TSS-23

Rapid diagnostic tests to detect mycobacterial lipoarabinomannan (LAM) antigen in urine, draft for comment

Technical specifications series for submission to WHO prequalification – diagnostic assessment

World Health Organization

TSS-23 Rapid diagnostic tests to detect mycobacterial lipoarabinomannan (LAM) antigen in urine, draft for comment

Technical specifications series for submission to WHO prequalification – diagnostic assessment

World Health Organization

Contents

Acknowledgments

Acknowledgements are due to the many experts whose contributions made this publication possible. The document was prepared in collaboration with Elisa Tagliani; Deirdre Healy and Julian Duncan; consultants to the Prequalification Unit – In Vitro Diagnostic Assessment Team, Ute Ströher, Prequalification Unit – In Vitro Diagnostic Assessment Team, Samuel Schumacher, Global TB Programme, World Health Organization (WHO), Geneva. This document was produced under the coordination and supervision of Ute Ströher and Irena Prat, Prequalification Unit – In Vitro Diagnostic Assessment Team, WHO, Geneva, Switzerland.

List of contributors

A technical consultation on WHO prequalification requirements was held from 24 to 25 January 2024.

Meeting participants: Valeria Alcon, In Vitro Diagnostic Division, Medical Devices Directorate, Health Canada, Canada; Patricia Hall-Eidson, International Laboratory Branch, Division of Global HIV and TB, US Centers for Disease Control and Prevention, United States of America (USA); Dinh Huong, National TB Reference Laboratory, Department of Microbiology, National Lung Hospital, Ha Noi, Vietnam; Mikashmi Kohli, TB Programme, FIND, Switzerland; Zee Ndlovu, South African Medical Unit (SAMU), Médecins Sans Frontières, South Africa; Subhadra Nandakumar, TB and Clinical Monitoring team, Division of Global HIV & TB, Global Health Center, US Centers for Disease Control and Prevention, USA; Shaheed Vally Omar, The National Institute for Communicable Diseases (NICD)/Division of the National Health Laboratory Service (NHLS), South Africa; Paulo Redner, Brazilian National Reference Laboratory for TB and other mycobacteria, Brazil; Siva Kumar Shanmugam, Department of Bacteriology, ICMR-National Institute for Research in Tuberculosis, India; Thomas Shinnick, Independent Consultant, USA; Seda Yerlikaya, Department of Infectious Disease and Tropical Medicine, Heidelberg University Hospital, Germany.

WHO Secretariat: Susie Braniff; Jean-Frédéric Flandin; Deirdre Healy; Anne-Laure Page; Irena Prat; Ute Ströher; Elisa Tagliani, In Vitro Diagnostics Assessment Team, Regulation and Prequalification Department; Carl-Michael Nathanson and Alexei Korobitsyn, Global TB programme.

This document has been developed with support from UNITAID, The Global Fund and the Bill & Melinda Gates Foundation.

Declarations of interests

All participants completed a Declaration of Interests form in advance of the meeting. Two of the participants declared interest in the topic under consideration. Mikashmi Kohli and Seda Yerlikaya declared significant interests connected with their employment and ongoing research support for manufacturers of TB diagnostics. It could not be excluded that the declared interests may be perceived as a potential conflict of interest. Therefore, while both persons mentioned above had been invited to participate in the meeting, they participated in the discussion as technical resource people.

All remaining experts were not considered by WHO to have declared any interest that may be perceived as a potential conflict with regard to the objectives of the meeting. All the declarations, together with any updates, were made known and available to all the participants at the beginning of the meeting. All the experts participated in their individual capacities and not as representatives of their countries, governments or organizations.

Abbreviations

CI	confidence interval
CLSI	Clinical and Laboratory Standards Institute
CV	coefficient of variation
IFU	instructions for use
IMDRF ToC	International Medical Device Regulators Forum Table of Contents
ISO	International Organization for Standardization
IVD	in vitro diagnostic
LAM	lipoarabinomannan
LOD	limit of detection
MTBC	*Mycobacterium tuberculosis* complex
POC	point of care
QA/QC	quality assurance/quality control
RDTs	rapid diagnostic tests
ROC	receiver operator characteristic
TB	tuberculosis
TGS	Technical guidance series
TSS	Technical Specification Series
US FDA	United States Food and Drug Administration
WHO	World Health Organization

A. Introduction

The document is developed for manufacturers who are interested in applying for WHO prequalification assessment, to assist in the compilation of their product dossier. The document summarizes the minimum analytical and clinical performance studies to be conducted for rapid diagnostic tests (RDTs) for the qualitative detection of mycobacterial lipoarabinomannan (LAM) antigen for point of care (POC) professional use in urine of HIV positive individuals.

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 a suggested method to undertake the testing, but it is not a requirement.

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.

For WHO prequalification purposes, manufacturers shall provide evidence in support of the clinical performance of an IVD to demonstrate that reasonable steps have been taken to ensure that a properly manufactured IVD, being correctly operated in the hands of the intended user, will detect the target analyte consistently and fulfil its indications for use.

Where possible, WHO analytical and clinical performance study requirements 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. A full list of the individual studies is provided in chapter E (Parts 1-2).

WHO prequalification requirements summarized in this document do not extend to the demonstration of clinical utility, 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 a given population or healthcare setting. To demonstrate clinical utility, a separate set of studies is required. Clinical utility studies usually inform programmatic strategy and are thus the responsibility of programme managers, ministries of health and other related bodies in individual WHO Member States. Such studies do not fall under the scope of WHO prequalification.

B. How to apply these specifications

For the purposes of WHO prequalification, immunoassays for the detection of mycobacterial LAM antigen shall comply with the specifications in Part 1 and Part 2 of this document.

The submission of the dossier must be according to TSS (Technical specification series) requirements and prequalification dossier instructions "Instructions for compilation of a product dossier)". [*1*]

C. Other WHO guidance documents

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

WHO prequalification documents:

Instructions for compilation of a product dossier (referred to as WHO document PQDx_018). [*1*]

- Technical guidance series for WHO prequalification – diagnostic assessment. [*2*]

WHO Global TB programme guidelines and policies:

- WHO consolidated guidelines on tuberculosis: module 3: diagnosis: rapid diagnostics for tuberculosis detection, 3rd ed. [*3*]
- WHO operational handbook on tuberculosis: module 3: diagnosis: rapid diagnostics for tuberculosis detection: web annex A: information sheets, 3rd ed. [*4*]
- High-priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting. [*5*].

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. lateral flow test);
- What the IVD medical device detects (e.g., LAM antigen);
- What the IVD medical device reports (e.g., qualitative test);
- Whether or not it includes automated components or it is intended to be used with a reader or automated instruments;
- The clinical indication and function of the IVD (e.g. aid in the diagnosis of active TB disease in individuals with signs or symptoms of TB);

- The intended testing population (e.g. HIV-positive adults, adolescents and children with signs and symptoms of TB or with advanced HIV disease or who are seriously ill irrespective of signs and symptoms of TB and with a CD4 cell count of less than 200 cells/mm^3 in inpatient settings, or a CD4 cell count of less than 100 cells/mm^3 in outpatient settings);
- The intended user (e.g. trained laboratory professionals[1], trained healthcare professionals or by trained healthcare workers/lay providers[2]);
- The intended operational setting (e.g., for professional use in a point of care and/or laboratory setting);
- The intended specimen type (urine and/or concentrated urine);
- Any limitation to the intended use.

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

For WHO prequalification submission, clinical performance studies shall be conducted using the specimen types (urine, concentrated urine) that are claimed in the instructions for use (IFU). Prequalified RDTs are likely to be used by laboratory professionals in low- and middle-income countries, or by healthcare workers/lay users trained in the use of the test at POC. Depending on the intended use of an immunoassay, analytical and clinical performance studies shall be designed to consider not only the diversity of knowledge and skills across the population of such individuals, but also the likely operational settings in which testing will occur.

Laboratory demonstration of equivalence between specimen types without evidence of clinical validation is insufficient. For example, studies that comprise the testing of left-over/repository specimens by research and development staff at a manufacturer's facility shall not, on their own, be considered sufficient to meet many of the clinical performance study requirements summarized in this document.

D.3 Applicability of supporting evidence to IVD under review

Analytical and clinical performance studies shall be undertaken using the specific, final (locked-down design) version of the immunoassay intended to be submitted for WHO prequalification. For WHO prequalification, design lock-down is the date that final documentation is signed off, including quality control and quality assurance specifications, and the finalized method is stated in the IFU. Where this is

[1] Medical technologists, medical laboratory technicians or similar, who have received a formal professional or paraprofessional certificate or tertiary education degree.

[2] Any person who performs functions related to healthcare delivery and has been trained to deliver specific services but has received no formal professional or paraprofessional certificate or tertiary education degree. Lay users do not include self-testing in the context of this document. *Consolidated guidelines on HIV testing services (2015)* https://iris.who.int/bitstream/handle/10665/179870/9789241508926_eng.pdf?sequence=1

not possible, a justification shall be provided; 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 [7]). If the method section of the IFU has been changed in any way, both the study protocol provided to a 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 verification and validation studies submitted for WHO prequalification assessment 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 inter-lot 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.

Where the manufacturer supplies instrumentation required to conduct the assay, safety and performance data shall be provided in the dossier for this instrumentation. If both a visual read and an automated digital read out version of the test can be used by end users, both modes shall be utilized in each study and results/performance reported. Closed system instruments and proprietary readers are eligible. For clinical performance studies, the target condition is active TB disease, which includes both pulmonary and extrapulmonary TB. For determining the true TB status of the study subjects, a microbiological reference standard shall be used, including at a minimum culture and molecular testing. In addition, clinical assessment for TB signs and symptoms, chest x-ray and/or biochemical marker testing can be utilized to assist in the confirmation of subjects unable to produce a sputum sample or in case of extrapulmonary TB. For WHO purposes the reference method should be to a level that is currently at a developed stage of technical capability based on the relevant consolidated findings of science, technology, and experience (commonly referred to as state of the art). Estimation (and reporting) of

IVD performance shall include the rate of invalid test results and the 95% confidence interval around the estimated values for key performance metrics, as appropriate. The cause of the invalid results should be reported if known such as sample issues (e.g. age of specimen, storage conditions, inadequate specimen volume), instrument error, operator error. For resolution of discrepant results, comparison with a similar device is insufficient. Data should be presented in a clear and understandable format.

It is acceptable to use contrived specimens for analytical performance studies unless otherwise specified in part 1. Preferably well characterized, purified *Mycobacterium tuberculosis* (MTB) LAM (e.g. from BEI resources) spiked into confirmed negative matrix of the claimed specimen type, however where indicated the use of mycobacterial culture is also permissible.

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, lot numbers etc.). Studies which may fall in this category are indicated in the appropriate chapters of part 1.

The performance of the IVD shall be established in all claimed specimen types unless otherwise noted in the table below.

Clinical studies shall be based on testing clinical specimens only sourced from population cohorts reflective of the intended use. The use of well-characterised repository specimens and panels may be acceptable if they are relevant to the IVD under assessment, taking into consideration storage conditions (including age of the specimen) and the stability of LAM antigen.

E. Table of requirements

WHO requires that a product dossier be submitted in the "Table of Contents" (ToC) format, described in the International Medical Device Regulators Forum (IMDRF) document IMDRF/RPS WG/N13 FINAL:2019 (Edition 3)[*8*]. In the tables below, the chapters and subheadings are labelled and numbered according to IMDRF ToC format. As the IMDRF ToC is comprehensive in nature, not all subheadings are required for WHO prequalification and are excluded. As a result, the subheading numbering in the tables below is not always continuous (e.g., 3.1.1, 3.1.3, etc). This has been done to maintain consistency between sections required in a product dossier for WHO prequalification assessment and the corresponding numbering defined in the IMDRF ToC format.

PART 1: IMDRF ToC CHAPTER 3 – ANALYTICAL PERFORMANCE AND OTHER EVIDENCE	
3.05	Analytical performance
3.05.01	Stability of specimen(s)
3.05.02	Validation of specimens
3.05.03	Metrological traceability of calibrator and control material values
3.05.04	Accuracy of measurement
3.05.04.02	Precision (repeatability and reproducibility)
3.05.05	Analytical sensitivity (limit of detection)
3.05.06	Analytical specificity
3.05.06a	Potentially interfering substances
3.05.06b	Cross-reactivity
3.05.07	High dose hook effect
3.05.09	Validation of assay cut-off
3.05.10	Validation of the assay procedure
3.05.10a	Validation of assay parameters
3.05.10b	Validation of the control line or dot
3.06	Other studies
3.06.04	Usability/human factors
3.06.04a	Flex/robustness studies
3.06.04b	Usability: label comprehension study including IFU
3.06.04c	Usability: result interpretation study
3.06.05	Stability of the IVD
3.06.05.01 &	
3.06.05.03	Claimed shelf-life and shipping stability
3.06.05.02	In use stability
PART 2: IMDRF ToC CHAPTER 4 – CLINICAL EVIDENCE	
4.02.03	Device specific clinical studies
4.02.03a	General requirement for clinical performance
4.02.03b	Clinical sensitivity
4.02.03c	Clinical specificity

Part 1: IMDRF ToC chapter 3: Analytical performance and other evidence

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
3.05.01 Stability of specimen(s)			
Specimen collection, storage and transport	1. Real time studies shall be determined for each specimen type taking into account: • Storage conditions (duration at different temperatures, temperature limits, freeze/thaw cycles); • Transport conditions (e.g., temperature and time from sample collection to arrival to the testing site); • Specimen collection and/or transfer devices intended to be used with the IVD. 2. Testing of a minimum of 10 specimens from different individuals (see note 3). 3. Clinical specimens shall be weakly reactive and include at least one negative sample (see note 4). 4. Testing shall be conducted using 1 lot.	1. In case the use of archived/stored specimens is considered for part 1 or 2 of this table, evidence of stability shall be demonstrated for the archiving conditions (e.g. repeated freeze/thaw cycles, temperature, duration). 2. Data generated by the manufacturer on other similar proprietary IVDs for the detection of the same analyte in the same specimen type may be submitted to support the specimen stability claims. 3. Specimens spiked with purified MTB LAM or mycobacterial culture are not accepted. 4. LAM concentration in specimens may be characterized as weakly reactive according to a calibrated, graduated colour chart or a semi-quantitative scoring system, or quantitative methods.	
3.05.02 Validation of specimens			
Matrix effect	1. If multiple specimen types are claimed (urine and concentrated urine), the manufacturer shall investigate a potential matrix effect. 2. 40 negative matrix (paired urine and concentrated urine) from individual donors spiked with the same amount (less than 5% v/v) of a known reactive specimen (e.g. purified MTB LAM, mycobacterial culture) shall be tested. A third of the positive	1. Urine and concentrated urine are considered different specimen types.	TGS-3 [*9*] CLSI EP35 [*10*]

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	specimens should be near (1-2 x) the RDT LOD (or cut-off if a reader is used) and the rest across the measuring range.		
3.05.03 Metrological traceability of calibrator and control material values			
Metrological traceability of calibrator and control values	1. The metrological traceability of the provided control material(s) to reference material shall be determined.	1. If a control material has an assigned concentration value, the metrological- (not commercial- nor documentary-) traceability to an accepted reference material should be demonstrated.	PQDx_018 [*1*] ISO 17511:2020 [*11*]
3.05.04 Accuracy of Measurement			
3.05.04.02 Precision (Repeatability & Reproducibility)			
Repeatability and reproducibility	1. Repeatability and reproducibility (see note 1) shall be estimated using a panel of spiked specimens (see note 2): • 1 negative; • 1 weakly reactive (approx. 1-2 x LOD (or cut-off if a reader is used)); • 1 medium reactive (approx. 2-3 x LOD (or cut-off if a reader is used)). 2. Each panel member shall be tested: • In 5 replicates per test; • Over 5 days (not necessarily consecutive) with 1 run per day (alternating morning/afternoon); • In 3 different lots (see note 3; at least 2 lots should be tested at each of the sites) • At each of 3 different sites; • By 3 different operators	1. Studies shall be statistically designed and analysed to identify and isolate the sources and extent of any variance. • Within or between -run, -lot, -day, -site, -users. • Users shall always be blinded to the expected results. 2. Where possible, the testing panel should be the same for all operators, lots, and sites. 3. The panel shall be prepared by spiking purified MTB LAM into confirmed negative matrix of the claimed specimen type. • The panels should be stored as per reference material package insert. 4. Each lot shall comprise different production (or manufacturing, purification, etc.) runs of critical reagents. 5. The effect of operator-to-operator variation on IVD performance may also be considered as a human	EN 13612:2002 [*12*] CLSI EP12 [*13*] U.S. FDA [*14*]

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	• If a reader is required to interpret the test results, a minimum of 3 different instruments (1 per site) shall be used. 3. The effect of operator-to-operator variation on IVD performance shall be included as part of the precision studies (see notes 5 and 6). Testing shall be conducted: • By users representative of intended users; • Unassisted; • Using only those materials provided with the IVD (e.g., IFU, labels and other instructional materials).	factor when designing robustness studies (see 3.06.04 Usability/human factors). 6. Operators' profiles shall be detailed in the study report (e.g. affiliation and skill level). 7. Results shall be reported as the proportion of specimens detected and in addition as graded band intensity results. 8. 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. This type of analysis is especially important for RDTs that may not have results with any numerical values.	
3.05.05 Analytical sensitivity			
Limit of detection (LOD)	1. The LOD of MTB LAM RDTs shall be determined relative to an accepted biological reference material (see note 1). 2. The determination should comprise a minimum of 24 replicate tests (3 replicates per dilution) of an 8-member dilution panel. 3. Testing shall be conducted using a minimum of 2 different lots. 4. LOD shall be estimated for all the claimed specimen types (e.g., urine, concentrated urine).	1. The source of reference biological material used shall be stated. 2. The LOD is defined as the lowest concentration of analyte (expressed in pg/mL) that can be consistently detected. Typically, in > 95% of samples tested under routine clinical laboratory conditions and in a defined specimen type. 3. Determination shall be according to an approved statistical method (e.g. see source document EP12 or EP17). 4. For qualitative assays, the logistic fit method is acceptable. 5. Each lot shall comprise different production (or manufacturing, purification, etc.) runs of critical reagents.	ISO 17511:2020 [*11*] CLSI EP12 [*13*] CLSI EP17-A2 [*15*]

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
3.05.06 Analytical specificity			
3.05.06a Potentially interfering substances	1. The potential for false results (false nonreactive and false reactive results) arising from interference from at least, but not limited to, the substances/conditions listed below shall be investigated (see note 1). 2. Testing shall be undertaken in both LAM antigen negative and reactive specimens (spiked using purified MTB LAM or characterized clinical specimens), unspiked or spiked with each potentially interfering substance at the highest concentration found in individuals. 3. Testing shall be performed in: • 1 lot (see note 4); • 3-5 replicates; • 1 specimen type (see note 8); • At least 100 specimens total.	1. The risk assessment conducted for the RDT should identify substances/conditions where the potential for interference and cross-reactivity can reasonably be expected with the analyte to be detected in the areas of intended use: • By conducting and documenting appropriate risk assessment, testing can be conducted on specimens spiked with the substances/conditions identified as likely to be significant and testing of potentially irrelevant substances/conditions avoided. • Not by simple reliance on published lists of such substances and conditions, which might be of limited relevance to this analyte. Under some circumstances stringent risk evaluation might eliminate the necessity to test requirements column (see paragraph above) but any such decision shall be documented in the submission to WHO and considered in the risk-benefit statements. • Any effect must be evaluated against the probability of that effect occurring, given the prevalence of that substance/condition in each of the population intended to be tested and the clinical significance of the effect.	EU Common specifications [*16*] CLSI EP07-A3 [*17*] CLSI EP37 [*18*] ISO 14971:2019 [*19*]
Endogenous substances	1. Substances/conditions expected to be found in the specimen types claimed e.g.: • Glucose; • Haemoglobin, blood, leukocytes; • Bilirubin, urobilinogen; • Urea; • Lipids; • Proteins; • Ketones, nitrates.		

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
Exogenous substances	1. Substances, relevant to the populations intended to be tested for example: • Antibacterial (including antituberculosis) drugs; • Anti-parasitic drugs (e.g. treatment for malaria, treponema); • Anti-viral/antiretroviral drugs; • Bovine serum albumin; • Acetylsalicylic acid; • Ascorbic acid; • Biotin (see note 7).	2. Any observed interference or cross-reactivity shall be investigated and performance limitations of the RDT reported in the IFU. 3. Results shall be reported with respect to each condition and not be reported as an aggregate of the total number of specimens tested in the study. 4. The lot used in this study shall be the same as one of the lots in 3.05.05 LOD studies. 5. The methods and concentrations used for interference studies shall be validated so that any effect of clinical importance would be detected. 6. Interference studies should be performed with LAM-positive specimens with an analyte response (MTB LAM antigen) near the LOD (not higher than 3 x LOD). MTB LAM concentration in clinical specimens may be characterized as weakly reactive according to a calibrated, graduated colour chart or a semi-quantitative scoring system. 7. For interference studies, if the technology of the test employs streptavidin, then biotin levels of up to 3500 ng/mL should be tested as part of this study. 8. For cross reactivity studies, where clinical specimens from individuals with the disease state to be tested are unavailable, a negative specimen shall be spiked with the organism of interest to a high concentration (a minimum of 10^5 plaque forming units/mL for viruses and 10^5 colony forming units/mL for bacteria).	
3.05.06b Cross-reactivity	The potential for false-positive results arising from cross-reactivity (see note 1) shall be determined in 3 to 5 replicates for each of the following microorganism (if applicable/based on a risk assessment: 1. At a minimum non-tuberculous mycobacterium clinically relevant to people living with HIV: • *M. avium;* • *M. kansasii;* • *M. intracellulare;* • *M. chelone;* • *M. abscessus;* • *M. gordonae;* • *M. fortuitum*. 2. Fungal infections, including candida and aspergillus; 3. Microorganisms causing urinary tract infections/sexually transmitted infections;		

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	4. Hepatitis B, C; 5. Other unrelated conditions known to cause cross-reactivity in MTBC immunoassays.	9. Testing shall be conducted in the claimed specimen type.	
3.05.07 High dose hook effect			
Prozone/ High dose hook effect	1. The potential for a high dose hook effect shall be investigated: • Spiking negative matrix (i.e., urine) with an increasing high purified MTB LAM concentration (approximately 10000 x LOD or until signal decreases); • In 3 lots. 2. If there is evidence of a prozone effect, this information shall be added to the IFU, and mitigation actions shall be described. 3. Testing shall be conducted in 1 specimen type.		TGS- 6 [*20*]
3.05.09 Validation of Assay Cut-off			
Establishment of reader cut-off	1. For RDTs provided with a reader, the way in which the reader has been designed to differentiate between reactive specimens and negative specimens shall be demonstrated and described in detail.	1. The statistical methods (e.g., receiver operator characteristic [ROC]) used to generate results and the testing performed to define a grey-zone/equivocal zone if applicable shall be described. 2. The cut-off shall be established prior to conducting any analytical and clinical performance studies.	
3.05.10 Validation of the assay procedure			
3.05.10a Validation of assay parameters	1. Evidence shall be provided on how any parameters specified in the IFU were determined, validated, and verified.	1. These parameters may be investigated as part of 3.06.04 Usability/Human factors or 3.06.05.02 In-use stability, below.	PQDx_018 [*1*]

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	2. The parameters specified in an IFU commonly include the following, but the actual requirement is assay dependent and must be ascertained for each IVD: • Allowable reading time (see note 2); • Time interval between opening the pouch and starting the assay; • Processing steps/timed steps; • Volumes, including numbers of drops; • Temperatures e.g., operating temperature range; • Humidity; • Steps to concentrate urine specimens (time, centrifugation speed, etc.). 3. Testing shall be conducted using 2 lots (1 freshly made lot and 1 lot of IVD towards the end of the assigned shelf life). 4. Specimen panel to be tested in triplicate shall be as follows (see note 3, 4): • 1 negative specimen; • 1 weakly reactive specimen (approx. 1-2 LOD (or cut-off if a reader is used)); • 1 medium reactive specimen (approx. 2-3 LOD (or cut-off if a reader is used)).	2. For RDTs where a reading interval is specified, validation data of the minimal and maximum allowable time shall be provided. 3. Pooled clinical specimens or contrived specimens (purified MTB LAM spiked into negative matrix) shall be used. 4. LAM concentration in clinical specimens may be characterized as weakly reactive according to a calibrated, graduated colour chart or a semi-quantitative scoring system.	
3.05.10b Validation of the control line or dot	1. The flow device shall have a control line. The nature of the control line shall be explained (see note 1).	1. The extent to which any control line corresponds to a valid test shall be validated. 2. The precise meaning of the control line must be stated in the IFU of the device, e.g., evidence of: • Reagent addition and flow;	

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
		• Specimen addition and flow; • Correct volumes being added; • Correct operation of the device; • Correct functionality of all reagents.	
3.06 Other Studies			
3.06.04 Usability/human factors			
3.06.04a Flex/robustness studies	1. The intent of this study is to demonstrate that no combination of small but defined variations in the parameters of the protocol will result in the IVD failing to meet any of the manufacturer's claims i.e., the assay is robust. 2. Specimen panel to be tested in triplicate shall be as follows: • 1 negative specimen; • 1 weakly reactive specimen (approx. 1-2 LOD (or cut-off if a reader is used)); • 1 medium reactive specimen (approx. 2-3 LOD (or cut-off if a reader is used)). 3. The influence of the following factors on expected results (both reactive and non-reactive) shall be considered based on the risk-assessment conducted, for example but not limited to: • Time between opening packaging or preparing reagents and starting the assay; • Specimen processing e.g. for concentrated urine; • Timing of processing steps; • Specimen volume including number of drops;	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 range of operational and environmental conditions consistent with intended use in resource limited settings. 2. The factors listed 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 device can be understood. For example, in addition to investigating deviations of temperature ranges should be investigated that exceed those of claimed operating conditions and which could cause test failure (incorrect/invalid results). 3. The resilience of label (e.g., strength of attachment, print stability, legibility over time, damp tolerance) shall be evaluated. 4. The impact of lighting: • On the visual reading of the control and test lines; • On labelling (fading).	ISO 14971:2019 [*19*] U.S FDA [*21*] IEC 62366-1:2015 [*22*] U.S. FDA [*23*]

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	• Reagent volume provided and used; • Specimen dilution/concentration factor; • Reading time; • Operating temperature, pressure, and humidity. 4. Ruggedness shall be considered based on the risk-assessment conducted, for example but not limited to the following conditions (see note 7): • RDT sturdiness including robustness of packaging and labelling. RDT in final packaging shall be subjected to drop-shock testing; • Permanence of component labels: print legibility, adhesiveness (see notes 3, 4); • Effects of lighting and humidity (see note 5); • Placement of the test device on non-level surface; • The effect of moving the test device while it is running (e.g., relocating to another surface or dropping it). 5. Review of instrumentation (if applicable and based on a risk assessment) including: • Ruggedness (see above); • Impact of dust and mould on componentry (e.g., optics if applicable). 6. Studies shall be conducted in a claimed specimen type.	5. The factors should be investigated using "designed experimentation" so that potential critical interactions between them can be understood e.g., the effect of low or high operating temperature with low or high volume of specimen at an incorrect reading time. 6. Some of these parameters/factors may be investigated as part of 3.05.10a Validation of assay parameters or 3.06.05.02 In-use stability. 7. For the purposes of this document, ruggedness means the ability to resist environmental shocks of a variety of kinds. 8. Pooled clinical specimens or contrived specimens (purified MTB LAM spiked into negative matrix) shall be used. 9. LAM concentration in clinical specimens may be characterized as weakly/medium reactive according to a calibrated, graduated colour chart or a semi-quantitative scoring system, or quantitative methods.	
3.06.04b Usability: Label comprehension	1. Testing shall be undertaken to assess the ability of intended users to correctly comprehend key messages from packaging and labelling:	1. Instructions for use and labelling shall be clear and easy to understand; use of pictorial instructional material is encouraged. If additional resources such	IEC 62366-1:2015 [*22*] U.S. FDA [*23*]

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
study (including IFU)	• Understanding key warnings, limitations and/or restrictions; • Proper test procedure; • Proper reader procedure (if included); • Test result interpretation; • Using only the information available to all users (IFU and any job aid). 2. Studies shall include: • At least 15 intended users including those whose native language may not be the language of the IFU if necessary; • In their usual working environment, not employees of the manufacturer; • From 2 geographically diverse populations to demonstrate comprehension of key messages in each user group.	as videos are provided, the information provided in the videos shall be the same as the information provided in the IFU. 2. Requirements listed may be investigated as a separate study or included as part of the results interpretation study and/or clinical study. 3. Testing may be conducted using questionnaire-based surveys.	EU IVD regulations [*24*]
3.06.04c Usability: Results interpretation study	1. Intended users shall interpret the results of contrived RDTs (e.g. static/pre-made tests) to assess their ability to correctly interpret pre-determined test results. 2. Contrived RDTs shall be made to demonstrate the following potential test results: • Non-reactive; • Range of invalid results; • Reactive; • Weakly reactive. 3. Testing subjects shall consist of:	1. The contrived tests shall be prepared by persons different from those reading the results. The tests shall be randomized prior to the users reading the results.	

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	• At least 15 intended users, including those whose native language may not be the IFU language; • In their usual working environment, not employees of the manufacturer; • From 2 geographically diverse populations to demonstrate correct interpretation of simulated test results.		
3.06.05 Stability of the IVD			
3.06.05.01 Claimed Shelf-life & 3.06.05.03 Shipping stability	1. Stability studies shall be conducted using the conditions expected in the environment of intended use 2. Lots shall be subjected to simulated "transport stress" before real time studies are undertaken on these lots 3. Lots shall be subject to simulated environmental stress conditions (e.g. temperature and humidity) 4. The effects of this simulated transport shall be documented separately and in addition to the real time studies 5. Real time shelf-life studies shall evaluate the storage temperature and humidity range 6. A minimum of 3 lots in final packaging shall be used (see note 1). 7. Testing in triplicate shall be undertaken using a panel of specimens of at least: • 1 negative specimen;	1. The lots used shall be manufactured to validated scale according to finalised protocols, including packaging, labelling, QA, and QC specifications and IFU method: • Each lot shall comprise different production (or manufacturing, purification, etc.) runs of critical reagents and ideally some of the reagents should be near the end of their assigned shelf lives; • The lot numbers of critical reagents and kit components in each lot of RDT shall be documented and reported. 2. If different reagent-container sizes are used in packs with different volumes of reagent (e.g., different volumes for packs with 25 or 50 individual devices), stability evidence (real time, open container, in-use) shall be obtained on all variants, even if the contents of the containers are identical. 1. Flow time and time to band development should be reported.	TGS-2 [*25*] Annex to TGS-2 [*26*] ISO 23640:2011 [*27*] CLSI EP25 [*28*] ASTM D4169-22 [*29*]

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	• 1 weakly reactive specimen (1-2 x LOD (or cut-off if a reader is used)); • 1 medium reactive specimen (2-3 x LOD (or cut-off if a reader is used)). 8. In addition, to address specificity a minimum of 100 negative specimens shall be tested at T=0 and at the end of the claimed shelf life. 9. Stability of labelling shall be determined (see chapter 3.06.04). 10. Lots shall be subject to simulated physical stress conditions (e.g. drop-shock, inversion, vibration, physical handling and stacking).	2. The number of invalid results and repeat testing with each lot shall be reported. 3. Claims for stability shall be based on the second-last successful data point from the least stable lot. 4. Accelerated studies do not replace the need for real time studies. 5. Clinical specimens are the preferred specimen type but with justification, contrived positive specimens (purified MTB LAM spiked into negative matrix) may be used. 6. LAM concentration in clinical specimens may be characterized as weakly/medium reactive according to a calibrated, graduated colour chart or a semi-quantitative scoring system, or quantitative methods.	
3.06.05.02 In-use stability (open pack/open vial)	1. There shall be evidence that once the device 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. 3. Liquid components, once opened, shall have a validated life and number of stated uses under environmental (including microbial) conditions expected. 4. Testing shall be conducted in at least 1 lot. 5. Testing in triplicate shall be undertaken using a panel of specimens of at least: • 1 negative specimen;	1. In-use stability of labile components shall be conducted using components in their final configuration.	

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	• 1 weakly reactive specimen (approx. 1-2 x LOD (or cut-off if a reader is used)); • 1 medium reactivity specimen (2-3 LOD (or cut-off if a reader is used)).		

296

297

298 **Part 2: IMDRF ToC chapter 4: Clinical evidence**

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
4.02.03 Device specific clinical studies			
4.02.03a General requirements for clinical performance studies	1. Clinical sensitivity and specificity shall be determined in accordance with claims made in the IFU. 2. Testing shall be conducted: • By the intended users representing relevant intended use settings (see note 1) • On specimens from all sections of the population for which claims are made (note 2) • Using specimens from different geographical settings (minimum of 3 settings in more than 1 WHO region) • On all claimed specimen types • Using at least 3 lots (see notes 5 and 6). 3. The true TB status of the study subjects shall be determined using at a minimum a microbiological reference standard (see notes 7 and 8). 4. Discrepant, invalid, and unexpected results shall be fully evaluated (see notes 14 to 17). 5. The procedure for selection of study subjects/specimens, how these represent an intended use population and how bias has been addressed shall be clearly described (see notes 2 and 4).	1. RDTs for LAM antigen detection are generally used by a variety of users including trained healthcare workers/lay providers in resource limited settings. This should be considered when preparing evaluation protocols. 2. The inclusion and exclusion criteria shall be clearly stated. 3. Up to 25% of the test specimens may be well-characterized archived specimens that have not undergone more than one freeze-thaw cycle, assuming that the impact of specimen storage has been validated (see 3.05.02). 4. Criteria for the selection of stored specimens shall be explained. Stored samples shall be randomized and blinded for testing (i.e., interspersed with an appropriate number of negative specimens). 5. The product code (not merely a product name), lot numbers and IFU version of index test shall be reported for each clinical site. 6. Approximately half of the specimens shall be tested on different lots at each site. 7. Clinical performance of the index test shall be determined against a microbiological reference standard including culture (sputum and/or other	

4.02.03b Clinical sensitivity	1. Testing of at least 300 confirmed MTBC-positive specimens from individual laboratory confirmed HIV positive study subjects with CD4 counts across the range representative of the intended use population. 2. The majority of specimens shall be freshly collected routine clinical urine specimens handled according to the IFU (see notes 2 and 3). 3. Fresh sputum/other relevant specimens shall be collected at the same time or within a short time interval of the claimed specimen type. 4. Blood shall be collected for CD4 cells count at the same time or within a short time interval of the claimed specimen type.	relevant specimen type) and molecular testing (sputum, concentrated urine, other relevant specimen type). 8. A positive TB case (active tuberculosis disease) is defined as a study subject with a positive culture (sputum /other extrapulmonary specimen) OR a positive molecular test (sputum OR urine OR other relevant extrapulmonary specimen). For subjects who are unable to produce sputum, additional clinical assessment (chest X-ray, biochemical markers) should be considered and justified. 9. Bacterial culture shall be followed by identification of the bacterial species in the positive culture using a molecular method approved by a stringent regulatory authority allowing to discriminate between MTBC and non-tuberculous mycobacteria. 10. The methods and specimen types used for culture and molecular testing shall be specified. 11. Estimates of clinical sensitivity and specificity shall be reported with 95% confidence intervals. 12. Clinical sensitivity and specificity shall be calculated for each specimen type and not for the aggregated data. 13. Study subjects should be classified, and results analysed according to • Presence of signs and symptoms compatible with TB • CD4 cell count (i.e. 0-100 cells/mm3, 101-200 cell /mm3, more than 200 cells/mm3); • Setting (i.e., inpatient, outpatient).	
4.02.03c Clinical specificity	1. Testing of at least 500 confirmed MTBC-negative specimens from individual laboratory confirmed HIV positive study subjects. 2. The majority of specimens shall be freshly collected routine clinical specimens handled according to the IFU (see notes 2 and 3).		

		14. Discrepant results should be resolved as much as possible, however performance characteristics shall be based on the original result. 15. Problematic specimens including those with unexpected results, but which otherwise meet selection criteria for the study, shall not be excluded from analysis. 16. Inconclusive results shall not be excluded from the denominator data for analysis. 17. All invalid test results shall be recorded and analysed separately in the final performance calculation.	

299

F. Source documents

1. World Health Organization. (2023). Instructions for compilation of a product dossier – IMDRF ToC. Prequalification of in vitro diagnostics. World Health Organization; https://iris.who.int/handle/10665/375773 Licence: CC BY-NC-SA 3.0 IGO
2. Technical Guidance Series: WHO prequalification of in vitro diagnostics (IVDs) (website) https://extranet.who.int/prequal/vitro-diagnostics/guidance-documents. Accessed 27 February 2024.
3. World Health Organization. (2024). WHO consolidated guidelines on tuberculosis: module 3: diagnosis: rapid diagnostics for tuberculosis detection, 3rd ed. World Health Organization. https://iris.who.int/handle/10665/376221. License: CC BY-NC-SA 3.0 IGO
4. World Health Organization. (2024). WHO operational handbook on tuberculosis: module 3: diagnosis: rapid diagnostics for tuberculosis detection: web annex A: information sheets, 3rd ed. World Health Organization. https://iris.who.int/handle/10665/376284. License: CC BY-NC-SA 3.0 IGO
5. World Health Organization. (2014). High priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting, 28-29 April 2014, Geneva, Switzerland. World Health Organization. https://iris.who.int/handle/10665/135617
6. World Health Organization. (2018). Compendium of WHO guidelines and associated standards: ensuring optimum delivery of the cascade of care for patients with tuberculosis, 2nd ed. World Health Organization. https://iris.who.int/handle/10665/272644. License: CC BY-NC-SA 3.0 IGO
7. World Health Organization. (2016). Reportable changes to a WHO prequalified in vitro diagnostic medical device. World Health Organization https://apps.who.int/iris/handle/10665/251915. License: CC BY-NC-SA 3.0 IGO
8. In Vitro Diagnostic Medical Device Market Authorization Table of Contents (IVD MA ToC). IMDRF/RPS WG/N13FINAL:2019 (Edition 3). International Medical Device Regulators Forum; 2019. https://www.imdrf.org/documents/vitro-diagnostic-medical-device-market-authorization-table-contents-ivd-ma-toc Accessed 15 January 2024.
9. World Health Organization. (2017). Principles of performance studies (TGS-3). World Health Organization. https://apps.who.int/iris/handle/10665/258985. License: CC BY-NC-SA 3.0 IGO
10. CLSI. EP35. Assessment of equivalence or suitability of specimen types for medical laboratory measurement procedures, first edition. Wayne, PA: Clinical and Laboratory Standards Institute; 2019. EP35Ed1 | Assessment of Equivalence or Suitability of Specimen Types for Medical Laboratory Measurement Procedures, 1st Edition (clsi.org)
11. ISO 17511:2020 In vitro diagnostic medical devices - Requirements for establishing metrological traceability of values assigned to calibrators, trueness control materials and human samples Geneva, Switzerland; 2020. ISO 17511:2020(en), In vitro diagnostic medical devices — Requirements for establishing metrological traceability of values assigned to calibrators, trueness control materials and human samples
12. EN 13612:2002. Performance evaluation of in vitro diagnostic medical devices. Brussels, Belgium; 2002. https://doi.org/10.3403/02556457
13. CLSI. EP12 Evaluation of qualitative, binary output examination performance, third edition. Wayne, PA: Clinical and Laboratory Standards Institute; 2023. EP12Ed3 | Evaluation of Qualitative, Binary Output Examination Performance, 3rd Edition (clsi.org)

14. 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. February 2020. https://www.fda.gov/media/109574/download
15. CLSI. EP17-A2 Evaluation of detection capability for clinical laboratory measurement procedures; approved guideline- second edition. Wayne, PA: Clinical and Laboratory Standards Institute; 2012. EP17A2 | Evaluation of Detection Capability for Clinical Laboratory Measurement Procedures, 2nd Edition (clsi.org)
16. Commission implementing regulation (EU) 2022/1107 of 4 July 2022 laying down common specifications for certain class D in vitro diagnostic medical devices in accordance with Regulation (EU) 2017/746 of the European Parliament and of the Council. O. J. E. U. 2022 L178/3 Implementing regulation - 2022/1107 - EN - EUR-Lex (europa.eu)
17. CLSI. EP07 Interference testing in clinical chemistry - third edition. CLSI EP07-A3. Wayne, PA; 2018. EP07 | Interference Testing in Clinical Chemistry (clsi.org)
18. CLSI. EP37 Supplemental tables for interference testing in clinical chemistry, first edition. CLSI EP37. Wayne, PA; 2018. EP37 | Supplemental Tables for Interference Testing in Clinical Chemistry (clsi.org)
19. ISO 14971:2019 Medical devices — Application of risk management to medical devices. Geneva: International Organization for Standardization; 2019 ISO 14971:2019 - Medical devices — Application of risk management to medical devices
20. World Health Organization. (2017). Panels for quality assurance and quality control of in vitro diagnostic medical devices (TGS-6). World Health Organization. https://iris.who.int/handle/10665/259408. License: CC BY-NC-SA 3.0 IGO
21. U.S. Food and Drug Administration Center for Devices and Radiological Health. Applying human factors and usability engineering to medical devices. Guidance for Industry and Food and Drug Administration Staff. February 2016. https://www.fda.gov/media/80481/download
22. IEC 62366-1:2015+AMD1:2020 CSV Consolidated version. Medical devices - Part 1: Application of usability engineering to medical devices. 2015. IEC 62366-1:2015+AMD1:2020 CSV | IEC Webstore
23. 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 February 2020 https://www.fda.gov/media/109582/download
24. 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. Regulation - 2017/746 - EN - Medical Device Regulation - EUR-Lex (europa.eu)
25. World Health Organization. (2019). Establishing stability of in vitro diagnostic medical devices. World Health Organization. https://apps.who.int/iris/handle/10665/259742. License: CC BY-NC-SA 3.0 IGO
26. World Health Organization. (2019). Establishing component stability for in vitro diagnostic medical devices: annex to TGS-2. World Health Organization. https://apps.who.int/iris/handle/10665/311345. License: CC BY-NC-SA 3.0 IGO
27. ISO 23640. In vitro diagnostic medical devices - Evaluation of stability of in vitro diagnostic reagents. Geneva: International Organization for Standardization; 2011. ISO 23640:2011 - In vitro diagnostic medical devices — Evaluation of stability of in vitro diagnostic reagents
28. CLSI. EP25 Evaluation of stability of in vitro medical laboratory test reagents, second edition. Wayne, PA: Clinical and Laboratory Standards Institute; 2023 EP25Ed2 | Evaluation of Stability of In Vitro Medical Laboratory Test Reagents, 2nd Edition (clsi.org)

29. ASTM. D4169-22, Standard practice for performance testing of shipping containers and systems, ASTM International, West Conshohocken, PA, 2022 D4169 Standard Practice for Performance Testing of Shipping Containers and Systems (astm.org)