EXPERT COMMITTEE ON BIOLOGICAL STANDARDIZATION
Geneva, 17 to 20 October 2017

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

Technical Specifications Series for WHO prequalification of in vitro diagnostic medical devices – Diagnostic Assessment

WHO Prequalification Team: Diagnostics

NOTE:

This document has been prepared for the purpose of inviting comments and suggestions on the proposals contained therein, which will then be considered by the Expert Committee on Biological Standardization (ECBS). Comments MUST be received by **18 September 2017** and should be addressed to the World Health Organization, 1211 Geneva 27, Switzerland, attention: Technologies, Standards and Norms (TSN). Comments may also be submitted electronically to the Responsible Officer: **Dr C M Nübling** at email: nueblingc@who.int.

Document Change			
File name	Technical Specifications Series for submission to WHO Prequalification – Diagnostic Assessment: Human immunodeficiency virus (HIV) rapid diagnostic tests for professional use and/or self-testing	**Replaces**	Technical Specifications Series for submission to WHO Prequalification – Diagnostic Assessment: Human immunodeficiency virus (HIV) rapid diagnostic tests for professional use and/or self-testing
Date	2017/03/08	**Date**	2016/12/05

Change	Location (Section, paragraph)	Nature of and /or Reason for Change
1	**Whole Document**	**Change to formatting, layout, and update of page numbers to add clarity**
2	**References**	**Update to reference formatting**

Table of contents

Acknowledgements

List of contributors

First round public comment were received from the following:

Ms 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; Mr G James, Institute for Clinical Pathology and Medical Research (ICPMR), New South Wales, Australia; Dr L Kestens, Institute of Tropical Medicine, Antwerp, Belgium; Ms D Lepine, IVDD Section, Medical Devices Bureau Health Canada, Ottawa, Canada; Dr A Reissinger, Paul Ehrlich Institut, Langen, Germany; and Dr M Nuebling and Ms M Perez Gonzalez, WHO, Geneva, Switzerland.

The draft technical specifications document was 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 public comments were received from the following:

Dr P Akolkar, Center for Biologics Evaluation and Research, United States Food and Drug Administration, MD, USA; Ms S Best, National Serology Reference Laboratory, Victoria, Australia; Dr J Duncan, London, United Kingdom; Epicentre and Médecins sans Frontières International Office: Dr C Kosack, Dr A Page and Ms E Tran, Geneva, Switzerland; Dr R Galli, bioLytical™ Laboratories Inc., Vancouver, Canada; Ms D Lepine, IVDD Section, Medical Devices Bureau Health Canada, Ottawa, Canada; Dr M Nuebling WHO, Geneva, Switzerland; OraSure Technologies Inc, PA, USA; Dr H Scheiblauer, Paul Ehrlich Institut, Langen, Germany; UNITAID/PSI HIV Self-Testing Africa (STAR) project consortium members: Dr E Cowan, Dr R Dacombe, Dr M Kumwenda, Dr M Neuman, Prof R Peeling, Dr M Taegtmeyer, Dr V Watson, Liverpool School of Tropical Medicine, Liverpool, United Kingdom.

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 in this document, 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 intended to be used as an aid to diagnosis of HIV infection. The current version of this document does not address IVDs that discriminate between the detection of HIV-1 and HIV-2 infection, IVDs intended as confirmatory tests, or the requirements for accompanying quality control material.

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

- "shall" indicates that the manufacturer is required to comply with the technical specifications.
- "should" indicates that the manufacturer is recommended to comply with the technical specifications but it is not a requirement.
- "may" indicates that the technical specifications are 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.

Minimum performance requirements for WHO prequalification are summarized in this document, and 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 prequalification has additional requirements.

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

B.   How to apply these specifications

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

An IVD intended both for **professional use and for self-testing** shall be supported by the studies outlined in Parts 1 and 2 of this document. In addition, the claim for self-testing shall 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**, shall 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 have already been satisfied, the additional claim for self-testing shall be supported by studies outlined in Part 3.

These requirements are summarized in Table 1.

Table 1 – Summary of requirements for submission for WHO prequalification based on the intended use of the IVD

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

This document should be read in conjunction with other relevant WHO guidance documentation, including:
- Technical Guidance Series for WHO Prequalification – Diagnostic Assessment
- Sample Product Dossiers for WHO Prequalification – Diagnostic Assessment
- Instructions for Compilation of a Product Dossier, WHO document PQDx_018.

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 shall 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 testing population for which the 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 prequalification submission, clinical performance studies should be conducted using the specimen types that are most likely to be used in resource-limited WHO Member States (e.g. capillary whole blood and oral fluid) and claimed in the instructions for use. If this is not possible, substantial data shall 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

---

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

individuals who self-test. Depending on the intended use of an RDT, performance studies shall 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 shall not, on their own, be considered sufficient to meet many of the performance requirements summarized in this document.

D.3 Applicability of supporting evidence to IVD under review

Performance studies shall be undertaken using the specific, locked-down version of the IVD intended to be submitted for WHO prequalification. Where this is not possible, a justification shall be provided and additional supporting evidence may also be required. This may occur in the case of minor variations to design where no negative impact on performance has been demonstrated.

Specific information is provided in Parts 1 and 2 of this document for the numbers of lots required for particular studies. Each lot should comprise different batches of critical components. It is a manufacturer's responsibility to ensure, via risk analysis of its IVD, that the minimum numbers of lots chosen for estimating performance characteristics 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 shall be determined using a suitable reference method, for which justification shall be provided. Estimation (and reporting) of IVD performance shall 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 shall 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

---

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 certification or tertiary education degree.

- 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 to optimize the presentation of an IVD and the understanding of 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 a particular order in which studies should be undertaken.

E.   Table of Requirements

PART 1	Establishing analytical performance characteristics
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
1.7.1.1	Endogenous
1.7.1.2	Exogenous
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 characteristics
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	Notes on testing requirements	Source documents
1.1 Specimen type			
1.1.1 Demonstration of equivalence between specimen types	For each claimed specimen type, testing in at least: • 25 analyte positive specimens • 25 analyte negative specimens	1. The relationship between IVD performance in claimed specimen types and reference materials used for analytical studies shall be established. The design of subsequent studies shall then take that relationship into account.	Technical Guidance Series for WHO Prequalification – Diagnostic Assessment **(1)**
1.1.2 Demonstration of equivalence of claimed anticoagulants	At least 25 positive and 25 negative specimens of each claimed anticoagulant. The equivalence of specimen types shall be determined for all claimed analytes (e.g. HIV-1 antibodies, HIV-2 antibodies, p24 Ag, as appropriate) (see comment 3).	2. 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 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 shall 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.* 3. 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 be provided if diluted or spiked specimens are used in the submitted studies. 4. Positive specimens (undiluted) shall be chosen so that the majority are near the IVD cut-off.	European Commission **(2)**
1.2 Specimen collection, storage and transport			
1.2.1 Specimen	Real time studies taking into account:	1. Evidence shall be provided which validates the maximum allowable time between specimen collection and its addition to	Technical Guidance Series for WHO

stability	• storage conditions (duration at different temperatures, temperature limits, freeze/thaw cycles) • transport conditions, where applicable • intended use (see comment 1) • specimen collection and/or transfer devices intended to be used with the IVD.	the IVD in the setting where testing takes place.	Prequalification – Diagnostic Assessment (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) shall be 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 shall be tested: • in 5 replicates • using 3 different lots • over 5 days (not necessarily consecutive) with one run per day (alternating morning/afternoon) • at each of 3 different testing sites. The effect of operator-to-operator variation on IVD performance should be included as part of the precision studies (see also Comment 8). Testing should be done: • by personnel representative of intended users • unassisted • using *only* those materials provided with the IVD (e.g. instructions for use, labels and other instructional materials).	1. E.g. within- or between-run, -lot, -day, -site, etc. 2. Precision shall 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 (Ag), as appropriate). 3. 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 shall include evidence of precision in, at a minimum, spiked whole blood specimens (negative whole blood spiked with highly-reactive plasma/serum specimens to produce an appropriate range of reactivities in the IVD). 5. Where possible, the testing panel should be the same for all operators, lots and sites. 6. Lots shall comprise different batches of critical components. 7. Results shall be statistically analyzed using analysis of variance (ANOVA) techniques to identify and isolate the sources and extent of any variance. In addition, the percentage of correctly-identified, incorrectly-identified and invalid results shall be tabulated for each specimen and be separately stratified according to site, lot, etc. This type of analysis is especially important for rapid tests that may not have any numerical values for ANOVA analysis. 8. The effect of operator-to-operator variation on IVD performance may also be considered as a human factor when designing robustness (flex) studies (see 1.10.1 Flex studies)	CLSI EP05-A3 (4) ISO 13612:2002 (5) CLSI EP12-A2 (6)

	and may be addressed as part of clinical studies in representative populations (see Part 2). 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 shall include: • 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) • at least 3 less common subtypes (other CRFs and unique recombinant forms (URFs) • panel of specimens with a range of analyte concentrations (e.g. antibody 'mixed-titre' panel).	1. Testing should be performed using more than 1 lot of the final design (locked-down). 2. All confirmed subtype-positive specimens shall be detected by the IVD. 3. All reasonable attempts shall be made to test rare subtypes. 4. For IVDs including a claim for detection of HIV Ag, appropriate specimens for the same subtypes shall 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 specimens shall be justified.	Health Products and Food Branch, Health Canada (7)
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 shall be provided. Performance studies shall 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 shall also be investigated.	1. The ranges of humidity tested for shall be risk-based, taking into consideration likely operational settings. 2. The intended operating temperature, upon which reading time has been validated, shall be clearly stated in the instructions for use.	WHO Prequalification – Diagnostic Assessment (8)
1.6 Analytical sensitivity			
1.6.1 Seroconversion	A minimum of 25 seroconversion panels shall be tested: • test at least 40 early seroconversion specimens (see comment 2) • all seroconversion specimens shall be reactive (see comment 3) • start with a negative bleed(s), and should have	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	European Commission (2) Health Products and Food Branch, Health Canada (7)

	narrow bleeding intervals.	3rd generation enzyme immunoassays – Indeterminate or negative by confirmatory assays.	CLSI EP12-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 shall 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).	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 shall be reported to the user in the instructions for use. 5. Optimally, testing should be conducted using more than one lot of the final design (locked-down).	
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, shall be described.	1. Where possible, a calibrated, graduated reading scale should be used for reliable differentiation of reactive and non-reactive specimens in validation studies.	WHO Prequalification – Diagnostic Assessment **(8)**
1.6.4 Measuring range	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 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 shall 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 shall be of a design not subject to prozoning. 2. An increase in signal upon dilution of a specimen implies a hook effect.	Health Products and Food Branch, Health Canada **(7)** Butch, AW **(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 shall be determined using (See Comment 1): • a 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.	1. The risk assessment conducted for an IVD shall identify substances where the potential for interference can reasonably be expected 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 shall be undertaken using both HIV-negative and HIV-positive specimens.	Health Products and Food Branch, Health Canada **(7)** European Commission**(2)** CLSI EP07-A2 **(11)**

	Testing shall be undertaken in both HIV-negative and HIV-positive specimens, unspiked or spiked, with each potentially interfering substance at physiologically relevant dosages.	3. In addition to the substances listed here, IVDs that are used to test oral fluid shall 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.
1.7.1.1 Endogenous	• Human antibodies to the expression system (for recombinants), e.g. anti-*Escherichia coli* (anti-*E.coli* positive), human anti-mouse antibody (HAMA) • recipients of multiple blood transfusions, pregnant (including multiparous) women • haemoglobin, lipids, bilirubin and protein • elevated Immunoglobulin G and Immunoglobulin M • rheumatoid factor • sickle-cell disease • other autoimmune conditions including systemic lupus erythematosus (SLE) and anti-nuclear antibodies. (ANA)	4. Any observed interference shall be investigated and performance limitations of the IVD reported in the instructions for use. 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.
1.7.1.2 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	The potential for false positive results arising from cross-reactivity (see Comment 1) shall be determined for a minimum of 100 specimens, including, where possible, at least 3-5 of each: • 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 cell virus-1 and -2 • bacteria/parasites, including: malaria, visceral leishmaniasis, tuberculosis and human African trypanosomiasis	1. The types of interferences tested for shall 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 shall be investigated and performance limitations of the IVD reported in the instructions for use.

	- influenza vaccine recipient - vaccine-induced HIV seropositivity - other unrelated conditions known to cause cross-reactivity in HIV IVDs.		
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 shall 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 shall 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 1: The nature of the procedural control (specimen addition or only reagent addition) shall be explained. NOTE 2: An external control specimen is one that is run in conjunction with the IVD, but is physically separate from it, 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 (or remove) the need for evidence of traceability to a validated reference material as described here.	WHO Prequalification – Diagnostic Assessment (8)
1.9 Stability			
1.9.1 IVD stability	Replicate testing shall be undertaken using a panel consisting for each claimed pathogen/analyte, of at least: 1 analyte non-reactive specimen 2 low-reactivity specimens, near assay cut-off (see comment 2) 1 medium reactivity specimen. Wherever possible, specimens chosen for the testing panel shall 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 shall 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) shall have been demonstrated; otherwise evidence of stability in these genotypes/subtypes will need to be provided. 3. Ideally, the stability testing panel shall 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	ISO 23640:2011 (12) CLSI EP25-A (13) WHO Technical Guidance Series for WHO Prequalification – Diagnostic Assessment (14) ASTM D4169-14 (15)
1.9.2	Real time, minimum of 3 lots of final design product transport stressed (simulated) *before* real time		

Shelf life	studies are undertaken • IVD in final packaging subjected to drop-shock testing.	infection (antibody maturation) so as to take into account the limitations of mimicking low IVD reactivity with a high avidity specimen.	
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).	4. Lots shall comprise different batches of critical components. 5. Determination of shipping stability shall be performed using simulated extreme stress conditions, ensuring that application of those conditions is consistent and controlled. 6. Claims for stability shall be based on the second-last successful data point from the least stable lot, with, if lots are different, a statistical analysis showing that the 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 shall be 12 months. 7. Accelerated studies do not replace the need for real time studies. 8. In-use stability of labile components shall be conducted using components in their final configuration.	
1.10 Flex studies			
1.10.1 Flex studies	The influence of the following factors on expected positive and negative results shall 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) • IVD sturdiness • lighting and humidity • operating temperature.	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 Prequalification – Diagnostic Assessment (8)

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

Aspect	Testing requirements	Comments	Source documents
2.1 Diagnostic sensitivity and specificity			
2.1.1 Diagnostic sensitivity and specificity	Diagnostic sensitivity and specificity shall 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 • using more than 1 master lot.	1. Prequalified HIV RDTs are generally used by lay providers and health care workers. For WHO 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 shall comprise those that are reactive only for each individual analyte (i.e. not dual HIV-1/HIV-2 positive, etc).	European Commission (2) Health Products and Food Branch, 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).	3. A separate specimen shall be collected prior to testing to establish the reference result. The testing algorithm used to determine the reference results shall 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, shall not be systematically excluded from analysis. 5. Consideration shall 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.	
2.1.3 Diagnostic specificity	Testing of: • at least 1000 HIV antibody/antigen negative specimens.	5. Lots (locked-down design) shall comprise different batches of critical components. 6. Where possible, all discrepant results (between assay under evaluation and the reference results) shall be repeated using the same lot, and then on all available lots and the variability noted. 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 IVD performance. 7. All indeterminate results shall be included in the denominator data for analysis. 8. All invalid test results shall be recorded.	

Aspect	Testing requirements	Comments	Source documents
		9. Estimates of diagnostic/clinical sensitivity and specificity shall be reported with 95% confidence intervals. 10. Results shall 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 a questionnaire to assess whether the key messages and instructions from packaging and labelling would be understood and easily followed by untrained intended users (i.e. self-testers).
- Results of the 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 shall comprise untrained subjects whose age, gender, level of education, literacy and additional, supplementary skills may 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 shall 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	Source documents
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 shall 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 shall be clear and easy to understand; use of pictorial instructional material is encouraged.	ISO 18113:2009 **(16)** ISO 15197:2013 **(17)** IEC 62366-1:2015 **(18)** MHRA **(19)** Poffenberger, K**(20)** FDA **(21)** European Commission **(2)** European Parliament **(22)** Center for Devices

Aspect	Testing requirements	Comments	Source documents
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 shall be made to 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.	1. The study group may include subjects recruited as part of the label comprehension study.	and Radiological Health, FDA (23) WHO (24) USAID and WHO (25) Center for Devices and Radiological Health, 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 may 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 shall 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 shall include use of a state of the art 4th generation immunoassay, 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 may be a high likelihood of bias at the community level when simple study population sample methodologies are applied. Efforts shall be made to avoid convenience sampling of people (participants) who already know they are HIV positive.	

References

1   Technical Guidance Series for WHO Prequalification – Diagnostic Assessment: Principles for Performance studies, TGS–3. Geneva: World Health Organization; 2016. (http://www.who.int/diagnostics_laboratory/guidance/160519_tgs3_draft_for_comment.pdf?ua=1 accessed 15 July 2016).

2   Commission Decision of 3 February 2009 amending Decision 2002/364/EC on common technical specifications for *in vitro*-diagnostic medical devices (2009/108/EC). O.J.E.U. 2009, L39/34-L39/49; (http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=uriserv:OJ.L_.2009.039.01.0034.01.ENG, accessed 18 July 2016)

3   Technical Guidance Series for WHO Prequalification – Diagnostic Assessment: Standards applicable to the WHO Prequalification of in vitro diagnostics, TGS–1. Geneva: World Health Organization; 2016 (http://www.who.int/entity/diagnostics_laboratory/guidance/161007_tgs1_standards_v3.pdf?ua=1, accessed 15 July 2016).

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

5   EN 13612:2002. Performance evaluation of in vitro diagnostic medical devices. Brussels: European Committee for Standardization; 2002.

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

7   Health Products and Food Branch, Health Canada. Guidance for Manufacturers of Human Immunodeficiency Virus (HIV) Test Kits intended to be used in the Laboratory., Ottawa: Health Canada; 2011 (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 – Diagnostic Assessment: Instructions for Compilation of a Product Dossier. WHO PQDx_018 v3, 27 August 2014. Geneva: World Health Organization; 2014 (http://www.who.int/entity/diagnostics_laboratory/evaluations/141015_pqdx_018_dossier_instructions_v4.pdf?ua=1, accessed 18 July 2016).

9   User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline – Second Edition. CLSI EP17-A2. Wayne, PA: Clinical and Laboratory Standards Institute; 2008.

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

11 Interference Testing in Clinical Chemistry. Approved Guideline - Second Edition. CLSI 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: International Organization for Standardization; 2011.

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

14   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://www.who.int/entity/diagnostics_laboratory/guidance/160613_tgs2_stability.pdf?ua=1, accessed 15 July 2016).

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

16   ISO 18113:2009. In vitro diagnostic medical devices – Information supplied by the manufacturer (labelling). Part 1: Terms, Definitions and General Requirements. Geneva: International Organization for Standardization; 2009.

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

18   IEC 62366-1:2015. Medical devices – Part 1: Application of usability engineering to medical devices. Geneva: International Electrotechnical Commission; 2015.

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

20   Poffenberger K. Update: Rapid HIV Test Approval Requirements and Standards [Slides]. Silver Spring, MD: Food and Drug Administration; 2000 (http://www.fda.gov/ohrms/dockets/ac/00/slides/3649s2.htm, accessed 15 July 2016).

21   OraQuick In-Home HIV Test. Summary of Safety and Effectiveness. FDA BP120001. Silver Spring, MD: Food and Drug Administration; 2012 (http://www.fda.gov/downloads/BiologicsBloodVaccines/BloodBloodProducts/ApprovedProducts/PremarketApprovalsPMAs/UCM312534.pdf; accessed 18 July 2016).

22   Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices. 2012 (http://data.europa.eu/eli/dir/1998/79/2012-01-11,accessed 18 July 2016).

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

24   Consolidated Guidelines on HIV Testing Services. Geneva: World Health Organization; 2015 (http://who.int/hiv/pub/guidelines/hiv-testing-services/en/; accessed 15 July 2016).

25   USAID and WHO. 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, University Research Co., LLC, and the World Health Organization, Bethesda, MD, and Geneva; 2009 (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   Center for Devices and Radiological Health, FDA. Applying Human Factors and Usability Engineering to Medical Devices., Silver Spring, MD: Food and Drug Administration; 2011 (http://www.fda.gov/downloads/MedicalDevices/.../UCM259760.pdf, accessed 23 November 2016).

= = =