TSS-24

In vitro diagnostic medical devices used for the qualitative detection of *Neisseria gonorrhoeae, Chlamydia trachomatis* and *Trichomonas vaginalis* nucleic acid

Technical specifications series for submission to WHO prequalification – diagnostic assessment

World Health Organization

TSS-24 In vitro diagnostic medical devices used for the qualitative detection of *Neisseria gonorrhoeae, Chlamydia trachomatis* and *Trichomonas vaginalis* nucleic acid

Technical specifications series for submission to WHO prequalification – diagnostic assessment

World Health Organization

Contents

Acknowledgements

Acknowledgements are due to the many experts whose contributions made this publication possible. The document was prepared in collaboration with E. Kersh, CDC, Atlanta, United States of America (USA); E. Tagliani, D. Healy and U. Ströher, Prequalification Unit – In Vitro Diagnostic Assessment Team, World Health Organization (WHO), Geneva; and technical and programmatic input from R. Peters and T. Wi, Global HIV, Hepatitis and STIs 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 was held from 20 to 23 August 2024.

Meeting participants: Sarah Alexander, United Kingdom Health Security Agency, London, United Kingdom; Manju Bala, Regional STD Teaching, Training and Research Centre, VMMC and Safdarjang Hospital, New Delhi, India;, Maria Luiza Bazzo, Federal University of Santa Catarina, Florianópolis, Brazil; Louise Causer, Kirby Institute, UNSW Sydney, Sydney, Australia; Tania Crucitti, Institute Pasteur, Antananarivo, Madagascar; Julian Duncan, London, United Kingdom; Amina El Kettani, Ministry of Health, Rabat, Morocco; Cecilia Ferreyra, FIND, Geneva, Switzerland; Matthew Hamill, Johns Hopkins University School of Medicine, Baltimore, USA, Celine Hayden, IVD Division, Health Canada, British Columbia, Canada;, Lauren Johnson, Division of Microbiology Devices, CDRH, FDA, Silver Spring, USA[1]; Ellen Kersh, Laboratory Reference and Research Branch, CDC, Atlanta, USA; Mohammed Majam, Ezintsha, Johannesburg, South Africa, Philippe Mayaud, London School of Hygiene & Tropical Medicine, London, United Kingdom; Enoch Nyarko, Food and Drugs Authority (FDA), Accra, Ghana; Richard Saunders, London, United Kingdom; Heiner Scheiblauer, Paul-Ehrlich Institut, Langen, Germany; Melanie Taylor, Division of HIV Prevention, CDC, Phoenix, USA; Barbara Van Der Pol, Infectious Diseases STD Laboratory, University of Alabama, Birmingham, USA; Catherine Villegas, MSF, Geneva, Switzerland; Cheng Wang, Guangdong Center for Skin Diseases and STI Control, Guangzhou, People's Republic of China.

WHO secretariat: Remco Peters and Teodora Wi, Global HIV, Hepatitis and Sexually Transmitted Infections Programmes;, Susie Braniff, Deirdre Healy, Anne-Laure Page, Irena Prat and Ute Ströher, WHO Regulation and Prequalification Department.

This document has been developed with support the Bill & Melinda Gates Foundation.

[1] Via teleconference

Declarations of interests

All participants completed a Declaration of Interests form in advance of the meeting. Six of the participants declared interest in the topic under consideration. Louise Causer, Cecilia Ferreyra, Philippe Mayaud, Matthew Hamill, Barbara Van der Pol and Julian Duncan declared significant interests connected with their (previous) employment and/or ongoing research support for manufacturers of STI diagnostics. It could not be excluded that the declared interests may be perceived as a potential conflict of interest. Therefore, while the above mentioned persons 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

ATCC	American Type Culture Collection
ANOVA	analysis of variance
CLSI	Clinical and Laboratory Standards Institute
Ct	cycle threshold
CT	*Chlamydia trachomatis*
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
LOD	limit of detection
NA	nucleic acid
NAT	nucleic acid technology
NG	*Neisseria gonorrhoeae*
POC	point of care
ROC	receiver operated curve
RPS WG	regulatory product submission working group (an IMDRF working group)
TGS	Technical guidance series
TSS	Technical specifications series
TV	*Trichomonas vaginalis*
US FDA	U.S. 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 that intend to seek WHO prequalification of qualitative, nucleic acid technology (NAT) for the detection of *Neisseria gonorrhoeae* (NG), *Chlamydia trachomatis* (CT) and *Trichomonas vaginalis* (TV) in a variety of clinical specimens including but not limited to urine, vaginal swabs, cervical/endocervical swabs, liquid PAP smears, urethral swabs, anorectal swabs, penile meatal swab and oropharyngeal swabs. This includes devices detecting one specific organism as well as devices that may detect two or more organisms in both symptomatic and asymptomatic individuals. Devices detecting additional organisms to the STIs listed above or including markers for antimicrobial resistance are not in the scope of this document.

For the purpose of 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.

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 (Part 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 purpose of WHO prequalification, NATs for the detection of NG, CT and TV (separately or combined) shall comply with the specifications in Part 1 and Part 2 of this document.

The submission of the dossier shall 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 the following WHO prequalification documents and WHO Global HIV, Hepatitis and Sexually Transmitted Infections programme guidelines and policies:

- Instructions for compilation of a product dossier (referred to as WHO document PQDx_018). (*1*)
- Technical guidance series for WHO prequalification – diagnostic assessment. (*2*)
- Laboratory and point-of-care diagnostic testing for sexually transmitted infections, including HIV; 2023. (*3*)
- The diagnostics landscape for sexually transmitted infections; 2023. (*4*)
- Consolidated guidelines on HIV, viral hepatitis and STI prevention, diagnosis, treatment and care for key populations (*5*)
- Guidelines for the management of symptomatic sexually transmitted infections; 2021. (*6*)
- FIND/WHO Target product profile for a rapid, low-cost diagnostic to distinguish gonorrhoea from Chlamydia infection at primary care. (*7*)

D. Performance principles for WHO prequalification

D.1 Intended use

An IVD submitted for WHO prequalification assessment 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, isothermal methods of nucleic acid amplification (iNAAT), transcription mediated amplification, helicase-dependent amplification, cross-priming amplification, etc.);
- what the IVD medical device detects (e.g., DNA of CT/NG/TV or rRNA of NG/CT/TV)
- Its function (e.g., screening, diagnosis, or aid to diagnosis);
- the specific disorder, condition or risk factor of interest that it is intended to detect, define or differentiate;
- whether or not it includes automated components or is intended to be used with automated instruments;
- what the IVD medical device reports (i.e., qualitative test);
- the type of specimen(s) required (e.g., urine, vaginal swabs, endocervical swabs, penile meatal and/or anorectal swabs);
- the collection method (e.g., patient-collected vaginal swabs, clinician-collected endocervical swabs);
- target population (e.g., sexually active population, pregnant people, etc.);

- the intended use setting and the intended user (e.g., laboratory professionals trained in the techniques of molecular IVDs, trained healthcare professionals in a healthcare, community setting or at point of care (POC)[2];
- any limitations to the intended use (e.g., identification or restrictions regarding age groups for testing or self-collection or other limiting characteristics).

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 that are claimed in the instructions for use (IFU).

Prequalified NAT IVDs for NG, CT, and/or TV in low- and middle-income countries are likely to be used by a range of users in different geographical regions:

- Laboratory professionals[3] either in centralised testing laboratories or at/near POC[4],
- Laboratory professionals in health care settings not experienced in nucleic acid testing,
- Health professionals/health care workers trained in the use of the test at or 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 (including where self-collection methods are indicated). 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, is 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

[2] Point-of-care (POC) in-vitro diagnostic testing refers to decentralized testing that is performed by a minimally trained healthcare professional near a patient and outside of laboratory testing facilities. It does not refer just to sample collection procedures. In some jurisdictions (e.g., European Union), the concept "near patient testing" is used instead of "point of care testing". Either term may be used in the intended use statement.

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

[4] Near POC require a small laboratory area for testing and their turnaround-times tend to be longer than some of the "true" POC tests (taken from Gonorrhea point of-care diagnostics technology and market Landscape report, Unitaid, 2024 (Section 4.4.5, page 24).

has been demonstrated (see WHO document PQDx_121 Reportable Changes to a WHO Prequalified In Vitro Diagnostic Medical Device) (*8*). If the method section of the IFU has been changed in any way, both the study 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 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(s) 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 numbers 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 a manufacturer's responsibility to ensure, via risk analysis of its IVD that the minimum numbers of lots chosen for estimating performance characteristics considers the variability in performance likely to arise from the interlot diversity of critical components and their formulation or from changes that could occur during the assigned shelf life of the IVD. Differences found between lots during the analytical and clinical performance studies shall be reported.

All instrumentation required when running the assay (from specimen processing to result interpretation) shall be specified and validated for the product under review.

For the purpose of this document the clinical status is defined as presence or absence of NG, CT, and TV infection and shall be determined using a state-of-the-art molecular reference method. For WHO purposes this should be a NAT assay that currently is 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). Justification for the choice of method shall be provided.

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. The cause of invalid results should be reported if available, such as specimen issues (e.g. age of the specimen, storage conditions), instrument error or operator error. Data should be presented in a clear and understandable format. 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. Studies which may fall in this category are indicated in the appropriate chapters of part 1.

If the validation of specimens (chapter 3.05.02) shows equivalency between specimen types, some analytical performance studies (as indicated in this document) may use a representative specimen type only.

Clinical performance studies shall be based on testing specimens only sourced from population cohorts reflective of the intended use population. Independent of the outcome of the equivalency study (chapter 3.05.02), all claimed specimen types need to be validated in the clinical performance study. 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 (e.g. including age of the specimen, temperature logs, freeze-thaw cycles if applicable);
- the stability of the nucleic acid target;
- selection bias.

Studies that comprise testing of left-over specimens by research and development staff at a manufacturer's facility shall not, on their own, be considered sufficient to meet the clinical performance study requirements outlined in this document. For analytical validation studies, where indicated, it is acceptable to use *C. trachomatis* (CT) and *N. gonorrhoeae* (NG) bacteria (***9***) or *T. vaginalis* (TV) parasites spiked into negative clinical matrix.

E. Table of Requirements

WHO requires that a product dossier is 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) (*10*). 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., 3.05.04, 3.05.09 etc). This has been done to maintain consistency between chapters 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
3.05.05a	Limit of detection
3.05.05b	Inclusivity
3.05.06	Analytical specificity
3.05.06a	Potentially interfering substances (endogenous and exogenous)
3.05.06b	Cross-reactivity
3.05.06c	Microbial interference
3.05.06d	Competitive interference
3.05.09	Validation of assay cut-off
3.05.10	Validation of the assay procedure
3.05.10a	Validation of the primer and probe choice
3.05.10b	Procedural control
3.05.10c	Whole system failure rate
3.06	Other studies
3.06.02	Software/firmware/programmed or programmable medical devices
3.06.02.08	Software verification and validation
3.06.04	Usability/human factors
3.06.04a	Flex studies/robustness
3.06.04b	Usability: Label comprehension study
3.06.04c	Usability: Result interpretation study
3.06.04d	Carry-over contamination
3.06.05	Stability of the IVD
3.06.05.01	Claimed shelf-life
3.06.05.02	In use stability
3.06.05.03	Shipping 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

162

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 Analytical Performance			
3.05.01 Stability of specimen(s)			
Specimen collection, storage and transport	1. Identify the different specimen types (e.g., vaginal, cervical/endocervical, anorectal, urethral, OP swabs or urine) that can be used with the IVD. 2. Real time specimen stability studies shall be conducted for each claimed specimen type, including swab type and transport/storage media that is specified by the manufacturer considering: • Storage conditions (for example, duration at different temperatures, temperature limits, freeze/thaw cycles etc.). • Transport conditions. • Intended use (see note 1). 3. The testing panel shall contain (see notes 3, 4) • A minimum of 10 discrete weak positive specimens approximately 1 – 2 x limit of detection (LOD) (see note 2) for each specimen type and target organism. 4. Testing shall be conducted in 1 lot.	1. Evidence shall be provided which validates the maximum allowable time between specimen collection and its processing or addition to the IVD in the setting where testing takes place. 2. The LOD is defined as the lowest bacterial or protozoan concentration in a sample volume that can be detected in 95% of tests (chapter 3.05.05). The concentration of the panel specimens shall be determined relative to the empirical LOD of the IVD under evaluation and the reference material utilized in the analytical validation studies. This applies to all studies in Part 1. 3. Contrived specimens prepared by spiking negative clinical matrix with quantified (genome copies/mL) representative strains of each target organism may be used. 4. In case the use of archived specimens is considered for chapter 4.02.03 of this document, evidence of stability in the conditions in which the specimens have been stored shall be demonstrated e.g., by re-testing a subset of specimens prior to and after storage using a suitable NAT to verify that the same result is obtained. 5. Acceptance criteria shall confirm that claimed specimen types transported, processed, and stored under recommended conditions will give expected results. • The manufacturer shall define what the acceptable deviation is when reporting their results in the study report.	CLSI MM13 (*10*)

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
		6. 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 period stated in the IFU.	
3.05.02 Validation of specimens			
Matrix effect	The relationship between IVD performance in claimed specimen types shall be established: 1. At least 25 positive and 25 negative specimens shall be tested for each claimed specimen type (see notes 1-3). 2. 1 replicate of each specimen shall be tested and the results compared between matrices. 3. Testing shall be conducted in 1 lot.	1. If multiple specimen types are claimed (e.g., vaginal, endocervical, anorectal swabs or urine) a matrix equivalency study should be conducted to establish the relationship between specimen type and IVD performance. 2. Specimens should be chosen that have low to moderate concentrations of each target organism. 3. Contrived specimens prepared by spiking negative clinical matrix from different individuals with quantified (genome copies/mL) representative strains of each target organism may be used. 4. The established relationship between IVD performance in claimed specimen types shall be considered in the design of subsequent studies. For example, if a manufacturer can demonstrate equivalency between $\geq$ 2 matrices or specimen types, only 1 representative specimen type/matrix needs to be tested in the following analytical studies: chapter 3.05.04 etc.	
Demonstration of equivalence between specimen	1. For each claimed collection device or method (e.g., swabs transported in liquid, dry swabs, liquid PAP smear) and/or claimed liquid collection media (specify brand), as appropriate, equivalence of performance shall be demonstrated.	1. The specimen collection device employed may differ depending on the specimen type. 2. Paired specimens shall be used.	

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
collection methods	2. Testing of: • 1 negative specimen • 1 low positive (1-2 x LOD) • 1 moderately positive specimen (3-5 x LOD) • 1 high positive specimen (7-10 x LOD) • In 5 replicates • Using 1 lot	3. For instruments that generate a visual result only, performance may be compared by testing serial dilutions of clinical CT/NG/TV positive specimens with a targeted level of analyte. 4. If equivalency is demonstrated between different collection devices or collection media, only 1 representative method needs to be tested in further analytical studies.	
3.05.03 Metrological traceability of calibrators and control material values			
Metrological traceability of calibrators and control material values	1. As applicable, the metrological traceability of the provided control and calibration material(s) to a certified reference material or a secondary material calibrated from it shall be determined	1. For any reference materials a detailed report from the supplier shall be provided.	
3.05.04 Accuracy of measurement			
3.05.04.02 Precision	1. Both repeatability and reproducibility (see note 1 and 2) shall be estimated using panels with defined bacterial/protozoan concentrations of each target organism. 2. The members of the repeatability and reproducibility testing panel shall include (see note 5): • 1 negative specimen • 1 low positive specimen (1-2 x LOD) • 1 high positive specimen (7-10 x LOD).	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. The testing panel should be the same for all operators, lots, and sites. 3. A run is defined depending on the IVD's throughput: if the platform can accommodate all specimens in a single run, i.e. in the same test	CLSI. EP05-A3 (*12*) CLSI EP12 (*13*)

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
	3. Testing shall be conducted in all claimed specimen types unless equivalency between different matrixes has been demonstrated (chapter 3.05.02).	plate, the specimens will be run together. If the assay can only accommodate a smaller set or a single specimen(s), a run will be defined as a testing session carried out on the same instrument/module on the same day. 4. Precision shall be determined utilizing the entire test system (specimen processing, nucleic acid extraction, detection). If more than one instrument is recommended for use in the IFU, precision shall be determined in all instruments. 5. Contrived specimens prepared by spiking negative clinical matrix with quantified representative strains of each target organism may be used. 6. Each lot shall comprise different production (or manufacturing, purification, etc.) runs of critical reagents, representative of routine manufacture. 7. To understand irregularities in results obtained, at least 2 lots should be tested at each of the 3 testing sites. 8. If operators are considered a significant source of test result variation (for example, with tests that have a significant proportion of manual manipulations or POC assays), then at least 2 intended users/site shall be used. 9. The number of invalid tests shall be reported. 10. Results shall be statistically analysed by ANOVA or other methods to identify and isolate the sources and extent of any variance if applicable.	
Repeatability	1. For **repeatability,** a panel of specimens composed of CT/NG/TV organisms (as applicable) spiked into the appropriate matrix shall be prepared using different combinations of negative, low, and high analyte concentrations. 2. Each panel member shall be tested: • in 2 replicate • over 20 days (not necessarily consecutive) with 1 to 2 runs per day • at 1 site		
Reproducibility	1. For **reproducibility**, each panel member containing individual organisms (e.g., CT or NG or TV, as applicable) as per point 2 above, shall be tested: • in 5 replicates • using 3 different lots (see notes 6, 7) • over 5 days (not necessarily consecutive) with 1 run in that day (alternating morning/afternoon) (see note 3). • at each of 3 different testing sites (see note 7). • Using 1 operator/site (see note 8).		

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
	• By operators representative of intended users. • Unassisted. • Using only those materials provided with the IVD (e.g. IFU, labels and other instructional material). 2. For low throughput assays multiple instruments per site may be used	11. The percentage of correctly-identified, incorrectly-identified and invalid results shall be tabulated for each specimen and be separately stratified according to each of site, lot, etc. 12. For instruments which generate Ct values (numerical values), • The coefficient of variation (CV) shall be calculated by analysing the underlying numeric values of the qualitative results (e.g., Ct values). • Defining an acceptable CV as <5% is strongly encouraged. • Acceptance criteria shall be defined that describe the maximum amount by which the underlying numeric values of the qualitative result (e.g., Ct value) can deviate before acceptable performance is said to be affected. 13. For instruments that generate a visual result only: • Results shall be reported as the proportion of specimens detected and in addition, graded band intensity results (if applicable). • Adding the proportion of hit rate per concentration level to the table can help clarify whether the expected hit rates were seen and which factors need investigation. 14. Widely disparate hit rates shall be investigated and explained.	

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
3.05.05 Analytical sensitivity			
3.05.05a Limit of detection	1. The LOD shall be estimated as the lowest concentration of target organism(s) detectable 95% of the time. 2. Testing 20-24 replicates of at least 8 serial dilutions of suitable biological reference materials (see note 2). The serial dilutions shall be chosen so that the microorganism concentrations span the LOD of the IVD (see notes 3). 3. The replicate testing shall be conducted: • On three different days (see note 4). • Using 2 lots. • At least 2 dilution series shall be tested. 4. The LOD shall be determined in all claimed specimen types	1. The LOD of the IVD shall be determined utilizing the entire test procedure from specimen processing, nucleic acid extraction, to detection. 2. At least two representative/reference strains of each target organism (e.g., for CT: D-K serovars, for TV: metronidazole sensitive and resistant strains) spiked into natural clinical matrix shall be used. 3. Results shall be reported in genome copies/mL 4. LOD shall be estimated by determining the 95% LOD with 95% confidence intervals (e.g. by probit analysis or an appropriate alternative). 5. For low through-put instruments, the number of testing days may be increased.	CLSI EP12 (*14*) CLSI EP17-A2 (*14*)
3.05.05b Inclusivity	1. The capacity of the IVD to detect clinically relevant and geographically diverse CT serovars (e.g. D-K serovars), representative NG strains (e.g. WHO reference strains) and/or TV protozoan strains (e.g., metronidazole resistant and sensitive TV strains, TV virus positive strains) should be demonstrated (see notes 1 and 2). 2. The potential impact of genetic variations on IVD performance shall be evaluated by testing dilution of the specimens: • at low concentration (1 – 2 x LOD)	1. The serovar (CT), and strain culture collection information/number (NG, CT and TV) shall be provided. 2. Testing of purified nucleic acid is only allowed in case of demonstrated unavailability of the whole organism.	WHO Reference strains (*9*)

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
	• at medium concentration (3 – 5 x LOD) • 3 replicates/dilution 3. The entire test procedure shall be utilized with the most challenging claimed specimen type. 4. Laboratory testing and *in silico* analysis shall be conducted.		
3.05.06 Analytical specificity			
3.05.06a Potentially interfering substances and medical conditions	The potential for false results (false negative and false positive results) arising from interference from at least, but not limited to, the substances/conditions listed below shall be determined (see note 1, 2, 3): 1. By testing negative specimens (see note 6) and low/weak positive specimens (negative specimens spiked with low concentrations (1 - 2 x LOD) of each target organism. 2. A minimum of 100 specimens shall be tested. 3. In triplicate.	1. The risk assessment conducted for an IVD shall identify substances at medically relevant levels for which the potential for interference can reasonably be expected for the analyte being detected 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 2. By conducting 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 3. 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 documented in any submissions to WHO and taken into account in the risk-benefit statements. 4. Any observed interference shall be further investigated and performance limitations of the IVD reported in the IFU 5. Results shall be reported with respect to each substance and not be reported as an aggregate of the total number of specimens tested in the study.	ISO 14971:2019 (*15*) CLSI EP07 (*16*) CLSI EP37 (*17*)
Endogenous and Exogenous	The interference of endogenous and exogenous substances likely found in the claimed specimen types/matrixes on the performance of the device shall be investigated. 1. Endogenous and exogenous substances shall be spiked at the highest levels found in individuals. 2. A list of the interfering substances tested, and the concentrations used shall be provided. 3. Substances to be tested include:		

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
	• mucin • human blood • lubricant gels • contraceptives (gel, film, foam) • powder deodorants • vaginal anti-itch cream • topical treatment for yeast infection • aciclovir • metronidazole • bilirubin (urine) • mouthwash (pharyngeal swab matrix) • cough medicines (pharyngeal swab matrix)	6. Prior to spiking, samples should be confirmed to be negative for each target organism prior to testing with a suitable PCR assay.	
3.05.06b Cross-reactivity	1. The manufacturer shall determine the potential for false positive results arising from cross-reactivity with: • Near neighbour species/strains (e.g. *N. meningitidis*, *N. lactamica* and *N. cinerea*). • Predominant normal microbiota that may be present in each of the claimed specimen types. • Microorganisms that may be present in each of the claimed specimen types. 2. *In silico* analysis shall be performed for all microorganisms. Any potential cross-reactivity identified through in silico analysis must be reported (see notes 4-6)	1. The risk assessment conducted for an IVD shall identify relevant microorganisms for which the potential for cross-reactivity can reasonably be expected for the analyte being detected and anatomical site/s in the areas of intended use 2. By conducting appropriate risk assessment, testing can be conducted on specimens spiked with the microorganisms identified as likely to be significant and testing of potentially irrelevant microorganisms avoided 3. Under some circumstances stringent risk evaluation may eliminate the requirement to test some of the items in the lists but any such	

<table>
<tr><th>IMDRF ToC chapter heading/aspect</th><th>Testing requirements</th><th>Notes on testing requirements</th><th>Source Documents</th></tr>
<tr>
<td></td>
<td>
3. Recommended microorganisms for cross-reactivity testing include (see note 1):
<table>
<tr><td>Achromobacter xerosis</td><td>Fannyhessae vaginae</td><td>Neisseria flava</td></tr>
<tr><td>Acinetobacter calcoaceticus</td><td>Flavobacterium meningosepticum</td><td>Neisseria flavescens</td></tr>
<tr><td>Acinetobacter Iwoffi</td><td>Fusobacterium nucleatum</td><td>Neisseria lactamica</td></tr>
<tr><td>Actinomyces israelii</td><td>Gardnerella vaginalis</td><td>Neisseria mucosa</td></tr>
<tr><td>Actinomyces pyogenes</td><td>Gemella haemolysans</td><td>Neisseria perflava</td></tr>
<tr><td>Aerococcus viridans</td><td>Giardia intestinalis</td><td>Neisseria polysaccharea</td></tr>
<tr><td>Aeromonas hydrophila</td><td>Haemophilus ducreyi</td><td>Neisseria sicca</td></tr>
<tr><td>Agrobacterium radiobacter</td><td>Haemophilus influenzae</td><td>Neisseria subflava</td></tr>
<tr><td>Alcaligenes faecalis</td><td>Herpes simplex virus I</td><td>Paracoccus denitrificans</td></tr>
<tr><td>Bacillus subtilis</td><td>Herpes simplex virus II</td><td>Peptostreptococcus anaerobius</td></tr>
<tr><td>Bacteriodes fragilis</td><td>Human papilloma virus 16</td><td>Peptostreptococcus productus</td></tr>
<tr><td>Bacteriodes ureolyticus</td><td>Kingella dentrificans</td><td>Plesiomonas shigelloides</td></tr>
</table>
</td>
<td>
decision shall be documented in any submissions to WHO and taken into account in the risk-benefit statements.
4. In silico analysis shall include multiple representative strains from GenBank sequence database for each organism.
5. The full sequence of each organism shall be analysed.
6. For all in-silico analyses, the accession numbers of the nucleotide sequences shall be provided.
7. Non-clinical matrix may be spiked with the organism of interest to a high concentration (a minimum 10^6 colony forming units/mL or 10^5 genome copies/mL for bacteria, 10^5 plaque forming units/mL for viruses, 10^6 cells/mL for parasites and yeasts).
8. Any observed cross-reactivity shall be further investigated and performance limitations of the IVD reported in the IFU.
9. Omissions from actual laboratory testing shall be supported by a well-documented justification that includes a due diligence attempt to obtain the organisms (and/or purified nucleic acid).
</td>
<td></td>
</tr>
</table>

IMDRF ToC chapter heading/aspect	Testing requirements			Notes on testing requirements	Source Documents
	Bifidobacterium adolescentis	*Kingella kingae*	*Prevotella spp*		
	Bifidobacterium brevi	*Klebsiella oxytoca*	*Propionibacterium acnes*		
	Branhamella catarrhalis	*Klebsiella pneumoniae*	*Proteus mirabilis*		
	Brevibacterium linens	*Lactobacillus acidophilus*	*Proteus vulgaris*		
	Campylobacter jejuni	*Lactobacillus brevis*	*Providencia stuartii*		
	Candida albicans	*Lactobacillus crispatus*	*Pseudomonas aeruginosa*		
	Candida glabrata	*Lactobacillus gasseri*	*Pseudomonas fluorescens*		
	Candida parapsilosis	*Lactobacillus iners*	*Pseudomonas putida*		
	Candida tropicalis	*Lactobacillus jensenii*	*Rahnella aquatilis*		
	Chlamydia pneumoniae	*Lactobacillus lactis*	*Rhodospirillum rubrum*		
	Chromobacterium violaceum	*Legionella pneumophila*	*Saccharomyces cerevisiae*		
	Citrobacter freundii	*Leuconostoc paramensenteroides*	*Salmonella minnesota*		
	Clostridium perfringens	*Listeria monocytogenes*	*Salmonella typhimurium*		

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents

Corynebacterium genitalium	*Micrococcus luteus*	*Serratia marcescens*
Corynebacterium xerosis	*Moraxella lacunata*	*Staphylococcus saprophyticus*
Cryptococcus neoformans	*Moraxella osloensis*	*Staphylococcus aureus*
Cytomegalovirus	*Morganella morganii*	*Staphylococcus epidermidis*
Deinococcus radiodurans	*Mycobacterium smegmatis*	*Streptococcus agalactiae*
Derxia gummosa	*Mycoplasma genitalium*	*Streptococcus bovis*
Dientamoeba fragilis	*Mycoplasma hominis*	*Streptococcus mitis*
Eikenella corrodens	*N. meningitidis* Serogroup A	*Streptococcus mutans*
Enterobacter aerogenes	*N. meningitidis* Serogroup B	*Streptococcus pneumoniae*
Enterobacter cloacae	*N. meningitidis* Serogroup C	*Streptococcus pyogenes*
Entercoccus avium	*N. meningitidis* Serogroup D	*Streptococcus salivarius*
Entercoccus faecalis	*N. meningitidis* Serogroup Y	*Streptococcus sanguis*
Entercoccus faecium	*N. meningitidis* Serogroup W135	*Streptomyces griseinus*

<table>
<tr><th>IMDRF ToC chapter heading/aspect</th><th>Testing requirements</th><th>Notes on testing requirements</th><th>Source Documents</th></tr>
<tr><td></td><td>
<table>
<tr><td>Erwinia herbicola</td><td>Neisseria cinerea</td><td>Trichomonas tenax</td></tr>
<tr><td>Erysipelothrix rhusiopathiae</td><td>Nesseria dentrificans</td><td>Ureaplasma urealyticum</td></tr>
<tr><td>Escherichia coli</td><td>Neisseria elongata</td><td></td></tr>
</table>
4. Samples shall be tested in triplicate
5. Using one claimed specimen type</td><td></td><td></td></tr>
<tr><td>3.05.06c Microbial interference</td><td>If in silico analysis reveals ≥80% homology between one of the primers or the probe to any sequence, there could be interference with amplification of the target gene(s) (even in the absence of cross-reactivity). In this case, the following study shall be considered (see note 2):
1. A microbial interference study with CT, NG, TV and the microorganisms that the test primer or probe have homology to.
2. Specimens shall be spiked at a low (3 x LOD) CT/NG/TV concentration and a high interferent level, to represent the worst-case scenario, with a minimum of 3 replicates.
3. If interference is observed at the level tested, an additional titration study should be performed to determine the highest microorganism interferent level the CT/NG test can tolerate.</td><td>1. Microbial interference studies aim at demonstrating that false negatives for the target organism/s will not occur in presence of other microorganisms.
2. Otherwise explain why the in silico results are irrelevant.</td><td></td></tr>
</table>

<table>
<tr><th>IMDRF ToC chapter heading/aspect</th><th>Testing requirements</th><th>Notes on testing requirements</th><th>Source Documents</th></tr>
<tr><td>3.05.06d Competitive interference</td><td>1. For devices that simultaneously detect two or more target organisms (e.g., CT/NG, CT/NG/TV) the manufacturer shall evaluate competitive interference
2. Testing may be conducted in stepwise manner, firstly by testing targets in the following combination (if applicable):
<table><tr><th>CT concentration</th><th>NG concentration</th><th>TV concentration</th></tr><tr><td>Low (at LOD)</td><td>High at least 10^6 cells/mL</td><td>Negative</td></tr><tr><td>Low (at LOD)</td><td>Negative</td><td>High at least 10^6 cells/mL</td></tr><tr><td>Negative</td><td>Low (at LOD)</td><td>High at least 10^6 cells/mL</td></tr><tr><td>High at least 10^6 cells/mL</td><td>Low (at LOD)</td><td>Negative</td></tr><tr><td>High at least 10^6 cells/mL</td><td>Negative</td><td>Low (at LOD)</td></tr><tr><td>Negative</td><td>High at least 10^6 cells/mL</td><td>Low (at LOD)</td></tr></table>3. Testing shall be performed in 5 replicates.</td><td>1. The manufacturer should assess the effect of all known clinically relevant co-infections that the test detects. If a target/analyte is expected to be present at a high level, the detection of another target present at low levels could potentially be impaired.</td><td>CLSI MM17 (18)</td></tr>
</table>

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
	4. If any interference is observed, additional testing of the targets combined, at a concentration near the individual LOD shall be performed. 5. Testing shall be conducted in one claimed specimen type		
3.05.09 Validation of assay cut-off			
Validation of assay cut-off	1. The assay cut-off shall be validated by testing the following testing panel (see note 1, 2 and 3): • 100 of each CT, NG or TV positive clinical specimens representative of low, medium and high bacterial loads (see note 1). • 500 CT/NG/TV negative specimens. • The testing panel shall include 10 positive and 10 negative specimens close to the cut-off for the target organisms. • The testing panel should contain different CT serovars, NG and/or TV strains to the extent possible. • Urine and vaginal swabs shall be tested (see note 4). 2. The manufacturer shall justify the positioning of the cut-off and describe the algorithm/method used to set the cut-off for the test, or in cases where the cut-off is set for each run or set of tests, the manufacturer shall describe the algorithm/method specified in the IFU or used by the instrument to set the cut-off.	1. Test specimens chosen shall display a range of numerical values (e.g. Ct values) that are representative of routine clinical cases. Specimens shall be characterized using a suitable state of the art NAT. 2. 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 the specimens used to validate the cut-off (so that the two processes are independent). 4. If there is a claim for oropharyngeal swab testing, then the assay cut-off shall also be validated for this specimen type.	

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
3.05.10 Validation of the assay procedure			
3.05.10a Validation of primer and probe choice	1. Evidence supporting the choice of critical reagents (primers and probes sequences) shall be provided.	1. The sequence of the primers and probes shall be provided. 2. A rationale for selection of primers and probes including target genes and specific sequences used shall be provided, including: • Justification for alignments made to generate consensus sequences or best-fit modifications made to existent 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.	
3.05.10b Procedural control	1. The product shall include at a minimum an exogenous internal control. Evidence for the acceptable range of underlying numerical values (e.g. Ct values) shall be provided. 2. If an endogenous internal control (housekeeping gene) is also used as part of the assay design, an acceptable range of underlying numerical values (e.g. Ct values) should be determined for each claimed specimen type (independent of the Ct value for the target organism(s)).	1. The design of the internal control shall be risk based, and a justification provided. 2. A positive control contains a defined amount of target sequence in a suitable matrix. 3. A negative control is a sample of a suitable matrix shown to be negative for the target 4. An exogenous internal control shall be added to each specimen before sample extraction so that all stages of the test, from extraction to final target detection, can be verified. An internal control consists of a defined nontarget sequence of the same type of nucleic acid as the target, which are extracted and amplified simultaneously with the test sample). Therefore, the test should be able to clearly identify and distinguish the amplified products (amplicons) of the internal control and the target	

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
3.05.10c Whole system failure rate	1. The potential for false negative results in low positive specimens shall be determined for each of the claimed specimen types (see note 1). 2. The testing panel shall be randomized and contain 10 contrived specimens for the target organisms (CT/NG/TV) at 1-2 x LOD (see note 2) 3. The testing panel shall be tested: • On 5 consecutive days (to give a total of 50 test results per target organism/specimen type). • Using 1 lot. • With 1 user.	1. The whole system failure rate shall cover the entire assay procedure, from specimen collection to result interpretation. 2. This may be conducted as part of precision studies if the minimum number of replicates are met. 3. Replicate contrived specimens should be prepared using a single specimen diluted in the appropriate matrix.	
3.06 Other Studies			
3.06.02 Software/Firmware/Programmed or programmable medical devices			
3.06.02.08 Software Verification and Validation	1. Software validation reports shall be available for submission if requested (see note 1).	1. Software validation to include as a minimum: • Verification of built-in fail-safe. • Verification of alert mechanisms. • Verification of qualitative/semi-quantitative result detection & interpretation. • Evidence to demonstrate that appropriate error codes are provided to the end user	IEC 62304: 2006/ Amd 1:2015 (*19*) US FDA (*20*, *21)*

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
3.06.04 Usability/human factors studies			
3.06.04a Flex studies/ robustness	1. Evidence is required to demonstrate that the conditions recommended in the IFU are validated and how they were verified. 2. Robustness (flex) studies shall be designed to challenge the system under conditions of stress to identify potential device deficiencies, including failures, and determine the robustness of the product. 3. The influence of the following factors on expected results (both detected and non-detected) shall be considered as applicable (see note 1, 2, 3): • Specimen collection and/or reagent volume (see note 2) • IVD instrument sturdiness (including the effect of non-level work surface) • Lighting, humidity and barometric pressure (simulating high altitude). • Handling contamination (e.g. from latex, powder, hand lotion, sweat, and/or soap, etc.). • Operating temperature. *4.* Instrumentation (both extraction and amplification) including: • Ruggedness (including the effect of vibration from other instruments) (see note 5).	1. Refer to WHO document PQDx_018 for other flex studies that may be relevant, taking into consideration the broad range of operational and environmental conditions consistent with intended use. 2. The risk assessment conducted for an IVD shall identify factors (including intended use setting and self-collection if applicable) which have potential to affect the performance of the assay. 3. The factors should be investigated in ways that not only reflect, but also exceed, likely operating conditions in lower- and middle-income countries so that the limitations of the device 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 and provides an indication of its reliability during normal usage. 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. 7. Contrived specimens prepared by spiking negative clinical matrix with quantified representative strains of each target organism may be used.	PQDx_018 (*1*) US FDA (*22*)

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
	• Impact of dust and mould on componentry (e.g. optics). • Impact of power/voltage fluctuation. 5. Testing to be performed in 1 lot. 6. The specimen panel shall contain: • 1 negative specimen • 1 low positive CT/NG/TV specimen (1 -2 x LOD) • Where different specimen types are claimed, flex studies shall be performed at least on urine and vaginal swabs.		
3.06.04b Usability: Label comprehension (including IFU) study – only applicable for (near) POC devices	1. Testing of subjects to assess ability of intended users to correctly comprehend key messages from packaging and labelling • Understanding key warnings, limitations and/or restrictions, including handling and storage and use of accessories • Proper test procedure • Test result interpretation 2. 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.	1. IFU and labelling should be clear and easy to understand. Use of pictorial instructional material is encouraged.	IEC 62366-1:2015 (*23*) European Parliament IVD regulations (*24*) US FDA (*22*)
3.06.04c Usability: Results	1. Intended users shall interpret a range of test results including control results to assess their ability to correctly	1. Study group may include subjects recruited as part of the label comprehension study	

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
interpretation study – only applicable for (near) POC devices	interpret pre-determined test results and error messages (see note 1 and 2). 2. Testing subjects to consist of at least 15 intended users including those whose native language may not be the language of the IFU.	2. The range of test results undergoing interpretation shall include dual positive (e.g., CT and NG positive) test results.	
3.06.04d Carry-over/ cross-contamination	The potential for false positive results due to carry-over shall be investigated. 1. The testing panel shall contain 40 alternating high-positive ($\geq 10^6$ EB/mL, CFU/mL or cells/mL) and negative CT/NG/TV specimens (see note 1) 2. Only 1 specimen type should be used. The most challenging specimen type shall be chosen as the test specimen. 3. The panel shall be tested: • In at least 5 different runs. • On 3 different days. • With at least 2 users. • Using 1 lot. 4. 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 ($\geq 10^6$ EB/mL, CFU/mL or cells/mL) and negative CT/NG/TV specimens • A total of 10 runs	1. Contrived specimens prepared by spiking negative clinical matrix with quantified representative strains of each target organism may be used. 2. All nucleic acid extraction and PCR platforms specified in the IFU shall be investigated.	Haeckel R (*25*) MM17 (*18*)

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
	• With at least 2 users • Using 1 lot		
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 evaluated for the shelf life of the test kit using the conditions expected in the environment of intended use. 2. Lots shall be subjected to simulated "transport stress", including environmental (e.g. temperature and humidity) and physical stress condition before real time studies are undertaken on these lots. 3. The effect of this simulated transport shall be documented separately and in addition to the real time studies. 4. Real time shelf-life studies shall evaluate the storage temperature and humidity range. 5. All kit configurations shall be tested (or a rationale provided if not) 6. At least 3 lots shall be tested (see note 4, 5) 7. The stability testing panel shall consist of the following contrived specimens (see note 3): • 1 negative specimen • 1 low CT/NG/TV positive specimen (1 – 2 x LOD) • 1 medium CT/NG/TV positive specimen (3 – 5 x LOD) 8. Each panel member shall be tested in triplicate at each time point/condition (see note 8)	1. For each target organism/s, when more than one part of the genome is targeted by primers, each region shall be monitored separately during stability evaluation 2. If the assay contains more than one of each primer/probe combination, then each primer/probe combination needs to be assessed for stability 3. Contrived specimens prepared by spiking negative clinical matrix with quantified representative strains of each target organism may be used 4. Each lot shall comprise different production (or manufacturing, purification, etc.) runs of critical reagents, representative of routine manufacture 5. The number of invalid tests with each kit lot shall be reported 6. 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 majority 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 7. Statistically designed experiments should be involved to allow evaluation of any interactions between environmental conditions.	TGS 2 (*26*) Annex TGS 2 (*27*) ISO 23640:2011 (*28*) CLSI EP25 (*29*) ASTM D4169 (*30*)

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
	9. Testing shall be performed using the most sensitive claimed specimen type. 10. All lots shall be subject to simulated physical stress conditions (e.g. drop-shock, inversion, vibration, physical handling and stacking).	8. The underlying numerical values (e.g. Ct values) of each replicate, condition and time point shall be provided 9. Determination of shipping stability shall be performed using simulated extreme stress conditions, ensuring that application of those conditions is consistent and controlled. 10. Accelerated studies do not replace the need for real time studies 11. Multiple instruments may be used to allow simultaneous testing at each time point	
3.06.05.02 In-use stability (open pack or open vial stability)	1. Minimum of 1 lot shall be tested using a stability testing panel composed of (see note 3): • 1 negative specimen • 1 low CT/NG/TV positive specimen (1 – 2 x LOD) • 1 medium CT/NG/TV positive specimen (3 – 5 x LOD) 2. Replicate testing of each panel member (see note 1) 3. Testing shall be conducted using the most challenging specimen type 4. All labile components (e.g. bulk volume buffers, single use reagent vials, sealed cartridges, control materials etc.) shall be evaluated 5. On-board stability shall be evaluated for an IVD used with an instrument	1. 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 2. In-use stability of labile components shall be conducted using components in their final configuration 3. Contrived specimens prepared by spiking negative clinical matrix with quantified representative strains of each target organism may be used 4. Consideration shall be given to operating temperature, humidity range and allowable freeze-thaw cycles of reagents/controls, as applicable	

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 sensitivity and specificity studies	Testing shall be conducted: 1. On specimens from all sections of the population for which claims are made. 2. Using specimens from • at least 2 different, geographically diverse regions (including at least one LMIC) • by a variety of intended users representing relevant intended use settings and • at different test settings (healthcare/laboratory) 3. Using at least 2 lots (see chapter D.3). 4. The reference method shall include state of the art NATs that detect two different target sequences for NG, CT, and/or TV. 5. Testing of all claimed specimen types is required with the IVD under evaluation (see note 1). 6. Reference testing shall be conducted using first catch male urine and vaginal swabs. For non-genital compartments, a head to head comparison using the claimed specimen type shall be performed. 7. Specimens with discrepant results shall be further evaluated. Where possible, follow-up testing shall be conducted (see note 4).	1. Clinical performance shall be established using specimens that correspond directly to claims made in the IFU 2. Specimens with unexpected results but which otherwise meet selection criteria for a study, shall not be systematically excluded from analysis 3. Problematic specimens, and those specimens with initial discrepant results shall not be excluded from the final analysis 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 by the IVD shall be included in the denominator data for analysis 6. All invalid results shall be recorded and evaluated in comparison to the reference result. 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. Up to 20% of the clinical specimens may be well-characterised archived specimens and the impact of storage/freezing shall be validated in 3.05.01.	

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
	8. The procedure for selection of study subjects, how these represent an intended use population and how bias has been addressed shall be clearly described. A prospective "all comers" study design is strongly recommended (see note 9).	8. Any archived specimens used in the study shall be tested in a randomized, blinded manner interspersed with an appropriate number of negative specimens 9. Criteria for the selection of subjects shall be explained (e.g. testing of consecutive patients). The specific status of the presence of absence of CT/NG/TV infection shall not be known before specimen collection. Subjects meeting the eligibility criteria shall not be excluded 10. Underlying numerical values (e.g., Ct values) for the assay under evaluation (including internal control results) and the reference assay shall be provided 11. Clinical performance study protocols shall describe how results of the IVD under evaluation and the reference method(s) will be established. 12. A clear description of the study design shall be provided. 13. The following basic information shall accompany each tested subject/specimen: • asymptomatic or symptomatic • types of symptoms (if available) • treatment status • gender • specimen type • specimen collection date • collection method and material • professionally-collected or self-collected	
4.02.03b Clinical sensitivity	1. For each targeted organism, a minimum of 100 prospective positive specimens collected from different subjects shall be tested per each claimed specimen type (e.g., vaginal swabs, endocervical swabs, oropharyngeal swabs, rectal swabs, urethral swabs, urine) for a symptomatic claim. 2. An additional 30 prospective positive specimens for each of the target organism(s) shall be tested per claimed specimen type from intended use population without symptoms. 3. The majority of specimens shall be freshly collected among the claimed specimen types and handled according to the IFU (see notes 7 - 9) 4. Where a claim is made for self-collection of specimens, demonstration of clinical performance shall be established (see note 12).		
4.02.03c Clinical specificity	1. At least 500 CT/NG/TV negative specimens collected from symptomatic individuals shall be tested per claimed specimen type.		

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source Documents
	2. Where a claim is made for self-collection of specimens, demonstration of clinical performance shall be established (see note 12).	• Product name, manufacturer and product code of the reference test(s) used 14. Clinical performance shall be stratified by gender, symptom status (symptomatic vs. asymptomatic), and specimen type. 15. Estimates of clinical sensitivity and specificity shall be reported with 95% confidence intervals	

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. (2023). Laboratory and point-of-care diagnostic testing for sexually transmitted infections, including HIV. World Health Organization. https://iris.who.int/handle/10665/374252. License: CC BY-NC-SA 3.0 IGO
4. World Health Organization. (2023). The diagnostics landscape for sexually transmitted infections. World Health Organization. https://iris.who.int/handle/10665/371498. License: CC BY-NC-SA 3.0 IGO
5. World Health Organization. (2022) Consolidated guidelines on HIV, viral hepatitis and STI prevention, diagnosis, treatment and care for key populations. Geneva: World Health Organization. https://iris.who.int/handle/10665/360601. Licence: CC BY-NC-SA 3.0 IGO.
6. World Health Organization. (2021). Guidelines for the management of symptomatic sexually transmitted infections. World Health Organization. https://iris.who.int/handle/10665/342523. License: CC BY-NC-SA 3.0 IGO
7. FIND/WHO Target product profile for a rapid, low-cost diagnostic to distinguish gonorrhoea from Chlamydia infection at primary care. Accessed 25 June 2024 https://cdn.who.int/media/docs/default-source/hq-hiv-hepatitis-and-stis-library/stis/target-product-profile-rapid-diagnostic-distinguish-gonorrhoea-from-chlamydia.pdf?sfvrsn=40ca213e_7
8. 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
9. Unemo M, Sánchez-Busó L, Golparian D, Jacobsson S, Shimuta K, Lan PT, Eyre DW, Cole M, Maatouk I, Wi T, Lahra MM. The novel 2024 WHO Neisseria gonorrhoeae reference strains for global quality assurance of laboratory investigations and superseded WHO N. gonorrhoeae reference strains-phenotypic, genetic and reference genome characterization. J Antimicrob Chemother. 2024 Aug 1;79(8):1885-1899. doi: 10.1093/jac/dkae176. PMID: 38889110; PMCID: PMC11290888.
10. In Vitro Diagnostic Device Regulatory Submission Table of Contents (IVD ToC). IMDRF/RPS WG/N13 FINAL:2024 (Edition 4). Regulated Product Submissions Table of Contents Working Group. International Medical Device Regulators Forum; 2024. https://www.imdrf.org/documents/vitro-diagnostic-medical-device-regulatory-submission-table-contents-ivd-toc Accessed 5 July 2024.
11. CLSI. MM13 Collection, Transport, Preparation, and Storage of Specimens for Molecular Methods, second edition. Clinical and Laboratory Standard Institute; 2020 https://clsi.org/standards/products/molecular-diagnostics/documents/mm13/
12. CLSI. EP05-A3 Evaluation of precision of quantitative measurement procedures, third edition. Wayne, PA: Clinical and Laboratory Standards Institute; 2014 https://clsi.org/standards/products/method-evaluation/documents/ep05/

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. 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)
15. 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
16. CLSI. EP07 Interference testing in clinical chemistry - third edition. CLSI EP07-A3. Wayne, PA; 2018. EP07 | Interference Testing in Clinical Chemistry (clsi.org)
17. CLSI. EP37 Supplemental tables for interference testing in clinical chemistry, first edition. Wayne, PA; 2018. EP37 | Supplemental Tables for Interference Testing in Clinical Chemistry (clsi.org)
18. CLSI. MM17 Validation and verification of multiplex nucleic acid assays, second edition. Wayne, PA; 2018. MM17: Validation of Multiplex Nucleic Acid Assays - CLSI
19. IEC 62304:2006/Amd 1:2015 Medical device software — Software life cycle processes. International Electrotechnical Commission; 2015 IEC 62304:2006/Amd 1:2015 - Medical device software — Software life cycle processes — Amendment 1 (iso.org)
20. U.S FDA General Principles of Software Validation; Final Guidance for Industry and FDA Staff. Issued 11 January 2002. General Principles of Software Validation; Final Guidance for Industry and FDA Staff Accessed 15 December 2023.
21. U.S FDA Content of Premarket Submissions for Device Software Functions Guidance for Industry and Food and Drug Administration Staff. Issued 14 June 2023. Content of Premarket Submissions for Device Software Functions | FDA Accessed 8 February 2024
22. 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
23. 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
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. Haeckel R. Proposals for the description and measurement of carry-over effects in clinical chemistry. Pure Applied Chemistry 1991; 63:301-306. https://doi.org/10.1351/pac199163020301
26. World Health Organization. (2019). Establishing stability of in vitro diagnostic medical devices (TGS-2). World Health Organization. https://apps.who.int/iris/handle/10665/259742. License: CC BY-NC-SA 3.0 IGO
27. 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

28. 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
29. 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)
30. 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)