WHO PREQUALIFICATION TEAM:
DIAGNOSTICS

World Health Organization

Technical Specifications Series for submission to WHO Prequalification – Diagnostic Assessment

TSS-4 In vitro diagnostic medical devices (IVDs) used for the detection of high-risk human papillomavirus (HPV) genotypes in cervical cancer screening

Technical Specifications Series for submission to WHO Prequalification – Diagnostic Assessment: In vitro diagnostic medical devices (IVDs) used for the detection of high-risk human papillomavirus (HPV) genotypes in cervical cancer screening

Published 03 April 2018

ISBN 978-92-4-151381-4

Contents

Acknowledgements

The document "*Technical Specifications Series for submission to WHO Prequalification – Diagnostic Assessment: In vitro diagnostic medical devices (IVDs) used for the detection of high-risk human papillomavirus (HPV) types in cervical cancer screening"* was developed with support from the Bill & Melinda Gates Foundation and Unitaid. The document was prepared in collaboration with S. Hojvat, Maryland, United States of America (USA), M. Lanigan, Geneva, Switzerland; D. Healy, and R. Meurant, Prequalification Team – Diagnostic Assessment, WHO, Geneva, Switzerland and with input and expertise from N. Broutet; H. De Vuyst, Human Reproduction Team, Department of Reproductive Health and Research, WHO, Geneva, Switzerland; H. Cubie, Global Health Academy, University of Edinburgh, Edinburgh, United Kingdom of Great Britain and Northern Ireland; K. Cuschieri, Scottish National HPV Reference Laboratory, United Kingdom; M. de la Luz Hernandez, R. Herrero, International Agency for Research on Cancer, Lyon, France.

This document was produced under the coordination and supervision of M. Lanigan, R. Meurant and I. Prat, Prequalification Team – Diagnostic Assessment, WHO, Geneva, Switzerland.

List of contributors

A technical consultation on "WHO Prequalification requirements for in vitro diagnostic medical devices (IVDs) used for the detection of high-risk Human Papillomavirus (HPV) types in cervical cancer screening" was held in Geneva, Switzerland from 22 – 23 May 2017.

Meeting participants: V. Alcon, In Vitro Diagnostic Devices Evaluation Division, Medical Devices Bureau, Therapeutic Products Directorate, Health Canada, Government of Canada, Canada; N. Almond, The National Institute for Biological Standards and Control, The Medicines and Healthcare Products Regulatory Agency, Potters Bar, United Kingdom; M. Almonte, International Agency for Research on Cancer, Lyon, France; M. Arbyn, Belgian Cancer Centre/Unit of Cancer Epidemiology, Scientific Institute of Public Health, Brussels, Belgium; N. Comella, Center for Devices and Radiological Health (CDRH), U.S. Food and Drug Administration, Silver Spring, Maryland, USA; H. Cubie, Edinburgh, United Kingdom[1]; S. de Sanjose, Catalan Institute of Oncology, Barcelona, Spain; J. Dillner, Karolinska Institute, Stockholm, Sweden; J. Gage, National Cancer Institute, Bethesda, USA; I. Heard, Hôpital Tenon, Paris, France; S. Hojvat, Independent consultant, Maryland, USA; J. Jeronimo, Independent consultant, Seattle, USA; S. Luciani, Pan American Health Organization. Washington, USA; M. Maza, Basic Health International, San Salvador, El Salvador; R. Murillo, Centro Oncologico Javeriano, Bogota, Columbia; M. Murtagh, The Murtagh Group, Woodside, California, USA; V. Pereira de Oliveira, Agência Nacional de Vigilância Sanitária (ANVISA), Brazil; A. Picconi, National Institute of Infectious Diseases, PAHO Regional WHO HPV Reference Laboratory, Buenos Aires, Argentina; M. Poljak, Institute of Microbiology and Immunology, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia; M. Schiffman, National Cancer Institute, Bethesda, USA[2]; R. Sethuraman, Health Sciences Authority, Singapore; E.

[1] Participated via web conferencing
[2] Participated via web conferencing

Unger, Centers for Disease Control and Prevention, Atlanta, USA; D. Wilkinson, The National Institute for Biological Standards and Control, The Medicines and Healthcare Products Regulatory Agency, Potters Bar, United Kingdom; A Williamson, University of Cape Town, Cape Town, South Africa; Y. Qiao, Peking Union Medical College, Beijing, China; L. Zaritsky, Center for Devices and Radiological Health (CDRH), U.S. Food and Drug Administration, Silver Spring, Maryland, USA.

Meeting participants with observer status: P. E. Castle, Albert Einstein College of Medicine, New York, USA; K. Cuschieri, Scottish National HPV Reference Laboratory, United Kingdom; A. Kaufmann, Charité-Universitätsmedizin, Berlin, Germany.

WHO Secretariat: H. Ardura; D. Healy; M. Lanigan; R. Meurant; M. Perez Gonzalez; I. Prat, Prequalification Team – Diagnostic Assessment Group; T. Zhou, Technologies Standards and Norms, Essential Medicines Programme; K. Blondeel; N. Broutet; I. Toskin, H. De Vuyst; Human Reproduction Team, Department of Reproductive Health and Research; WHO, Geneva, Switzerland; M. de la Luz Hernandez; R. Herrero, International Agency for Research on Cancer, Lyon, France.

First round of comments were received from the following: V. Alcon, In Vitro Diagnostic Devices Evaluation Division, Medical Devices Bureau, Therapeutic Products Directorate, Health Canada, Government of Canada, Canada; N. Almond, The National Institute for Biological Standards and Control, The Medicines and Healthcare Products Regulatory Agency, Potters Bar, United Kingdom; M. Arbyn, Belgian Cancer Centre/Unit of Cancer Epidemiology, Scientific Institute of Public Health, Brussels, Belgium; H. Cubie, Global Health Academy, University of Edinburgh, Edinburgh, United Kingdom; K. Cuschieri, Scottish National HPV Reference Laboratory, United Kingdom; S. de Sanjose, Catalan Institute of Oncology, Barcelona, Spain; J. Gage, National Cancer Institute, Bethesda, USA; I. Heard, Hôpital Tenon, Paris, France; J. Jeronimo, Independent consultant, Seattle, USA; V. Pereira de Oliveira, ANVISA, Brazil; A. Picconi, National Institute of Infectious Diseases, PAHO Regional WHO HPV Reference Laboratory, Buenos Aires, Argentina; R. Sethuraman, Health Sciences Authority, Singapore; E. Unger, Centers for Disease Control and Prevention, Atlanta, USA; W. Urassa, WHO, Geneva; D Wilkinson, The National Institute for Biological Standards and Control, The Medicines and Healthcare Products Regulatory Agency, Potters Bar, United Kingdom; A. Williamson, University of Cape Town, Cape Town, South Africa; L. Zaritsky, Center for Devices and Radiological Health (CDRH), U.S. Food and Drug Administration, Silver Spring, Maryland, USA.

The draft technical specifications document was posted on the WHO website for public consultation on 18 September 2017 for a 2-month commenting period. 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.

Public comments were received for consideration from AdvaMed, Washington D.C. USA; N. Marlowe, Associate Director, Global Scientific Affairs, Research &

Development, Abbott Molecular, Des Plaines, Illinois, USA; Cepheid AB, Röntgenvägen; Sweden; In Vitro Diagnostic Devices Evaluation Division, Medical Devices Bureau, Therapeutic Products Directorate, Health Canada, Government of Canada; ISO/TC 212 Secretary on behalf of ISO/TC 212, International Standards Organization, Geneva, Switzerland; M. Kondratovich, Associate Director for Clinical Studies, Office of In Vitro Diagnostics and Radiological Health (OIR), Center for Devices and Radiological Health (CDRH), U.S. Food and Drug Administration, Silver Spring, Maryland, USA; N. Gachuhi, Women's Health Lead, Global Health Technologies, Good Global Washington, USA; P. W. Kamau, Pharmacy and Poisons Board, Nairobi, Kenya; M. B. Rumaney, Goodwood, Western Cape, South Africa; V. Tsu, Director, Cervical Cancer Prevention Project, Associate Director, Reproductive Health Global Program, PATH, Seattle, Washington, USA; QIAGEN, Germantown, Maryland, USA.

Abbreviations

ANOVA	analysis of variance
CIN	cervical intraepithelial neoplasia
HPV	human papillomavirus
IFU	instructions for use
IVD	in vitro diagnostic medical device
IU	international units
POC	point of care
ROC	receiver operated curve
spp.	species
TSS	Technical Specifications Series
US FDA	U.S. Food and Drug Administration
v/v	volume per volume
WHO	World Health Organization
w/v	weight per volume

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 tests for the detection of human papillomavirus (HPV). This document is relevant to IVDs that detect HPV genotypes that are associated with cervical cancer *(1)*. Although this does not exclude those IVDs that may claim to detect other HPV associated cancers (e.g. anal cancer), IVDs will only be prequalified on the basis of evidence that pertains specifically to detection of HPV types associated with *cervical* cancer.

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, are aligned with published guidance, standards and/or regulatory documents. Although references to source documents are provided, in some cases WHO Prequalification has additional requirements. A full list of the individual studies is provided in Section D.

- Part 1 lists the analytical studies that are required to assess the ability of the IVD to measure the relevant analyte(s).
- Part 2 lists the clinical studies that are required to support the clinical performance of an IVD, and 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.

Clinical utility studies, 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 each population or healthcare setting, do not fall under the scope of WHO Prequalification and are not included in this document. Clinical utility studies usually inform programmatic strategy and are thus the responsibility of programme managers, ministries of health and other relevant bodies in individual WHO Member States.

B. Other guidance documents

This document should be read in conjunction with other 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/

C. Performance principles for WHO Prequalification

C.1 Intended use

An IVD submitted for WHO 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 nucleic acid from specified high-risk HPV genotypes; to differentiate detection of HPV genotypes; etc. as appropriate);
- the testing population for which functions are intended and the ages of individuals for which this test has clinical relevance (e.g. females over 30 years of age at risk of cervical cancer);
- the intended operational setting and user (e.g. for professional use in a laboratory setting, or point-of-care[3] (POC));
- the intended specimen type(s), collection media and collection device(s) and/or method(s), and whether the specimen can be self-collected; and
- indication(s) for use (e.g. primary screening of women over 30 years of age).

WHO cervical cancer screening guidelines[4] have been published and help define the role that any HPV test will play in patient management. They are also useful in defining an intended use statement for IVDs used for the detection of high risk HPV genotypes.

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

For WHO Prequalification submission, clinical studies shall be conducted using the specimen types, collection method and collection material (e.g. self-collected or health care provider-collected swabs or collection media such as suspensions of exfoliated cervical cells in cytology and/or other media) that are claimed in the instructions for use (IFU).

Prequalified IVDs in low- and middle-income countries are likely to be used by:

- laboratory professionals[5] either in centralised testing laboratories or at POC,
- healthcare workers[6] trained in the use of the test at POC.

Similarly, it is anticipated that laboratory professionals or healthcare workers may test either self-collected or healthcare provider-collected specimens at POC.

---

3 Point-of-care in-vitro diagnostic testing (POC) refers to decentralized testing that is performed by a minimally trained healthcare professional near a patient and outside of central laboratory testing facilities. It does not refer just to sample collection procedures.

4 http://apps.who.int/iris/bitstream/10665/94830/1/9789241548694_eng.pdf

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

6 Any person who performs functions related to healthcare delivery and has received a formal professional or paraprofessional certification or tertiary education degree.

The complexity of a test shall be clearly elucidated in the IVD IFU and reflected in the risk analysis.

Depending on the intended use of an IVD, analytical and clinical studies shall be designed to consider not only the diversity of knowledge and skills across the population of IVD users, but also the likely operational settings in which testing will occur. It is a manufacturer's responsibility to ensure that the risk assessment for an IVD reflects the intended operational settings, including laboratory or service delivery complexity, user expertise and testing population.

C.3 Applicability of supporting evidence to IVD under review

Analytical and clinical studies shall be undertaken using the specific, final (locked-down design) version of the IVD intended to be submitted for WHO Prequalification. For WHO Prequalification, design lock-down is the date that final documentation is signed off, including quality control and quality assurance specifications, and the finalized method is stated in the IFU. Where this is not possible, a justification shall be provided; additional supporting evidence may also be required. This may occur in the case of minor variations to the design where no negative impact on performance has been demonstrated.

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 production (or manufacturing, purification, etc.) runs of critical reagents, representative of routine manufacture. It is a manufacturer's responsibility to ensure, via risk analysis of its IVD that the minimum numbers of lots chosen for estimating performance characteristics considers the variability in performance likely to arise from the diversity of key components and their formulation.

The true HPV nucleic acid status and, if applicable, HPV genotype shall be determined using a suitable reference standard. For WHO purposes this should be a standard that currently is at a developed stage of technical capability based on the relevant consolidated findings of science, technology and experience (commonly referred to as state of the art). Justification for the choice of standard shall be provided such as a laboratory based assay that has undergone comprehensive pre-market assessment by a stringent regulatory authority.

Estimation (and reporting) of IVD performance shall include the rate of invalid test results along with the confidence interval. The cause of the invalid results should be reported if available (such as sample inadequacy and instrument error). Data should be presented in clear and understandable format (for examples see reference *(6)*).

For analytical studies, it is acceptable to use contrived specimens (e.g. cell lines or plasmid spiked into normal human cells in culture medium or negative clinical matrix, as appropriate) which have been prepared in a validated and standardized manner. Where contrived specimens are used, for anything other than a negative clinical matrix background, specimen type equivalence shall be demonstrated, for example by conducting a paired sample "limit of detection" study. The material chosen should use the entire assay system from specimen preparation to interpretation. Dilutions of a high-concentration specimen may be used, if they are in an appropriate clinical matrix.

For certain analytical studies it may be acceptable to generate evidence of performance in a single HPV genotype (e.g. HPV 16) as a surrogate for performance in other high-risk genotypes (e.g. where potential interference with detection of HPV 16 is used to infer

potential interference with other genotypes). A justification for this approach, including a relevant demonstration of the relationship between genotypes, shall be provided.

For quantitative assays, additional requirements may apply such as linearity studies. Please contact WHO Prequalification for more information on these requirements.

Clinical studies shall be based on testing human specimens only sourced from population cohorts reflective of the intended use.

The use of well-characterised repository specimens may be acceptable if they are relevant to the IVD under assessment, taking into consideration:

- the collection device (swab, brush or other collection device; whether specimens have been provider- or self-collected); collection media; storage conditions (including age of the specimen);
- the nucleic-acid target and its stability; and/or
- any requirement for testing in fresh specimens only.

D. Table of Requirements

PART 1	**Establishing analytical performance characteristics**
1.1	**Specimen type and collection**
1.1.1	Demonstration of equivalence between contrived specimens and clinical specimens
1.1.2	Demonstration of equivalence between specimen collection methods
1.2	**Specimen collection, storage and transport**
1.2.1	Specimen collection, processing and stability
1.3	**Precision of measurement**
1.3.1	Repeatability, reproducibility
1.4	**Performance panels**
1.4.1	Genotype panels
1.5	**Carry-over contamination**
1.5.1	Carry-over contamination
1.6	**Analytical sensitivity**
1.6.1	Analytical sensitivity
1.6.2	Validation of assay cut-off
1.7	**High dose hook effect**
1.7.1	High dose hook effect
1.8	**Analytical specificity**
1.8.1	Potentially interfering substances
1.8.1.1	Endogenous
1.8.1.2	Exogenous
1.8.2	Cross-reactivity
1.9	**Metrological traceability of calibrators and control material values**
1.9.1	Metrological traceability of calibrators and control material values
1.10	**Stability**
1.10.1	Shelf-life (including transport stability)
1.10.2	In-use stability (open pack or open vial stability)
1.11	**Flex studies**
1.11.1	Flex studies/robustness
PART 2	**Establishing clinical performance characteristics**
2.1	**Diagnostic sensitivity and specificity**
2.1.1	Diagnostic sensitivity
2.1.2	Diagnostic specificity
2.2	**Qualification of usability for self-collection and/or POC testing**
2.2.1	Label comprehension (including IFU)
2.2.2	Results interpretation

Part 1 Establishing analytical performance characteristics

Aspect	Testing requirements	Notes on testing requirements	Source Documents
1.1 Specimen type and collection			
1.1.1 Demonstration of equivalence between contrived specimens and clinical specimens	Equivalence between contrived specimens and clinical specimens shall be demonstrated to enable contrived specimens to be used instead of clinical specimens in certain analytical HPV test analytical studies. For example: Performance can be compared by testing; • serial dilutions of clinical HPV positive cervical specimens with a targeted level of analyte; and • serial dilutions of contrived specimens derived from either, HPV infected cell lines, HPV genomic DNA plasmids or RNA transcripts in sample collection buffer with targeted levels of analyte (see note 4). Equivalence shall be determined for each reportable type or type group (e.g. HPV 16, HPV 18, etc.), as appropriate. For assays that detect more than 1 genotype, but do not differentiate the genotype, the justification as to whether to study each genotype shall be risk- and evidence-based.	1. The relationship between IVD performance in claimed specimen types and contrived specimens used for analytical studies shall be established. The design of subsequent studies shall then take that relationship into account. 2. If there is no equivalence between claimed specimen types and contrived specimens then the impact that this will have on each subsequent performance claim shall be fully understood and described. 3. Specimens should be chosen that are considered to have low to moderate concentrations of the analyte (include levels of analyte at relevant medical decision points and near the cut-off decision point) 4. Equivalence may be established as part of analytical sensitivity.	Technical Guidance Series for WHO Prequalification – Diagnostic Assessment TGS-3*(2.)*
1.1.2 Demonstration of equivalence between specimen collection methods methods	For each claimed collection device or method (e.g. health care provider collection: brush/spatula; self-collection: tampon, etc.) and/or claimed liquid collection media (specify brand), as appropriate (see note 2). Performance may be compared by testing serial dilutions of clinical HPV positive cervical specimens with a targeted level of analyte.	1. Equivalence between specimen collection methods may be established as part of clinical studies. 2. The specimen collection method may use different specimen types, e.g. vaginal specimen collection using a tampon, or cervical specimen collection using a brush. 3. Paired specimens should be used.	

Aspect	Testing requirements	Notes on testing requirements	Source Documents
	Equivalence shall be determined for each reportable type or type group (e.g. HPV 16, HPV 18, etc.), as appropriate) For assays that detect more than 1 genotype, but do not differentiate the genotype, the justification as to whether to study each genotype shall be risk- and evidence-based.		
1.2 Specimen storage and transport			
1.2.1 Specimen collection, processing and stability	Real time studies taking into account: • storage conditions (duration at different temperatures, temperature limits, freeze/thaw cycles) • transport conditions (see note 1) • intended use (see note 1) • specimen collection and/or transfer devices intended to be used with the IVD (see 1.1.2)	1. Recommended storage conditions should be established using a panel of real clinical cervical specimens. 2. Evidence shall be provided which validates the maximum allowable time between specimen collection and its processing or addition to the IVD in the setting where testing takes place.	
1.3 Precision of measurement			
1.3.1 Repeatability, reproducibility	Both repeatability and reproducibility (see note 1) should be estimated using panels with defined analyte levels (see note 2) of several of each of the following: • Negative specimens • Low positive specimens (with a concentration of analyte just above the assay cut-off such that results of repeated tests of this sample are positive approximately 95% of the time) • Moderately positive specimens (with a concentration at which one can anticipate positive results approximately 100% of the time e.g., approximately two to three times the concentration of the assay cut-off) If applicable, estimation of precision shall include the	1. E.g. within- or between-run, -lot, -day, -site, etc. 2. For HPV screening assays, precision shall be determined at a minimum, for HPV 16 or HPV 18. For genotyping assays in addition to HPV 16 and HPV 18, other claimed high-risk HPV genotypes may be analyzed as a pool, provided a justification for such an approach is given. 3. The testing panel shall be the same for all operators, lots and sites. 4. The number of invalid tests shall be reported. 5. Each lot should comprise different production (or manufacturing, purification, etc.) runs of critical reagents, representative of routine manufacture. 6. Results shall be statistically analyzed by ANOVA to identify and isolate the sources and extent of any variance. In addition, the percentage of correctly-	CLSI EP05-A3 *(3)* EN 13612:2002 *(4)* CLSI EP12-A2 *(5)* U.S. FDA *(6)*

Aspect	Testing requirements	Notes on testing requirements	Source Documents
	effect of extraction efficiency. Each panel member shall be tested: • in 5 replicates • using 3 different lots • over 5 days (not necessarily consecutive) with 1 run in that 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 note 8). Testing should be done: • by personnel representative of intended users • unassisted • using only those materials provided with the IVD (e.g. IFU, labels and other instructional materials).	identified, incorrectly-identified and invalid results shall be tabulated for each specimen and be separately stratified according to each of site, lot, etc. 7. To understand irregularities in results obtained, at least 2 lots should be tested at each of the 3 testing sites. 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 1.11.1 Flex studies) and may be addressed as part of clinical studies in representative populations (see Part 2). 9. Panel members derived from cell lines and/or real clinical specimens should be processed as real specimens from suspension in a relevant collection media step to extraction step. 10. Contrived specimens and or well characterized real clinical specimens can be used to make any member of the precision/reproducibility panels. 11. Alternative methods used to establish repeatability and reproducibility performance of the assay shall be discussed with WHO in advance of dossier submission.	
1.4 Performance panels			
1.4.1 Genotype panels	Testing of the IVD in suitable performance panels (see note 1) or clinical specimens that represent the targeted genotypes to show diversity of variations within genotypes (e.g. proficiency panels or similar comprising relevant, high-risk genotypes).	1. Suitable performance panels should consist of members which have been well characterized by other recognized HPV molecular assays (see section C3) or *in-silico* analysis. The panel may comprise a combination of both clinical and contrived specimens. 2. Testing shall be performed using more than 1 lot of the final design (locked-down). 3. All confirmed positive specimens should be detected	Technical Guidance Series for WHO Prequalification – Diagnostic Assessment TGS-3 *(2)* WHO Prequalification – Diagnostic Assessment PQDx_018 *(7)*

Aspect	Testing requirements	Notes on testing requirements	Source Documents
		by the IVD. For low positive specimens with a concentration of analyte just above the assay cut-off, results of repeated tests of specimens may be positive approximately 95% of the time.	
1.5 Carry-over contamination			
1.5.1 Carry-over contamination	The potential for carry-over contamination or similar shall be investigated using a minimum of 5 multiple runs of alternating high-positive and negative specimens.	1. For IVDs where specimen volumes are compartmentalized, the potential for carry-over contamination may be performed as part of flex studies (see section 1.11 Flex Studies); otherwise, a stringent carry-over study shall be performed.	European Commission decision on CTS *(8)* Haeckel R *(9)*
1.6 Analytical sensitivity			
1.6.1 Analytical Sensitivity	Analytical sensitivity as expressed by limit of detection, may be determined by testing serial dilutions (at least 8 dilutions, 15-20 replicates tested per dilution) of suitable biological reference materials (e.g. WHO 1st International Standard for HPV) or a secondary standard calibrated from it (see note 1). Material with well characterized copy number should only be used in cases where material with an assigned value in International Units (IU) is not available. Analytical sensitivity shall be estimated by determining the lowest concentration for which the rate of detection is 95%, where possible. (See also note 3, as well as section 1.9 "Metrological traceability of calibrators and control material values").	1. Analytical sensitivity shall be determined for each claimed genotype, where a suitable biological reference material exists. 2. Testing should be conducted using more than 1 lot of the final design (locked-down). Where this is not possible, justification shall be provided. 3. Analytical sensitivity may be demonstrated in HPV negative clinical specimen matrix but shall use the entire assay system from specimen preparation to interpretation. Alternative matrix (e.g. collection fluid containing a negative cell line background) may be used provided the equivalency is demonstrated between the clinical specimen matrix and the alternative matrix. 4. Analytical sensitivity studies may include the use of: • dilutions of claimed genotype plasmids in pooled HPV negative patient specimens collected in a relevant collection medium • dilutions of cell lines (for DNA detection assays) or RNA in vitro transcripts or plasmid stocks in pooled HPV negative patient specimens or collection fluid containing a negative cell line background	European Commission decision on CTS *(8)* CLSI EP17-A2 *(11)*

Aspect	Testing requirements	Notes on testing requirements	Source Documents
		5. The estimate of analytical sensitivity should be confirmed by separately testing an additional 20 replicates. 6. If reporting in copies/mL, the manufacturer shall ensure that the copy number is determined with standard nucleic acid quantitation method (e.g. A260).	
1.6.2 Validation of assay cut-off	HPV assays are generally qualitative and do not use a numerical value of assay cut-off. Nevertheless, the way in which: • a screening HPV IVD was designed to differentiate positive high-risk HPV genotype specimens from negative high-risk genotype HPV specimens shall be demonstrated; and • for genotyping assays to differentiate accurately which high-risk HPV genotype was found positive in a screening HPV assay, shall be demonstrated.	1. Selection of the appropriate clinical cut-off can be justified by the relevant levels of sensitivity and specificity that are based on Receiver Operating Curve (ROC) analysis of pilot studies, testing well characterized clinical cervical samples and including information on central histology results (cervical intraepithelial neoplasia (CIN) 2 and CIN3).	
1.7 High dose hook effect			
1.7.1 High dose hook effect	For each claimed analyte, the potential for competitive inhibition arising from high analyte (nucleic acid) concentration, where applicable, shall be determined: • using multiple, strong positive 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. The manufacturer shall provide a justification where high dose hook effect is not applicable (e.g. due to the design of the assay). 2. If the assay detects a large group of genotypes, but does not differentiate between them, then testing may be conducted in those genotypes that are closely related or those that are of highest risk. For genotyping assays, all targeted genotypes shall be tested. HPV DNA in plasmids or viral transcripts may be used for these studies. 3. Specimens shall be chosen that have a high analyte concentration, as determined using an IVD method other than the IVD intended to be prequalified. This second method shall be of a design not subject to competitive inhibition.	

Aspect	Testing requirements	Notes on testing requirements	Source Documents
		4. An increase in signal upon dilution of a specimen implies competitive interference. 5. If there is evidence of competitive interference, this information should be added to the IFU and mitigation actions identified.	
1.8 Analytical specificity			
1.8.1 Potentially interfering substances	The potential for false results (false negatives and false positives) arising from interference by the substances/conditions listed below shall be determined • For at least one of the most clinically relevant HPV genotypes (see note 1). • Each substance/condition represented, with a minimum of 3 specimens per substance/condition Testing shall be undertaken in both HPV-negative and HPV-positive specimens, unspiked or spiked, with each potentially interfering substance.	1. The risk assessment conducted for an IVD shall identify substances at medically relevant levels for which the potential for interference can reasonably be expected for the analyte being detected (e.g. HPV 16). 2. HPV genotypes may be tested as plasmids or in vitro transcripts. 3. The potential for interference may be addressed using either contrived laboratory specimens or as part of clinical studies in representative populations (see Part 2). • The potential for interference shall be investigated for each claimed specimen type, specimen collection medium and/or collection device, as appropriate. 4. HPV positive specimens used in the studies should have low reactivity. 5. Any observed interference shall be further investigated and performance limitations of the IVD reported in the IFU. 6. 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.	U.S. FDA *(6)* European Commission decision on CTS *(8)* CLSI EP07-A2 *(11)*
1.8.1.1 Endogenous	• Leukocytes ($1x10^6$ cells/ml) • Whole blood (human) >2% v/v • Cervical mucus 0.5-5% (w/v)		
1.8.1.2 Exogenous	• Acetic acid • Contraceptive and feminine hygiene products, including, contraceptive jelly, anti-fungal cream, spermicide, anti-yeast medication, vaginal lubricant, vaginal deodorant spray, intravaginal hormones		
1.8.2 Cross-reactivity	Determination of the potential for false positive results arising from cross-reactivity (see note 1) for a minimum of 100 specimens, including, where possible, at least 3-5 each of: • Potential cross-reacting organisms found in the	1. The types of conditions tested for should be risk-based, taking into consideration the operational setting as well as the intended users for the analyte being detected. 2. Where either the scientific literature and/or risk	

Aspect	Testing requirements	Notes on testing requirements	Source Documents
	reproductive tract • Targeted and non-targeted HPV genotypes from the same species group • Other unrelated conditions known to cause cross-reactivity in HPV IVDs.	analysis identifies the potential for false results in co-infected individuals (e.g. decreased sensitivity or specificity), further investigation shall be undertaken using both HPV-negative and HPV-positive specimens. 3. *In silico* analysis may be considered for some organisms identified as potential risks for cross-reactivity and/or competitive inhibition. These may include: • Non-HPV viral infections, including: human immunodeficiency virus, hepatitis B, C infection, adenovirus, cytomegalovirus, herpes simplex virus 1 and 2 • Bacteria/yeast/parasites, including: *Lactobacillus acidophilus, Staphylococcus epidermidis, Staphylococcus aureus, Enterococcus faecalis, Streptococcus pyogenes, Streptococcus agalactiae, Corynebacterium* spp., *Escherichia coli, Enterococcus* spp., *Clostridium* spp., *Peptostreptococcus* spp., *Klebsiella* spp., *Enterobacter* spp., *Proteus* spp., *Pseudomonas* spp., *Bacteroides* spp., *Bifidobacterium* spp., *Fusobacterium* spp., *Chlamydia trachomatis* and *Neisseria gonorrhoeae*, *Candida albicans, Trichomonas vaginalis.* 4. Complete information about the *in-silico* analysis should be provided including detail of the database used, sequence used for the analysis, and its relevance to the product. 5. Any observed cross-reactivity shall be further investigated and performance limitations of the IVD reported in the IFU.	

Aspect	Testing requirements	Notes on testing requirements	Source Documents
1.9 Metrological traceability of calibrators and control material values			
1.9.1 Metrological traceability of calibrators and control material values	As applicable, the traceability of an assay-specific quality control specimen to a validated reference material shall be demonstrated (e.g. WHO 1st International Standard for HPV) or a secondary standard calibrated from it. Material with well characterized copy number should only be used in cases where material with an assigned value in International Units (IU) is not available.	1. The IVD should include accurate external positive and negative controls that are processed as specimens through all steps of testing. Similarly, an internal amplification control should be incorporated into the assay design to check for integrity of reagents, sample adequacy (cellularity and preservation of target) and the presence of inhibitors. 2. Appropriate to the IFU, the positive external control should contain target nucleic acid(s) for an appropriate genome(s) in plasmid DNA at levels approximately two times the estimated cut-off in an appropriate buffer or sample transport media. 3. If reporting in copies/mL, the manufacturer shall ensure that the copy number is determined with standard nucleic acid quantitation method (e.g. A260)	WHO *(12)*
1.10 Stability			
	Replicate testing shall be undertaken using a panel consisting, for each claimed analyte reported type or type group, of at least: • 1 negative specimen • 2 low positive specimens (with a concentration of just above the assay cut-off, such that results of repeated tests of this sample are positive approximately 95% of the time) • 1 moderately positive specimen (with a concentration at which one can anticipate positive results approximately 100% of the time e.g., approximately two to three times the concentration of the clinical cut-off)	1. Where nucleic acid reagents (primers, probes, etc.) are essentially identical in design, such that their chemical lability can be considered identical, stability may be determined using a representative genotype (e.g. HPV 16). 2. Each lot should comprise different production (or manufacturing, purification, etc.) runs of critical reagents, representative of routine manufacture. 3. The number of invalid tests with each kit lot shall be reported. 4. Determination of shipping stability shall be performed using simulated extreme stress conditions, ensuring that application of those conditions is consistent and controlled.	ISO 23640:2011 *(13)* CLSI EP25-A *(14)* Technical Guidance Series for WHO Prequalification – Diagnostic Assessment TGS-2 *(15)* ASTM D4169-14 *(16)*

Aspect	Testing requirements	Notes on testing requirements	Source Documents
1.10.1 Shelf-life (including transport stability)	• Real time studies using a minimum of 3 lots of final design product • Lots are "transport stressed" (simulated) before real time studies are undertaken • IVD in final packaging subjected to drop-shock testing	5. 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 can be 12 months. 6. Accelerated studies do not replace the need for real time studies. 7. In-use stability of labile components shall be conducted using components in their final configuration.	
1.10.2 In-use stability (open pack or open vial stability)	• Minimum of 1 lot, using panel(s) compiled as described above • Testing of all labile components (e.g. buffers vials, sealed cartridges, etc.; see note 7) • On-board stability shall be tested for an IVD used with an instrument.		
1.11 Flex studies			
1.11.1 Flex studies/ robustness	The influence of the following factors on expected results (both reactive and non-reactive) shall be considered as applicable: • Specimen and/or reagent volume • Buffer pH • IVD instrument sturdiness (including the effect of non-level work surface) • Lighting, humidity and barometric pressure (simulating high altitude), • Handling contamination (e.g. from latex, powder, hand lotion, sweat, and/or soap, etc. as appropriate) • Operating temperature • Instrumentation (both extraction and amplification) including: o Ruggedness (including the effect of vibration from other instruments), o Impact of dust and mould on componentry (e.g. optics)	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. 2. The factors listed opposite should be investigated in ways that not only reflect, but also exceed, likely operating conditions in lower- and middle-income countries so that the limitations of the device can be understood. For example, in addition to investigating deviations of temperature within those claimed in the IFU (in the middle and at both lower and upper extremes of a claimed temperature range), temperature ranges should be investigated that exceed those of claimed operating conditions and which cause test failure (incorrect/invalid results).	WHO Prequalification – Diagnostic Assessment PQDx_018 *(7)*

Aspect	Testing requirements	Notes on testing requirements	Source Documents
	o Consistency of power supply o Software validation (including verification of built-in fail-safe and alert mechanisms)		

Part 2 Establishing clinical performance characteristics

Aspect	Testing requirements	Notes on testing requirements	Source documents
2.1 Diagnostic sensitivity and specificity			
	Testing shall be conducted: • at different, geographically diverse settings representing screening populations (minimum of 2 regions; see note 3) • by a variety of intended users (see note 1) • using more than 1 lot • using specimens from unvaccinated subjects representing the ages of women at risk of cervical cancer, and from lesions representing different stages of severity (see note 4). The procedure for selection of study specimens, how these represent a screening population and how bias has been addressed shall be clearly described.	1. Prequalified HPV IVDs are likely to be used at POC by a diversity of users. For WHO Prequalification purposes, these should be considered as the intended users in addition to trained laboratory professionals. 2. Clinical performance shall be established using specimens that correspond directly to claims made in the IFU. Clinical studies shall include details of collection medium and collection device, etc., as appropriate. Specimens may include those collected from colposcopy clinics, provided sources of collection are reported so that the potential for sampling bias can be understood (see also note 7). 3. Where a claim is made for self-collection of specimens, demonstration of clinical performance may be established by pair-wise comparison of matched self-collected and provider-collected known HPV positive specimens from a colposcopy clinic and similarly matched HPV negative specimens from different locations which are screening women for cervical cancer /HPV infection. 4. The HPV vaccine status of women whose specimens are obtained shall be reported, if available. 5. Testing should include some specimens that are positive for two or more claimed HPV genotypes. However, determination of sensitivity shall be reported for individual genotypes, where applicable. 6. Investigators, patients and clinicians, including those conducting colposcopy and histology, should be blinded to a patient's HPV status until colposcopy, histology is complete. 7. CIN2+ status shall be determined either by: • Examination of pathology in conjunction with determination of p16 immunochemistry status, or	U.S. FDA *(6)* European Commission decision on CTS *(8)* Meijer et al. *(17)* Arbyn et al. 2015 *(18)* Arbyn et al. 2016 *(19)*
2.1.1 Diagnostic sensitivity	Testing of: • At least 60 well characterized, clinically confirmed CIN2+ specimens		
2.1.2 Diagnostic specificity	Testing of: • At least 100 well characterized, clinically confirmed non-CIN2+ specimens		

Aspect	Testing requirements	Notes on testing requirements	Source documents
		• Review by a panel of at least three expert pathologists. 8. Where possible, biopsy specimens should be available for future review. 9. Problematic specimens, those with unexpected results but which otherwise meet selection criteria for a study, shall not be systematically excluded from analysis. 10. Each lot should comprise different production (or manufacturing, purification, etc.) runs of critical reagents, representative of routine manufacture. 11. All results that are indeterminate by the IVD shall be included in the denominator data for analysis. 12. All invalid results shall be recorded and evaluated in comparison to the reference result. Invalid results should be reported as individual categories (e.g. internal control failure, extraction failure, etc.) and not aggregated. Invalid results should be analyzed separately in the final performance calculations. 13. Estimates of diagnostic sensitivity and specificity shall be reported with 95% confidence intervals.	

Aspect	Testing requirements	Notes on testing requirements	Source documents
2.2 Qualification of usability for self-collection and/or POC testing			
2.2.1 Label comprehension (including IFU)	Testing of subjects to assess ability of intended users to correctly comprehend key messages from packaging and labelling that relate to self-collection and/or POC testing: • Understanding key warnings, limitations and/or restrictions, including correct use of self-collection methods and equipment (POC and self-collection). • Proper test procedure (POC only) • Test result interpretation (POC only). Studies shall include at least 15 intended users, users including those whose native language may not be the language of the IFU if necessary, to demonstrate comprehension of key messages in each population described above.	1. Intended only for IVDs which include a claim for POC testing and/or self-collection of specimen for subsequent testing by a laboratory/healthcare professional. 2. IFU and labelling should be clear and easy to understand. Use of pictorial instructional material is encouraged. 3. For IVDs which include a claim for specimen collection and testing during pregnancy, consideration shall be given to the potential risk associated with specimen collection materials (e.g. brush, etc.).	European Parliament IVD regulations *(20)* Backinger CL and Kingsley PA *(21)*
2.2.2 Results interpretation	For POC tests, intended users to interpret the results of contrived IVDs (e.g. static/pre-made tests, or similar) to assess their ability to correctly interpret pre-determined test results. Contrived tests should be made to demonstrate the following potential test result interpretations, as appropriate: • Not-detected. • Range of invalid results. • Reactive. • Weak reactive. Testing subjects to consist of at least 15 intended users including those whose native language may not be the language of the IFU if necessary from at least two geographically diverse populations to demonstrate correct interpretation of simulated test results.	1. Intended for IVDs which include quantitative outputs (e.g. cycle times, viral load) that must be interpreted by a user in a POC setting. 2. Study group may include subjects recruited as part of the label comprehension study.	

References

1. Schiffman M, Clifford G, Buonaguro F. Classification of weakly carcinogenic human papillomavirus types; addressing the limits of epidemiology at the borderline. Infectious Agents and Cancer. 2009; 4:1-8

2. Technical Guidance Series for WHO Prequalification – Diagnostic Assessment. Principles for Performance studies, TGS–3. Geneva: World Health Organization; 2016 http://apps.who.int/iris/bitstream/10665/258985/1/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf?ua=1 accessed 09 October 2017.

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

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

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

6. Food and Drug Administration. Establishing the Performance Characteristics of In Vitro Diagnostic Devices for the Detection or Detection and Differentiation of Human Papillomaviruses. Guidance for Industry and Food and Drug Administration Staff. 2017. Available at: https://www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-gen/documents/document/ucm458179.pdf f, accessed 13 February 2018

7. 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 09 September 2017.

8. 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/eli/dec/2009/108(1)/oj, accessed 15 March 2017).

9. Haeckel R. Proposals for the description and measurement of carry-over effects in clinical chemistry. Pure Applied Chemistry 1991; 63:301-306.

10. 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.

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

12. Manual for the preparation of secondary reference materials for in vitro diagnostic assays designed for infectious disease nucleic acid or antigen detection: Calibration to WHO International Standards; Geneva: World Health Organization; 2016. Available at: http://www.who.int/biologicals/expert_committee/WHO_Manual_Calibration_of_secondary_standards_final_mn.pdf?ua=1 accessed 02 June 2017.

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

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

15. Technical Guidance Series for WHO Prequalification – Diagnostic Assessment. Establishing stability of in vitro diagnostic medical devices, TGS–2. Geneva: World Health Organization; http://apps.who.int/iris/bitstream/10665/259742/1/WHO-BS-2017.2304-eng.pdf?ua=1/ accessed 08 October 2017.

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

17. Meijer, C. J. L. M. Guidelines for human papillomavirus DNA test requirements for primary cervical cancer screening in women of 30 years and older. Int J Cancer. 2009; 124(3): 516–520.

18. Arbyn M, Snijders PJ, Meijer CJLM, Berkhof H, Cuschieri K, Kocjan BJ , Poljak M. Which high-risk HPV assays fulfil criteria for use in primary cervical cancer screening? Clin Microbiol Infect 2015; 21: 817-26.

19. Arbyn M, Depuydt C, Benoy I, Bogers J, Cuschieri K, Schmitt M et al. VALGENT: a protocol for clinical validation of human papillomavirus assays. J Clin Virol 2016; 76 (Suppl 1): S14-S21.

20. Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in-vitro diagnostic medical devices. O. J. E. U. 2017 L 117/176 http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R0746&from=DE accessed 08 September 2017.

21. Backinger CL and Kingsley PA. Write It Right: Recommendations for Developing User Instructions for Medical Devices Used in Home Health Care. HHS Publication FDA 93-4258; 1993. Silver Spring, MD: Food and Drug Administration; 1993 Available at https://www.fda.gov/downloads/MedicalDevices/.../ucm070771.pdf, accessed 15 September 2017

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 978-92-4-151381-4