TSS-27 Syphilis rapid diagnostic tests for professional use and/or self-testing

Technical specifications series for submission to WHO prequalification – diagnostic assessment

World Health Organization

TSS-27 Syphilis rapid diagnostic tests for professional use and/or self-testing

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 T. Crucitti, Institute Pasteur, Antananarivo, Madagascar and J Duncan, London, United Kingdom; D. Healy, E. Tagliani and U. Ströher, Prequalification Unit – In Vitro Diagnostic Assessment Team; 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 Ute Ströher and Irena Prat, Prequalification Unit – In Vitro Diagnostic Assessment Team, WHO headquarters, Geneva, Switzerland.

List of contributors

A technical consultation for the second edition to increase the scope to include Syphilis rapid diagnostic tests for self-testing 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 ; 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 from the Bill & Melinda Gates Foundation and UNITAID.

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 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.

The first edition was prepared in collaboration with T. Crucitti, Centre Pasteur du Cameroun, Yaoundé, Cameroon; R. J. S. Duncan, London, United Kingdom of Great Britain and Northern Ireland; D. Healy, M. Lanigan, U. Ströher, Prequalification Team – Diagnostic Assessment, WHO, Geneva and with technical and programmatic input from M. Taylor and I. Toskin, Department of Reproductive Health and Research, Human Reproduction team, WHO, Geneva. This document was produced under the coordination and supervision of D. Healy, U. Ströher and I. Prat, Prequalification Team – Diagnostic Assessment, WHO, Geneva, Switzerland.

A technical consultation for the first edition of this document was held in Geneva, Switzerland from 20 to 21 August 2018. The meeting participants included: R. Ballard, Atlanta, United States of America (USA); K. Bialas, Scientific Reviewer, Division of Microbiology Devices, Office of In Vitro Diagnostics and Radiological Health, U.S. Food and Drug Administration, Silver Spring, Maryland, USA[1]; S. Braniff, National Serology Reference Laboratory (NRL), Victoria, Melbourne, Australia; C. Bristow, UC San Diego School of Medicine, San Diego, California, USA; A. Chua, Leader, Diagnostics Network Working Group, Médecins Sans Frontières International, Geneva, Switzerland; T. Crucitti, Centre Pasteur du Cameroun, Yaoundé, Cameroon; R. J. S. Duncan, London, United Kingdom of Great Britain and Northern Ireland; N. Gerald, Scientific Reviewer, Division of Microbiology Devices, Office of In Vitro Diagnostics and Radiological Health, U.S. Food and Drug Administration, Silver Spring, Maryland, USA[2]; A. Kijo, Tanzania Food and Drugs Authority (TFDA), Dar Es Salaam, Tanzania; A. L. Page, Epidemiologist, Epicentre, Paris, France; B. Vetter, Foundation for Innovative New Diagnostics (FIND), Geneva, Switzerland. The WHO Secretariat included D. Healy; M. Lanigan; I. Prat; U. Ströher; W. Urassa, Prequalification Team – Diagnostic Assessment Group, Regulation of Medicines and other Health Technologies; M. Taylor and I. Toskin, Department of Reproductive Health and Research, Human Reproduction team.

The first edition draft technical specifications document was posted on the WHO website for public consultation on 19 September 2018 for a period of two months. Various stakeholders, including manufacturers submitting to WHO Prequalification of IVDs, IVD manufacturing industry associations, various national and international regulatory bodies, and IVD standards organizations were informed of the consultation in order to solicit feedback. First edition public comments were received for consideration from the Asian Harmonization Working Party (AHWP) Working group 2, China, Hong Kong SAR; S. Best, Melbourne, Australia; Division of Microbiology Devices, Office of In Vitro Diagnostics and Radiological Health, U.S. Food and Drug Administration, Silver Spring, Maryland, USA; C. Hayden, R&D Manager, bioLytical® Laboratories Inc, Richmond, British Columbia, Canada; In Vitro Diagnostic Devices Evaluation Division, Medical Devices Bureau, Therapeutic Products Directorate, Health Canada, Government of Canada, Canada; ISO/TC 212 Secretary on behalf of ISO/TC 212, International Standards Organization, Geneva, Switzerland; MedTech Europe, Brussels, Belgium; W. Schmandt, Siemens Healthcare Diagnostics Products GmbH, Marburg, Germany.

[1] Participated via web conferencing

[2] Participated via web conferencing

Abbreviations

ANA	anti-nuclear antibodies
CLIA	chemiluminescence Immunoassays
CLSI	Clinical and Laboratory Standards Institute
CV	coefficient of variation
EDTA	ethylenediaminetetraacetic acid
EIA	enzyme immunoassay
IFU	instructions for use
IgG, IgM	Immunoglobulin G, M
ISO	International Standards Organisation
IVD	in vitro diagnostic
LOD	limit of detection
MSM	men who have sex with men
POC	point of care
PrEP/PEP	pre-exposure prophylaxis/post-exposure prophylaxis
RDT	rapid diagnostic test
RPR	rapid plasma reagin
STI	sexual transmitted infection
TPHA	Treponema pallidum haemagglutination assay
TPPA	Treponema pallidum particle agglutination assay
TpN	*T. pallidum* Nichols (polypeptide)
TGS	Technical guidance series
TSS	Technical specification series
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 outlines the minimum analytical and clinical performance studies to be conducted for rapid diagnostic tests (RDTs) used to detect syphilis infection by self-testing or professional use. For the purpose of this document, RDTs are lateral-flow or flow-through immune-chromatographic antibody detection tests, which rely on the capture of dye-labelled antibodies or antigens to produce a visible band or dot on a strip of nitrocellulose, often encased in plastic housing, referred to as cassettes.

The document is relevant to qualitative RDTs that:

- detect antibodies to *T. pallidum* (for professional use and self-testing) or
- detect antibodies to *T. pallidum* and to non-treponemal antigens (for professional use only).

A *T. pallidum* antibody RDT submitted for WHO prequalification is expected to detect all disease stages with the exception of congenital syphilis or neurosyphilis, unless claimed;

A non-treponemal antibody RDT is expected to detect syphilis related reagin antibodies.

The requirements outlined in this document do not include those necessary to demonstrate that the IVD could be used for confirmatory testing, nor the requirements for any accompanying quality control material. However, if quality control material is provided with the assay, it should demonstrate that the IVD is functional and performs as claimed. (7)

Minimum performance requirements for WHO prequalification are summarized in this document and apply equally to RDTs intended solely for the detection of syphilis and to those tests manufactured where syphilis detection comprises one component of a multi-detection system (e.g. a HIV/syphilis dual-detection RDT).

For this document, the verbal forms identified below are defined as follows:

- “shall” indicates that the manufacturer is required to comply with the technical specifications. 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.
- “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.

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 section D and parts 1-3.

Detailed numbers for specimens to be tested are provided for each study in part 1, 2 and 3 of this document. These are the minimum numbers that are necessary to meet WHO prequalification requirements. The final study numbers chosen by the manufacturer need to be evaluated based on the risk assessment of the RDT under evaluation.

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.

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 health care 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, the following requirements apply:

- Part 1, Part 2; antibody detection RDTs for professional use;
- Part 1, Part 2, and Part 3: antibody detection RDTs for self-testing.

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 guidance documents

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

WHO prequalification documents

- Technical guidance series for WHO prequalification – diagnostic assessment; (*2*)
- Instructions for compilation of a product dossier, WHO document PQDx_018. (*1*)
- WHO syphilis guidelines
- Laboratory diagnosis of sexually transmitted infections, including human immunodeficiency virus (*3*)
- WHO guideline on syphilis screening and treatment for pregnant women. (*4*)
- WHO Updated recommendations for the treatment of *Neisseria gonorrhoeae*, *Chlamydia trachomatis* and *Treponema pallidum* (syphilis), and new recommendations on syphilis testing and partner services. (*5*)

C.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:

IVDs for professional testing

- the assay type and what is detected (e.g., lateral-flow assay for the detection of antibodies to *T. pallidum*, and/or detection of non-treponemal antibodies related to syphilis infection);
- the function of the IVD (e.g. screening for surveillance or case management among the sexually active population for symptomatic or asymptomatic *T. pallidum* infections, as an aid to diagnosis of syphilis infection by detection of antibodies to *T. pallidum*; as an aid for detection of non-treponemal antibodies related to syphilis infection);

- the clinical indication of the IVD (e.g. aid in the diagnosis syphilis);
- what the IVD medical device reports (e.g., qualitative test);
- the intended testing population e.g. sexually active population, special populations (e.g. pregnant people, adolescents);
- the intended user (e.g. trained laboratory professional or by health care workers/lay providers);
- the intended operational setting (e.g. for professional use in a laboratory setting, or point of care[3] (POC));
- any limitation to the intended use (e.g. that the antibody test cannot differentiate between active disease and treated infection; exclusion of blood donor screening, neonatal screening, testing of cerebrospinal fluid (neurosyphilis));
- the type of specimen required.

IVD for self-testing

- the assay type and what is detected (e.g., lateral-flow assay for the detection of antibodies to *T. pallidum*)
- the clinical indication of the IVD (e.g., aid in the diagnosis of syphilis infection);
- what the IVD medical device reports (e.g., qualitative test);
- the intended testing population e.g. sexually active population, special populations (e.g. pregnant people, adolescents);
- the intended user (e.g. self-testers, lay users);
- the intended operational setting (e.g., home setting);
- any limitation to the intended use (e.g. individuals with a confirmed history of syphilis, exclusion of blood donor screening, neonatal screening, testing of cerebrospinal fluid (neurosyphilis));
- the type of specimen required (e.g., whole capillary blood).

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

Depending on the intended use of the RDT, analytical and clinical performance studies shall be designed to consider the diversity of knowledge and skills of potential RDT users, and the operational settings in which testing is likely to occur. It is a manufacturer's responsibility to ensure that the risk assessment and subsequent validation studies for an RDT reflect the intended operational settings, including service delivery complexity and involve the user population expected to conduct the test.

Prequalified syphilis RDTs in low- and middle-income countries are likely to be used by

[3] 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.

- laboratory professionals[4] either in centralized testing laboratories (although access may be limited) or at POC;
- healthcare worker or lay providers[5] trained in the use of the test at POC;
- lay users self-testing in their home environment/private setting.

For WHO prequalification submission, clinical performance studies shall be conducted using each specimen type (e.g. capillary whole blood, venous whole blood, serum, plasma) claimed in the instructions for use (IFU). Note that the specimen type that is most likely to be used in resource limited WHO Member States at POC is capillary whole blood. Testing of cerebral spinal fluid is not included in the scope of this document; however, the performance of all specimen types claimed by the manufacturer shall be demonstrated. Laboratory demonstration of equivalence between specimen types without evidence of clinical validation is insufficient (with exception of anticoagulants). For example, studies that comprise testing of left-over or repository specimens by research and development staff at a manufacturer's facility shall not, on their own, be considered sufficient to meet most of the performance requirements for WHO prequalification.

C.3. Applicability of supporting evidence to RDT under review

Analytical and clinical performance studies shall be undertaken using the specific, final (locked-down design) version of the RDT intended to be submitted for WHO prequalification. 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; additional supporting evidence may also be required. This may occur in the case of minor variations to the design where no negative impact on performance has been demonstrated (see WHO document (*6*)).

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 and Part 3 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 should 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 the RDT that the minimum numbers of lots chosen for estimating performance characteristics reflect the variability in performance likely to arise from the interlot diversity of critical components and their formulation or from changes that occur during the assigned shelf-life of the IVD.

Differences found between lots during the analytical and clinical performance studies shall be reported.

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

[5] 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 (taken from World Health Organization. (2020). Consolidated guidelines on HIV testing services, 2019 World Health Organization)

The true *T. pallidum* status of a specimen shall be determined using suitable reference methods for which justification shall be provided; comparison with a similar device is insufficient for resolution of discrepant specimens (e.g. other method from the same manufacturer or other method using the same antigens provided by the same supplier). 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 RDT 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.

Analytical studies shall include testing for all specific characteristic factors (e.g. relevant epitopes) for which detection is claimed. For certain analytical studies it may be acceptable to use contrived specimens (e.g. where non-reactive human specimens have been spiked with those containing analyte specific antibodies). All reasonable attempts should be made to use clinical specimens (unless otherwise stated) and justification should be provided where contrived specimens are used in the submitted studies.

For IVDs that include a claim for detection of multiple analytes on one strip (e.g., syphilis antibodies and HIV antibodies), evidence of performance must be provided for each claimed analyte. It should be noted that data generated for a related single analyte test are not sufficient to support the performance claims of an IVD submitted for prequalification.

For professional use quantitative assays, additional requirements may apply. Contact WHO prequalification for more information on these requirements.

D. 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:2024 (Edition 4) (*7*). 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
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 reading times
3.05.10c	Validation of the controls
3.05.10d	Validation of the control line
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: results interpretation study
3.06.04	Usability – self tests: see part 3
3.06.05	Stability of the IVD
3.06.05.01 &	Claimed shelf-life
3.06.05.03	Shipping stability
3.06.05.02	In use stability
3.08	**Other evidence**
3.08.01	Testing in performance panels and other TSS-specific evidence
3.08.01a	Mixed titre panels
3.08.01b	Performance panels
Part 2: IMDRF ToC chapter 4 – Clinical evidence	
4.02	**Overall clinical evidence summary**
4.02.03	Device specific clinical studies
4.02.03a	General requirements for clinical performance
4.02.03b	Clinical sensitivity
4.02.03c	Clinical specificity
4.07	**Other clinical evidence**
4.07.01	Observed untrained user study - see part 3 (only applies to self-tests)
Part 3: Qualification of usability – applicable for RDTs intended for self-testing	

231	3.06.04d	Label comprehension study (including IFU) for self-tests
232	3.06.04d	Results interpretation study for self-tests
233	4.07.01	Observed untrained user study for self-tests

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. Real time studies accounting for: • storage conditions (duration at different temperatures, temperature limits, freeze/thaw cycles); • transport conditions, where applicable (see note 1); • intended use; • specimen collection and/or transfer devices intended to be used with the RDT.	1. Evidence shall be provided which validates the maximum allowable time between specimen collection, processing of the specimen and its addition to the RDT in the setting where testing takes place. 2. Unless all specimens are expected to be processed as fresh samples within a specified time frame, the RDT performance shall be established for each different storage condition and at the beginning and end of the stated period. 3. In case the use of archived specimens is considered for Part 2 evidence of stability shall be demonstrated.	
3.05.02 Validation of specimens			
Specimen types	1. For each claimed specimen type, testing in paired specimens shall be undertaken in at least: • 50 treponemal positive specimens (see note 1 & note 3); • 50 treponemal negative specimens. 2. If equivalence is claimed between different anticoagulants, testing shall be conducted in at least: • 25 positive specimens of each claimed anticoagulant;	1. Specimens confirmed positive by either treponemal assays enzyme immunoassay (EIA), chemiluminescence immunoassay (CLIA) and/or Treponema pallidum particle agglutination assay (TPPA) or non-treponemal assays (such as rapid plasma reagin (RPR) assays). 2. The relationship between RDT performance in claimed specimen types and materials used for analytical studies shall be established. The design of subsequent	TGS-3 (*8*)

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	• 25 negative specimens of each claimed anticoagulant. 3. The equivalence of specimen types shall be determined for all claimed analytes (e.g. non-treponemal antibodies, *T. pallidum* specific antibodies etc., as appropriate). 4. If there is no equivalence between claimed specimen types, then the impact that this will have on each subsequent performance claim shall be fully understood and described (see note 4, 5). 5. If an RDT is intended for testing whole blood or capillary whole blood and some aspects of performance have been established using serum or plasma specimens, then • the relationship between analytical sensitivity in serum/plasma to that of the same characteristic in whole blood shall be understood (see notes 4,5); • paired specimens shall be used for RDTs intended to test capillary and venous blood (see note 6).	studies shall then take that relationship into account. 3. Positive specimens (undiluted) should be chosen so that the majority is weakly reactive for the respective analyte (near the RDT limit of detection (LOD[6])) and that different stages of infection are included. 4. Where a significant difference in performance exists between specimen types, equivalence may need to be investigated as part of a larger clinical study (See Part 2). 5. Demonstration of the comparability of specimen types may be achieved by comparing RDT results between end-point dilution series of several positive whole blood specimens titrated into compatible (blood group type) whole blood and compared with the serum from those same specimens titrated into serum. 6. All reasonable attempts should be made to use clinical specimens giving responses close to the LOD for capillary and venous blood.	

[6] measured quantity value, obtained by a given measurement procedure, for which the probability of falsely claiming the absence of a component in a material is β, given a probability α of falsely claiming its presence. NOTE 1 IUPAC recommends default values for α and β equal to 0,05. NOTE 2 The term analytical sensitivity is sometimes used to mean detection limit, but such usage is now discouraged. (taken from ISO 18113-1:2022)

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
3.05.03 Metrological traceability of calibrator and control material values			
Metrological traceability of control material values	1. If a control material has an assigned concentration value, the metrological- (not commercial- nor documentary-) traceability to an accepted international standard, if available, shall be demonstrated		PQDx_018 (*1*) ISO 15198:2004 (*9*) ISO 17511:2020 (*10*)
3.05.04 Accuracy of measurement			
3.05.04.02 Precision (repeatability and reproducibility)	1. Both repeatability (within-condition) and reproducibility (between-condition) (See note 1) shall be determined for each analyte for which detection is claimed (e.g. non-treponemal and/or *T. pallidum* antibodies as appropriate). 2. The panel of spiked specimens shall include at least: • 1 non-reactive specimen; • 1 weak reactivity positive specimen (approx. 1-2 x RDT's LOD); • 1 medium reactivity positive specimen (approx. 2-3 x RDT's LOD); • the panel shall include whole blood specimens if claimed. • Each panel member shall be tested: • in at least 5 replicates; • using 3 different lots (see note 6); • over 5 days (not necessarily consecutive) with 1 run per day (alternating morning/afternoon);	1. Within- or between-run, -lot, -day, -site, etc. 2. Studies shall be statistically designed and analysed to identify and isolate the sources and extent of any variance. • The extent of variance (due to manufacturing, test procedures or environment) which could nullify any claim shall be identified and the power of testing shall be sufficient to identify any such variance. 3. Where possible, the testing panel should be the same for all operators, lots, and sites. 4. Each lot shall comprise different production (or manufacturing, purification, etc.) runs of critical reagents. 5. 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 any numerical values.	EN 13612:2002 (*11*) CLSI EP12-A2 (*12*)

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	• at each of 3 different testing sites. (see note 8) 3. The effect of operator-to-operator variation on RDT performance should be included as part of the precision studies (see also note 8). Testing should be done: • by personnel representative of intended users. • unassisted; • using *only* those materials provided with the RDT (e.g. IFU, labels and other instructional materials).	6. To understand irregularities in results obtained, at least 2 lots should be tested at each of the 3 testing sites (3 different lots are required to be tested overall in the 3 testing sites). 7. The effect of operator-to-operator variation on RDT performance may also be considered as a human factor when designing flex studies (see 3.06.04a Flex studies) and may be addressed as part of clinical studies in representative populations (see Part 2). 8. Users shall be selected based on a pre-determined and contextually appropriate level of education, with literacy and auxiliary skills that will challenge the usability of the RDT and reflect the diversity of intended users and operational settings. These characteristics shall be detailed in the submission.	
3.05.05 Analytical sensitivity			
Analytical sensitivity	1. Analytical sensitivity shall be determined using the two approaches outlined below using a minimum of 2 lots (see note 5). Analytical sensitivity shall be established by determining the lowest concentration for which the probability of detection is ≥95%. • Analytical sensitivity shall be determined relative to the available international standards or to secondary standards metrologically traceable to them.	1. For the international standards, the result shall be expressed in international units as analytical end-point sensitivity with its associated metrological uncertainty. 2. The version of the international standard used shall be stated. 3. A widely used test system that is used in large clinical laboratories and in reference laboratories in more than 1 country shall be used.	European common specifications (*13*) CLSI EP17 (*14*) WHO Technical Report Series, No. 1004, 2017 Annex 6. (*15*)

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	2. Analytical sensitivity shall be determined for at least 5 specimens collected from individual patients during the primary phase of syphilis infection in comparison with a widely used test system for that analyte (e.g. TPPA, RPR) (notes 3 and 4).	4. Patients may be selected by testing primary lesions for the presence of *T. pallidum* nucleic acids and confirming subsequent blood specimens for the presence of *T. pallidum* or non-treponemal antibodies. 5. The lots shall comprise different production (or manufacturing, purification, etc.) runs of critical reagents.	
3.05.06 Analytical specificity			
3.05.06a Potentially interfering substances	1. The potential for false results (false-negatives and false-positives) arising from interference from, at least, the substances/conditions listed below shall be determined (see note 1). • a minimum of 100 specimens; • each substance represented by at least 5 to 10 specimens from different individuals. 2. Testing shall be undertaken in both treponemal negative and -positive specimens spiked with each potentially interfering substance at physiologically relevant dosages	1. The risk assessment conducted for the RDT should identify substances/conditions where 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 (and overlook those which might be of relevance). • By conducting appropriate risk assessment, testing can be performed on the substances or conditions identified as likely to be significant and testing of potentially irrelevant substances/conditions can be avoided. 2. Interference studies should be performed with specimens with an analyte response near the RDT LOD. • The methods and concentrations used must be validated so that any effect of clinical importance would be detected.	CLSI EP07-A3 (*16*) CLSI EP37-A (*17*) ISO 14971:2019 (*18*) European common specifications (*13*)
Endogenous	1. Antibody interference • heterophile antibodies such as human antibodies to the expression system (for recombinants), e.g. anti-*Escherichia coli* (anti-*E.coli* positive); • human anti-animal antibodies e.g. anti-mouse;		

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	• autoantibodies including systemic lupus erythematosus, anti-nuclear antibodies (ANA), rheumatoid factor, anti-phospholipid antibodies (Hughes syndrome); 2. High titres of potentially interfering antibodies such as in patients with • recent infection; • immunization; • pregnant and multiparous women. 3. Biochemical interference • haemolysis or haemoglobin; • hyperglobulinaemia; • cholesterol, triglycerides and bilirubin; • sickle-cell disease; • thyroiditis.	3. Any observed interference shall be investigated and performance limitations of the RDT reported in the IFU. 4. 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. • Any effect must be evaluated against the probability of that effect occurring and causing clinically significant issues in the population tested in resource limited settings. 5. Evaluation of endogenous interfering substances may be addressed as part of the clinical studies but the number of specimens of each type evaluated shall be in accord with the requirement in this section.	
Exogenous	1. Relevant medicines, including: • antiparasitic, antibacterial, antimalarial, antiretroviral, antiviral (including for hepatitis C, B, cytomegalovirus) and anti-tuberculosis medications; • common over-the-counter analgesic medications (aspirin, paracetamol).		
3.05.06b Cross-reactivity	1. The potential for false-positive results arising from cross-reactivity (see note 1) shall be determined for a minimum of 100 specimens, including, where possible, at least 5-10 of each (See **Annex 1** for the full list): • viral/bacterial/parasitic infections;	1. The types of conditions/disease tested for shall be risk-based, taking into consideration the operational setting as well as the intended users for the analyte being detected in the areas of intended use and not simply rely on published lists of such cross-reactivity which might be of limited	

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	• sexually transmitted infections; • infections by other spirochaetes (Borrelia, Leptospira, periodontal disease causing spirochaete); • immunization; • other unrelated conditions known to cause cross-reactivity in *T. pallidum* RDTs.	relevance in resource limited settings. (see 3.05.06a, note 1) 2. Cross-reactivity with other *T. pallidum* subspecies causing nonvenereal treponematoses is known and is not required to be demonstrated unless specificity solely to *T. pallidum* is claimed. 3. Any observed cross-reactivity shall be investigated and performance limitations of the RDT reported in the IFU. 3.05.06a, note 3 4. For studies of interference by cytomegalovirus and Epstein-Barr-Virus it is most important to use specimens containing immunoglobulin M (IgM) as these are well known to cause clinically relevant interference while the corresponding immunoglobulin G (IgG) specimens do not.	
3.05.07 High dose hook effect			
High dose hook effect	1. For each claimed analyte the potential for a prozone/high dose hook effect shall be determined: • using multiple, highly-reactive specimens (minimum of 20); • using at least 2 different concentrations (diluted by at least a factor of 10).	1. Specimens shall be chosen that have a high antibody titre as determined using a method other than the RDT intended to be prequalified e.g. using an EIA. This second method shall be of a design not subject to high dose hook effect. 2. An increase in signal upon dilution of a specimen implies a hook effect. 3. At least 3 different lots should be tested.	Butch, A.W. (*19*)

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
3.05.09 Validation of assay cut-off			
Establishment of reader cut-off	1. In *T. pallidum* assays provided with a reader, the way in which the reader has been designed to differentiate positive specimens from negative specimens shall be demonstrated.	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. 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 (refer to section below *3.05.10b Validation of reading times*); • time interval between opening packaging or preparing reagents and starting the assay; • processing steps/timed steps; • volumes, including numbers of drops; • temperatures e.g. operating temperature range (see note 4); • humidity.	1. These parameters may be investigated as part of 3.06.04 Usability/Human factors, below. 2. The intent of parameter validation 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. • "designed experiments" – changing more than one parameter at once – are more appropriate than single changes with all other parameters held constant.	

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	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: • 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)).		
3.05.10b Validation of reading times	1. For RDTs where a reading interval is specified (i.e. time when result can first be read; time beyond which result should not be read), validation of critical time points shall be provided. The study shall use panels of at least: • 1 non-reactive specimen; • 1 weak reactivity positive specimen (approx. 1-2 x RDT's LOD); • 1 medium reactivity positive specimen (approx. 2-3 x RDT's LOD); • the panel shall include whole blood and anticoagulated plasma (e.g. ethylenediamine tetraacetic acid (EDTA)) if claimed. 2. Performance studies shall be conducted at the extremes of the intended operational temperature range; the effect of humidity on reading times shall also be investigated (see note 1).	1. The ranges of temperature and humidity validated shall be risk-based, taking into consideration likely operational settings in resource limited settings. 2. The intended operating temperature range within which the reading time has been validated, shall be clearly stated in the IFU. 3. The studies should take into account possible differences between use of freshly made devices and those stored until near the end of their assigned shelf-lives under the conditions expected in resource limited settings and being used under those conditions. 4. Some of these aspects could be evaluated within the flex studies (3.06.04a)	

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
3.05.10c Validation of controls	1. If control materials (positive controls) are provided, the control materials should be validated as showing that if the RDT would not meet the claims, that the positive control will indicate the failure		
3.05.10d Validation of the control line	1. The flow device shall have a control line. • The nature of the control line shall be explained (see notes 1, 2).	1. The extent to which any control line or dot 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; • specimen addition and flow; • correct volumes being added; • correct operation of the device; • correct functionality of all reagents (e.g. stability).	
3.06 Other studies			
3.06.04 Usability/human factors			
3.06.04a Flex/robustness studies	1. The influence of the following factors on expected results (both reactive and non-reactive) should be considered as below (this list is not exhaustive): 2. Any numerical factor in the IFU method provided and/or identified by risk assessment such as: • specimen and/or reagent volume; • specimen dilution factor;	1. Refer to WHO document PQDx_018 "Instructions for Compilation of a Product Dossier" for other flex studies that may be relevant, taking into consideration the broad range of operational and environmental conditions consistent with intended use in resource limited settings. 2. 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	WHO Prequalification – Diagnostic Assessment PQDx 018 (*1*) TGS-6 (*20*) U.S. FDA CLIA Waiver guidance (*21*, *22*)

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	• operating temperature, pressure and humidity; • time between opening packaging or preparing reagents and starting the assay; • any mixing, rotating or incubating times, temperatures; • reading time: both the time after starting any incubations and the time for which the result is stable. 3. Ruggedness (see note 5): • 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 and durability, adhesiveness; • effects of lighting and humidity (see note 4); • residual volumes and characteristics of liquids (potential evaporation, pH changes, microbial growth, antimicrobial efficacy). 4. Instrumentation (if applicable and based on a risk assessment) including: • ruggedness; • impact of dust and mould on componentry (e.g. optics).	operating temperature with low or high volume of specimen at an incorrect reading time. • The panel to be used should be similar to that used for shelf-life studies, see 3.06.05 below. 3. The factors listed opposite 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 RDT can be understood. For example, in addition to investigating deviations of temperature within those claimed in the IFU, temperature ranges should be investigated that exceed those of claimed operating conditions and which could cause test failure (incorrect/invalid results). 4. The impact of lighting can be twofold – i.e. the impact of lighting on packaging e.g. fading, and the sufficiency of lighting to read the test lines. 5. For the purposes of this document, ruggedness means the ability to resist environmental shocks of a variety of kinds.	
3.06.04b Usability: label comprehension study (including IFU)	1. Testing shall be undertaken to assess the ability of intended users to correctly comprehend key messages from packaging and labelling:	1. Requirements listed may be investigated as separate studies or included as part of clinical studies. 2. Testing may be conducted using questionnaire based surveys.	

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	• understanding key warnings, limitations and/or restrictions; • proper test procedure; • 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.		
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. Contrived tests shall be made to demonstrate the following potential test results: • non-reactive; • range of invalid results; • reactive; • weak reactive. 2. Testing subjects shall consist of: • 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.		

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
3.06.04d Usability: self-testing	For RDTs intended for **self-testing**, please refer to **Part 3**		
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 (temperature, humidity); 2. Replicate testing shall be undertaken using a panel of spiked specimens consisting of at least: • a sufficient number of non-reactive specimens (see note 1); • at least 1 specimen for each analyte and each epitope used or detected by the RDT (approx. 1 - 2x LOD) (see note 2). • A minimum of 3 lots in final packaging; (see note 4). 3. Lots shall be subjected to simulated "transport stress" before real time studies are undertaken on these lots. This mimics the real situation • Transport stress includes physical stress conditions (e.g. drop-shock, inversion, vibration, physical handling and stacking) and environmental stress conditions (temperature, humidity see note 6). 4. If different reagent-container sizes are used in packs with different volumes of reagent (e.g., different volumes for packs with one or 25 or 50 individual devices),	1. The testing panels should include whole blood (to verify correct functioning of the device, such as flow, clearance of debris, lack of autoagglutination, no background interference) and serum or plasma as the specimen matrix. 2. The testing panel shall include members to monitor all claimed critical epitopes, for example TpN47, TpN17 and TpN15 and reagin antibodies if claimed • Each of these epitopes play a role in detecting syphilis in different stages of the infection. It is necessary to have a panel member to monitor each epitope system present (and possibly each stage of infection), even if poly-epitope-fusion proteins are used. This may be avoided if the manufacturer can demonstrate that each epitope system is equally stable within the fusion protein. 3. Specimens to be diluted should represent a range of stages of infection (antibody maturation) to take into account the limitations of mimicking low RDT reactivity by dilution of high avidity specimens. 4. Each lot shall comprise different production (or manufacturing, purification, etc.) lots of critical reagents.	ISO 23640:2011 (*23*) CLSI EP25 (*24*) TGS-2 (*25*) ASTM D4169 (*26*)

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	stability evidence (real time, open container, in-use) shall be obtained on all variants, even if the contents of the containers are identical.	5. The numbers of invalid tests per lot shall be reported. 6. Consideration should be given to cyclic changes of temperature over ranges likely to be met during transport to users in resource limited settings 7. Claims for stability shall be based on the second-last successful data point from the least stable lot. 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. 8. Accelerated studies do not replace the need for real time studies.	
3.06.05.02 In-use stability (open pack or open vial stability)	1. There shall be evidence that once the RDT 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 include all labile components (e.g. buffers, cassettes, etc.) (see note 1); 3. Liquid components, once opened, shall have a validated life and number of stated uses under environmental (including microbial) conditions expected. 4. Minimum of 1 lot, using panel below:	1. In-use stability of all components shall be conducted using components stored in their final configuration. 2. Statistically designed experiments should be involved to allow evaluation of any interactions between environmental conditions 3. Most aspects of in-use stability may be considered as part of “flex” studies (see 3.06.04 Flex studies)	

IMDRF ToC chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	• a sufficient number of non-reactive specimens (see 3.06.05.01 note 1); • at least 1 specimen for each analyte and each epitope used or detected by the RDT (approx. 1 - 2x RDT's LOD).		
3.08 Other evidence			
3.08.01 Testing in performance panels and other TSS-specific evidence			
3.08.01a Mixed titre panels	1. Testing of specimen panels with a range of analyte concentrations (e.g. non-treponemal and *T. pallidum* antibody 'mixed titre' panels)		
3.08.01b Performance panels	1. Testing of the RDT shall be undertaken using, as appropriate: • suitable performance panels which include all claimed critical epitopes, as available (see note 1); • with a minimum of 1 lot.	1. Each epitope plays a role in detecting syphilis in different stages of the infection. It is necessary to monitor each epitope system present (and each stage of infection), even if polyepitope-fusion proteins are used. • This may be done using specimens characterized by a line immunoassay.	

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 Clinical sensitivity and specificity study: general requirements	1. Clinical sensitivity and specificity shall be determined for each claimed specimen type. 2. Testing shall be conducted: • in different geographical settings (minimum of 2 regions) including high and low prevalence settings; • by a variety of intended users (see note 1) in the intended testing settings (e.g. decentralized at point of care use, laboratories in hospital setting); • using at least 2 different lots (see note 3).	1. Prequalified RDTs are generally used by lay providers. For WHO prequalification purposes, these should be considered as the intended user rather than a trained laboratory professional. 2. The performance shall be evaluated on the intended use population(s) in the environment of expected use in resource limited settings. 3. Each of the two lots shall comprise different production (or manufacturing, purification, etc.) runs of critical reagents. 4. Criteria for the selection of archived specimens shall be explained. 5. An effort should be made to include low RPR titre/early stage specimens. 6. Negative and positive archived specimens should be blinded to the user. 7. A separate specimen shall be collected prior to testing to establish the reference result. The laboratory where the clinical evaluation is occurring shall confirm all reactive specimens by the following methods: • Non-treponemal positive specimens confirmed with RPR; • Treponemal positive specimens screened with CLIA or EIA and if positive, confirmed with TPPA.	European common specifications (*13*) TGS-3 (*8*) STARD (*27*)
4.02.03b Clinical sensitivity	1. Testing of at least 500 confirmed positive specimens: • consolidation of results from archived specimen collections and clinical evaluation studies is permissible (note 4); • however, at least 50% of the results from which the diagnostic sensitivity is calculated must be from freshly taken, unfrozen routine specimens of the types claimed (e.g. capillary blood, venous blood, serum, plasma); • at least 50 pregnant women;		

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	• at least 50 specimens from a low prevalence setting.	8. Specimens with results discrepant between the confirmed laboratory result and the RDT under evaluation should be further evaluated: • For *T. pallidum* assays: a state of the art IgG and IgM syphilis immunoassay (LIA/EIA), with positive specimens further evaluated using TPPA/TPHA; • For non-treponemal assays: at least an IgM anti-treponemal immunoassay and a blot-type assay; • All discrepant specimens shall be repeated using the same lot of RDT, and then on all available lots and any variability noted; • Characterization of the donor of a specimen is acceptable evidence in the case of primary syphilis (e.g. detection of organisms by dark field microscopy). 9. Problematic specimens, those with unexpected results but which otherwise meet selection criteria for a study, shall not be excluded from analysis. 10. Performance characteristics shall be reported using initial results only. The results of further testing of specimens with discrepant results shall be reported separately as additional information about RDT performance. 11. All invalid test results and indeterminate results shall be recorded and reported as the invalid rate. 12. Estimates of diagnostic/clinical sensitivity and specificity shall be reported with 95% confidence intervals.	
4.02.03c Clinical specificity	1. The specimens for specificity studies shall be: • as per requirements in 4.02.03a; • unselected, other than as being syphilis negative; • archived specimens shall not exceed 20%; • if the RDT claim is for diagnostic use, blood bank specimens will be insufficient – the expected environment would be STI clinics or POC settings. • At least 1000 specimens shall be tested. • At least 2 different lots (and ideally three) shall be used; • At least 100 pregnant women (to include at least 20 multiparous women).		

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
		13. Results shall be expressed separately for each specimen type and for each specimen type per intended use (no aggregation of results).	
4.07 Other clinical evidence			
4.07.01 Observed untrained user study for self-tests	1. Only applicable to RDTs intended for self-testing, please refer to **Part 3**		

PART 3 Qualification of usability – applicable for RDTs intended for self-testing

Note: The information must be provided in the dossier under the IMDRF ToC Chapter headings indicated in the table below.

Purpose of qualification of usability for self-testing: Assessment of product design, labelling and usability of rapid diagnostic tests (RDTs) for self-testing by analysis of the following:

- Results of a questionnaire to assess whether the key messages and instructions from packaging and labelling are understood and easily followed by untrained intended users (i.e. self-testers) including whether the test is appropriate for use by that individual (understanding the limitations of the assay, for example that intended users understand that a second, professional test is needed to confirm syphilis or acute treponema infection, that the test may be negative in case of recent exposure, that the test is not intended for individuals that had been previously diagnosed with syphilis).
- Results of the interpretation of test-results study in untrained intended users (i.e. self-testers) of simulated RDTs (e.g. pre-made and with contrived results).
- Test results and interpretations when the assay is performed by untrained intended users (i.e. self-testers), in an observed setting.
- Evaluation of the actions taken by lay users performing the entire test procedure, from specimen collection to result interpretation (e.g., an understanding of the meaning of a negative test and that a positive test requires professional confirmation).

Additional points:

- The standard of care protocol at the study site and setting must be used to guide referral of participants to further testing or clinical management as needed. Any self-test participant who receives a reactive result should be linked to further testing and care according to the standard of care protocol at the study site.
- For RDTs that detect more than 1 analyte (e.g., HIV antibodies and syphilis antibodies), an appropriate study design shall be chosen. In the tables below, the number of study participants required for results interpretation and labelling comprehension studies for combined RDTs are specified. Additional interpretation parameters may also be required depending on the combined RDT based on the risk assessment (e.g., correct understanding of the test result outcome in HIV positive individuals on ART when using a combined HIV/syphilis RDT)
- For each of the studies outlined below the study group shall represent the intended testing population/intended users (sexually active people, e.g., MSM, sex workers, PrEP/PEP users, adolescents, pregnant people) and comprise untrained subjects whose age,

level of education, literacy, geography (urban, rural), can challenge the usability and performance of the IVD. Deviations from the requirements below must be justified and addressed in a risk assessment.

- Manufacturers are encouraged to conduct a human factors study prior to starting the full validation.
- These assessment activities will determine the changes needed to optimize the IVD for use by self-testers. Changes may range from minor (simplification of instructions for use without change to the method) to major (e.g. change to the method of specimen collection). The impact of any change on safety and performance shall be determined for both professional and lay use.
- Results from any one of the studies outlined below may indicate that assay redesign is necessary. This may in turn result in a need to revalidate the IVD or to perform additional specific performance studies and to update the risk analysis.
- Use by lay users may necessitate changes in the packaging, including changes to volumes of liquids, the printing on the outer and inner packaging and the number of devices and hence to the size of the individual primary and secondary packaging. In these cases, the effect on stability and transport ruggedness shall be assessed (refer to Part 1: Analytical performance and other evidence).

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
a) 3.06.04d Label comprehension study (including IFU) for self-tests	1. Testing (e.g. questionnaire-based) shall be undertaken to assess the ability of intended users to correctly comprehend key messages from packaging and labelling using only the information available to all users (e.g. IFU, job aids, online video) (note 1, 2). 2. The following parameters shall be investigated: • correct self-selection (whether users understand if it is appropriate for them to undertake testing) • understanding key warnings, limitations and/or restrictions to the use of the test;	1. Requirements listed here may be investigated as a separate study or combined with b) results interpretation studies for self-testing. 2. The intended user for self-testing shall represent the intended testing population, (i.e., sexually active population including sex workers, MSM, PrEP/PEP users, pregnant people, adolescents). 3. Some of the study group shall be naive to self-testing with finger-prick capillary blood. • Subjects who are trained in laboratory procedures/have experience with laboratory techniques shall be excluded.	EU IVD Regulation (*28*) Backinger C (*29*) U.S. NIH (*30*)

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	• correct test procedure; • test result interpretation. • appropriate follow up actions after a reactive or non-reactive result. 3. Studies shall include (see notes 2, 3 and 6): • for single analyte RDTs, at least 200 subjects representative of all intended users or at least 300 subjects for dual testing (HIV/syphilis) • including subjects whose native language is not the language of the IFU; • in their usual working or home environment and not employees of the manufacturer; • from 2 geographically diverse populations to demonstrate comprehension of the key messages by each user group.	• Subjects who use capillary blood glucose meters shall be excluded. 4. Instructions for use and labelling shall be clear and easy to understand; • Instruction material shall include pictures, quick guides or job aids; Use of electronically accessible instructional material is encouraged. 5. Label comprehension shall include the comprehensive evaluation of the actions the user identifies to be taken following each possible test result e.g.: • Do the intended users understand that if they have had syphilis, they should not use this test? They should seek medical care. • Do the intended users understand that if they have HIV, they should not use combined HIV/syphilis test? They should consider a syphilis only test. • Do the intended users understand that a positive test result means they might have syphilis and that they should seek medical care? • Do the intended users understand that a positive test means that confirmatory testing is required? • Do the intended users identify that a negative test does not necessarily mean they do not have syphilis? They should test again and/or seek medical care.	
b) 3.06.04d Usability: Results interpretation study for self-tests	1. Intended users shall interpret the results of contrived RDTs (e.g., static/pre-made devices) to assess their ability to visually read and correctly assign pre-determined test results. 2. Contrived RDTs shall be made to demonstrate the following potential test results (see note 7): • non-reactive;		

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	• range of invalid results including failure of the viewing area to clear correctly, broken, indistinct or absent control lines; • reactive; • weak reactive (the colour intensity of the line shall be faint and resemble a real clinical test line with reactivity close to that of the assay LOD). 3. Testing shall consist of the following subjects (see notes 2, 3 and 6): • for single analyte RDTs: at least 200 subjects • for combined detection (HIV/syphilis): at least 300 subjects • including those whose native language is not the IFU language; • in their usual working or home environment, and not employees of the manufacturer; • from 2 geographically diverse populations to demonstrate correct interpretation of simulated test results.	• Do the intended users understand they should not use this test to monitor treatment? 6. The study group for both the label comprehension and results interpretation study • Shall include intended users (notes 2, 3) • Shall include diverse demographic profiles (see additional points in the introductory section of Part 3) • Should include any other characteristics that can challenge the usability of the IVD by the intended users (e.g., small font size for subjects with vision impairment). • The relative percentage of each group of intended users shall be reported. 7. For RDTs intended to detect both syphilis and HIV antibodies, all potential test results shall be evaluated for each analyte individually and in combination.	
c) 4.07.01 Observed untrained user study for self-tests	1. Testing by at least 900 subjects from 2 diverse areas with different demographics, including at least one LMIC (see note 1): • at least 40 self-testing subjects who are reactive on the RDT (see note 2); 2. Each subject shall self-collect the test specimen and perform the test using only those materials	1. Study subjects shall not be the same as those who have participated in the usability label comprehension/results interpretation studies for self-tests. 2. Subjects with a positive reference result shall be linked to clinical management.	

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	provided with the IVD (e.g. IFU, labels and other instructional materials (see note 3, 5). 3. The self-testing shall be observed by a trained laboratory or health care professional. The observing professional shall not tutor nor interact with the subject conducting the test but shall note errors and other observations about the self-tester (see notes 4). 4. The concordance between the subject's self-test result and interpretation of the same result by the trained professional (observer) shall be reported. • The observing professional shall interpret the user's test result in a blinded fashion and within the validated reading time stated in the instructions for use. 5. A second sample shall be collected using the same means of collection as the self-testing subject and tested on the same RDT under evaluation by a laboratory professional with extensive experience using RDTs. • Taking the sample from a different finger. • Concordance between professional result and self-testing result on the RDT shall be reported. 6. A third specimen (venous blood) shall be collected for reference testing to allow the estimation of the clinical performance of the	3. Accessories and components provided with the self-test kit shall be used in the observed untrained user study • The laboratory professional performing the study shall not replace the lancets or sample transfer devices. 4. Particular attention should be paid to documenting the subjects' compliance with each of the factors raised during a documented risk assessment (ISO 14971) of the self-testing process, e.g. but not limited to: • paying attention to the instructions before starting; • checking that the IVD is within the expiration date; • correct sampling technique and preparation of specimen once collected (e.g. cleansing of sampling site, any massage required, application and subsequent safe disposal of sampling devices); • application of correct volumes to the IVD; • use of a timing device to follow the stated reading time as in the IFU; • safe disposal of the RDT and accessories (e.g. swabs, liquids, lancets, transfer devices); • correct interpretation of the result, including of any invalid test. 5. Study subjects shall not be provided additional training by observing how health care professionals collect the specimen for the comparator test (collect the specimen for the	

IMDRF ToC Chapter heading/aspect	Testing requirements	Notes on testing requirements	Source documents
	self-test and for participant care purposes. See reference test requirements in 4.02.03 note 7.	professional test *after* the subject has self-sampled). 6. Manufacturer are encouraged to conduct the study in a setting that reflects the intended use setting (i.e., home environment) in a small subset of participants.	

272

E. Annex 1: List of conditions that should be evaluated for cross-reactivity.

Viral infections

- HIV;
- hepatitis B infection;
- hepatitis C infection;
- acute hepatitis A infection;
- acute cytomegalovirus (IgM);
- acute Epstein-Barr virus (IgM)

Sexually transmitted infections

- herpes simplex virus 2;
- Chlamydia trachomatis;
- human papillomavirus;
- trichomoniasis

Bacteria/parasites

- infections caused by genus *Borrelia* (Lyme disease);
- malaria;
- visceral leishmaniosis;
- tuberculosis;
- brucellosis;
- leptospirosis;
- leprosy

Immunity/autoimmunity

- lupus erythematosus

Immunization

- mRNA-vaccine-induced false-seropositivity;
- Influenza-vaccine-induced false-seropositivity;
- HIV-vaccine-induced false-seropositivity.

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. https://extranet.who.int/prequal/vitro-diagnostics/technical-guidance-series

3. Laboratory diagnosis of sexually transmitted infections, including human immunodeficiency virus

4. WHO guideline on syphilis screening and treatment for pregnant women https://www.who.int/publications/i/item/9789241550093

5. World Health Organization. (2024). Updated recommendations for the treatment of Neisseria gonorrhoeae, Chlamydia trachomatis and Treponema pallidum (syphilis), and new recommendations on syphilis testing and partner services. World Health Organization; https://iris.who.int/bitstream/handle/10665/378213/9789240090767-eng.pdf Licence: CC BY-NC-SA 3.0 IGO.

6. 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

7. 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.

8. World Health Organization. (2017). Principles of performance studies. World Health Organization. https://apps.who.int/iris/handle/10665/258985. License: CC BY-NC-SA 3.0 IGO

9. ISO 15198:2004 Clinical laboratory medicine -- In vitro diagnostic medical devices -- Validation of user quality control procedures by the manufacturer. Geneva: International Organization for Standardization; 2004.

10. 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

11. EN 13612:2002. Performance evaluation of in vitro diagnostic medical devices. Brussels, Belgium; 2002. https://doi.org/10.3403/02556457

12. User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline – Second Edition. CLSI EP12-A2. Wayne, PA: Clinical and Laboratory Standards Institute; 2008.

13. 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 http://data.europa.eu/eli/reg_impl/2022/1107/oj Accessed 17 July 2024 ;

14. Evaluation of Detection Capability for Clinical Laboratory Measurement Procedures; Approved Guideline—Second Edition CLSI EP17-A2. Wayne, PA: Clinical and Laboratory Standards Institute; 2012.

15. WHO Technical Report Series, No. 1004, 2017. Annex 6. WHO manual for the preparation of secondary reference materials for in vitro diagnostic assays designed for infectious disease nucleic acid or antigen detection: calibration to WHO International Standards. Geneva, World Health Organization; 2016.

16. Interference Testing in Clinical Chemistry. Approved Guideline - Second Edition. CLSI EP07-A3. Wayne, PA: Clinical and Laboratory Standards Institute; 2018.

17. Supplemental Tables for Interference Testing in Clinical Chemistry, 1st Edition. CLSI EP37. Wayne, PA: Clinical and Laboratory Standards Institute; 2018.

18. ISO 14971:2019 Medical devices -- Application of risk management to medical devices. Geneva: International Organization for Standardization;2019.

19. Butch AW. Dilution protocols for detection of hook effects/prozone phenomenon. Clin Chem. 2000; 46:1719-1721.

20. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment: TGS-06 Panels for quality assurance and quality control of in vitro diagnostic medical devices. Geneva: World Health Organization; 2017 https://apps.who.int/iris/bitstream/handle/10665/259408/WHO-EMP-RHT-PQT-2017.10-eng.pdf?sequence=1

21. U.S. Food and Drug Administration Center for Devices and Radiological Health. Recommendations for Dual 510(k) and CLIA Waiver by Application Studies. Draft Guidance for Industry and Food and Drug Administration Staff. Issued November 29, 2018.

22. U.S. Food and Drug Administration Center for Devices and Radiological Health. Select Updates for Recommendations for Clinical Laboratory Improvement Amendments of 1988 (CLIA) Waiver Applications for Manufacturers of In Vitro Diagnostic Devices. Draft Guidance for Industry and Food and Drug Administration Staff. Issued November 29, 2018.

23. ISO 23640:2011. In vitro diagnostic medical devices - Evaluation of Stability of In Vitro Diagnostic Reagents. Geneva: International Organization for Standardization; 2011.

24. Evaluation of Stability of In Vitro Diagnostic Reagents; Approved Guideline. CLSI EP25-B. Wayne, PA: Clinical and Laboratory Standards Institute; 2009.

25. Technical Guidance Series for WHO Prequalification – Diagnostic Assessment. Establishing stability of an in vitro diagnostic for WHO Prequalification, TGS–2. Geneva: World Health Organization; 2016 http://apps.who.int/iris/bitstream/handle/10665/259742/WHO-BS-2017.2304-eng.pdf?sequence=1 accessed 20 November 2018.

26. Standard Practice for Performance Testing of Shipping Containers and Systems. ASTM D4169-14. West Conshohocken, PA: ASTM International; 2014.

27. Cohn J.F. et al., STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open. 2016 6 e012799. doi: 10.1136/bmjopen-2016-012799

28. Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices. O. J. E. U. 2017 L 117/176 http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2017:117:TOC Seen 4 April 2024

29. Backinger C, Kingsley P. Recommendations for Developing User Instruction Manuals for Medical Devices Used in Home Health Care. HHS Publication FDA 93-4258. Silver Spring, MD: Food and Drug Administration; 1993. Recommendations for Developing

User Instruction Manuals for Medical Devices Used in Home Health Care (qualysinnova.com) Accessed 21 June 2023

30. U.S. National Institutes of Health (U.S. Access Board) Best Practices for the Design of Accessible COVID-19 Home Tests, 2023, https://www.access-board.gov/tad/radx/ Accessed 4 April 2024