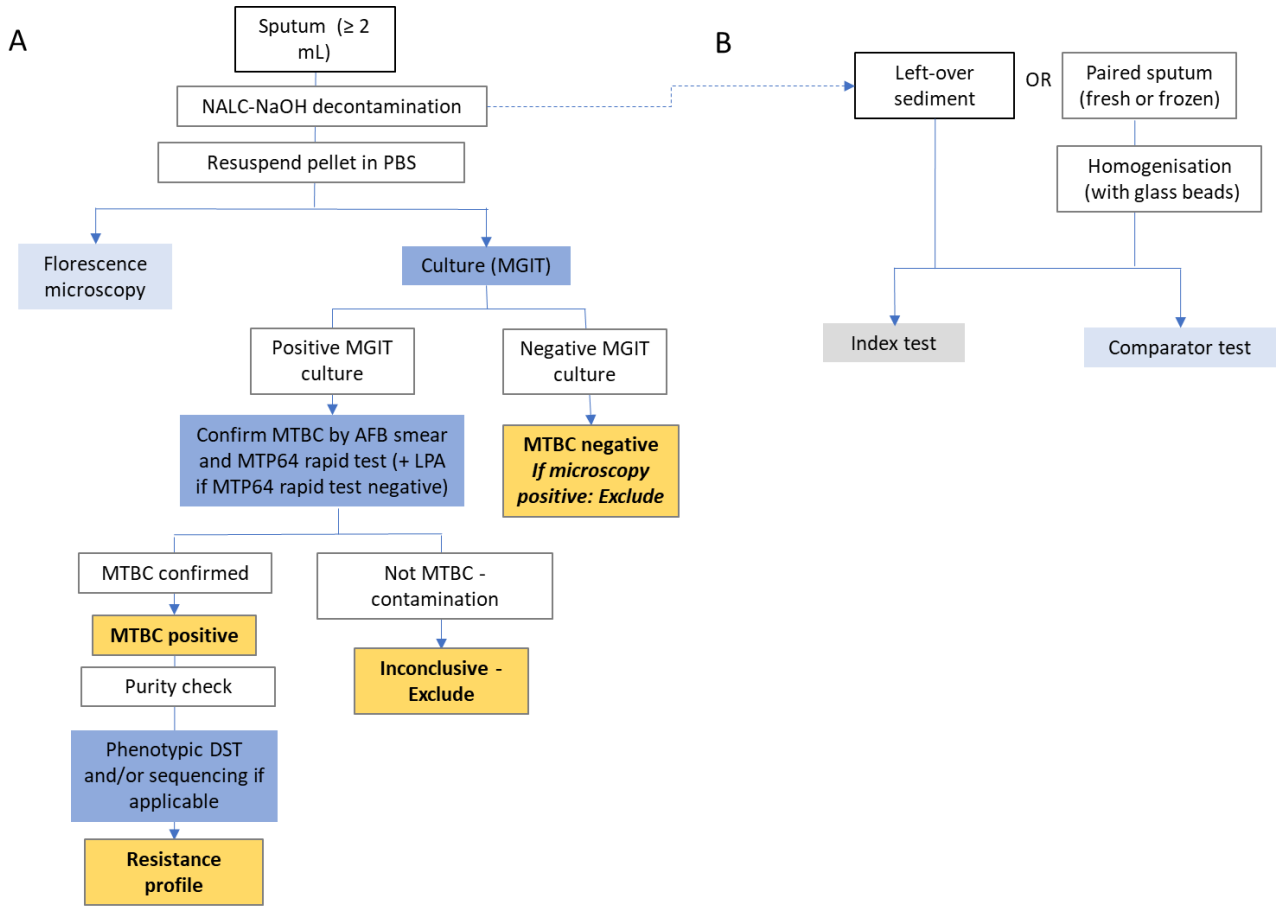

Figure 1. Characterization (A) and testing (B) algorithm

AFB: acid fast bacilli; MTBC: M. tuberculosis complex; DST: drug susceptibility testing

5.1.5 Data collection

The following information will be collected:

- specimen type
- collection date
- collection site (country)
- for specimens previously characterized: characterization data (results and dates of smear microscopy, MGIT culture, DST results, comparator assay results if available); number of previous freeze-thaw cycles
- if available and acceptable with ethics requirements
    - age
    - sex
    - previous history of TB treatment
    - HIV status.

5.2 Analytical performance panels

5.2.1 General description of the panels

The panels used for the analytical performance evaluation are described in the table below.

Table 2. Specimen requirements for analytical evaluation

Analyte / strain	Concentration	Number of replicates
Limit of Detection (LoD)		
WHO International Standard for *M. tuberculosis* (H37Rv) DNA for NAT-based assays[c]	0.5 log_{10}-fold serial dilutions with: • Two concentrations above the manufacturer's claimed LoD for MTBC detection • One concentration at the stated LoD for MTBC detection • Two concentrations below the LoD for MTBC detection	24 each= 120
If applicable, for assessment of LoD for drug resistance detection: WHO International Standard for *M. tuberculosis* (H37Rv) DNA for NAT-based assays[c, d]	0.5 log_{10}-fold serial dilutions with: • Two concentrations above the manufacturer's claimed LoD for drug resistance detection • One concentration at the stated LoD for drug resistance detection • Two concentrations below the LoD for drug resistance detection	24 each= 120
Reproducibility		
M. tb, sensitive (H37Rv)	Approx. 10^3 cfu/mL	40
M. tb, resistant[e]		40
Negative		40
Inclusivity, exclusivity		
M. bovis	Approx. 10^3 cfu/mL	3
M. africanum		3
M. avium	Approx. 10^5 cfu/mL	3
M. kansassii		3
M. intracellulare		3
Resistance detection (*6*), as applicable according to intended use statement		
RIF resistance: rpoB_S450L; rpoB_D435V; rpoB_H445Y; rpoB_H445D; rpoB_D435Y; rpoB_S450W; rpoB_L452P; rpoB_H445L; rpoB_S450F; rpoB_L430P; rpoB_H445R; rpoB_V170F; rpoB_I491F	Approx. 10^4 cfu/mL	3 each
INH resistance: katG_S315T; inhA_c-777t (fabG1_c-15t)	Approx. 10^4 cfu/mL	3 each
Cross-contamination / carry-over		
MTBC	≥ 10^5 cfu/mL	20
Negative specimens	0	20

This panel may be adapted on a case-by-case basis according to the manufacturer's claims

[c] https://www.nibsc.org/documents/ifu/20-152.pdf

[d] *Although the WHO International Standard is prepared from a drug susceptible strain, it will be used to assess the capacity of the test to provide a conclusive result for drug resistance*

[e] *Well-characterized resistant strain with the most common mutations for the relevant drugs*

5.2.2 Specimen preparation

For the assessment of the limit of detection, the WHO International Standard will be diluted into MTBC negative or artificial sputum to the concentration indicated in Table 2.

The strains used for assessment of reproducibility, inclusivity/exclusivity, resistance detection and carry-over may be commercially acquired or locally prepared, well-characterized (at least by phenotypic DST and sequencing) strains. For assessment of reproducibility, the strains will be diluted into MTBC negative or artificial sputum to the concentration indicated in Table 2. For inclusivity/exclusivity, resistance detection and cross-contamination, mycobacterial strains may be diluted into 7H9 medium at the concentrations indicated in Table 2. Concentrations (cfu/mL) presented in Table 2 will be estimated by adjusting the bacterial suspension density to the McFarland standards.

Note: if the manufacturer has an objection to the use of strains diluted into 7H9 medium, the manufacturer is requested to inform WHO as soon as possible and in any case before the evaluation commences.

6. Laboratory testing

6.1 Review of the instructions for use

Each product under evaluation will be used in accordance with the IFU issued by the manufacturer. The evaluating site will send a copy of the IFU to WHO upon delivery of the reagents and prior to the commencement of the laboratory evaluation. The IFU must be reviewed against the IFU submitted to WHO as part of the application or pre-submission. If the IFU has been updated since this time, it is the onus of the manufacturer to submit to WHO a letter detailing changes made prior to the start of the laboratory evaluation. Records of the version used must be kept.

6.2 DNA extraction

Wherever DNA extraction is integrated and automated, or where only one extraction method is recommended for use with the product, this method will be used.

Where several extraction methods are validated for use with the product, one extraction method (per specimen type, if applicable) will be used in this evaluation. If a simple method that does not necessitate purchase of an extraction kit is validated, then this method will be used by default (e.g. use of thermal lysates).

6.3 Clinical performance

Each specimen from the clinical performance panel will be tested once on the assay under evaluation.

The test will be repeated once if the initial test is invalid, as defined in the IFU, or any types of test failure, if sufficient specimen is available. If the test cannot be repeated, or if the repeat result is still invalid, the result will be recorded for the analysis of invalid rate, but an additional specimen will be included to reach the sample size.

Specimens will be randomized and blinded so that the laboratory staff performing the assay under evaluation are not aware of the reference result. Blinding is not considered critical if the result is generated by the instrument and not subjectively read.

6.3.1 Resolution of discrepant results

Specimens with results that are consistent with the reference testing results undergo no further testing. Discrepant results, defined as results that do not agree with results obtained with the reference assay, will be retested by the same operator on the assay under evaluation if sufficient specimen is available.

Specimens showing discrepant resistance results between the test under evaluation and DST will be further characterized by next generation sequencing.

6.4 Analytical performance

6.4.1 Limit of detection

Each dilution of the WHO International Standard, as described in Table 1, will be tested 24 times. The 24 replicates will be performed over at least three days by at least two users and, for low-throughput instruments, on at least three different instruments, or sets of instruments if applicable (e.g., DNA preparation and amplification instruments). The WHO International Standard for *M. tuberculosis* for nucleic acid amplification techniques will be used for this purpose (*7*).

At least 20 valid results must be obtained for each concentration.

The dilutions used may be adapted from the concentrations proposed in Table 2, to ensure having at least one concentration with a hit rate above 95% and 2 concentrations with hit rates between 10% and 90%.

For tests that include a claim for drug resistance testing, the LoD for drug resistance, considered as the capacity of the test to provide a conclusive drug resistance (ie. drug resistance detected or not detected) will be assessed. As the claimed LoD for MTB detection and drug resistant detection may differ, this could require a different set of concentrations, as described in Table 2. Wherever concentrations overlap, the same dataset may be used for both assessments.

6.4.2 Precision

Three specimens will be used (see Table 2). Within-run and within-laboratory reproducibility will be assessed by measuring, at a minimum, eight replicates of these three specimens in the same run over five different days, performed by at least two different users, using at least two different reagent lots and, for low-throughput instruments, on at least three different instruments.

If there are two or more invalid results for the same specimen in the same run, then the run should be repeated for this specimen. Invalid results should be reported.

6.4.3 Inclusivity and exclusivity

Representative MTBC and non-tuberculosis mycobacteria (NTM) strains will be tested in triplicate for inclusivity and exclusivity verification.

6.4.4 Resistance detection

For assays with a claim for detection of drug resistance, the applicable specimens from the resistance detection panel (Table 2) will be tested in triplicates.

6.4.5 Cross-contamination / carry-over

The experiment will allow the determination of the well-to-well or vial-to-vial cross-contamination rate of high throughput platforms or potential carry over in low throughput instruments. This will be assessed by

alternating one high positive with one negative specimen, as described in Table 2, and repeating this sequence twenty times. For high-throughput assays, this will be done by alternating the high positive and negative specimens in the same plate/run. For low-throughput assays, each sequence of high positive followed by negative specimen should be done on the same instrument. If more than one instrument is used, each run (i.e same instrument and same day) should include a minimum of two sets of alternating high-positive and negative specimens.

6.5 Interpretation and recording of test results

The interpretation of results for each assay under evaluation is made strictly according to the manufacturer's instructions within the IFU. Invalid runs and/or test results are recorded on the data collection sheets as per good documentation practice.

Wherever possible, all test results are saved and exported directly from the instrument to standardized test result worksheets in Microsoft Excel spreadsheets (see Section 13 Other documents required) for further data analysis.

Prior to analyses, the databases will be checked, either by comparison of double data entry (with Excel Pro) or by comparing the printout of the Excel spreadsheet with the digital or print-out record, and successful check completion documented.

7. Quality control

7.1 Competency panels

A competency panel of routine specimens comprised of at least five well-characterized specimens including positive and negative specimens must be run successfully for each assay by each operator before the evaluation commences. This may be the same panel as that used at the time of assay demonstration by the manufacturer or for training purposes. They may be left-over from EQA panels.

7.2 Internal quality control

Internal procedural controls should be incorporated into the design of most assays by the manufacturer. These may take the form of extraction, and/or amplification, and/or detection controls, as indicated in the IFU. Any internal quality control must be valid as per manufacturer's instructions.

7.3 Test kit controls

Manufacturer-supplied positive and negative test kit controls will be run at the frequency indicated in the IFU, in each test run when applicable and at the start of each testing session when applicable. Where positive and negative test kit controls are not supplied by the manufacturer, the external quality control specimen will act at the control specimen.

7.4 External quality control specimen

The evaluating site will supply a previously validated external quality control (QC) specimen which is tested in single for each test run or once at the start of the day per instrument for single use devices. The QC specimens include a low positive specimen and negative specimen. The QC specimen will be made by the evaluating site or acquired commercially and validated by the site.

7.5 Limits of acceptability

All results on test kit controls and the QC specimen are documented. Should the QC sample not give the expected result, the run will be considered invalid, in which case the run will be repeated, and troubleshooting should occur for instruments using single use devices. Such problems should be recorded on the data sheets. The PI will be responsible for carefully checking all data entry forms for legibility, accuracy, and completeness.

8. Analysis of data

8.1 Invalid runs and invalid individual results

The number of invalid test runs (i.e., where positive and/or negative kit or external controls do not give the expected result) will be recorded as the absolute number of invalid runs and as a percentage of the total number of runs performed for the entire evaluation using all specimens.

The number of individual invalid specimen results, as well as any other types of readings (e.g., error) indicating a failure preventing from providing a result, depending on the platform under evaluation, will also be recorded. They will be presented, separately for invalid results and errors, as a percentage of the total number of specimens tested for the entire evaluation, as well as separately for the clinical and analytical part of the evaluation.

In addition, when applicable, invalid results and errors will be classified by cause (e.g., operator, specimen, instrument-related or assay-related, error codes, etc.).

If results on both processed and unprocessed sputum specimens are available, then the proportion of invalid results and errors will also be presented separately.

8.2 Clinical Performance

8.2.1 Sensitivity and specificity for the detection of M. tuberculosis complex

For the estimation of sensitivity and specificity of the assay under evaluation for the detection of *M. tuberculosis* complex, the results of the assay under evaluation will be compared to the results of the reference method on the main sample (100 MTBC-positive and 200 MTBC-negative specimens).
Estimation of sensitivity and specificity will be conducted separately for each specimen type, if applicable.

If results are available for both processed and unprocessed specimens are available, these will also be described separately. However, the sensitivity and specificity analysis will be performed on pooled data.

Table 3. 2x2 table for calculation of sensitivity and specificity for detection of MTBC

		Results of reference testing		
		+	–	
Results of assay under evaluation	+	**a** **True positives**	b False positives	a+b
	–	c False negatives	**d** **True negatives**	c+d
		a+c	b+d	

Sensitivity will be calculated as the proportion of true positive results detected by the assay under evaluation compared to all positives by the reference method (see Table 3) and expressed as a percentage.

$$\boldsymbol{Sensitivity} = \frac{\boldsymbol{a}}{\boldsymbol{a+c}}$$

Specificity will be calculated as the proportion of true negative specimens identified by the index method compared to all negatives by the reference method (see Table 3). Specificity will be expressed as a percentage.

$$\boldsymbol{Specificity} = \frac{\boldsymbol{d}}{\boldsymbol{b+d}}$$

Sensitivity and specificity will be estimated with their exact binomial 95% confidence intervals. The 95% confidence intervals are calculated in order to assess the level of uncertainty introduced by sample size.

8.2.2 Discrepant results

The initial and re-testing results of specimens with discrepant results (as defined in section 6.2) will be described in the report. However, only the initial result will be used for calculations described above.

8.2.3 Description of results stratified by smear microscopy result

In addition to the calculation of clinical sensitivity and specificity overall, the results will be described after stratification by smear microscopy result, i.e., the proportion of specimens detected among smear-negative and smear-positive culture-positive specimens.

8.2.4 Comparison of clinical sensitivity and specificity of the assay under evaluation and the comparator method

The results of the assay under evaluation will also be compared to those obtained with the comparator method, using two 2x2 tables presented below (Table 4): one comparing the results among true positive specimens (to compare sensitivity) and the other comparing the results among true negative specimens (to compare specificity).

Table 4. 2x2 table for comparison with comparator method

		Results of comparator method		
		+	–	
Results of assay under evaluation	+	**a**	b	a+b
	–	c	**d**	c+d
		a+c	b+d	

The difference in sensitivity and specificity with the comparator method and 95% confidence interval will be estimated using the Tango score method (*8*).

8.2.5 Sensitivity and specificity for drug resistance, if applicable

For the estimation of sensitivity and specificity of the assay under evaluation for the detection drug resistance, the results of the assay under evaluation will be compared to the results of the reference method (*i.e.* phenotypic or genotypic drug resistance results, as described in 5.1.4) using the 2x2 table presented below (Table 5). The samples for this assessment will include all MTBC-positive specimens (100 MTBC-positive specimens and additional drug resistant and susceptible specimens)

Estimation of sensitivity and specificity will be conducted separately for each drug, if applicable.

For each drug, the proportion of indeterminate/inconclusive resistance result will be described. Indeterminate resistance results will be excluded for the calculation of sensitivity and specificity.

Table 5. 2x2 table for calculation of sensitivity and specificity for drug resistance

		Results of reference testing		
		Resistant	Susceptible	
Results of assay under evaluation	Resistant	**a** **True resistant**	b False resistant	a+b
	Susceptible	c False susceptible	**d** **True susceptible**	c+d
		a+c	b+d	

8.3 Analytical performance

8.3.1 Limit of detection

The limit of detection (LoD) is the lowest concentration of analyte that can be consistently detected in 95% of specimens tested under routine laboratory conditions and in a given specimen matrix. It defines the analytical sensitivity.

The concentrations will be reported in IU/mL, as calculated directly from the assigned value of the WHO International Standard. The detection rate at each concentration will be reported. For the LoD for drug resistance, the detection rate will reflect the proportion of tests with a conclusive drug resistance result. The LoD will be estimated using the Probit analysis on the results of the 24 replicates of the LoD panel.

8.3.2 Precision

Hit rates (i.e., the proportion of expected results) will be calculated for each specimen.

In addition, if Ct values are provided, the standard deviation and CV% will be calculated.

8.3.3 Inclusivity and exclusivity

The number and proportion of MTBC and NTM correctly detected or not detected, according to the claim, will be described.

8.3.4 Resistance detection

The number and proportion of resistance-conferring mutations consistently detected will be described for the panel of strains described in Table 2.

8.3.5 Cross-contamination or carry-over

If applicable, the cross-contamination (or carry-over) experiment will allow the determination of the well-to-well, vial-to-vial or device-to-device cross-contamination rate of the platform. The proportion of false-positive results among the 20 negative specimens tested will be reported.

9. Technician's appraisal

The operational characteristics and ease of use of the assay under evaluation will be assessed by the technician(s) who performed the testing according to a standardized form (Annex 1). These assessments, along with other selected assay characteristics, contribute to an overall appraisal of each assay's suitability for use in small laboratories. Special attention should be paid to the IFU in order to evaluate whether these instructions are sufficient for WHO Member State end-users. Comments on the IFU must be made in the report if it does not meet an acceptable standard for any of the following criteria: clarity, presentation, content, safety instructions.

10. Ethical considerations

10.1 Compliance with International Standards

This protocol was submitted for review to the World Health Organization (WHO) Ethical Review Committee for ethical clearance and was approved in February 2023 (ERC.0003891). Amendments to the protocol was also submitted for review. When required by local regulations, this protocol will be submitted to the national and institutional ethical review boards of the PELs. Any substantial change to the protocol must be approved by all the bodies that have approved the initial protocol, prior to being implemented, unless it is due to participant's safety concerns. The evaluation will be carried out according to the principles stated in the Declaration of Helsinki as amended in 2013 and any further updates, all applicable national and international regulations and according to the most recent *International Conference on Harmonization of technical requirements for registration of pharmaceuticals for human use – Good Clinical Practice* (ICH-GCP) guidelines.

10.2 Specimen collection

As stated in section 4, specimens used in this evaluation may belong to the following categories:

10.2.1 Left-over specimens

Residual specimens from routine patient care may be used in the evaluation. These specimens may be part of an already existing panel of specimens (e.g. sputum sediments), or freshly collected. Prior to collection, participants are made aware of the potential secondary use of their specimens through written informed consent according to local regulations. Alternatively, informed consent can be waived by local ethics committees when specific conditions are met (see 10.3.1). In some cases, left-over specimens from research studies can be used, provided the original study-protocol and informed consent clearly outlined the secondary use of specimens for this purpose.

10.2.2 Purposeful collection of new specimens

Specimens may be collected from participants, specifically for the purpose of this evaluation (i.e., not as part of ongoing routine clinical practice). Participants will be asked to provide written informed consent (Annex 2) before collection of the specimen.

10.2.3 Commercially acquired specimens

Commercially available panels or pooled human specimens (e.g., sputum) intended for dilutions can be requested. These panels are presumed to be collected in an ethical manner and no additional informed consent will be sought.

10.2.4 Non-clinical specimens

The evaluation may use cultured mycobacteria. No additional informed consent will be sought for validation of tests using non-human specimens or isolates. These specimens can be derived from routine clinical isolates, provided they cannot be traced back to the original host and were acquired in an ethical manner. Stock specimens may need to be diluted in pathogen-negative human sputum or other specimen types. Informed consent may still apply for these specimens, depending on their origins.

10.3 Informed Consent

10.3.1 Left-over clinical specimens

Written informed consent for secondary use of left-over specimens will be acquired from participants prior to collecting the specimen, except if waivers conditions apply (see below). The participant, or legal guardian for infants, will be provided with general information about the secondary use of their specimen, both on paper and by staff. In case of illiteracy, the information will be read to the participant and a fingerprint will be used as a signature.

The evaluating laboratories are allowed to use their own version of informed consent forms (ICFs). As per good ethical practice, it is advisable that ICFs contain at least the following information in clear understandable language respective to the target audience: *the possibility of specimen storage and secondary use for specified purposes; that participation is voluntary and the participant is free to withdraw without consequences; what (if any) compensation the participant will receive and any other benefits related to the participant; any foreseeable risks involved in participating; provisions made to respect and preserve the participants privacy and confidentiality; that the protocol/ICF was reviewed by relevant ethical bodies.*

When using left-over specimens from routine care for non-research purposes, the need for written informed consent can be waived by local ethical review boards, provided one of the following conditions are met:

<u>Presumed consent</u>

Prior to collection, specimen donors are made aware of the potential secondary use of their specimen for research purposes. This may be done through several channels, including: clearly visible pamphlets and posters at the collection site, available information on the institute's website and personal communication through the treating physician or nurse. The participant will be made aware of the right to refuse and opt out without any consequence for the quality of care.

Anonymization

The evaluating laboratories may choose to anonymize data wherever reasonably possible, thus removing any chance of reidentification and maximizing confidentiality. It is possible for laboratories to obtain a waiver for local ethical review when using fully anonymized left-over samples from routine practice for non-research purposes. It is the responsibility of the PEL to check the requirements for ethical and regulatory approval with their own institution.

10.3.2 Purposeful collection of new specimens

Where purposeful collection of new specimens is used, written informed consent will be acquired from participants prior to collecting the specimen. The written informed consent is provided in Annex 2 and will be translated into local language(s) by the PEL. The participant will be provided with all necessary information on the purpose of the evaluation, both on paper and by staff. In case of illiteracy, the information will be read to the participant in the presence of a witness and the fingerprint of the participant together with the witness's signature will be recorded. Two copies of the informed consent form will be completed, one will be kept by the staff at the evaluating laboratory and one by the participant.

10.4 Risk-Benefit assessment

There will be no direct individual benefits to the participants providing specimens. The results obtained during this evaluation will be used as part of WHO prequalification assessment. There will thus be broader benefits to people living with tuberculosis by contributing to the selection of well-performing tuberculosis molecular tests.

There is a minor risk of breach of confidentiality. This will be minimized by collecting as little personal information as is necessary and ensuring proper data protection according to general data protection regulations.

10.5 Storage of data and specimens

10.5.1 Confidentiality

Alongside collection of the specimen, the following information will be collected for all specimens: specimen type, collection date, collection site and characterization data, if applicable. In addition, the following information may be collected if acceptable according to local ethics requirements: age, gender, TB treatment history, and HIV status. All non-essential personal identifiers (including direct identifiers such as name, address, etc.) will be removed and specimens will receive a unique identification number at the collection site. In short, all collected data will be pseudonymized and transformed in order to preserve participant's privacy. When a link with the original identifiable information is no longer required, this link will be destroyed, and data will be anonymized.

10.5.2 Data storage

Personal data will be handled and stored according to the European general data protection regulations (GDPR) or local alternatives where applicable. Any documents containing the names and/or signatures of participants (e.g. consent forms) will be kept separately from all other evaluation documents containing participant data. All evaluation documents will be stored in lockable rooms or cabinets with access limited to evaluation staff.

Names of the participants will not appear on any reports or publications resulting from this evaluation.

After the evaluation, all source data, data analysis records and all correspondence will be retained at the testing laboratory for five years. Anonymized specimens may be stored according to local policy.

10.6 Results and incidental findings policy

The evaluation will not interfere with the clinical care the participant would normally receive. None of the results generated in this evaluation from evaluated assays will be used as a replacement for the gold-standard tests currently in use by the reference (or diagnostic) laboratory. As the purpose is to evaluate new assays, incidental findings are unlikely to occur. Should such a finding of significant clinical importance occur, the subsequent actions (e.g. feedback to the patient) shall be considered on a case-by-case basis by the evaluating laboratory and its medical staff according to their incidental findings policy. Where anonymized left-over specimens are used, incidental findings cannot be communicated to the patients.

11. Report preparation and dissemination

The preliminary data analysis and drafting of the report will be carried out by the PEL according to pre-defined report templates (see Section 14, Other documents required).

For evaluations coordinated by WHO (option 1), the draft report will be shared with the WHO PQ team. WHO will verify the data analysis and review the draft report and send the approved draft report to the authorized contact designated by the manufacturer for comment.

For evaluations commissioned by the manufacturer (option 2), the data and initial draft report will be shared simultaneously with the WHO PQ team and the manufacturer in copy. In both cases, WHO will verify data analysis and review the draft report and send the approved draft report to the authorized contact designated by the manufacturer for comment.

Manufacturers will have one month right of reply after WHO has officially shared the draft report. The final report will be prepared by WHO considering the manufacturer's comments. If the manufacturer has not provided any comment after one month, the draft report shared with the manufacturer will be considered the final report. A copy of the final report will be sent to the authorized contact designated by the manufacturer and to the PEL.

If the assay under evaluation successfully meets all WHO prequalification requirements and is prequalified, a summary of these data will be published in the WHO Public Report for the prequalification assessment of the assay.

The PEL(s) conducting the evaluation will inform local and national authorities of the evaluation through seminar/dissemination and will share the WHO Public Report in case of positive outcome of the prequalification assessment.

WHO reserves the right to publish the results of the evaluation, irrespective of the outcome. In this case, WHO will share the manuscript with the manufacturer for comments at least 30 days before submission, but WHO will have ultimate authority over the version submitted. Authors will include contributors from WHO and the PEL.

Any publication by WHO of the results of these evaluations and the WHO recommendations derived therefrom will, however, be accompanied by the following disclaimer:

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

WHO and the Performance Evaluation Laboratory, do not warrant or represent that the evaluations conducted with the tuberculosis test kits referred to in this document are accurate, complete and/or error-free. WHO and the Performance Evaluation Laboratory disclaim all responsibility for any use made of the data contained herein, and shall not be liable for any damages incurred as a result of its use. This document must not be used in conjunction with commercial or promotional purposes.

12. Materials and supplies

Manufacturers will provide the products and any equipment necessary for training, competency assessment, quality control and the evaluation free of charge.

An indicative number of tests to be performed for this evaluation for one specimen type is indicated in Table 6 below. This number will be increased if additional specimen types are assessed.

Table 6. Number of tests required for the evaluation

	Minimum number of tests required for the evaluation*
Clinical performance	400
LoD	120
Precision	120
Inclusivity, exclusivity	15
Resistance panel	45
Cross-contamination/carry-over	40
Total	740
Total + ~25% (for training, controls and possible repeats)	925

** The number of tests may vary depending on the claims (additional specimen types or drug resistance)*

13. Roles and responsibilities

13.1 Responsibilities of the Performance Evaluation Laboratory

i. If required by national authorities, obtaining ethical clearance for the evaluation;
ii. Ensure availability and maintenance of all specimen panels necessary for the evaluation;
iii. Conducting the performance evaluation in accordance with this protocol and good laboratory practice;
iv. Informing WHO of any unforeseen event that could have an impact on the evaluation or its timeline;
v. Preparation of draft report v1 of the laboratory evaluation;
vi. Advising WHO on operational characteristics of assays evaluated;
vii. Archiving all source data, data analysis records and all correspondence for a period of at least ten years.

13.2 Responsibilities of WHO- Prequalification of In Vitro Diagnostics Team

i. Technical advice to the PI;
ii. Technical and administrative management of the laboratory evaluation (option 1);
iii. Verification of analysis and draft report;
iv. Communication of the approved draft report to manufacturer and seeking comments from the manufacturer;
v. Preparation and dissemination of the final report;
vi. Formal contacts with the manufacturers.

13.3 Responsibilities of the manufacturer

i. Reviewing the protocol and contacting WHO in case of questions about applicability of the protocol to the specific product;
ii. Providing the appropriate number of test kits and instruments free-of-charge for the evaluation;
iii. Ensuring that kits and instruments are shipped under appropriate conditions and in time for the commencement of the evaluation;
iv. Providing training on the use of the instrument(s) and assay, in agreement with WHO;
v. Providing comments to WHO on the draft performance evaluation report within one month;
vi. For option 2 evaluations, selection of the PEL, agreement on terms and conditions of the evaluation in line with the conditions set forth in the Letter of Agreement with WHO, and funding of the evaluation.

14. Other documents and tools required

Report Templates

IVD/TP/4/P23a Template report for the performance evaluation of nucleic acid tests for the diagnosis of tuberculosis

Data entry spreadsheet

IVD/TP/4/P23b Template data entry spreadsheet for performance evaluation of nucleic acid test for the diagnosis of tuberculosis

15. References

1. World Health Organization. Overview of the WHO Prequalification of in Vitro Diagnostics Assessment. Version 8. [Internet]. Geneva: World Health Organization; 2018. Available from: https://apps.who.int/iris/bitstream/handle/10665/259403/WHO-EMP-RHT-PQT-2017.02-eng.pdf;jsessionid=387D24F53BDCAA88F4DBA8BFD1B8DBF1?sequence=1

2. World Health Organization. Laboratory biosafety manual, 4th edition [Internet]. Geneva: World Health Organization; 2020. Available from: https://www.who.int/publications/i/item/9789240011311

3. World Health Organization. Tuberculosis Laboratory Biosafety Manual [Internet]. World Health Organization. Geneva: World Health Organization; 2012. p. 1–60. Available from: https://apps.who.int/iris/bitstream/handle/10665/77949/9789241504638_eng.pdf?sequence=1

4. The Global Laboratory Initiative a Working Group of the Stop TB Partnership. Mycobacteriology Laboratory Manual [Internet]. 2014 [cited 2022 Dec 15]. Available from: https://stoptb.org/wg/gli/assets/documents/gli_mycobacteriology_lab_manual_web.pdf

5. World Health Organization. WHO operational handbook on tuberculosis. Module 3: Diagnosis - Rapid diagnostics for tuberculosis detection 2021 update [Internet]. Module 3: Diagnosis Rapid diagnotics for tuberculosis diagnosis. 2021. 176 p. Available from: https://apps.who.int/iris/bitstream/handle/10665/340256/9789240022614-eng.pdf

6. World Health Organization. Catalogue of mutations in Mycobacterium tuberculosis complex and their association with drug resistance [Internet]. Geneva; [cited 2022 Aug 31]. Available from: https://www.who.int/publications/i/item/9789240028173

7. Ho MM. The development and establishment of a heat inactivated preparation of Mycobacterium tuberculosis (H37Rv) as the first international standard for nucleic acid amplification techniques. Biologicals. 2023 May;82:101667. doi: 10.1016/j.biologicals.2023.101667. Epub 2023 Mar 31. PMID: 37004276.

8. Tango T. Equivalence test and confidence interval for the difference in proportions for the paired-sample design. Stat Med. 1998;17(8):891–908.

16. Document revision history

Version	Date	Reason for revision / summary of changes	Prepared by
1.0	16 December 2022	For WHO ERC submission	AL Page, N Ismail, CM Nathanson, A Korobitsyn, S Schumacher External reviewers: P Hall, SV Omar, S Tahseen, M Ruhwald, M Kohli, L Rigouts, A Chua, H Nguyen
1.1	27 January 2023	Clarification requested by WHO ERC: add name and contact details of institution and IRB on ICF. Minor editing.	AL Page
2.0	29 September 2023	*Note: previous versions of the protocol were not used for PQ evaluations.* Addition of assessment of sensitivity and specificity for drug resistance: additional sample size (for a total of 50 drug resistant and 100 drug susceptible specimens for each drug), amendment of section	AL Page / E Tagliani

		8.2.5 on analysis of drug resistance results. Removal of LoD assessment using a drug resistant strain and addition of concentrations for assessment of LoD for drug resistance detection (Table 2). Change in concentrations for analytical panel (Table 2) from concentrations expressed in relation to LoD to concentrations in cfu/mL. Addition that strains for inclusivity/exclusivity, resistance detection and cross-contamination will be diluted in 7H9 medium. Addition of section on DNA extraction (6.2). Addition that protocol was approved by WHO ERC. Change to new PQT/IVD document reference numbers. Minor editing.	
2.1	March 2024	Addition ISBN and copyright page. Minor editing	AL Page / JF Flandin

17. Annexes

17.1 Annex 1. Operational characteristics and ease of use

Indicate name of instruments assessed (e.g., extraction and amplification units)

Instrument 1:		Number of units used for the evaluation	
Instrument 2 (if applicable):		Number of units used for the evaluation	

If only 1 instrument, then indicate N/A for instrument 2 and all questions on instrument 2 below
If several units of the same instruments were used for the evaluation, these should be considered as only one instrument.

Table A1.1 Operational characteristics

1	**Assay characteristics**		
1.1	Need to reconstitute reagents		Yes/No
1.2	Total number of steps for one specimen*		*Each action required to obtain a result for one specimen (excluding specimen collection, instrument management, maintenance/calibration)*
1.2.1	Number of steps requiring timing		
1.2.2	Number of steps requiring precision pipetting		
1.3	Number of steps for instrument management		
1.3.1	Number of daily steps for instrument management (excluding maintenance) **		*Each action required daily or per run to set up and shut down the instrument*
1.3.2	Number of steps for maintenance		
1.4	Number of tests per run		*N/A if single-test instrument*
1.5	Time from start to completion for one test		minutes

1.6	Time from start to completion for one run		*N/A if single-test instrument*
1.7	Operator hands-on time for one run (/one test for single-test instruments)		minutes
1.8	For molecular assays: is extraction automated?		Yes/No
1.9	For molecular assays: are extraction and amplification integrated?		*Integrated = no manual step between extraction and amplification*
1.10	Kit controls		
1.10.1	Controls provided by manufacturer (in the kit / separately)?		In the kit / Purchased separately / Not provided
1.10.2	Frequency of controls recommended in the IFU		E.g. Each run / Each day / Not specified
1.10.3	Number of controls		
1.10.4	Type of controls		*E.g. High positive, low positive, negative*
2	**Specimen**		
2.1	Type of specimen collection device provided in the kit		*N/A if none provided*
2.2	Validated specimen types		
2.3	Specimen volume(s) for the assay		
2.3.1	If applicable, minimum volume needed in tube		
2.4	Maximum time between specimen collection and testing (for each specimen type) recommended in the IFU		*Also specify conditions (e.g., temperature) if applicable*
3	**Results and data management**		
3.1	Specimen information entered manually or by scanning a bar code		Manually / Bar code

3.2	Result display		On device only / Printed / External computer
3.2.1	If printed, is printer provided?		
3.2.2	If printer provided, printer cartridges format		Standard format / Specific for printer
3.3	Compatibility / interfacing with Information Management Systems		
3.4	Language options (software)		*Include all languages available*
3.5	Additional monitoring & evaluation functionalities		On instrument / ability to export data for reporting
3.6	Data export		Wireless data transfer / cable / USB / Other
3.6.1	Is a software needed for viewing extracted data?		
4	**Kit storage**		
4.1	Number of tests per kit		
4.2	Kit dimensions		*width / depth /height (cm)*
4.3	Recommended storage temperatures for the kit		
4.4	If specified, recommended storage humidity		
4.5	Recommended storage conditions of kit components after opening		
4.6	Shelf-life of kit components after opening		
5	**Instrument(s) and infrastructure**		
5.1	For instrument 1		
5.1.1	Dimensions		
5.1.2	Weight		

5.1.3	Type of instrument		Benchtop / Freestanding / Portable
5.2	If applicable for instrument 2		
5.2.1	Dimensions		
5.2.2	Weight		
5.2.3	Type of instrument		Benchtop / Freestanding / Portable
5.3	Power sources		Main power / Battery / Solar power
5.4	Need for stable electricity/UPS)		
5.5	Additional requirements, if applicable (e.g., weight to surface area ratio)		
5.6	Operating temperature		
5.7	Is equipment sensitive to dust?		
5.8	If specified, altitude and humidity specifications		
5.9	What are the requirements for separation of workspace (e.g. specimen processing, extraction, amplification)		
5.10	Is distilled/deionized water required?		
5.11	General laboratory equipment required to perform the assay but not provided		E.g., precision pipette, vortex, centrifuge
5.12	Consumables required to perform the assay but not provided		E.g. tips, aerosol-barrier tips, bleach, ethanol
6	**Biosafety and waste disposal**		
6.1	What are the biosafety requirements?		BSL2 / BSL3 / Biosafety cabinet / Other
6.2	Are there any safety concerns for the user (*outside of infectious specimen handling)*		Yes / No

6.2.1	Biological hazards		*List hazards identified*
6.2.2	Chemical hazards		*List hazards identified*
6.3	Waste volume produced per run (solid / liquid)		
6.4	Does waste require additional treatment before disposal and/or specific disposal procedures (*in addition to usual laboratory waste disposal procedures*)		
6.5	Does disposal of the consumables and waste pose a substantial risk for people?		
6.6	Does disposal of the consumables and waste pose a substantial risk for the environment?		
7	**Installation, calibration, maintenance, and troubleshooting**		
7.1	Does installation require vendor or engineer?		
7.2	Minimal frequency of calibration		Daily / weekly / monthly/ yearly / No need
7.2.1	External support needed for calibration?		
7.3	Minimal frequency of maintenance (by instrument type if applicable)		Daily / weekly / monthly/ yearly / No need
7.3.1	External support needed for maintenance?		
7.3.1	Time required for maintenance		
7.4	Number of breakdowns/ blockages during evaluation period		

** Steps for one specimen: each action required to obtain a result for one specimen (excluding specimen collection, instrument management, maintenance/calibration), e.g. add specimen to the cartridge, close the cartridge, scan/type specimen ID, load the cartridge on the instrument, press start (5 steps) OR scan/type specimen ID, load the specimen collection tube into the instrument, press start (3 step)*

*** Steps for instrument management: each action required daily or per run to set up and shut down the instrument, e. g. switch on instrument, log in, maintain supplies, maintain reagents, discard liquid waste, discard solid waste, archive results, switch off instrument (8 steps)*

Table A1.2 Ease of use

		Not applicable	Strongly disagree	Disagree	Agree	Strongly agree
1	**Instruction for use**					
1.1	The IFU for the assay is clear					
1.2	If applicable, pictures/diagrams are clear					
1.3	The IFU contains all important information*					
1.4	Safety instructions are clear					
1.5	The manual for the instrument 1 is clear					
1.6	If applicable, the manual for the instrument 2 is clear					
2	**Kit packaging and labelling**					
2.1	Kit labelling is clear and consistent					
2.2	Safety labelling is sufficient					
2.3	Kit packaging is of good quality					
2.4	All reagents were provided in sufficient quantities					
3	**Assay set-up**					
3.1	If applicable, specimen preparation is simple					
3.2	Loading specimen into the plate/cartridge/sample processing unit is simple					
4	**Use of instruments**					
4.1	The use of instrument 1 is simple					
4.2	If applicable, the use of instrument 2 is simple					
4.3	The steps for trouble shooting, error codes and steps to resolve are clearly documented					
4.4	The software is user-friendly					
4.6	Maintenance of the instrument(s) is simple and rapid					
4.7	The instrument(s) is/are) robust (not susceptible to breakdowns)					
4.8	If needed, technical support was provided in a timely manner					
5	**Results and interpretation**					
5.1	The test results are presented clearly					

5.2	Interpretation of the test result is simple and does not require any additional calculation/transformation by the user					
5.3	Output (print-out, export format) is user-friendly					
5.4	Results export or transmission is user-friendly					
6	**Overall appraisal**					
6.1	Overall, the test is easy to use					
6.2	The test can be used in a laboratory with limited facilities					
6.3	The test can be used in non-laboratory settings (for point-of-care tests)**					
6.4	The test can be used by trained non-laboratory staff (for point-of-care tests)**					

** Refer to TGS-5 Designing instructions for use for in vitro diagnostic medical devices*
https://apps.who.int/iris/bitstream/handle/10665/259737/WHO-EMP-RHT-PQT-TGS5-2017.05-eng.pdf;sequence=1
*** For laboratory-based assays, report "Not applicable"*

If disagree or strongly disagree, report item number and describe

n°	Comment

General appraisal – advantages and disadvantages of the platform/assay	
Advantages	
Disadvantages	

17.2 Annex 2. Informed consent for purposeful collection of specimens

Informed Consent
Performance evaluation of molecular tests for the detection of tuberculosis (with or without resistance detection)

To be used only in case of purposeful collection of specimens

To be translated to local language(s) by the PEL

[Name of Principle Investigator]: <Insert name>

[Name of Organization]: <Insert name>

This Informed Consent Form has two parts:

Information Sheet (to share information about the evaluation with you)

Certificate of Consent (for signatures if you agree to take part)

You will be given a copy of the full Informed Consent Form.

PART I: Information Sheet

Introduction

I am ……., working for [Name of the hospital]. We are doing an evaluation of tests for the diagnosis of tuberculosis (TB). I am going to give you information and invite you to participate in this evaluation. You do not have to say right now whether or not you are willing to participate in the evaluation. Before you decide, you can talk to anyone you feel comfortable with about the evaluation. There may be some words that you do not understand. Please ask me to stop as we go through the information and I will take time to explain. If you have questions later, you can ask me, the study staff.

Purpose of the evaluation

Tuberculosis (TB) is an important public health issue. It is important to diagnose TB and detect possible resistance to common drugs as early as possible to start effective treatment quickly. Some tests for TB diagnosis, called molecular tests, can diagnose TB and detect resistance to TB drugs in one or two days, which can accelerate TB diagnosis and initiation of appropriate treatment compared to other methods such as culture, which take weeks. But their ability to correctly diagnose TB has to be verified before being approved by the World Health Organization (WHO). We aim to evaluate the performance of diagnostic tests for the diagnosis of tuberculosis and detection of resistance to common TB drugs.

Type of evaluation

This evaluation process will involve comparing the results of the TB molecular test compared to the results obtained using culture, which is considered the gold standard for the diagnosis of tuberculosis.

Participant selection

We are inviting patients suspected of having tuberculosis attending this hospital to participate in the evaluation to assess the performance of TB diagnostic tests. In total, specimens from approximately 400 individuals will be included in this evaluation.

Voluntary participation

Your participation in this evaluation is entirely voluntary. It is your choice whether to give a specimen or not. Whether you choose to participate or not, all the medical services you receive at this hospital/clinic will continue and nothing will change. You may change your mind later and stop participating even if you agreed earlier.

Procedures and Protocol

We will ask you to provide a sputum [extrapulmonary] specimen. This will happen once, and you will not be asked to come back for this evaluation.

The specimen will be sent to [name of the laboratory], where it will be used with the TB molecular test under evaluation and to the standard established tests for the diagnosis of tuberculosis and detection of resistance.

Risks

We do not expect that any harm will happen to you because of joining this evaluation.

Benefits

If you participate in this evaluation, you will not have immediate individual benefits, but your participation is likely to help WHO to identify suitable molecular tests for the diagnosis of tuberculosis which can be used in the future in order to help with the control of TB.

Reimbursements

You will not be given any money or gifts to take part in this evaluation.

Confidentiality

The information that will be collected from this evaluation will be kept confidential. Information about you that will be collected during the evaluation will be put away and no one but the evaluation team will be able to see it. Any information about you will have a number on it instead of your name. Only the team in charge of this evaluation will know what your number is and we will lock that information up with a lock and key.

Sharing the Results

Once the evaluation is finished, a report will be written by [name of the laboratory] and WHO. You may contact the laboratory if you would like to know the results of the evaluation. Confidential information will not be shared in this report. We may also publish the results of this evaluation in scientific journals so that other interested people may learn from our research. Your name will not appear in the report or publication, which will be only about the overall results.

Right to Refuse

You do not have to take part in this evaluation if you do not wish to do so and refusing to participate will not affect your treatment at this hospital/clinic in any way. You will still have all the benefits that you would otherwise have at this hospital/clinic.

Who to Contact

If you have any questions you may ask them now or later, even after the evaluation has started. If you wish to ask questions later, you may contact any of the following:

Name:

Address:

Phone number:

Email:

This proposal has been reviewed and approved by the <Insert name>, which is a committee whose task it is to make sure that evaluation participants are protected from harm. If you wish to find about more about the IRB or about your rights as a participant, please contact:

Name:

Address:

Phone number:

Email:

You can ask me any more questions about any part of this evaluation, if you wish to.

Do you have any questions?

PART II: Certificate of Consent

I have read the foregoing information, or it has been read to me. I have had the opportunity to ask questions about it and any questions that I have asked have been answered to my satisfaction. I consent voluntarily to participate in this evaluation.

Name of Participant _____ Date _____
(Day/month/year)

Signature of Participant OR Thumbprint

If illiterate

I have witnessed the accurate reading of the consent form to the potential participant, and the individual has had the opportunity to ask questions. I confirm that the individual has given consent freely.

Name of witness _____ Date _____
(Day/month/year)

Signature of witness

Statement by the evaluation team leader/person taking consent
I have accurately read out the information sheet to the potential participant, and to the best of my ability made sure that the participant understands that a sputum [extrapulmonary] specimen will be collected and used for the evaluation of a molecular test for the diagnosis of TB.

I confirm that the participant was given an opportunity to ask questions about the evaluation, and all the questions asked by the participant have been answered correctly and to the best of my ability. I confirm that the individual has not been coerced into giving consent, and the consent has been given freely and voluntarily. A copy of this ICF has been provided to the participant.

Name of Evaluation team lead/person taking the consent _____

Date _____(Day/month/year)

Signature of Evaluator /person taking the consent

In Vitro Diagnostics Assessment Team

Prequalification Unit
World Health Organization
Avenue Appia 20
1211 Geneva 27
Switzerland
www.who.int

9789240091054