WHO PREQUALIFICATION TEAM:
DIAGNOSTICS

World Health Organization

Technical Specifications Series for submission to WHO Prequalification – Diagnostic Assessment

TSS-1

Human Immunodeficiency Virus (HIV) rapid diagnostic tests for professional use and/or self-testing

Technical Specifications Series for submission to WHO Prequalification – Diagnostic Assessment: Human Immunodeficiency Virus (HIV) rapid diagnostic tests for professional use and/or self-testing

ISBN 978-92-4-151174-2

Suggested citation. Technical Specifications Series for submission to WHO Prequalification – Diagnostic Assessment: Human Immunodeficiency Virus (HIV) rapid diagnostic tests for professional use and/or self-testing. Geneva: World Health Organization; 2016. Licence: CC BY-NC-SA 3.0 IGO.

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

Table of contents

Acknowledgements

The document TSS-1 Technical specifications for WHO prequalification of Human Immunodeficiency Virus (HIV) rapid diagnostic tests for professional use and/or self-testing was developed with support from the Bill & Melinda Gates Foundation and UNITAID. The first draft was prepared in collaboration with Dr M Lanigan, Geneva, Switzerland and R Meurant, WHO and with input and expertise from Dr H Scheiblauer, Paul Ehrlich Institute, Langen, Germany; D Healy; H Ardura, Dr R Baggaley; Dr C Case; Dr C Figueroa; D Healy; Dr M Nuebling; and A Sands, WHO, Geneva, Switzerland. This document was produced under the coordination and supervision of Robyn Meurant and Irena Prat, Prequalification Team, WHO, Geneva, Switzerland.

List of contributors

First round of public comment participants include the following:
S Best, National Serology Reference Laboratory, Victoria, Australia; Dr T Crucitti, Institute of Tropical Medicine, Antwerp, Belgium; Dr J Duncan, London, United Kingdom; Dr F Gruszka, E-Meddia, Paris, France; Dr C Hill, CA, USA; G James, Reference Laboratories Section, Centre for Infectious Diseases and Microbiology Laboratory Services Department, Institute for Clinical Pathology and Medical Research (ICPMR), Westmead, New South Wales, Australia; Dr L Kestens, Institute of Tropical Medicine, Antwerp, Belgium; D Lepine, IVDD Section, Medical Devices Bureau Health Canada, Ottawa, Canada; Dr A Reissinger, Paul Ehrlich Institut, Langen, Germany; Dr M Nuebling and M Perez Gonzalez, WHO, Geneva, Switzerland.

The draft technical specifications document has been posted on the WHO website for public consultation on 15 September 2016. 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. A two month response period was provided.

Second round of public comments were received for consideration from the following:
Dr P Akolkar, Center for Biologics Evaluation and Research, United States Food and Drug Administration, MD, USA; S Best, National Serology Reference Laboratory, Victoria, Australia; Dr J Duncan, London, United Kingdom; R Galli, bioLytical™ Laboratories Inc., Vancouver, Canada; D Lepine, IVDD Section, Medical Devices Bureau Health Canada, Ottawa, Canada; Dr M Nuebling; Dr H Scheiblauer, Paul Ehrlich Institut, Langen, Germany; STAR consortium members: M Taegtmeyer, V Watson, R Dacombe, E Cowan, M Newman, M Kumwenda, R Peeling, LSTM, Pembroke Place, Liverpool, United Kingdom; OraSure Technologies, Inc, PA, USA; C Kosack, Laboratory Working Group Leader, A Page, Epidemiologist at Epicentre, E Tran, International Coordinator for MD/IVD quality, Médecins sans Frontières International Office, Geneva, Switzerland.

A. Introduction

The purpose of this document is to provide technical guidance to in vitro diagnostic medical device (IVD) manufacturers that intend to seek WHO prequalification of rapid diagnostic tests (RDTs) for the detection of Human immunodeficiency virus (HIV).

Minimum performance requirements for prequalification are summarized herein, and apply equally to RDTs intended solely for HIV detection, and to those tests where HIV detection comprises one component of a multi-detection assay (e.g. a HIV/syphilis dual-detection RDT). This document applies to RDTs the intended use of which is as an aid to diagnosis of HIV infection. The current version of this document does not address IVDs that discriminate between detection of HIV-1 and HIV-2 infection, IVDs intended as confirmatory tests, or the requirements for accompanying quality control material. These additional requirements will be addressed in a later version of this document.

Where possible, WHO performance conditions are aligned with published guidance, standards and/or regulatory documents. Although references to source documents are provided, in some cases WHO has additional requirements.

For prequalification purposes, manufacturers must provide evidence in support of the clinical performance of an IVD which can 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 and fulfil its indications for use.

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 prequalification.

B. How to apply these specifications

For the purposes of prequalification, an IVD intended for **professional use only** (by a laboratory professional, healthcare worker or trained lay provider) should be supported by studies outlined in Parts 1 and 2 of this document.

An IVD intended both for **professional use and for self-testing** should be supported by not only the studies outlined in Parts 1 and 2 of this document, but in addition, the claim for self-testing should be supported by studies that qualify the usability of the IVD among a broad range of self-testing users, as outlined in Part 3.

An IVD intended for **self-testing only**, should be supported by studies outlined in Parts 1, 2 and 3.

For an IVD with an **intended use that has been amended to include self-testing**, and for which performance in professional use is already established, and Parts 1 and 2 of this document can be satisfied, the additional claim for self-testing should be supported by studies outlined in Part 3.

Intended Use	Parts of the TSS to be fulfilled
Professional Use	Parts 1 and 2
Self-testing	Parts 1, 2 and 3
Prequalified professional use IVD with additional claim for self-testing	Part 3, on the provision that any adaptations made do not impact the established safety and

performance

C. Other guidance documents

It is recommended that this document is read in conjunction with other relevant WHO guidance documentation, including:

- WHO Prequalification Team: Diagnostics "Technical Guidance Series (TGS)"
- Sample Product Dossiers for WHO prequalification
- WHO document PQDx_018 "Instructions for Compilation of a Product Dossier"

These documents are available at: http://www.who.int/diagnostics_laboratory/evaluations/en/

D. Performance principles for WHO prequalification

D.1 Intended use

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

- The function of the IVD (e.g. to detect antibodies to HIV-1, HIV-2 and/or HIV p24 antigen, etc.) and whether it is qualitative, semi-quantitative or quantitative;
- The intended testing population for which functions are intended (e.g. detection of susceptible individuals) and the intended operational setting (e.g. for use in near-patient testing); and
- Clinical indication (e.g. aid to diagnosis of HIV infection).

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

For WHO purposes, clinical performance studies should be conducted using the specimen types that are both claimed in the instructions for use and most likely to be used in resource-limited WHO Member States (e.g. capillary whole blood and oral fluid). If this is not possible, substantial data should be presented to show the equivalence between specimen types used in performance studies.

Prequalified RDTs in low- and middle-income countries are likely to be used by laboratory professionals[1] and at point-of-care by healthcare workers, trained lay providers[2] or by individuals who self-test. Depending on the intended use of an RDT, performance studies must be designed to take into account not only the diversity of knowledge and skills across the population of RDT users, but also the likely operational settings in which testing will occur. For example, studies that comprise the testing of left-over/repository specimens by research and development staff at a manufacturer's facility would, on their own, be considered insufficient to meet many of the performance requirement summarised in this document.

D.3 Applicability of supporting evidence to IVD under review

Performance studies must be undertaken using the specific IVD intended to be submitted for prequalification. These studies should include the final, locked-down version of the IVD tested. Where this is not possible a justification must be provided, and additional supporting evidence may

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

[2] Any person who performs functions related to healthcare delivery and has been trained to deliver specific services but has received no formal professional or paraprofessional certificate or tertiary education degree.

also be required. This can occur as minor variations to design where no negative impact on performance has been demonstrated.

Where applicable, several different lots of the IVD should be used, such that each lot comprises different batches of critical components. Specific information is provided in Parts 1 and 2 of this document for the numbers of lots required for particular studies. However, it is a manufacturer's responsibility to ensure, via risk analysis of its IVD that the numbers of lots chosen for estimating performance characteristics takes into account the variability in performance likely to arise from the diversity of key components and their formulation.

The true HIV status of a specimen must be determined using a suitable reference method, justification for which must be provided. Estimation (and reporting) of IVD performance must include the rate of invalid test results.

For certain analytical studies, it may be acceptable to use contrived specimens (e.g. where normal human specimens have been spiked with those containing HIV antibodies). Although all reasonable attempts should be made to use natural specimens, justification should be provided where contrived specimens are used in the submitted studies. Clinical studies should be based on testing in natural specimens only.

For IVDs that include a claim for detection of multiple analytes, evidence of performance must be provided for each claimed analyte. It should be noted that, depending on the design of an IVD, evidence generated in a similar, related product will usually not be considered sufficient by WHO to support performance claims in an IVD submitted for prequalification.

> *Example: an IVD designed to detect HIV antibodies only, and the same IVD designed for dual detection of HIV and syphilis. It is unlikely that performance evidence presented for the HIV-only IVD would be acceptable to support performance claims for the dual-detection IVD.*

For an IVD with an intended use that has been expanded to include self-testing, changes are usually required to improve the usability of the IVD for this new testing population. Such changes may include modification of:

- Instructions for use (e.g. simplification of instructions to reflect new intended users),
- Buffer vials,
- Collection procedures, and/or
- Reading times, etc.

It is a manufacturer's responsibility to verify through testing (as summarised in Parts 1 and 2 of this document) that any changes do not have an adverse impact on critical safety and performance characteristics of an IVD. Usability studies are undertaken so as to optimise the presentation of an IVD and, in turn, understanding by self-testing users. The minimum reporting requirements summarized in Part 3 of this document are not intended to be an exhaustive list or to indicate any particular order that studies should be undertaken in.

E. Table of Requirements

PART 1	**Establishing analytical performance**
1.1	**Specimen type**
1.1.1	Demonstration of equivalence between specimen types
1.1.2	Demonstration of equivalence of claimed anticoagulants
1.2	**Specimen collection, storage and transport**
1.2.1	Specimen stability
1.3	**Precision of measurement**
1.3.1	Repeatability, reproducibility
1.4	**Performance panels**
1.4.1	Genotype panels
1.5	**Validation of reading time**
1.5.1	Validation of reading time
1.6	**Analytical sensitivity**
1.6.1	Seroconversion
1.6.2	Limit of detection for HIV-1 p24 Ag, where appropriate
1.6.3	Validation of assay cut-off
1.6.4	Measuring range
1.7	**Analytical specificity**
1.7.1	Potentially interfering substances
1.7.2	Cross-reactivity
1.8	**Metrological traceability of calibrators and control material values**
1.8.1	Metrological traceability of calibrators and control material values
1.9	**Stability**
1.9.1	IVD stability
1.9.2	Shelf life
1.9.3	In-use stability
1.10	**Flex studies**
1.10.1	Flex studies
PART 2	**Establishing Clinical Performance**
2.1	**Diagnostic sensitivity and specificity**
2.1.1	Diagnostic sensitivity and specificity
2.1.2	Diagnostic sensitivity
2.1.3	Diagnostic specificity
PART 3	**Qualification of usability (self-testing)**
3.1	**Qualification of usability (self-testing)**
3.1.1	Label comprehension study
3.1.2	Results interpretation study
3.1.3	Observed untrained user study

Part 1 Establishing analytical performance characteristics

Aspect	Testing requirements	Comments	References
1.1 Specimen type			
1.1.1 Demonstration of equivalence between specimen types	At least 25 positive and 25 negative specimens of each claimed type. Ideally, positive specimens (undiluted) should be chosen so that the majority are near the IVD cut-off.	1. The relationship between IVD performance in claimed specimen types and reference materials used for analytical studies should be established. The design of subsequent studies must then take that relationship into account. If performance is not equivalent between claimed specimen types then the impact this will have subsequent performance claims (e.g. clinical performance, analytical sensitivity) must be fully understood and described. 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). *Example: an IVD intended for testing whole blood for which seroconversion sensitivity is estimated using panels of serum/plasma specimens.* – *The relationship between seroconversion sensitivity in serum/plasma to that of the same characteristic in whole blood must be understood.* – *This might be achieved by comparing titres between end-point dilution series of matched specimen types (whole blood vs serum/plasma) from a set of positive patients.* 2. The equivalence of specimen types must be determined for all claimed analytes (e.g. HIV-1 antibodies, HIV-2 antibodies, p24 Ag, as appropriate). 1. In some cases, it may be acceptable to use diluted or spiked specimens. This approach is acceptable in early development work, but all reasonable attempts should be made to use natural specimens. Justification should	WHO TGS-3 (1) EU CTS (2)

Aspect	Testing requirements	Comments	References
		be provided if diluted or spiked specimens are used in the submitted studies.	
1.1.2 Demonstration of equivalence of claimed anticoagulants	At least 25 positive and 25 negative specimens of each claimed anticoagulant.		
1.2 Specimen collection, storage and transport			
1.2.1 Specimen stability	Real time studies taking into account: • storage conditions (duration at different temperatures temperature limits, freeze/thaw cycles). • transport conditions, where applicable. • intended use (see comment). • specimen collection and/or transfer devices intended to be used with the IVD.	1. Particular attention should be paid to the length of time likely to elapse between specimen collection and its addition to the IVD in the setting this IVD may be used.	WHO TGS-1 (3)
1.3 Precision of measurement			
1.3.1 Repeatability, reproducibility	Both repeatability (within-condition – see comment 1) and reproducibility (between-condition – see comment 1) estimated using panels of at least: • 1 analyte-negative specimen. • 1 low reactivity positive specimen (near assay cut-off). • 1 medium reactivity positive. Each panel member tested: • in five replicates; • using 3 different lots; • over 5 days (not necessarily consecutive) with one run in that day (alternating morning/afternoon); and • at 3 different testing sites.	1. E.g. within- or between-run, -lot, -day, -site, etc. 2. Precision must be determined for each pathogen and/or analyte for which detection is claimed (e.g. HIV-1 antibody, HIV-2 antibody, HIV-1 p24 antigen, as appropriate). 3. Ideally, the testing panel should be composed of natural (i.e. undiluted) specimens. Where this is not feasible, stock specimens that are to be diluted should represent a range of stages of infection (antibody maturation) in order to take into account the limitations of mimicking low IVD reactivity with a high avidity specimen. 4. IVDs which include whole blood as a specimen type must include evidence of precision in at least spiked whole blood specimens (negative whole blood spiked with highly-reactive plasma/serum specimens to produce an appropriate range of reactivities in the IVD).	CLSI EP05-A3 (4) ISO 13612:2002 (5) CLSI EP12-A2 (6)

Aspect	Testing requirements	Comments	References
	The effect of operator-to-operator variation on IVD performance is to be included as part of the precision studies (see also Comment 8). Testing should be done: • by personnel representative of intended users; • unassisted; and • using *only* those materials provided with the IVD (e.g. instructions for use, labels and other instructional materials).	5. Where possible, the testing panel should be the same for all operators, lots and sites. 6. Lots should be composed of different batches of critical components. 7. Results must be statistically analyzed by ANOVA to identify and isolate the sources and extent of any variance. In addition, the percentage of correctly-identified, incorrectly-identified and invalid results should be tabulated for each specimen and be separately stratified according to each of site, lot, etc. This type of analysis is especially important for rapid tests which may not have any numerical values for ANOVA analysis. 8. The effect of operator-to-operator variation on IVD performance is also to be considered as a human factor when designing robustness (flex) studies (see1.10.1 Flex studies). 9. Users should be selected based on a pre-determined and contextually appropriate level of education, literacy and auxiliary skills that will challenge the usability of the IVD and reflect the diversity of intended users and operational settings.	
1.4 Performance panels			
1.4.1 Genotype panels	Testing of WHO International Reference Preparations and/or commercial HIV genotype panels including: • all HIV-1 subtypes (e.g. A,B,C,D,G, etc.) HIV-2, HIV-1 group O, and common circulating recombinant forms (CRFs); • at least 10 each of the most common subtypes (Subtype C, Subtype A, Subtype B, CRF02_AG, CRF01_AE and Subtype G); and • at least 3 less common subtypes (other CRFs and unique recombinant forms (URFs). • panel of specimens with a range of analyte	1. Testing should be performed using more than 1 final (locked-down) design lot. 2. All confirmed subtype-positive specimens should be detected by the IVD. 3. All reasonable attempts should be made to test rare subtypes. 4. For IVDs including a claim for detection of HIV Ag, appropriate specimens for the same subtypes must also be included in the testing panel. Use of panels of viral-like-particles (VLPs) or viral cultures may be considered acceptable however their use in place of characterized	Health Canada (7)

Aspect	Testing requirements	Comments	References
	concentrations (e.g. antibody 'mixed-titre' panel).	specimens should be justified.	
1.5 Validation of reading time			
1.5.1 Validation of reading time	For IVDs 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 must be provided. Performance studies should be conducted at each of three temperatures (at the mid-point and two extremes of the claimed operating range); the effect of humidity on reading times should also be investigated.	1. The ranges of humidity tested for should be risk-based, taking into consideration likely operational settings. 2. The intended operating temperature, upon which reading time has been validated, must be clearly stated in the instructions for use.	WHO (8)
1.6 Analytical sensitivity			
1.6.1 Seroconversion	Minimum of 25 seroconversion panels: • test at least 40 early seroconversion specimens (see comment 2). • all seroconversion specimens must be reactive (see comment 3). • start with a negative bleed(s), and should have narrow bleeding intervals.	1. Panels should have been collected at short intervals to cover the seroconversion period and should also cover the whole window period. 2. Early seroconversion: p24 Ag and/or HIV RNA-positive, Not recognized by all of European Conformity (CE)-marked 3rd generation enzyme immunoassays, Indeterminate or negative by confirmatory assays. 3. Seroconversion: p24 Ag and/or HIV RNA-positive, Recognized by all of European Conformity (CE)-marked 3rd generation enzyme immunoassays, Indeterminate or positive by confirmatory assays. 4. Seroconversion sensitivity should be reported to the user in the instructions for use. 5. Optimally, testing should be conducted using more than one final design (locked-down) lot.	EC CTS (2) Health Canada (7) EP17-A2 (9)
1.6.2 Limit of detection for HIV-1 p24 Ag, where appropriate	Analytical sensitivity estimated as the concentration of HIV-1 p24 Ag at the assay cut-off. The determination should comprise a minimum of 15-20 replicate tests of an 8-member dilution panel of a suitable biological reference material (e.g. WHO International Standard HIV-1 p24 Ag, NIBSC code 90/636).		
1.6.3 Validation of assay cut-off	HIV RDTs are generally qualitative/semi-quantitative and do not use a numerical value of assay cut-off. Nevertheless the way in which the IVD was designed, in order to differentiate positive specimens from negative specimens, should be described.	1. Where possible, the use of a calibrated, graduated reading scale is recommended for reliable differentiation of reactive and non-reactive specimens in validation studies.	WHO (8)

Aspect	Testing requirements	Comments	References
1.6.4 Measuring range	For each claimed analyte, the potential for a prozone/high-dose hook effect should be determined: • Using multiple, highly-reactive specimens (minimum of 20). • Using at least two different concentrations (diluted by at least a factor of 10). • By testing of several replicates by the same operator on the same day.	1. Specimens should be chosen that have a high analyte concentration, as determined using an IVD method other than the IVD intended to be prequalified e.g. enzyme immunoassay. This second method must be of a design not subject to prozoning. 2. An increase in signal upon dilution of a specimen implies a hook effect.	Health Canada (7) Butch, A.W. (10)
1.7 Analytical specificity			
1.7.1 Potentially interfering substances	The potential for false results (false negatives and false positives) arising from interference from at least the substances/conditions listed below should be determined (See Comment 1). • Minimum of 100 specimens (either naturally occurring or spiked to a low reactivity). • Each substance/condition represented, where possible, by at least 3-5 specimens from different individuals. Testing should be done in both HIV-negative and HIV-positive specimens, unspiked or spiked, with each potentially interfering substance at physiologically relevant dosages.	1. The risk assessment conducted for an IVD should identify substances for which the potential for interference is reasonably foreseeable for the analyte being detected (e.g. HIV-1/2 antibodies and/or HIV-1 p24 Ag). 2. Where either the scientific literature and/or risk analysis identifies the potential for false results in co-infected individuals (e.g. decreased sensitivity or specificity), further investigation should be undertaken using both HIV-negative and HIV-positive specimens. 3. In addition to the substances listed here, IVDs that are used to test oral fluid should take into account the effect of oral infections, such as Candida, as well as tobacco, mouthwash, concomitant medications, dental fixtures, toothpaste food or drink (consumed immediately prior to testing), consumption of alcohol and teeth brushing. 4. Any observed interference should be investigated and performance limitations of the IVD reported in the instructions for use. Results should be reported with respect to each condition and not be reported as an aggregate of the total number of specimens tested in the study.	Health Canada (7) EC CTS (2) CLSI EP07-A2 (11)
Endogenous	• Human antibodies to the expression system (for recombinants), e.g. Anti-*Escherichia coli* (anti-*E.coli* positive), Human anti-mouse antibody (HAMA)		

Aspect	Testing requirements	Comments	References
	• Recipients of multiple blood transfusions, pregnant (including multiparous) women. • Haemoglobin, lipids, bilirubin and protein. • Elevated Immunoglobulin G (IgG) and Immunoglobulin M (IgM). • Rheumatoid factor. • Sickle-cell disease. • Other autoimmune conditions including Systemic lupus erythematosus (SLE) and anti-nuclear antibodies (ANA.)		
Exogenous	• Relevant medicines, including: antiparasitic, antimalarial, antiretroviral and anti-tuberculosis medications. • Common over-the-counter anti-inflammatory medications (aspirin, paracetamol, ibuprofen). • Ethanol, caffeine.		
1.7.2 Cross-reactivity	Determination of the potential for false positive results arising from cross-reactivity (see Comment 1) for a minimum of 100 specimens, including, where possible, at least 3-5 each of: • Non HIV viral infections, including: hepatitis B, C infection, acute hepatitis A infection, cytomegalovirus, acute Epstein–Barr virus, varicella zoster virus, Yellow fever virus post-immunization measles, influenza A and B, tick borne encephalitis. • Other retroviruses, including: Human T-lymphotropic virus-1 and -2. • Bacteria/parasites, including: malaria, visceral leishmaniasis tuberculosis and human African trypanosomiasis. • Influenza vaccine recipient. • Vaccine-induced HIV seropositivity. • Other unrelated conditions known to cause cross-reactivity in HIV IVDs.	1. The types of interferences tested for should be risk-based, taking into consideration the operational setting as well as the intended users for the analyte being detected (e.g. HIV-1/2 antibodies and/or HIV-1 p24 Ag). 2. Any observed interference should be investigated and performance limitations of the IVD reported in the instructions for use.	

Aspect	Testing requirements	Comments	References
1.8 Metrological traceability of calibrators and control material values			
1.8.1 Metrological traceability of calibrators and control material values	The traceability of an assay-specific quality control specimen to a validated reference material must be demonstrated (e.g. WHO International Standard HIV (antibody), 1st International Reference Panel; WHO International Standard HIV-1 P24 Antigen).	1. HIV RDT kits may not include external quality control specimens, but the IVD should have a procedural control. The extent to which a control band corresponds to a valid test (identification of and traceability to a suitable reference) should be demonstrated. NOTE: The nature of the procedural control (specimen addition or only reagent addition) should be explained. NOTE 2: an external control specimen is one that is run in conjunction with the IVD, but is physically separate from =, for example, an RDT test cassette. 2. In some jurisdictions there is a requirement for use of a 'National Testing Panel' for lot release and IVD validation. Such a national requirement does not obviate the need for evidence of traceability to a validated reference material as described here.	WHO (8)
1.9 Stability			
1.9.1 IVD stability	Replicate testing of a panel consisting of at least, for each pathogen/analyte claimed: • 1 analyte non-reactive specimen; • 2 low-reactivity specimens, near assay cut-off (see comment 2); and • 1 medium reactivity specimen. Specimens chosen for the testing panel panel should, wherever possible, include panel members that reflect the main specimen types intended for use with the IVD (e.g. capillary whole blood/oral fluid, as appropriate).	1. The testing panel must include all claimed analytes and include whole blood specimens and/or oral fluid specimens, as appropriate, in accordance with intended use (for example to verify proper flow, no background interference and account for other variables). 2. Where detection of multiple genotypes and/or subtypes is claimed, equivalent performance (e.g. sensitivity and specificity) must have been demonstrated, otherwise evidence of stability in these genotypes/subtypes will need to be provided. 3. Ideally, the stability testing panel should be composed of natural (i.e. undiluted) specimens. Where this is not feasible, stock specimens to be diluted should represent a range of stages of infection (antibody maturation) so as to take into account the limitations of mimicking low IVD reactivity with	ISO 23640:2011 (12) CLSI EP25-A (13) WHO TGS-2 (14) ASTM D4169-14 (15)

Aspect	Testing requirements	Comments	References
		a high avidity specimen. 4. Lots must be comprised of different batches of critical components. 5. Determination of shipping stability must be performed using simulated extreme stress conditions, ensuring that application of those conditions is consistent and controlled. 6. Claims for stability must be based on the second-last successful data point from the least stable lot, with, if lots are different, a statistical analysis showing that the bulk of lots will be expected to meet the claimed life. For example: for testing conducted at 3, 6, 9, 12 and 15 months, if stability was observed at 15 months, then the maximum stability claim can be 12 months. 7. Accelerated studies do not replace the need for real time studies. 8. In-use stability of labile components should be conducted using components in their final configuration.	
1.9.2 Shelf life	• Real time, minimum of 3 lots of final design product. • Transport stressed (simulated) *before* real time studies are undertaken. • IVD in final packaging subjected to drop-shock testing.		
1.9.3 In-use stability	• Minimum of 1 master lot, using panel(s) compiled as above. • Testing of all labile components (e.g. buffers vials, sealed cartridges, etc.; see Comment 8).		
1.10 Flex studies			
1.10.1 Flex studies	The influence of the following factors on expected positive and negative results should be considered: • Specimen and/or reagent volume. • Buffer pH. • Reading time (i.e. the interval between when the first and last readings can be taken).	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.	WHO **(8)**

Aspect	Testing requirements	Comments	References
	• IVD sturdiness. • Lighting and humidity. • Operating temperature.		

Part 2 Establishing clinical performance characteristics (professional use and/or self-testing)

Aspect	Testing requirements	Comments	References
2.1 Diagnostic sensitivity and specificity			
2.1.1 Diagnostic sensitivity and specificity	Diagnostic sensitivity and specificity should be determined for each claimed specimen type. Testing should be conducted: • at different geographical settings (minimum of 2 regions); • by a variety of intended users; and • using more than one master lot.	1. Prequalified HIV RDTs are generally used by lay providers and health care workers. For prequalification purposes, these should be considered as the intended user, rather than a trained laboratory professional. 2. Where an IVD is intended to detect multiple analytes without differentiating which analyte is detected, specimens chosen for the testing panel must comprise those that are reactive only for each individual analyte (i.e. not dual HIV-1/HIV-2 positive, etc). 3. A separate specimen should be collected prior to testing to establish the reference result. The testing algorithm used to determine the reference results should include a state of the art 4th generation immunoassay (EIA), with all initially reactive specimens reflexed for full characterization of the HIV status. 4. Problematic specimens, those with unexpected results but which otherwise meet selection criteria for a study, should not be systematically excluded from analysis. 5. Consideration should be given to the influence of antiretroviral medications present in a specimen on the serostatus of such specimens, and how this might affect specimen selection. 5. Lots (locked-down design) should be comprised of different batches of critical components. 6. Where possible, all discrepant results (between assay under evaluation and the reference results) should be repeated using the same lot, and then on all available lots and the variability noted. Performance characteristics must be reported using initial results, only. The results of further testing of specimens with discrepant results must be reported separately as additional information about IVD performance. 7. All indeterminate results must be included in the denominator data	EC CTS **(2)** Health Canada **(7)**
2.1.2 Diagnostic sensitivity	Testing of: • At least 400 specimens confirmed HIV-1 antibody positive. • At least 100 specimens confirmed HIV-2 antibody positive (where HIV-2 detection is claimed; see Comment 2). • At least 50 specimens confirmed HIV p24 Ag positive (where Ag detection is claimed; see Comment 2).		
2.1.3 Diagnostic specificity	Testing of: • At least 1000 HIV antibody/antigen negative specimens.		

Aspect	Testing requirements	Comments	References
		for analysis. 8. All invalid test results must be recorded. 9. Estimates of diagnostic/clinical sensitivity and specificity should be reported with 95% confidence intervals. 10. Results must be expressed separately for each specimen type and for each specimen type per intended use (no aggregation of results).	

Part 3 Qualification of usability (self-testing)

PURPOSE: Assessment of product design, instructions for use and usability of RDTs for self-testing by analysis of the following:

- Results of questionnaire to assess whether key messages and instructions from packaging and labelling would be understood and easily followed by untrained intended users (i.e. self-testers).
- Results of interpretation of test-results by untrained users (i.e. self-testers) of simulated RDTs (e.g. pre-made and with contrived results).
- Test results and interpretations when assay is performed by untrained intended users (i.e. self-testers).
- For each of the studies summarized below the study group should comprise untrained subjects whose age, gender, level of education, literacy and additional, supplementary skills can challenge the usability of the IVD in intended users and in unfavourable operational settings (e.g. poor lighting).
- 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) to major. The impact of any change on safety and performance must be determined.
- Results from any one of the stages summarized 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.

Aspect	Testing requirements	Comments	References
3.1 Qualification of usability (self-testing)			
3.1.1 Label com-prehension study	Questionnaire-based testing of subjects, representative of end users, to assess ability of intended users to correctly comprehend key messages from packaging and labelling: • Proper self-selection (whether or not users understand if it is appropriate for them to undertake testing). • Understanding key warnings, limitations and/or restrictions. • Proper test procedure. • Test result interpretation. Questionnaire to be administered to at least 200 subjects, representative of end users, in order to demonstrate comprehension of key messages.	1. Instructions for use and labelling should be clear and easy to understand; use of pictorial instructional material is encouraged.	ISO 18113:2011 (16) ISO 15197:2013(en) (17) IEC 62366-1:2015 (18) MHRA (19) FDA (20); example of Summary of Safety and Effectiveness (21) EC CTS (2) European Directive 98/79/EC (22) FDA CLIA Waiver Requirements (23) WHO HIV testing
3.1.2 Results interpretation study	A minimum of 400 subjects to interpret the results of contrived IVDs (e.g. static/pre-made tests) to assess their ability to correctly interpret pre-determined test results. Contrived tests should be made to	1. The study group may include subject recruited as part of the label comprehension study.	

Aspect	Testing requirements	Comments	References
	demonstrate the following potential test results: • Non-reactive. • Range of invalid results. • Reactive. • Weak reactive. • Testing subjects to consist of at least 200 self-testers from two high-prevalence (>5%), geographically diverse populations and at least 200 self-testers from a low-prevalence (<5%) population to demonstrate correct interpretation of simulated test results.		guidelines **(24)** How to use an RDT **(25)** FDA **(26)**
3.1.3 Observed untrained user study	Testing by at least 900 self-testing subjects comprising: at least 200 self-testers in each of two high-prevalence (>5%), geographically diverse population and at least 500 self-testers from a low-prevalence (<5%) population. • Each subject to self-collect test specimen and perform test according to only those materials provided with the IVD (e.g. instructions for use, labels and other instructional materials). • Each such test to be observed by trained laboratory or healthcare professional. The observing professional does not tutor or interact with subject conducting test, but notes errors and other observations about the self-tester. Observation could also be conducted by way of video recording of self-testing. • Observing professional also interprets the test result, in a blinded fashion and within the validated reading time stated in the instructions for use.	1. A separate venous whole blood specimen should be collected prior to testing to establish the reference results for HIV-1 status (and HIV-2 where detection is claimed). The testing algorithm used to determine the reference results should include use of a state of the art 4th generation immunoassay (EIA), with all initially reactive specimens reflexed for confirmation of the HIV status. 2. For WHO purposes the term 'professional use' encompasses a diversity of skills, training and experience and does not necessarily imply 'highest standard of skills, training and experience'. It may be a useful step in development of usability to compare performance between self-testers with that of healthcare workers, lay providers, and laboratory technicians. However, concordance observed between types of users may mask poor performance within each user group. Consequently, such comparisons do not replace the need for comparisons to 'clinical truth' by establishment of reference results for each subject. 3. There can be a high likelihood of bias at the community level when simple study population sample methodologies are applied. Efforts should be made to avoid convenience sampling of people (participants) who already know they are HIV positive.	

References

1 WHO Prequalification Team – Diagnostic Assessment. Technical Guidance Series (TGS). Principles for Performance studies TGS–3. Geneva: World Health Organization; 2016. Available at: http://www.who.int/diagnostics_laboratory/guidance/160613_tgs3_principles_for_performance_studies.pdf?ua=1, accessed 15 July 2015.

2 European Commission. 2009/108/EC: Commission Decision of 3 February 2009 amending Decision 2002/364/EC on common technical specifications for in vitro-diagnostic medical devices. Official Journal of the European Union; L39/34-L39/49. Available at: http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=uriserv:OJ.L_.2009.039.01.0034.01.ENG, accessed 18 July 2016.

3 WHO Prequalification Team – Diagnostic Assessment. Technical Guidance Series (TGS). Standards applicable to the WHO Prequalification of in vitro diagnostics TGS–1. Geneva: World Health Organization; 2016. Available at: http://www.who.int/entity/diagnostics_laboratory/guidance/160321_tgs1_standards_v2.pdf?ua=1, accessed 15 July 2016.

4 CLSI. Evaluation of Precision of Quantitative Measurement Procedures; Approved Guideline - Third Edition. CLSI document EP05-A3. Wayne, PA: Clinical and Laboratory Standards Institute; 2014.

5 ISO 13612:2002. Performance evaluation of in vitro diagnostic medical devices. International Organization for Standardization/British Standards Organization; 2002.

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

7 Health Canada. Guidance for Manufacturers of Human Immunodeficiency Virus (HIV) Test Kits intended to be used in the Laboratory. Health Products and Food Branch; 2011. Available at: http://hc-sc.gc.ca/dhp-mps/md-im/applic-demande/guide-ld/md_gd_hiv_im_ld_vih-eng.php, accessed 15 July 2016.

Note: The above guidance does not apply to HIV test kits which are used for patient management, or HIV test kits intended to be used outside the laboratory i.e. at the point of care and/or for home use. A separate guidance titled "Draft Guidelines for HIV Simple/Rapid Test Kit" is available for manufacturers of near-patient HIV test kits upon request from Health Canada's Medical Devices Bureau (MDB), e-mail: device_licensing@hc-sc.gc.ca.

8 WHO Prequalification Team – Diagnostic Assessment. Instructions for Compilation of a Product Dossier. WHO document PQDx_018 v3, 27 August 2014. Available at: http://www.who.int/entity/diagnostics_laboratory/evaluations/141015_pqdx_018_dossier_instructions_v4.pdf?ua=1, accessed 18 July 2016.

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

10 Butch, A. W. Dilution protocols for detection of hook effects/prozone phenomenon. Clin Chem. 2000; 46:1719-21.

11 CLSI. Interference Testing in Clinical Chemistry. Approved Guideline - Second Edition. CLSI document EP07-A2. Wayne, PA: Clinical and Laboratory Standards Institute; 2005.

12 ISO 23640:2011. In vitro diagnostic medical devices - Evaluation of stability of in vitro diagnostic reagents. Geneva, Switzerland: International Organization for Standardization; 2011.

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

14 WHO Prequalification Team: Diagnostics. Technical Guidance Series (TGS). Establishing stability of an in vitro diagnostic for WHO Prequalification TGS–2. Geneva: World Health Organization. 2016. Available at: http://www.who.int/diagnostics_laboratory/guidance/160613_tgs2_stability.pdf?ua=1, accessed 15 July 2016.

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

16 BS EN ISO 18113:2011.In vitro diagnostic medical devices – Information supplied by the manufacturer (labelling). - Part 1: Terms, Definitions and General Requirements . International Organization for Standardization/British Standards Organization; 2011.

17 ISO 15197:2003(en). In vitro diagnostic test systems – Requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus. International Organization for Standardization; 2003.

18 ISO/IEC 62366-1:2015. Medical devices – Part 1: Application of usability engineering to medical devices. International Organization for Standardization/ International Electrotechnical Commission; 2015.

19 MHRA. Guidance for notified bodies on the regulation of IVDs for self-testing. Competent Authority (UK) Medicines and Healthcare products Regulatory Agency. London, United Kingdom; 2012. Available at: https://www.gov.uk/government/publications/regulation-of-ivds-for-self-testing, accessed 15 July 2016.

20 FDA. Update: Rapid HIV Test Approval Requirements and Standards [Slides]. Kimber Poffenberger. Food and Drug Administration, Blood Products Advisory Committee; 2000. Available at: http://www.fda.gov/ohrms/dockets/ac/00/slides/3649s2.htm, accessed 15 July 2016.

21 FDA. OraQuick In-Home HIV Test. Summary of Safety and Effectiveness. Food and Drug Administration. BP120001. Available at: http://www.fda.gov/downloads/BiologicsBloodVaccines/BloodBloodProducts/ApprovedProducts/PremarketApprovalsPMAs/UCM312534.pdf; accessed 18 July 2016.

22 European Parliament, Council of the European Union. Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices. Official Journal of the European Communities, L 331/1; 2009. Available at: http://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:31998L0079; accessed 18 July 2016.

23 FDA. Guidance for Industry and FDA Staff: Recommendations for Clinical Laboratory Improvement Amendments of 1988 (CLIA) Waiver Applications for Manufacturers of In Vitro Diagnostic Devices. Center for Devices and Radiological Health, Food and Drug Administration. 2008. Available at: http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm079632.htm, accessed 18 July 2016.

24 WHO, Department of HIV. Consolidated Guidelines on HIV Testing Services. Geneva: World Health Organization, 2015. Available at: http://who.int/hiv/pub/guidelines/hiv-testing-services/en/; accessed 15 July 2016.

25 How to use a rapid diagnostic test (RDT): A guide for training at a village and clinic level (Modified for training in the use of the Generic Pf Test for falciparum malaria). The USAID Quality Assurance Project (QAP), University Research Co., LLC, and the World Health Organization (WHO), Bethesda, MD, and Geneva. Available at: http://www.wpro.who.int/malaria/NR/rdonlyres/23DD7DCB-48C4-4CFF-BD45-332F0BE3DCC7/0/generic_PfPan_training_manual_web.pdf, accessed 15 July 2016.

26 FDA. Guidance for Industry and FDA Staff. Applying Human Factors and Usability Engineering to Medical Devices. Center for Devices and Radiological Health, Food and Drug Administration. 2011. Available at: http://www.fda.gov/downloads/MedicalDevices/.../UCM259760.pdf, accessed 23 November 2016.

List of related WHO Publications of related interest

WHO Prequalification Team- Diagnostic Assessment. Technical Guidance Series for WHO Prequalification – Diagnostic assessment (available online)

WHO Prequalification Team – Diagnostic Assessment. Instructions for Compilation of a Product Dossier. WHO document WHO/PQDx_18 (available online)

The Technical Specifications Series for submission to WHO Prequalification – Diagnostic Assessment set out appropriate performance evaluation criteria to meet prequalification requirements. Each Technical Specification provides information on the minimum performance requirements for WHO prequalification that should be met by a manufacturer to ensure that their in vitro diagnostic medical device is safe and performs optimally.

Further information on these and other WHO publications can be obtained from WHO Press, World Health Organization, 1211Geneva 27, Switzerland

http:/www.who.int/bookorders

tel.: +41 22 791 3264; fax: +41 22791 4857; email bookorders@who.int

ISBN 978-92-4-151174-2