WHO PREQUALIFICATION TEAM:
DIAGNOSTICS

World Health Organization

Technical Specifications Series for submission to WHO Prequalification – Diagnostic Assessment

TSS-17

In vitro diagnostic (IVD) medical devices used for the qualitative detection of *Mycobacterium tuberculosis* complex DNA and mutations associated with drug-resistant tuberculosis

In vitro diagnostic (IVD) medical devices used for the qualitative detection of Mycobacterium tuberculosis complex DNA and mutations associated with drug-resistant tuberculosis
(Technical specifications series for submission to WHO prequalification – diagnostic assessment, TSS17)

ISBN 978-92-4-005586-5 (electronic version)
ISBN 978-92-4-005587-2 (print version)

Table of Contents

Acknowledgements

This document has been developed with support from the Bill & Melinda Gates Foundation, The Global Fund, UNITAID and FIND. The document was prepared in collaboration with S. Schumacher, Global Tuberculosis Programme, WHO; E. Tagliani, San Raffaele Scientific Institute, Milan, Italy; D. Healy; U. Ströher, Prequalification Unit – In Vitro Diagnostic Assessment Team, WHO, and technical and programmatic input from N. Ismail, Global TB Programme, WHO, Geneva. This document was produced under the coordination and supervision of U. Ströher and I. Prat, Prequalification Unit – In Vitro Diagnostic Assessment Team, WHO, Geneva, Switzerland.

List of contributors

A technical consultation on WHO prequalification requirements for IVDs used for the qualitative detection of MTBC and DR-TB was held online from 22 to 24 November 2021.
Meeting participants: D. Gomez, Médecins Sans Frontières, Amsterdam, The Netherlands; P. de Haas, KNCV Tuberculosis Foundation, The Netherlands; P. Hall, International Laboratory Branch, Division of Global HIV and TB US Centres for Disease Control and Prevention (CDC), Atlanta, United States of America (USA); S. Hojvat, Virginia, USA; F. Ismail, The National Institute for Communicable Diseases (NICD)/ Division of the National Health Laboratory Service (NHLS), South Africa; M. Kohli, TB Programme, FIND, Geneva, Switzerland; N. Kumarasamy, VHS-Infectious Diseases Medical Centre, Chennai Antiviral Research and Treatment, Chennai, India; F. Maurer, German National Reference Centre for Mycobacteria at Research Centre Borstel, Germany; J.C. Muhwa, International Union Against Tuberculosis and Lung Diseases and Kenya Medical Research Institute, Kenya; P. Redner, Brazilian National Reference Laboratory for TB and other mycobacteria, Fiocruz, Brazil; M. Ruhwald, TB Programme, FIND, Geneva, Switzerland; T Shinnick, USA; E. Tagliani, San Raffaele Scientific Institute, Milan, Italy; S. Tahseen, National TB Reference Laboratory, Pakistan.

WHO Secretariat: L. Feldcamp; D. Healy; J. Kealy; M. Lanigan; A.L. Page; U. Ströher, In Vitro Diagnostics Assessment Team, Regulation and Prequalification Department; N. Ismail, S. Schumacher; C. Nathanson, A. Korobitsyn, Global TB Programme, WHO.

Public comments were received for consideration from BD, New Jersey, USA; the TC WG2 secretary on behalf of Global Harmonization Working Party (GHWP) (formerly Asian Harmonization Working Party (AHWP)) TC working group 2; Eiken Chemical Co., Ltd., Tokyo, Japan; Hain Lifescience GmbH (Bruker), Nehren, Germany; FIND, Geneva, Switzerland; Health Sciences Authority (HSA), Singapore; ISO/TC 212 Secretary on behalf of ISO/TC 212, International Standards Organization, Geneva, Switzerland; Laura Feldcamp, Ottawa, Canada.

Abbreviations

BCG	Bacille Calmette-Guérin
CFU	colony-forming units
DR	drug-resistant
DST	drug susceptibility testing
GTB	Global Tuberculosis Programme in WHO
Hr-TB	isoniazid-resistance TB
IFU	instructions for use
IS	International Standard
IVD	in vitro diagnostic
LAMP	loop-mediated isothermal amplification
LJ	Löwenstein–Jensen medium
LMIC	low and middle income countries
LOD	limit of detection
MTBC	Mycobacterium tuberculosis complex
MDR	multidrug-resistant
MGIT	Mycobacteria Growth Indicator Tube
NAT	nucleic acid amplification technology
NALC-NaOH	*N*-Acetyl-L-cysteine-sodium hydroxide
PFU	plaque forming units
PLHIV	people living with HIV
POC	point of care
RR-TB	rifampicin-resistance TB
TB	tuberculosis
XDR	extensively drug-resistant
WGS	whole genome sequencing
TSS	Technical Specification Series
US FDA	United States Food and Drug Administration
WHO	World Health Organization

A. Introduction

The purpose of this document is to provide technical guidance to in vitro diagnostic (IVD) medical device manufacturers of IVD closed / semi-closed systems[1] that fall under an existing WHO policy recommendation and intend to seek WHO prequalification of qualitative nucleic acid tests for the detection of *Mycobacterium tuberculosis* complex (MTBC) and resistance to first and/or second line anti-TB drugs (i.e., for detection of rifampicin-resistance TB (RR-TB), multidrug-resistant TB (MDR-TB), isoniazid-resistance TB (Hr-TB), pre- extensively drug-resistant (pre-XDR) and XDR-TB) in pulmonary and extrapulmonary specimens (if claimed).

For this document, the verbal forms used follow the usage described below:

- "shall" indicates that the manufacturer is required to comply with the technical specifications;
- "should" indicates that the manufacturer is recommended to comply with the technical specifications, but it is not a requirement;
- "may" indicates that the technical specifications are suggested methods to undertake the testing, but not requirements.

A documented justification and rationale shall be provided by the manufacturer when the WHO prequalification submission does not comply with the required technical specifications outlined in this document.

Minimum performance requirements for WHO Prequalification are summarized in this document, and where possible, are aligned with published guidance, standards and/or regulatory documents. Although references to source documents are provided, in some cases WHO Prequalification has additional requirements.

Clinical utility studies, i.e., the effectiveness and/or benefits of an IVD, relative to and/or in combination with other measures, as a tool to inform clinical intervention in each population or healthcare setting, do not fall under the scope of WHO Prequalification and are not included in this document. Studies on feasibility, acceptability, and equity as well as health economic studies are used to inform programmatic strategy and included in the development of WHO Global Tuberculosis Programme (GTB) class-based recommendations. These aspects are relevant for programme managers, ministries of health and other bodies in individual WHO Member States.

[1] An IVD closed /semi-closed system is a combination of reagents, calibrators and quality control materials that share a common intended purpose; and are to be used only in combination with each other as components of a single assay. Closed systems are usually dedicated for use with a single instrument, while semi-closed systems comprise different instruments used in combination as components of a single assay (e.g., one instrument for DNA extraction, one for DNA detection) https://www.tga.gov.au/publication/including-ivd-medical-devices-artg

B. How to apply these specifications

For the purposes of WHO prequalification, nucleic acid tests for detection of MTBC and resistance to first-line and second-line anti-TB drugs (qualitative tests) shall comply with the specifications in Part 1 and Part 2 of this document.
Part 3 only applies if, according to the instructions for use (IFU), testing is performed by health care providers in near point-of-care (POC) settings.

C. Other guidance documents

This document should be read in conjunction with other relevant WHO guidance documentation, including:

WHO prequalification documents[2]:

- Technical Guidance Series for WHO Prequalification – Diagnostic Assessment
- Instructions for Compilation of a Product Dossier, WHO document PQDx_018

WHO GTB guidelines:

- WHO Target product profile for next-generation drug-susceptibility testing at peripheral centres[3]
- High priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting, 28-29 April 2014, Geneva, Switzerland[4]
- WHO consolidated guidelines on tuberculosis. Module 3: Diagnosis - Rapid diagnostics for tuberculosis detection 2021 update[5]
- WHO consolidated guidelines on tuberculosis. Module 4: treatment - drug-resistant tuberculosis treatment[6]
- Technical Report on critical concentrations for drug susceptibility testing of medicines used in the treatment of drug-resistant tuberculosis [7]
- Technical Report on critical concentrations for drug susceptibility testing of isoniazid and the rifamycins (rifampicin, rifabutin and rifapentine)[8]
- Tuberculosis laboratory biosafety manual[9]
- Technical manual for drug susceptibility testing of medicines used in the treatment of tuberculosis[10]
- Catalogue of mutations in *Mycobacterium tuberculosis* complex and their association with drug resistance[11]

[2] Available at: https://extranet.who.int/pqweb/vitro-diagnostics/guidance-documents
[3] https://www.who.int/publications/i/item/9789240032361
[4] https://apps.who.int/iris/handle/10665/135617
[5] https://www.who.int/publications/i/item/9789240029415
[6] https://www.who.int/publications/i/item/9789240007048
[7] https://apps.who.int/iris/handle/10665/260470
[8] https://www.who.int/publications/i/item/technical-report-on-critical-concentrations-for-drugsusceptibility-testing-of-isoniazid-and-therifamycins-(rifampicin-rifabutin-and-rifapentine)
[9] https://www.who.int/publications-detail-redirect/9789241504638
[10] https://apps.who.int/iris/handle/10665/275469
[11] https://www.who.int/publications/i/item/9789240028173

D. Performance principles for WHO prequalification

D.1 Intended use

An IVD intended for WHO prequalification shall be accompanied by a sufficiently detailed intended use statement. This should allow an understanding of at least the following:

- The type of assay (e.g., real-time PCR, reverse hybridization/line probe assays, LAMP etc.) and mechanism of MTBC detection and resistance detection.
- What the assay detects, e.g., DNA of MTB or MTBC species; DNA of MTBC species and detection of MTBC genomic changes associated with resistance to one or more anti-TB drugs; detection of MTBC genomic changes associated with resistance to one or more anti-TB drugs
- What the IVD reports (e.g., qualitative test)
- Whether it is automated or not[12]
- The clinical indication and function of the IVD (e.g., the diagnosis of TB disease in individuals with presumptive TB, people at increased risk for drug-resistant TB (DR-TB)).
- The target population for which the functions are intended (see WHO GTB guidelines). (e.g., individuals with signs and symptoms of pulmonary or extra-pulmonary TB)
- The intended use environment (e.g., for professional use in a laboratory setting, and/or near POC[13])
- The intended user (e.g., trained laboratory professional or health care provider).
- The intended pulmonary and/or extra-pulmonary specimen types (see WHO GTB guidelines[14]) (e.g., raw sputum, concentrated sputum sediments, bronchial alveolar lavage, cerebrospinal fluid, stool, lymph node aspirate, etc.).
- Any limitation to the intended use (e.g., only for patients who have not received antituberculosis therapy, if test result is negative and clinical symptoms persist follow-up is recommended, resistance could still be caused by mutations outside the regions interrogated by this assay (rare event), or by mutations present at a frequency below the limit of detection of this assay (e.g., hetero-resistant strains) etc.

D.2 Diversity of specimen types, users and testing environments and impact on required studies

For WHO prequalification submission, supportive evidence (analytical and clinical) must be provided for all claimed specimen types, unless otherwise stated in section E) Table of requirements.

[12] Fully-automated IVDs are sample-to-result systems that do not require hands-on processing steps. If only a part of the testing process is automated, the IVD would be considered semi-automated.

[13] In some jurisdictions, the concept “near patient testing” is used instead of “point of care testing”. Either term may be used in the intended use statement.

[14] https://www.who.int/publications/i/item/9789240029415

Prequalified NAT IVDs for TB in low- and middle-income countries (LMICs) are likely to be used by a range of users in different geographical regions:

- laboratory professionals either in centralised testing laboratories or at/near POC,
- health care providers trained in the use of the test near POC.

Depending on the intended use of an IVD, analytical and clinical performance studies shall be designed to consider not only the diversity of knowledge and skills across the population of IVD users, but also, the likely operational settings in which testing will occur. It is a manufacturer's responsibility to ensure that the risk assessment for an IVD reflects the intended operational settings, including laboratory or service delivery complexity, user expertise, training received and test population.

D.3 Applicability of supporting evidence to IVD under review

Analytical and clinical performance studies shall be undertaken using the specific, final (locked-down) version of the assay intended to be submitted for WHO prequalification assessment. For WHO prequalification, design lock-down is the date that final documentation, including quality control and quality assurance specifications are signed off and the finalized method is stated in the IFU. Where this is not possible, a justification shall be provided, and additional supporting evidence may also be required. This may occur in the case of minor variations to design where no impact on performance has been demonstrated (see WHO document PQDx_121 Reportable Changes to a WHO Prequalified In Vitro Diagnostic Medical Device).[15] If the test procedure of the IFU has been changed in any way, both the protocol provided to laboratory for clinical performance studies as outlined in Part 2 of this document and that in the final version of the IFU intended for users shall be provided with the submission for WHO prequalification assessment.

The version of the IFU used for performance evaluations submitted to WHO prequalification shall be stated. If the test procedure in the IFU is changed in any way after completing performance verification and validation studies the change shall be reported to WHO, including a rationale for the change, and an explanation of why the study results support the claimed performance.

Specific information is provided in this document for the minimum number of lots required for each study. Where more than one lot is required, each lot shall comprise different production (or manufacturing, purification, etc.) runs of critical reagents, representative of routine manufacture. It is the manufacturers responsibility to ensure, via risk analysis of the IVD, that the minimum numbers of lots chosen for estimating performance characteristics reflect the variability in performance likely to arise from the interlot diversity of critical components and their formulation or from changes that occur during the assigned shelf-life of the IVD. Differences found between lots during the analytical and clinical performance studies shall be reported. For

[15] http://apps.who.int/iris/bitstream/handle/10665/251915/WHO-EMP-RHT-PQT-2016.01-eng.pdf;jsessionid=30D5BF0B09FFDA3B38A1698E65C8B496?sequence=1

clinical performance studies, the true status of TB disease as well as the true resistance status, shall be determined using a suitable reference method.

Estimation (and reporting) of IVD performance shall include the rate of invalid test results and the two-sided 95% confidence interval around the estimated values for key performance metrics, as appropriate. The cause of the invalid results should be reported if available, such as sample issues (e.g., age of specimen, storage conditions, inadequate specimen volume), instrument error or operator error. Data should be presented in clear and understandable format.

It is unlikely that clinical specimens will be available in the volumes required for all analytical performance studies. Therefore, it is acceptable to use contrived specimens, for example, well - characterized clinical strains or isolates that have been sequenced, spiked into the appropriate matrix, i.e., a matrix that has been claimed in the intended use of the IVD (e.g., MTBC negative respiratory specimens), and which has been prepared in a validated and standardized manner for such studies. The use of artificial sputum specimens may be acceptable for certain analytical performance studies but must be discussed with WHO in advance unless specifically stated in the table of requirements. Evidence shall be provided that any negative clinical or artificial matrix used is truly MTBC-negative prior spiking. In addition, dilutions of a high-concentration clinical specimen (where MTB concentration has been pre-quantified by colony forming unit (CFU) enumeration) may be used, if they are in an appropriate matrix for certain studies, e.g., stability studies. The material chosen should use the entire assay system from specimen preparation to interpretation.

For analytical performance studies described in part 1 it may be also possible to carefully design protocols that will generate useful data for more than one of the required studies, provided the specific criteria for each requirement are met by the study (e.g., number of replicates, concentration of analyte, specimen types, etc.). For example, precision testing and whole system failure testing could be combined in a single study.

The clinical performance studies shall be determined for all specimen types claimed in the intended use of the IFU. In some analytical performance studies (where indicated) it is acceptable to use one specimen type.

Clinical performance studies shall be based on testing human specimens only sourced from population cohorts reflective of the intended use. The use of well-characterised repository and/or biobank specimens and panels may be acceptable if they are relevant to the IVD under assessment, taking into consideration (however not limited to):

- storage conditions (e.g., including age of the specimen, temperature logs, freeze-thaw cycles if applicable)
- the stability of the nucleic acid target
- selection bias.

E. Table of requirements

PART 1	ANALYTICAL PERFORMANCE AND OTHER EVIDENCE
1.1	**Stability of specimen(s)**
1.1.1	Specimen collection, storage, and transport
1.2	**Validation of specimens**
1.2.1	Demonstration of equivalence of specimen types
1.3	**Metrological traceability of calibrators and control material values**
1.3.1	Metrological traceability of calibrators and control material values
1.4	**Precision**
1.4.1	Repeatability and reproducibility
1.5	**Analytical sensitivity**
1.5.1	Limit of detection for MTBC
1.5.2	Limit of detection for resistance
1.5.3	Analytical reactivity (inclusivity)
1.6	**Analytical specificity**
1.6.1	Potentially interfering substances and medical conditions
1.6.1.1	Endogenous
1.6.1.2	Exogenous
1.6.2	Cross-reactivity
1.6.2.1	Mycobacteria
1.6.2.2	Fungi
1.6.2.3	Viruses
1.6.2.4	Bacteria
1.7	**Validation of assay cut-off**
1.7.1	Validation of assay cut-off
1.8	**Validation of the assay procedure**
1.8.1	Validation of primer and probe choice
1.8.2	Whole system failure rate
1.8.3	Carry-over contamination
1.9	**Usability/human factors**
1.9.1	Flex studies/robustness
1.9.2	Software validation
1.10	**Stability of the IVD**
1.10.1	General requirements
1.10.2	Claimed shelf-life including shipping stability
1.10.3	In-use stability (open pack or open vial stability)
PART 2	**CLINICAL EVIDENCE (CLINICAL PERFORMANCE CHARACTERISTICS)**
2.1	**Clinical sensitivity and specificity**
2.1.1	General requirements for clinical sensitivity and specificity studies
2.1.2	Clinical sensitivity for MTBC
2.1.3	Clinical specificity for MTBC
2.1.4	Clinical sensitivity for DR-TB

2.1.5	Clinical specificity for DR-TB
PART 3	**QUALIFICATION OF USABILITY FOR NEAR POINT OF CARE TESTING BY HEALTHCARE PROFESSIONALS**
3.1	**Qualification of usability for near POC testing by healthcare professionals**
3.1.1	Label comprehension (including IFU)
3.1.2	Results interpretation

Part 1: Analytical performance and other evidence

Aspect	Testing requirements	Notes on testing requirements	Source documents
1.1 Stability of specimens(s)			
1.1.1 Specimen collection, storage and transport	1. Real time studies shall be conducted for each claimed specimen type considering: • storage conditions (for example, duration at different temperatures, temperature limits, freeze/thaw cycles etc.) • transport conditions • intended use (see note 1) • specimen collection and/or transfer devices intended to be used with the IVD 2. The testing panel shall contain (see note 3, 4) • 10 low positive specimens for each specimen type claimed • Specimens shall be approx. 3 x 95% limit of detection (LOD) for MTBC detection (see note 2) 3. Testing shall be conducted in 1 lot	1. Evidence shall be provided which validates the maximum allowable time between specimen collection, processing of the specimen and its addition to the IVD in the setting where testing takes place. 2. The 95% LOD is defined as the minimum concentration of bacterium, expressed as CFU/ml or genomic DNA copy numbers/mL, in a sample volume that can be detected in 95% of tests. (Section 1.5.1) 3. Contrived specimens may be used (i.e., MTBC strains may be spiked in clinical sputum). Evidence must be provided that any clinical sputum used for spiking is MTBC negative. Artificial sputum is not an acceptable specimen matrix. 4. In case the use of archived specimens is considered for Part 2 of this document, evidence of stability in the conditions in which the specimens have been stored shall be demonstrated. 5. Unless all specimens are expected to be processed as fresh samples within a specified time frame, the IVD performance shall be established for each storage condition at the beginning and end of the stated period	TGS-3 [1]
1.2 Validation of specimens			
1.2.1 Demonstration of equivalence between specimen types	1. The relationship between IVD performance in claimed specimen types shall be established: • At least 50 MTBC positive and 50 MTBC negative specimens shall be tested for each claimed specimen type (see note 3) 2. If testing in raw sputum is claimed, and some aspects of performance have been obtained/established using	1. The relationship between IVD performance in claimed specimen types and in the materials used for analytical performance studies shall be established. The design of subsequent studies shall then take that relationship into account. 2. If multiple pulmonary specimen types are claimed, (e.g., raw and/or processed sputum, BAL, other pulmonary specimens), then equivalence shall be	TGS-3 [1] European Commission

Aspect	Testing requirements	Notes on testing requirements	Source documents
	MTBC strains spiked into artificial or processed sputum, then • the relationship between analytical sensitivity in spiked artificial sputum to that of the same characteristic in raw or processed sputum shall be understood (see notes 3 and 4). 3. Similarly, the relationship between analytical sensitivity in MTBC spiked into artificial sputum and clinical sputum shall be understood by end point dilution studies (see note 4) 4. 1 replicate of each specimen for each specimen type shall be tested 5. 1 lot shall be used for testing	demonstrated using paired specimens; in other words, demonstrated in each specimen type for all specimens • If equivalency can be demonstrated between the different pulmonary specimen types claimed, and processed and raw sputum, then clinical performance is required to be established in only 1 (or a combination) of the equivalent pulmonary specimen types 3. Contrived specimens may be used for extrapulmonary specimens (i.e., MTBC strains may be spiked into clinical or artificial specimen matrix where relevant) 4. Demonstration of the comparability of specimen types may be achieved by comparing NAT results between end-point dilution series of several MTBC positive sputum specimens spiked into clinical sputum vs. artificial sputum	decision on CTS [2] IMDRF IVD MA TOC [3] Georghiou SB [5]
1.3 Metrological traceability of calibrator and control material values			
1.3.1 Metrological traceability of calibrator and control values	1. The traceability of provided calibrator and control materials to a validated reference material shall be demonstrated (e.g., WHO International Standard) or a secondary standard calibrated from it.	1. The version of the IS used shall be stated 2. Appropriate to the IFU, the low, positive, external control should contain target nucleic acid at levels approximately 3 x LOD. 3. An internal control shall be added to each sample before sample extraction so that all stages of the test, from extraction to final target detection, can be verified. An internal control consists of defined nucleic acid sequences, which are extracted and amplified simultaneously with the test specimen. Therefore, the test should be able to clearly identify the amplified products (amplicons) of the internal control and the target.	PQDx_018 [6] ISO 15198 [7] ISO 17511 [8] European Pharmacopoeia [9]
1.4 Precision			

Aspect	Testing requirements	Notes on testing requirements	Source documents
1.4.1 Repeatability and reproducibility	1. Both repeatability (see note 1) and reproducibility (see note 2) shall be determined for MTBC and drug resistance detection using a panel of spiked specimens including (see note 7 and 8): • 1 negative specimen • 1 MTBC wild-type low positive specimen (approx. 3 x LOD) • 2 medium wild-type MTBC positive specimen (approx. 5-7 x LOD) • 1 DR low positive MTBC specimen (approx. 3 x LOD) (see note 7) • 2 DR medium positive MTBC specimen (approx. 5-7 x LOD) (see note 7) 2. For **repeatability**, each panel member shall be tested (see notes 6 & 11): • in at least 2 replicates of each panel member per run • over 20 days (not necessarily consecutive) with 1 or 2 runs per day • at 1 site • using 1 lot 3. For **reproducibility**, each panel member shall be tested (see note 6 & 11): • at each of 3 different testing sites • using 3 lots per site • in 3 replicates of each panel member per run • over 5 days (not necessarily consecutive) with 2 runs per day • using 1 operator per site (see note 4) • The effect of operator-to-operator variation on IVD performance should be included as part of the precision studies 4. Testing should be done:	1. Within-run (same operator, same measuring system, same operating conditions, and same location) 2. Between-run, -lot, -day, -site, -operator 3. Where possible, the testing panel should be the same for all operators, lots, and sites 4. The effect of operator-to-operator variation on IVD performance shall be included as part of the precision studies 5. Each lot shall comprise different production (or manufacturing, purification, etc.) runs of critical reagents 6. The nucleic extraction/purification component shall be considered for estimating precision 7. Contrived specimens may be used (i.e., MTBC strains with specific/most common mutations in the target genes spiked into a clinical matrix claimed in the IFU) for repeatability and reproducibility studies 8. If the assay is not intended for DR-TB detection, the panel should include 2 different MTBC species (e.g., M. *tuberculosis* and M. *bovis*) 9. If the assay is intended for DR-TB detection, include DR specimens at the concentrations specified for each DR-TB (i.e. RR-TB, Hr-TB, MDR-TB, TB resistant to fluoroquinolones) 10. Results shall be statistically analysed by ANOVA or other methods to identify and isolate the sources and extent of any variance. In addition, the percentage of correctly-identified, incorrectly-identified and invalid results shall be tabulated for each specimen and be separately stratified according to each site, lot, etc. 11. Alternative study designs to establish repeatability and reproducibility of the assay are acceptable, but shall be discussed with WHO prior to dossier submission	TGS 3 [1] EN 13612:2002 [10] CLSI EP12-A2 [11]

Aspect	Testing requirements	Notes on testing requirements	Source documents
	• by operators representative of intended users in addition to members of manufacturer's staff • unassisted • using only those materials provided with the IVD (e.g., instructions for use, labels, and other instructional materials)		
1.5 Analytical sensitivity			
1.5.1 Limit of detection for MTBC	Analytical sensitivity for M. *tuberculosis* and M.*bovis* shall be estimated as the concentration of bacteria detectable 95% of the time, otherwise known as limit of detection (LOD). 1. For M.*tuberculosis (*H37Rv*)*, the determination shall comprise 24 replicate tests (8 replicate tests on each of 3 days) of a minimum 8-member 0.5$\log_{10}$ dilution panel of a suitable biological reference material (e.g., WHO International Standard or a secondary standard calibrated against it) 2. For M.*bovis* (e.g., BCG) where no IS exists, LOD shall be determined (as described above) using well characterized positive clinical specimens of known concentration (see notes 1 & 2) 3. The replicate testing shall be conducted on 3 different days (see note 4) 4. using 2 lots (see notes 5 & 6) 5. at least 2 dilution series shall be tested	1. Well characterized MTBC strains of known concentration (expressed as CFU/mL) shall be spiked into each claimed MTBC negative specimen type. 2. The LOD may also be presented as genomic DNA copy numbers/mL for each dilution provided that the relationship between CFU/mL and genomic DNA copy numbers/mL is previously established. 3. Analytical sensitivity shall be estimated by determining the 95% LOD with 95% confidence intervals (e.g., by probit analysis). 4. For low through-put instruments, the number of testing days may be increased 5. Each lot shall comprise different production (or manufacturing, purification, etc.) runs of critical reagents 6. Interlot variation shall be evaluated by appropriate statistical means	European Commission decision on CTS [2] CLSI EP17 [12] WHO Technical Report Series, No. 1004, 2017 Annex 6. [13] Miotto P [4]
1.5.2 Limit of detection for resistance	Analytical sensitivity for resistance detection shall be estimated as the lowest number of colony forming units (CFU) per specimen that can be reproducibly distinguished from negative specimens with 95% confidence 1. LOD shall be determined using well characterized positive clinical specimens of known concentration (see note 1,2,3)	1. Relevant DR strains for which the claim is made (expressed as CFU/mL) shall be spiked into each claimed MTBC negative specimen type (e.g., raw and/or processed sputum, and each claimed extrapulmonary specimen) 2. The LOD may also be presented as genomic DNA copy numbers/mL for each dilution provided that the	WHO [14]

Aspect	Testing requirements	Notes on testing requirements	Source documents
	2. The determination shall comprise 24 replicate tests (8 replicate tests on each of 3 days) of a minimum 8-member 0.5log_{10} dilution panel 3. The replicate testing shall be conducted on three different days (see note 6) 4. using 2 lots (see notes 7 & 8) 5. at least 2 dilution series shall be tested	relationship between CFU/mL and genomic DNA copy numbers/mL is previously established 3. DR strains shall be characterized by sequencing as a minimum 4. If the assay detects resistance to more than 1 target drug, the LOD for each target drug in addition to a composite resistance LOD, defined as the highest LOD among the tested target, shall be reported 5. Analytical sensitivity shall be estimated by determining the 95% LOD with 95% confidence intervals (e.g., by probit analysis) 6. For low through-put instruments, the number of testing days may be increased 7. Each lot shall comprise different production (or manufacturing, purification, etc.) runs of critical reagents 8. Interlot variation must be evaluated by appropriate statistical means	
1.5.3 Analytical reactivity (inclusivity)	1. The capacity of the device to detect the MTBC species and/or genetic variations associated with drug resistance as claimed in the IFU shall be demonstrated (see notes 1 & 2) 2. If a claim for MTBC detection is made, the following species shall be tested: *M. tuberculosis, M. bovis, M. bovis BCG, M. africanum, M. microti, and M. caprae* (see note 2) • using a minimum of 20 artificial sputum samples spiked with relevant MTBC species (see note 1) 3. If a claim for detection of other Mycobacteria is made, all the claimed species shall be tested • using a minimum of 20 artificial sputum samples spiked with the relevant bacteria (see note 1)	1. The concentration of MTBC isolates used in inclusivity studies shall be at levels at or near the specific LOD and must be confirmed by plating and counting bacterial CFUs or by quantitative PCR (i.e., CFU/mL or genomic DNA copy numbers/mL) 2. The selection of specific MTBC strains with relevant genetic variations associated with DR must be selected to support the claims in the IFU. This should be done by testing strains that carry the most prevalent mutations, including associated or interim associated resistance mutations capturing the majority (at least 80%) of observed resistance mechanisms globally for each of the assay target drug is acceptable 3. Different whole organism mutant isolates shall be tested at resistance LOD levels and testing concentration confirmed by plating and counting	WHO [14] FDA [15]

Aspect	Testing requirements	Notes on testing requirements	Source documents
	4. The capacity of the device to detect the human-adapted lineages including *M. tuberculosis* sensu stricto (Lineage 1, 2, 3, 4 and 7) and *M. africanum* (Lineage 5 and 6) shall be demonstrated (see notes 1 & 3) • using a minimum of 20 artificial sputum samples spiked with relevant MTBC strains at levels at or near the specific LOD 5. Analytical reactivity for DR MTBC isolates from the most representative *M. tuberculosis* sensu stricto lineages (i.e., Lineage 1-4) should be conducted (see notes 2 & 3) • using a minimum of 20 artificial sputum specimens spiked with relevant MTBC mutant strains at levels at or near the specific resistance LOD	bacterial CFUs or by quantitative PCR (i.e., CFU/mL or genomic DNA copy numbers/mL) 4. The use of cell lysates or, if not available, genomic DNA, from strains (either wild-type or mutant) that are rare or difficult to study is an acceptable alternative only if a minimum of 20 artificial sputum specimens spiked with whole organisms have been tested as per requirements • Cell lysates from rare species can be tested at a titre higher than LOD/resistance LOD if testing for the most common MTBC isolates is conducted as in note 1 and 2 • Genomic DNA from rare species can be spiked into artificial sputum or appropriate matrix at a titre higher than LOD/resistance LOD if testing for the most common MTBC isolates is conducted as in note 1 and 2	
1.6 Analytical specificity			
1.6.1 Potentially interfering substances	The potential for false results arising from interference from the substances/conditions listed below shall be determined (see note 1): 1. by testing both MTBC-negative and MTBC-positive specimens (see note 2), unspiked or spiked with each potentially interfering substance 2. for each interfering substance/condition, 3 replicates of MTBC positive and 3 replicates of MTBC negative specimens shall be tested (see note 2) 3. using a minimum of 100 specimens (see note 3) 4. Competitive target amplification shall be assessed by testing simulated mixed infections, where the possible interfering organism is present at high concentration and the target organism (MTBC) level is low (see note 4)	1. The risk assessment conducted for the IVD shall identify substances/conditions where the potential for interference can reasonably be expected in the areas of intended use and not simply rely on published lists of such compounds and conditions which might be of limited relevance in resource limited settings • by conducting appropriate risk assessment, testing can be performed on the substances or conditions identified as likely to be significant and testing of potentially irrelevant substances/conditions can be avoided • under some circumstances stringent risk evaluation may eliminate the requirement to test some of the items in the lists but any such decision shall be	European Commission decision on CTS [2] CLSI EP07-A3 [16] CLSI EP37-A [17] ISO 14971 [18] FDA [15]

Aspect	Testing requirements	Notes on testing requirements	Source documents
	• Interfering organisms may include other Mycobacteria such as M. *avium*, M. *intracellular*, M. *kansasii*, M. *malmoense*, M. *abscessus* 5. Capacity to detect drug resistance conferring mutations in specimens containing both drug resistant and susceptible MTBC strains (i.e., hetero-resistance) shall be assessed by testing mixtures at various ratios (e.g., 0%, 10%, 15%, 20%, 25%, 50%, 60%, 75%, 90%, and 100%) of bacterial cells carrying mutant and wild-type sequences	documented in any submissions to WHO and considered in the risk-benefit statements 2. Contrived specimens may be used diluted in sputum specimens or artificial sputum matrix 3. Interference studies shall be performed with specimens with a response near the LOD (approx. 3 x LOD) 4. Testing shall be performed with competitive non-MTBC organism in vast excess (e.g., >10^6 CFU/mL) to simulate a high clinical load and low target (MTBC) level near the LOD (approx. 3 x LOD) 5. The methods and concentrations used shall be validated so that any effect of clinical importance would be detected 6. Any observed interference shall be investigated and performance limitations of the IVD reported in the IFU 7. Results shall be reported with respect to each condition or substance and not be reported as an aggregate of the total number of specimens tested in the study 8. Any effect must be evaluated against the probability of that effect occurring and causing clinically significant issues in the population tested in resource limited settings 9. Evaluation of endogenous interfering substances may be partially addressed as part of the clinical performance studies but the number of specimens of each type evaluated shall be in accordance with the requirement in this section 10. Other organisms may be needed to be tested in the relevant specimen type when non-pulmonary specimens are claimed. An individual risk assessment shall be conducted for these specimen types	
1.6.1.1 Endogenous	The interference of endogenous substances on the performance of the assay shall be investigated Endogenous substances shall be spiked at the highest medically relevant concentration, for: 1. Blood (human) 2. Mucus 3. White blood cells (human) 4. Human DNA 5. Gastric acid		
1.6.1.2 Exogenous	The interference of exogenous substances on the performance of the assay shall be investigated. Exogenous substances shall be spiked at the highest medically relevant concentration, for: 1. Medicines, relevant to the populations intended to be tested, including: • treatment of viral infections (e.g., influenza, SARS-CoV-2, HIV) • systemic antibacterial drugs (Tobramycin, Amoxicillin, Levofloxacin) • nasal antibiotic ointments (Mupirocin) • anti-fungal drugs (Pentamidine) 2. Tuberculosis drugs: • Isoniazid, Rifampicin, Ethambutol, Pyrazinamide		

Aspect	Testing requirements	Notes on testing requirements	Source documents
	• Second-line drugs (e.g., fluoroquinolones) 3. Physiological and hypertonic saline 4. Oral expectorants 5. Oral anaesthetics 6. Oral analgesics 7. Nasal sprays, gels, and corticosteroids 8. Mucin: Bovine submaxillary gland, type I-S 9. Tobacco/Nicotine 10. Germicidal mouthwash 11. Inhaled bronchodilators		
1.6.2 Cross-reactivity	1. The potential for false-positive results arising from cross-reactivity (see note 1) shall be determined for: • a minimum of 100 specimens, including, at least 3 – 5 replicates of each microorganism listed below (see note 4) 2. Testing of other unrelated conditions known to cause cross-reactivity in MTBC NATs shall also be included where appropriate 3. *In silico* analysis shall be performed for all microorganisms. Any potential cross-reactivity identified through *in silico* analysis must be reported (see notes 5 & 6) 4. For microorganisms that may not be obtained through reasonable efforts, cross-reactivity may be evaluated only by *in silico* analysis (without laboratory wet-testing) (see notes 5 & 6)	1. The types of conditions/disease tested for shall be risk based, taking into consideration the operational setting as well as the intended users and not simply rely on published lists of such cross-reactivity which might be of limited relevance in resource limited settings • See 1.6.1, Note 1 2. Contrived specimens may be used diluted in sputum specimens or artificial sputum matrix 3. Any observed cross-reactivity shall be investigated and performance limitations of the IVD reported in the IFU. 4. Studies should be conducted using a concentration of a minimum 10^6 CFU/mL for mycobacteria, fungi, and bacteria; a minimum of 10^5 plaque forming units (PFU)/mL or genomic equivalents/mL for viruses 5. For all in-silico analyses, the accession numbers of the nucleotide sequences shall be provided 6. If *in silico* analysis reveals a greater than 80% homology between any of the assay MTB-complex target primers and/or probes and other mycobacterial species (e.g., *M. celatum and M. kumamotonense*), or other common oral and respiratory tract commensals and pathogens, selected fungi, and selected viruses, a	
1.6.2.1 Mycobacteria	1. *M. abscessus* 2. *M. leprae* 3. *M. asiaticum* 4. *M. lentiflavum* 5. *M. avium* 6. *M. malmoense* 7. *M. bohemicum* 8. *M. massiliense* 9. *M. celatum* 10. *M. marinum*		

Aspect	Testing requirements	Notes on testing requirements	Source documents
	11. *M. chelonae* 12. *M. mucogenicum* 13. *M. flavescen* 14. *M. phlei* 15. *M. fortuitum* 16. *M. scrofulaceum* 17. *M. gastri* 18. *M. shimodii* 19. *M. genavense* 20. *M. simiae* 21. *M. goodii* 22. *M. smegmatis* 23. *M. gordonae* 24. *M. szulgai* 25. *M. haemophilum* 26. *M. terrae complex* 27. *M. immunogenum* 28. *M. thermoresistibile* 29. *M. intracellulare* 30. *M. triviale* 31. *M. kansasii* 32. *M. ulcerans* 33. *M. kumamotonense* 34. *M. xenopi*	microbial interference study with MTBC and the microorganism for which the test primers/probe(s) have homology to, should be conducted. Otherwise, a justification as to why the performance of the assay would not be impacted by the presence of a causative agent of clinically significant co-infection or an explanation as to why the *in-silico* results are clinically irrelevant should be provided • This shall be discussed with WHO in advance of submission 7. Other organisms may be needed to be tested in the relevant specimen type when non-pulmonary specimen matrices are claimed. An individual risk assessment shall be conducted for these specimen types	
1.6.2.2 Fungi	1. *Candida albicans* 2. *Candida glabrata* 3. *Candida krusei* 4. *Cryptococcus neoformans* 5. *Candida parapsilosis* 6. *Candida tropicalis* 7. *Aspergillus fumigatus* 8. *Blastomyces dermatitidis* 9. *Histoplasma capsulatum* 10. *Penicillium spp.* 11. *Rhizopus spp.* 12. *Scedosporium spp.*		
1.6.2.3 Viruses	1. *Adenovirus* 2. *HIV* 3. *Human Influenza Virus (Types A and B)* 4. *Human Metapneumovirus* 5. *Human Parainfluenza Virus (Types 1, 2, 3, 4)* 6. *Respiratory Syncytial Virus Rhinovirus* 7. *Rubella Virus* 8. *Rubeola Virus* 9. *Mumps Virus*		

Aspect	Testing requirements	Notes on testing requirements	Source documents
	10. *Varicella Zoster Virus*		
1.6.2.4 Bacteria	1. *Actinomyces israelii* 2. *Acinetobacter baumannii* 3. *Acinetobacter calcoaceticus* 4. *Bacteroides fragilis* 5. *Bacillus cereus* 6. *Bacillus subtilis* 7. *Burkholderia cepacia* 8. *Chlamydophila (Chlamydia) pneumoniae* 9. *Citrobacter freundii* 10. *Clostridium spp.* 11. *Corynebacterium diptheriae* 12. *Corynebacterium jeikeium* 13. *Corynebacterium pseudodiptheriticum* 14. *Eikenella corrodens* 15. *Enterobacteriaceae (including ESBL and KPC producers)* 16. *Enterobacter aerogenes* 17. *Enterobacter cloacae* 18. *Enterococcus faecalis* 19. *Enterococcus faecium* 20. *Escherichia coli* 21. *Fusobacterium spp* 22. *Haemophilus influenzae* 23. *Haemophilus parainfluenzae* 33. *Moraxella catarrhalis* 34. *Mycoplasma pneumoniae* 35. *Neisseria sicca* 36. *Neisseria meningitidis* 37. *Neisseria gonorrheae* 38. *Neisseria mucosa* 39. *Neisseria lactamica* 40. *Pediococcus spp.* 41. *Peptostreptococcus spp.* 42. *Proteus mirabilis* 43. *Proteus vulgaris* 44. *Pseudomonas aeruginosa* 45. *Serratia marcescens* 46. *Staphylococcus aureus* 47. *Staphylococcus epidermidis* 48. *Staphylococcus haemolyticus* 49. *Staphylococcus lugdunensis* 50. *Streptococcus equi* 51. *Streptococcus pneumoniae* 52. *Streptococcus pyogenes* 53. *Streptococcus agalactiae*		

Aspect	Testing requirements	Notes on testing requirements	Source documents
	24. *Haemophilus parahemolyticus* 25. *Kingella kingae* 26. *Klebsiella pneumoniae* 27. *Klebsiella oxytoca* 28. *Lactobacillus spp.* 29. *Legionella pneumophila* 30. *Legionella micdadei* 31. *Leuconostoc spp.* 32. *Listeria monocytogenes* 54. *Streptococcus salivarius* 55. *Stenotrophomonas maltophilia* 56. *Streptomyces anulatus* 57. *Veillonella spp.* 58. *Viridans Group Streptococcus (a minimum of 5 different species)* 59. *Yersinia enterocolitica* 60. *Nocardia farcinica* 61. *Nocardia brasiliensis* 62. *Nocardia otitidiscaviarum* 63. *Rhodococcus equi* 64. *Tsukamurella spp.*		
1.7 Validation of the assay cut-off			
1.7.1 Validation of assay cut-off	1. The assay cut-off shall be verified by testing the following testing panel (see notes 1 & 2): • 100 MTBC positive clinical sputum specimens with low, medium, and high bacterial loads • 1000 MTBC negative sputum specimens • The testing panel shall include 10 positive and 10 negative specimens close to the cut-off • The testing panel shall contain different species (e.g., M. *tuberculosis*, M. *bovis*/BCG, M. *africanum*) 2. The manufacturer shall justify the positioning of the cut-off and describe the algorithm/method used to set the cut-off for the assay, or in cases where the cut-off is set for each run or set of tests, the manufacturer	1. Concentration of the specimens shall be provided in CFU/mL and if possible be quantified with a suitable MTBC DNA quantitative assay 2. Contrived specimens (high positive clinical specimens diluted in an appropriate negative matrix) may be used if natural clinical specimens of the required concentration are not available 3. Specimens used to establish the cut-off shall be different from specimens used to verify the cut-off so that the 2 processes are independent	

Aspect	Testing requirements	Notes on testing requirements	Source documents
	shall describe the algorithm/method specified in the IFU or used by the instrument to set the cut-off.		
1.8 Validation of the assay procedure			
1.8.1 Validation of primer and probe choice	1. For each claimed analyte, evidence supporting the choice of the critical reagents (primer and probe sequences and all *in silico* analysis) shall be provided	1. A rationale for selection of primers and probes including specific sequences used to be provided, including: • Justification for alignments made to generate consensus sequences, or best-fit modifications made to existing sequences e.g., to permit maximum homology to several strains, and • Information on size, GC content, melting temperatures, hairpin or other secondary structures if any, and the nucleotide position on the genome map of the primers and probes 2. Data should be provided to demonstrate that the primers and or probes chosen are effective for all claimed MTBC species and strains, as well as all resistance mutations and combinations of resistance mutations detected	IMDRF IVD MA ToC [3] FDA [15]
1.8.2 Whole system failure	The potential for false negative results in low positive specimens shall be determined: 1. Testing shall be done on a panel (see note 1) consisting of 100 replicates of negative sputum spiked with a MTBC strain (approx. 3 x LOD) 2. The panel shall be tested • on 5 consecutive days (20 tests/day) • using 1 lot • by 1 user	1. This may be conducted as part of precision studies if the minimum number of replicates are met (see Section 1.4)	
1.8.3 Carry-over contamination	1. The potential for carry-over contamination or similar shall be investigated by testing a panel of 40 alternating high positive (e.g., 10^6 CFU/mL) and negative MTBC specimens (see note 1): • at least 5 different runs	1. Contrived specimens prepared by spiking MTBC strains into MTBC negative clinical or artificial sputum may be used for these studies	European Commission

Aspect	Testing requirements	Notes on testing requirements	Source documents
	• on 3 different days • at least 2 users • using 1 lot 2. For testing platforms that can only accommodate a single specimen, testing shall be conducted on a single instrument: • at least 4 tests per run • using alternating high positive and negative specimens • a total of 10 runs • at least 2 users • using 1 lot		decision on CTS [2] Haeckel R [19]
1.9 Usability/human factors			
1.9.1 Flex studies	Evidence is required to demonstrate that the conditions recommended in the IFU have been validated, and a description of how they were verified shall be provided. The influence of the following factors on expected results (both MTBC negative and MTBC positive including both DR and susceptible strains) shall be considered: 1. Specimen and/or reagent volume 2. IVD instrument sturdiness (including the effect of non-level work surface) 3. Lighting, humidity, and barometric pressure (simulating high altitude) 4. Handling contamination (e.g., from latex, powder, hand lotion, sweat, and/or soap, etc.) 5. Operating temperature 6. Instrumentation (extraction, amplification & interpretation) including: • Ruggedness (including the effect of vibration from other instruments) (see note 4) • Impact of dust and mould on componentry (e.g., optics)	1. Refer to WHO document PQDx_018 “Instructions for compilation of a product dossier” for other flex studies that may be relevant, taking into consideration the broad range of operational and environmental conditions consistent with intended use 2. Contrived specimens of the appropriate clinical matrix may be used 3. The factors listed opposite should be investigated in ways that not only reflect, but also exceed, likely operating conditions in low- and middle-income countries so that the limitations of the assay can be understood 4. For the purposes of this document, ruggedness means the ability to resist environmental shocks of a variety of kinds 5. Robustness testing generally takes the form of statistically designed experiments to evaluate the effect of simultaneous “small but deliberate changes” in method parameters	PQDx_018 [6] U.S FDA [20, 21] IEC 62366-1:2015 [22]

Aspect	Testing requirements	Notes on testing requirements	Source documents
	• Impact of power/voltage fluctuations 7. Testing to be performed in 1 lot 8. Where different specimen types are claimed, sputum specimens shall be used 9. Studies investigating the impact of specimen volume shall be conducted in all specimen types claimed 10. The testing panel shall contain (see note 2): • 1 negative specimen • 1 low positive specimen (approx. 3 x LOD)	6. Since assay and analyser parameters are locked down in a closed system and cannot be changed, there should be evidence that these parameters have been optimized	
1.9.2 Software validation	Software validation (including verification of built-in fail-safe and alert mechanisms)	1. If software is utilized for amplification, detection, and calculation of results, validation of such software for the intended function should be provided	FDA [15] FDA [23]
1.10 Stability of the IVD			
1.10.1 General requirements	1. Testing shall be undertaken using a stability testing panel consisting of at least (see note 3): • 1 negative specimen • 1 MTBC low (approx. 3 x LOD) and 1 medium (5 – 7 x LOD) positive specimen • 1 DR-TB low (approx. 3 x LOD) and 1 medium (5 – 7 x LOD) positive specimen with a DR mutation in each of the gene targets detected • Replicate testing (n=3) shall be undertaken at each time point (see note 2) 2. Only 1 specimen type is required to be tested	1. Claims for stability shall be based on the second-last successful data point from the least stable lot, with, if lots are different, a statistical analysis showing that the bulk of lots will be expected to meet the claimed life. For example: for testing conducted at 3, 6, 9, 12 and 15 months, if stability was observed at 15 months, then the maximum stability claim can be 12 months. 2. Justification for the number of replicates shall be based on the stability study set up, statistical analysis of the data and a prior knowledge of the assay's performance. 3. Contrived specimens may be used 4. Each lot should comprise different production (or manufacturing, purification, etc.) runs of critical reagents	

Aspect	Testing requirements	Notes on testing requirements	Source documents
1.10.1 Shelf-life (including shipping stability)	The following conditions shall be investigated: 1. Conditions to mimic extremes of conditions (temperature, humidity, pressure) exposed to during transport 2. Storage temperature and humidity range 3. Operating temperature and humidity range 4. Testing in a minimum of 3 lots 5. Lots shall be subjected to simulated "transport stress" before real time studies are undertaken on these lots. This mimics the real situation (see note 3)	1. Accelerated studies do not replace the need for real time studies. 2. Statistically designed experiments should be involved to allow evaluation of any interactions between environmental conditions 3. Determination of shipping stability shall be performed using simulated extreme stress conditions, ensuring that application of those conditions is consistent and controlled 4. Multiple instruments may be used to allow simultaneous testing at each time point	ISO 23640 [24] CLSI EP25-A [25] TGS-2 [26] ASTM D4169-16 [27]
1.10.2 In-use stability (open pack/ open vial)	1. There shall be evidence that once the unit is removed from its primary packaging, it is stable at the expected temperature and humidity ranges for a defined period of time at the beginning and end of its assigned shelf-life 2. Testing shall be performed: • for all labile components (e.g., buffers vials, sealed cartridges etc.) (see note 1) • using a minimum of 1 lot 3. On-board stability shall be tested for an assay used with an instrument	1. In-use stability of labile components shall be conducted using components in their final configuration	

Part 2: Clinical evidence (clinical performance characteristics)

Aspect	Testing requirements	Notes on testing requirements	Source documents
2.1 Clinical sensitivity and specificity			
2.1.1 General requirements for sensitivity and specificity studies	1. Clinical sensitivity and specificity shall be determined in all claimed specimen types, including different pulmonary and extrapulmonary clinical specimens as per intended use (see note 1) 2. Testing shall be conducted: • in a minimum of 2 lots • in 3 different geographical settings (in more than 1 WHO region) • in low to high TB burden countries • by a variety of intended users in the intended testing settings • on specimens from all sections of the population for which claims are made. The manufacturer shall consider relevant sub populations (e.g., PLHIV, children, previously treated TB patients etc.) 3. For drug resistance detection only (reflex testing), a composite reference method/algorithm relying on phenotypic drug susceptibility testing in liquid media and sequencing shall be used 4. The sampling strategy shall ensure sufficient volume or specimen numbers to ensure index and reference testing are both completed (e.g., for both near POC and reference laboratory testing) 5. Sputum shall be used as the reference specimen for all pulmonary specimens investigated as part of the study 6. Non-determinate, discrepant and unexpected results shall be fully evaluated 7. The procedure for selection of study specimens, how these represent an intended use population and how	1. Specimens used for estimation of specificity and sensitivity shall include: • raw and/or processed (as claimed) freshly taken, unfrozen routine specimens collected and stored as described in the IFU • if validation of specimen types (section 1.2) demonstrate equivalency between the different pulmonary specimens claimed in the IFU, then the number of pulmonary specimens tested may be aggregated • appropriately stored, well characterized specimens may also be used for clinical evaluation testing if necessary assuming that such storage conditions have been validated during analytical performance studies (see Section 1.1) • criteria for the selection of archived specimens shall be explained. In addition, any archived specimens used in the study shall be tested in a randomized, blinded manner interspersed with an appropriate number of negative specimens • bacterial culture shall not be performed on frozen specimens 2. The protocol (study design) used shall be fully described and be reflected in the final IFU. • inclusion criteria should be clearly stated (e.g., patients with signs and symptoms of TB, presentation of the disease (pulmonary or extrapulmonary), patients at high risk of MDR-TB, PLHIV, etc.)	FDA [15]

Aspect	Testing requirements	Notes on testing requirements	Source documents
	bias has been addressed shall be clearly described (see note 2)	• criteria for excluding patients from the study shall be clearly stated (e.g., TB patients undergoing treatment for more than 2 weeks, TB patient under a certain age, insufficient volume, etc.) • the patients should be classified, and results analysed accordingly (e.g., concomitant infections, age, gender, TB treatment history). 3. All specimens shall be characterized using a reference method/algorithm • all specimens shall be subjected to characterization of their TB status • the MTBC detection algorithm shall include smear and liquid culture. In addition, it is recommended that all specimens are tested by a molecular method approved by a stringent regulatory authority. • the method and sample type used for smear microscopy shall be specified • it is recommended that liquid bacterial culture is followed by identification of the MTBC bacteria in the positive culture using a TB specific immunochromatographic test, or a molecular method approved by a stringent regulatory authority. • the DR-TB testing algorithm shall include smear and culture (preferably liquid culture), followed by phenotypic-based drug susceptibility testing as well as sequencing to confirm mutation detection • liquid bacterial culture shall be the reference method to estimate clinical performance of MTBC DNA detection assays, while a composite reference including phenotypic-based drug susceptibility testing and sequencing shall be the reference method to estimate the performance of MTBC drug resistance detection assays. The methodology used shall be fully described	
2.1.2 Clinical sensitivity for MTBC	Testing of: 1. at least 300 confirmed MTBC culture-positive specimens from individual patients, • at least 90 (30%) of which are smear-negative, culture-positive for MTBC (see notes 1, 8) 2. The sensitivity of the IVD shall be presented both overall and separately for smear positive and smear negative patients enrolled		FDA [15]
2.1.3 Clinical specificity for MTBC	Testing of: 1. at least 300 MTBC culture-negative specimens (see notes 1, 8) from individual patients from a symptomatic population as claimed		
2.1.4 Clinical sensitivity for DR-TB	Testing of: 1. at least 80 confirmed DR-TB specimens from individual patients for each drug claimed in the IFU (see notes 8, 9 & 10) • at least 20% shall be from MDR-and pre-XDR TB cases if such a claim is made in the IFU. • resistant specimens do not have to be mutually exclusive 2. The sensitivity of the IVD must be presented both overall and separately for smear positive and smear negative patients enrolled (see note 3)		Georghiou SB [5]
2.1.5 Clinical specificity for DR-TB	Testing of: 1. at least 150 confirmed drug-susceptible MTBC specimens from individual patients for each drug claimed in the assay • Susceptible specimens do not have to be mutually exclusive		

Aspect	Testing requirements	Notes on testing requirements	Source documents
		• clinical sensitivity and specificity shall be stratified by smear result. • manufacturers are recommended to discuss with WHO in advance of WHO submission if a different testing algorithm has been used. 4. Discrepant results should be resolved as much as possible, however performance characteristics shall be based on the original result 5. All results that are indeterminate for drug resistance by the IVD shall be included in the denominator data for MTB detection performance 6. All invalid test results shall be recorded. Invalid results should be reported as individual categories (e.g., internal control failure, extraction failure, etc.) and not aggregated. Invalid results should be analysed separately in the final performance calculations 7. Estimates of clinical sensitivity and specificity shall be reported with 95% confidence intervals 8. The mutation pattern detected shall be reported and interpreted according to ref. 14 (Catalogue of mutations in Mycobacterium tuberculosis complex and their association with drug resistance. Geneva: World Health Organization; 2021.") 9. At least 50% of the results from which the clinical sensitivity and specificity is calculated shall be from freshly collected patient specimens 10. Applicable to rifampicin, isoniazid, ethambutol, pyrazinamide, fluoroquinolones. For bedaquiline, linezolid, clofazimine, delamanid and pretomanid, contact WHO for sample size discussion.	

Part 3: Qualification of usability for near point of care testing by healthcare professionals

Aspect	Testing requirements	Notes on testing requirements	Source documents
3.1 Qualification of usability for near POC testing by healthcare professionals			
3.1.1 Label comprehension (including IFU)	1. Testing of subjects to assess ability of intended users to correctly comprehend key messages from packaging and labelling that relate to near POC testing: • Understanding key warnings, limitations and/or restrictions, including correct collection methods and equipment • Proper test procedure • Test result interpretation • Studies shall include at least 15 intended users including those whose native language may not be the language of the IFU if necessary, to demonstrate comprehension of key messages in the user population	1. IFU and labelling should be clear and easy to understand. Use of pictorial instructional material is encouraged	European Parliament IVD regulations [28] U.S. FDA [20, 21] IEC 62366-1:2015 (22) Backinger CL and Kingsley PA (29)
3.1.2 Results interpretation study	1. For near POC tests, intended users shall interpret the results of contrived IVDs to assess their ability to correctly interpret pre-determined test results and error messages. Contrived tests should be prepared to demonstrate all potential test result interpretations 2. Testing subjects to consist of at least 15 intended users including those whose native language may not be the language of the IFU if necessary, from at least two geographically diverse populations to demonstrate correct interpretation of simulated test results and error messages	1. Study group may include subjects recruited as part of the label comprehension study	

F. Source documents

1. Technical Guidance Series for WHO Prequalification – Diagnostic Assessment: Principles of performance studies, TGS-3. Geneva: World Health Organization; 2017. Licence: CC BY-NC-SA 3.0 IGO.
2. 2009/886/EC: Commission Decision of 27 November 2009 amending Decision 2002/364/EC on common technical specifications for in vitro diagnostic medical devices (notified under document C(2009) 9464) (Text with EEA relevance).
3. In Vitro Diagnostic Medical Device Market Authorization Table of Contents (IVD MA ToC). IMDRF;2019. http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-190321-ivd-mdma-toc-n13.pdf accessed 28 October 2021.
4. Miotto P, Tessema B, Tagliani E, Chindelevitch L, Starks AM, Emerson C, Hanna D, Kim PS, Liwski R, Zignol M, Gilpin C, Niemann S, Denkinger CM, Fleming J, Warren RM, Crook D, Posey J, Gagneux S, Hoffner S, Rodrigues C, Comas I, Engelthaler DM, Murray M, Alland D, Rigouts L, Lange C, Dheda K, Hasan R, Ranganathan UDK, McNerney R, et al. A standardised method for interpreting the association between mutations and phenotypic drug resistance in Mycobacterium tuberculosis. Eur. Respir. J. [Internet] 2017 [cited 2019 Jun 7]; 50: 1701354 Available from: http://www.ncbi.nlm.nih.gov/pubmed/29284687.
5. Georghiou SB, Schumacher SG, Rodwell TC, Colman RE, Miotto P, Gilpin C, Ismail N, Rodrigues C, Warren R, Weyer K, Zignol M, Arafah S, Cirillo DM, Denkinger CM. Guidance for Studies Evaluating the Accuracy of Rapid Tuberculosis Drug-Susceptibility Tests. J. Infect. Dis. 2019; .
6. WHO Prequalification – Diagnostic Assessment: Instructions for Compilation of a Product Dossier. WHO PQDx_018 v3, 27 August 2014. Geneva: World Health Organization; 2014. https://apps.who.int/iris/bitstream/handle/10665/329995/WHO-MVP-EMP-RHT-PQT-2014.01-eng.pdf?ua=1 accessed 12 November 2021
7. ISO 15198:2004 Clinical laboratory medicine -- In vitro diagnostic medical devices -- Validation of user quality control procedures by the manufacturer. Geneva: International Organization for Standardization; 2004.
8. ISO 17511:2003 In vitro diagnostic medical devices -- Measurement of quantities in biological samples -- Metrological traceability of values assigned to calibrators and control materials. Geneva, Switzerland; 2003.
9. Nucleic Acid Amplification Techniques. Eur. Pharmacopoeia 2008; 20: 20621.
10. EN 13612:2002. Performance evaluation of in vitro diagnostic medical devices. Brussels, Belgium; 2002.
11. User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline – Second Edition. CLSI EP12-A2. Wayne, PA; 2008.
12. Evaluation of Detection Capability for Clinical Laboratory Measurement Procedures, 2nd Edition. CLSI EP17-A2. Wayne, PA; 2020
13. WHO. WHO Technical Report Series, No. 1004, 2017. Annex 6. WHO manual for the preparation of secondary reference materials for in vitro diagnostic assays designed for infectious disease nucleic acid or antigen detection: calibration to WHO International Standards. Geneva, Switzerland; 2016.
14. Catalogue of mutations in *Mycobacterium tuberculosis* complex and their association with drug resistance. Geneva: World Health Organization; 2021. Licence: CC BY-NC-SA 3.0 IGO.
15. Class II Special Controls Guideline: Nucleic Acid-Based In Vitro Diagnostic Devices for the Detection of Mycobacterium tuberculosis Complex and Genetic Mutations Associated with Mycobacterium

tuberculosis Complex Antibiotic Resistance in Respiratory Specimens - Guideline for Industry and Food and Drug Administration Staff. 2014.
16. CLSI. Interference Testing in Clinical Chemistry. Approved Guideline - Second Edition. CLSI EP07-A3. Wayne, PA; 2018.
17. Supplemental Tables for Interference Testing in Clinical Chemistry, 1st Edition. CLSI EP37. Wayne, PA; 2018.
18. ISO 14971:2007 Medical devices -- Application of risk management to medical devices. Geneva, Switzerland; 2007.
19. Haeckel R. Proposals for the description and measurement of carry-over effects in clinical chemistry. Pure Appl. Chem. 1991.
20. U.S. Food and Drug Administration Center for Devices and Radiological Health. Recommendations for Dual 510(k) and CLIA Waiver by Application Studies. Guidance for Industry and Food and Drug Administration Staff. 2020.
21. U.S. Food and Drug Administration Center for Devices and Radiological Health. Recommendations for Clinical Laboratory Improvement Amendments of 1988 (CLIA) Waiver Applications for Manufacturers of In Vitro Diagnostic Devices. Guidance for Industry and Food and Drug Administration Staff. 2020
22. IEC 62366-1:2015. Medical devices -Part 1: Application of usability engineering to medical devices. 2015.
23. U.S. Food and Drug Administration Center for Devices and Radiological Health. Content of Premarket Submissions for Device Software Functions. Draft Guidance for Industry and Food and Drug Administration Staff. 2021
24. ISO 23640:2011. In vitro diagnostic medical devices - Evaluation of stability of in vitro diagnostic reagents. Geneva, Switzerland; 2011.
25. Evaluation of Stability of In Vitro Diagnostic Reagents; Approved Guideline. CLSI EP25-A. Wayne, PA; 2009.
26. Technical Guidance Series for WHO Prequalification – Diagnostic Assessment: Establishing stability of in vitro diagnostic medical devices. Geneva: World Health Organization; 2018 Licence: CC BY-NC-SA 3.0 IGO.
27. ASTM. Standard Practice for Performance Testing of Shipping Containers and Systems. ASTM D4169-14. West Conshohocken, PA; 2014.
28. Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in-vitro diagnostic medical devices. O. J. E. U. 2017 L 117/176. 2017.
29. Backinger CL and Kingsley PA. Write It Right: Recommendations for Developing User Instructions for Medical Devices Used in Home Health Care. HHS Publication FDA 93-4258;. Silver Spring, MD: Food and Drug Administration; 1993 Available at: https://www.qualysinnova.com/download/files/write-it-right.pdf accessed 28 October 2021