PROTOCOL FOR THE LABORATORY EVALUATION OF NUCLEIC ACID BASED HIV VIRAL LOAD TESTING TECHNOLOGIES
© World Health Organization 2017

Some rights reserved. This work is available under the Creative Commons Attribution Non-Commercial-Share Alike 3.0 IGO license (CC BY-NC-SA 3.0 IGO; https://creativecommons.org/licenses/by-nc-sa/3.0/igo).

Under the terms of this license, you may copy, redistribute and adapt the work for noncommercial purposes, provided the work is appropriately cited, as indicated below. In any use of this work, there should be no suggestion that WHO endorses any specific organization, products or services. The use of the WHO logo is not permitted. If you adapt the work, then you must license your work under the same or equivalent Creative Commons license. If you create a translation of this work, you should add the following disclaimer along with the suggested citation: “This translation was not created by the World Health Organization (WHO). WHO is not responsible for the content or accuracy of this translation. The original English edition shall be the binding and authentic edition”.

Any mediation relating to disputes arising under the license shall be conducted in accordance with the mediation rules of the World Intellectual Property Organization (http://www.wipo.int/amc/en/mediation/rules).

Cataloguing-in-Publication (CIP) data: CIP data are available at http://apps.who.int/iris. Sales, rights and licensing. To purchase WHO publications, see http://apps.who.int/bookorders. To submit requests for commercial use and queries on rights and licensing, see http://www.who.int/about/licensing.

Third-party materials: If you wish to reuse material from this work that is attributed to a third party, such as tables, figures or images, it is your responsibility to determine whether permission is needed for that reuse and to obtain permission from the copyright holder. The risk of claims resulting from infringement of any third-party-owned component in the work rests solely with the user.

General disclaimers: The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of WHO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers’ products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

WHO and the WHO Prequalification Evaluating Laboratory, do not warrant or represent that the evaluations conducted with the HIV test kits referred to in this document are accurate, complete and/or error-free. WHO and the WHO Prequalification Evaluating Laboratory disclaim all responsibility for any use made of the data contained herein, and shall not be liable for any damages incurred as a result of its use. This document must not be used in conjunction with commercial or promotional purposes.
TABLE OF CONTENTS

1. INTRODUCTION .................................................................................................................. 5
2. INTENDED AUDIENCE ........................................................................................................ 5
3. STUDY OBJECTIVES ........................................................................................................... 5
   3.1. Overall Objectives ......................................................................................................... 5
   3.2. Specific Objectives: ...................................................................................................... 6
4. WHO PREQUALIFICATION EVALUATING LABORATORIES ........................................... 6
   4.1. Training, performance evaluation and supervision ...................................................... 6
   4.2. Safety ........................................................................................................................... 7
   4.3. Storage of assays ......................................................................................................... 7
5. STUDY DESIGN ................................................................................................................... 7
6. SPECIMEN PANELS ............................................................................................................ 8
   6.1. Analytical Performance Specimen Reference Panel .................................................... 8
   6.2. Clinical Performance Specimen Reference Panel ...................................................... 9
   6.3. Characterization of the Specimen Panels .................................................................... 10
   6.4. Laboratory testing ....................................................................................................... 10
   6.5. Recording test results ................................................................................................. 10
7. QUALITY CONTROL AND INTERPRETATION OF TEST RESULTS .......................... 11
   7.1. Test kit controls ........................................................................................................... 11
   7.2. Internal quality control .............................................................................................. 11
   7.3. External copy control specimens ............................................................................... 11
   7.4. Proficiency panels ...................................................................................................... 11
   7.5. Limits of acceptability ............................................................................................... 11
   7.6. Interpretation of results ............................................................................................. 11
8. ANALYSIS OF DATA .......................................................................................................... 12
   8.1. Invalid runs .................................................................................................................. 12
   8.2. Invalid individual specimen results (invalid IQC/calibrator) ...................................... 12
   8.3. Performance characteristics from WHO specimen reference panel ....................... 12
      8.3.1. Analytical performance ......................................................................................... 12
         8.3.1.1. Precision of measurement .............................................................................. 12
         8.3.1.1.1. Repeatability (within-run variation) and Within-laboratory/Between-run precision ........................................ 12
         8.3.1.1.2. Inter-instrument variation .......................................................................... 12
         8.3.1.2. Linearity ........................................................................................................... 12
         8.3.1.3. Limit of detection ........................................................................................... 13
         8.3.1.4. Robustness ..................................................................................................... 13
      8.3.2. Clinical Performance ............................................................................................ 13
         8.3.2.1. Trueness of measurement .............................................................................. 13
         8.3.2.2. Rate of misclassification ............................................................................... 14
         8.3.2.3. Sensitivity and Specificity ............................................................................. 14
      Confidence intervals ....................................................................................................... 15
9. OPERATIONAL CHARACTERISTICS ............................................................................. 15
10. REPORT PREPARATION ................................................................................................... 15
11. MATERIALS AND SUPPLIES ......................................................................................... 16
12. ROLES AND RESPONSIBILITIES ................................................................................... 16
   12.1. Responsibilities of the WHO Prequalification Evaluating Laboratory ...................... 16
   12.2. Responsibilities of WHO ........................................................................................ 16
1. Introduction

1.1. Prequalification of In Vitro Diagnostics

The World Health Organization (WHO) Prequalification of In Vitro Diagnostics is coordinated through the Prequalification-Diagnostics Assessment Team (PQDx). The aim of the WHO Prequalification of In Vitro Diagnostics Assessment is to promote and facilitate access to safe, appropriate and affordable diagnostics of good quality in an equitable manner. Focus is placed on products for high burden diseases and their suitability for use in resource-limited settings.

The WHO prequalification of in vitro diagnostics process includes three main components:

- Review of a product dossier;
- Performance evaluation of the product;
- Inspection of the manufacturing site(s).

The performance evaluation will be conducted by a WHO Prequalification Evaluating site following a choice of two different mechanisms described here. Performance evaluations conducted by a laboratory in List 1 will be coordinated and cost covered by WHO. Performance evaluations conducted by a laboratory in List 2 will be coordinated and cost incurred by the manufacturer.

This protocol describes the procedures required to perform an evaluation of HIV nucleic-acid based viral load assays submitted for WHO prequalification assessment. This protocol is not intended to replace validation and verification studies that need to be conducted by the manufacturer in order to fulfil WHO prequalification product dossier requirements.

This protocol was developed in collaboration with the National Health Laboratory Services HIV PCR LAB, Charlotte Maxeke Johannesburg Academic Hospital, South Africa and the Centers for Disease Control and Prevention, Division of Global HIV & TB, International Laboratory Branch, Atlanta, USA.

2. Intended audience

This document is intended to provide WHO Prequalification Evaluating Laboratories and manufacturers with the WHO performance evaluation procedure.

3. Study Objectives

3.1. Overall Objectives

The overall objective of the performance evaluation is:
To evaluate the performance and operational characteristics of commercially available nucleic acid based viral load technologies intended for the quantification of HIV viral load in HIV-infected individuals.

3.2. Specific Objectives:
The specific objectives of the evaluation are to:

- Assess the precision, of selected technologies;
- Assess linearity, limit of detection and robustness (if applicable);
- Assess subtype detection and specificity;
- Compare the performance of the technologies and to determine agreement between the different technologies and against the reference method;
- Evaluate the operational characteristics of the technologies, such as ease of performance, specimen type utility, suitability for use in extreme conditions (high/low temperatures, high humidity), suitability for use in countries with limited infrastructure (no/limited electricity, no/limited clean water, inadequate means of waste disposal;
- Report and disseminate the findings of the evaluation in a timely manner.

4. WHO Prequalification Evaluating Laboratories

The performance evaluation will be exclusively conducted by a WHO Prequalification Evaluating Laboratory. These laboratories have successfully undergone assessment through the WHO Alternative Laboratory Evaluation Mechanism which includes:

- Submission of an Expression of Interest (EoI) by the laboratory,
- Stage 1 audit of the laboratory (assessment of EoI and specific quality management system (QMS) documentation),
- Stage 2 (on-site) audit to assess compliance with WHO requirements

The list of WHO Prequalification Evaluating Laboratories can be accessed here.

The laboratory shall hold the following certification for quality management within the laboratory: ISO17025 (General requirements for the competence of testing and calibration laboratories), ISO15189 (Medical laboratories: Particular requirements for quality and competence) or equivalent.

The person(s) listed in the EoI letter to WHO will act as the Principal Investigator (PI) for the work performed by the WHO Prequalification Evaluating Laboratory.

4.1. Training, performance evaluation and supervision

The following issues are key to minimizing error and maximizing the value of this evaluation:

- Only personnel having received specific training for this evaluation will be employed in the evaluation;
• Accurate record keeping is crucial to the success of the evaluation and the PI will be responsible for ensuring that all data required for the evaluation are recorded as agreed.
• Worksheets should be prepared and tubes, test devices or plates labelled prior to commencement of any run / assay;
• Because objective, machine-generated, permanent results for some of the technologies available may not be feasible, it is essential that the PI emphasizes the need for accurate recordkeeping;
• To minimize the risk of error, results will be directly exported from the platform wherever possible. If this is not the case, results should be entered by one staff member and verified by another.

4.2. Safety
HIV, hepatitis B and hepatitis C and other viruses are transmissible by blood and body fluids. Therefore, all types of specimens (including venous and capillary whole blood, serum/plasma, oral fluid, etc.) must be handled as potentially infectious. Appropriate precautions to minimize infectious hazards must be taken at all stages from the collection of specimens to the disposal of used materials from the laboratory. The WHO Guidelines on HIV Safety Precautions, and Guidelines for the Safe Transport of Specimens (WHO/EMC/97.3) and the site’s guidelines on laboratory safety should be carefully followed by the laboratory staff.

4.3. Storage of assays
All reagents must be stored as indicated in the “Instructions for Use” (IFU) document also known as package insert. Calibrated thermometers are placed at each location where reagents and specimens are stored, i.e. ambient, refrigerator and freezer. Temperatures are recorded daily on temperature logs. The lot numbers of the test kits received/used and their expiry dates shall be recorded.

5. Study Design
Only selected performance characteristics will be evaluated; this performance evaluation does not intend to duplicate all validation and verification studies already conducted by the manufacturer to support its claims. The study will be conducted in two separate stages, one investigating analytical performance and the other to investigate clinical performance. Laboratories may choose to conduct either one of two stages or the entire evaluation depending on their resources.

Analytical aspects that will be evaluated include the following performance characteristics:
• Precision of measurement
  – Intra-assay variation (within-run),
  – Inter-assay variation (within-days)
  – Inter-instrument variation (for point of care technologies with very low-throughput)
• Linearity
• Limit of detection
• Robustness
The evaluation of clinical performance will assess the assay’s agreement against the reference method using clinically derived specimens. In particular the following will be assessed:

- Trueness of measurement: sensitivity including bias, misclassification rate and specificity

The evaluation will also include an assessment of the assays' operational characteristics in view of their anticipated use in resource-limited settings. This assessment will include but is not limited to the following characteristics:

- Skills and training requirements
- Maintenance and calibration requirements
- Specimen requirements
- Electricity and water requirements
- Equipment required (including equipment provided and ancillary equipment that is required but not provided)
- Storage requirements for reagents
- Shelf life of reagents (upon the time of manufacture)
- Time to result and hands-on time required (including number of steps)
- Laboratory logistics, including equipment footprint.

6. Specimen Panels

6.1. Analytical Performance Specimen Reference Panel

At a minimum, four five stock specimens representing the most commonly occurring subtypes A, CRF02_AG, B, and C, and D will be used to construct the panel of specimens for the analytical stage of evaluation. Each stock specimen will be diluted in defibrinated normal human plasma that is negative for HIV, HBV and HCV. Details regarding the panel are shown in Table 1. These specimens may be commercially acquired.

Table 1 - Specimen Requirements for First Stage Evaluation

<table>
<thead>
<tr>
<th>Subtype</th>
<th>Concentration (copies/ml)</th>
<th>Minimum Number of replicates</th>
<th>Minimum Total replicates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeatability (within-run variation) and Within-laboratory/Between-run precision assessment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>$10^7$, $10^8$</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>C</td>
<td>$10^7$, $10^9$</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>Linearity assessment*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>$10^6$, $10^5$, $10^4$, $5 \times 10^3$, $10^3$, $5 \times 10^2$ and $10^2$</td>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td>C</td>
<td>$10^6$, $10^5$, $10^4$, $5 \times 10^3$, $10^3$, $5 \times 10^2$ and $10^2$</td>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td>A/G</td>
<td>$10^6$, $10^5$, $5 \times 10^3$, $10^3$, $5 \times 10^2$ and $10^2$</td>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td>A</td>
<td>$10^6$, $10^5$, $5 \times 10^3$, $10^3$, $5 \times 10^2$ and $10^2$</td>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td>Limit of Detection (LOD)*</td>
<td>10⁴, 10³, 5x10³, 10² and 10</td>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------------------</td>
<td>---</td>
<td>----</td>
</tr>
<tr>
<td>B, WHO International Reference Preparation</td>
<td>Two-fold serial dilutions with: Two concentrations above the manufacturer’s stated LOD One concentration at the stated LOD Two concentration below the LOD</td>
<td>24</td>
<td>120</td>
</tr>
<tr>
<td>Robustness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B or C Subtype</td>
<td>10⁶</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Negative specimens</td>
<td>0</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

*Concentrations may be modified to suit the limit of detection of the assay (see 8.3.1.3).

Given that some of the experiments will share the same dilution for specific specimens, whenever possible and depending on the assay under evaluation, specimens may be accommodated on the platform to maximise the throughput and avoid the need to run the same dilution of specimen twice for different purposes, e.g. 10⁴ of subtype B could be tested once in a manner that would allow the calculated value to be used for multiple purposes. All specimens for the first stage of the evaluation will be prepared as a single use aliquot with a volume corresponding to the specific requirements of each platform.

6.2. Clinical Performance Specimen Reference Panel
Clinically-derived specimens should comprise specimens collected from HIV-infected individuals currently on antiretroviral therapy (ART) and those not on ART. The panel will also include specimens from known HIV-negative individuals and individuals showing viral suppression. Table 2 represents the minimum number of specimens required.

Table 2 - Total number of specimens required

<table>
<thead>
<tr>
<th>Viral load – HIV positive specimens</th>
<th>Minimum number of specimens</th>
<th>n/0.85*</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1000 copies/ml</td>
<td>78</td>
<td>92</td>
</tr>
<tr>
<td>1000-4999 copies/ml</td>
<td>59</td>
<td>69</td>
</tr>
<tr>
<td>≥5000 copies/ml</td>
<td>182</td>
<td>214</td>
</tr>
<tr>
<td>Negative specimens</td>
<td>90</td>
<td>106</td>
</tr>
<tr>
<td>Total</td>
<td>409</td>
<td>481</td>
</tr>
</tbody>
</table>

*Adjustment calculation to account for an assay failure rate of less than 15%
6.3. Characterization of the Specimen Panels

The QC stock specimens used to prepare the various concentrations (Table 1) will be prepared from cell culture supernatants and diluted in negative human plasma and characterised the same way as the clinically derived HIV positive specimens. Clinically derived HIV positive and HIV negative specimens for the clinical performance evaluation will be characterized using the COBAS® AmpliPrep/COBAS® TaqMan® (CAP/CTM) HIV-1 Test v2.0 or the subsequent version of the product. The result obtained using this test will serve as the reference result; no additional testing will be required. Discrepant resolution plan is described in 8.3.2.3. An alternative reference method may be selected by the WHO Prequalification Evaluating Laboratory with prior agreement from WHO.

HIV negative whole blood and plasma specimens required for the preparation of contrived specimens in the analytical stage shall be tested using either nucleic acid testing (NAT) methods intended for blood donor screening or by a qualitative NAT assay for HIV, HBV and HCV. Negative specimens required for the clinical performance stage of the evaluation must be tested using at a minimum, a third generation serology assay followed by a WHO prequalified laboratory-based quantitative nucleic acid testing (NAT) method for HIV. Negative specimens included in the panel will be limited to those with a non-reactive serology result followed by “Target Not Detected” result on the NAT assay. An alternative characterization method may be selected by the WHO Prequalification Evaluating Laboratory with prior agreement from WHO.

6.4. Laboratory testing

Each product under evaluation will be used in accordance with the instructions for use (IFU) issued by the manufacturer. The evaluating site will send a copy of the IFU to WHO upon delivery of the reagents and prior to the commencement of the laboratory evaluation. The IFU must be reviewed against the IFU submitted to WHO as part of the application or pre-submission form. If the IFU has been updated since this time, it is the onus of the manufacturer to submit to WHO a letter detailing changes made prior to the start of the laboratory evaluation. Records of the version used must be kept.

The interpretation of results for each assay under evaluation is made strictly according to the manufacturer’s instructions within the IFU. Invalid runs and/or test results are recorded on the data collection sheets.

6.5. Recording test results

Wherever possible, all test results are saved and exported directly from the instrument to standardized test result worksheets in Microsoft Excel spreadsheets for further data analysis.

A technician’s appraisal is made of each assay under evaluation and is completed by the operator performing the testing. This appraisal is comprised of questions addressing ease of the procedure, reading of results, clarity of IFU, as well as records of any specific difficulties encountered during the evaluation.
technician’s appraisal table will be made available to WHO Prequalification Evaluating Laboratories as a separate document.

7. Quality control and interpretation of test results

7.1. Test kit controls
If applicable, manufacturer-supplied positive and negative test kit controls will be run as indicated in the IFU with each test run. Where positive and negative test kit controls are not supplied by the manufacturer, an external copy control specimen will act at the control specimen, (see later section 7.3).

7.2. Internal quality control
Internal procedural controls, incorporated into the design of the assay by the manufacturer, must be valid as per manufacturer’s instructions.

7.3. External copy control specimens
A well characterised HIV positive copy control and a HIV negative control should be run regularly. These specimens will be supplied by the WHO Prequalification Evaluating Laboratory or sourced from a commercial entity. Results from the copy control will be recorded and provided in the laboratory evaluation report.

7.4. Proficiency panels
A proficiency panel of routine specimens comprised of high and undetectable specimens must be run successfully for each assay by each operator before the evaluation commences. This may be the same panel as that used at the time of assay demonstration by the manufacturer or for training purposes.

7.5. Limits of acceptability
All results on test kit controls (if applicable) and the copy controls are documented. Should the copy control specimen fall between ± 2 and 3 SD, the result will be investigated. Should the copy control specimen lie outside ± 3SD, the run will be considered invalid, in which case the run will be repeated after resolution of the problem leading to this result. Such problems should be recorded. The PI will be responsible for carefully checking all data entry forms for legibility, accuracy and completeness.

7.6. Interpretation of results
The interpretation of results for each assay under evaluation is made strictly according to the manufacturers’ instructions within the IFU. Invalid test results and errors will also be recorded. Recording will be done either directly from the instrument printer (if applicable) or manually in the data collection sheet. In order to avoid transcriptional errors, a digital record will be kept of the results in the latter case.
8. Analysis of data

8.1. Invalid runs
The number of invalid test runs will be recorded as the absolute number of invalid runs and as a percentage of the total number of runs performed for the entire evaluation using all specimens. Other types of readings indicating an invalid run may be possible depending on the platform under evaluation. These will also be recorded.

8.2. Invalid individual specimen results (invalid IQC/calibrator)
The number of individual invalid specimen results is recorded. They are presented as a percentage of the total number of specimens tested per platform for the entire evaluation.

8.3. Performance characteristics from WHO specimen reference panel
The following methods are used to calculate the performance characteristics for each assay under evaluation.

8.3.1. Analytical performance

8.3.1.1. Precision of measurement
Estimation of precision will require, at a minimum the testing of the specimens described in section 6.1. The mean, standard deviation (SD) and percentage coefficient of variation (% CV) will be calculated and compared.

8.3.1.1.1. Repeatability (within-run variation) and Within-laboratory/Between-run precision
Within-run variation and within-laboratory/between-run precision assessments will be carried out as part of the same experiment. Variation will be assessed by measuring, at a minimum, five replicates of four specimens (two subtypes, two different concentrations) in the same run over five different days. A run will be defined depending on the assay's throughput: if the platform can accommodate all specimens in a single run, i.e. in the same test plate, the replicates will be run together. If the assay can only accommodate a smaller set of specimens, a run will be defined as a testing session carried out on the same instrument/module.

8.3.1.1.2. Inter-instrument variation
Where possible, inter-instrument variation will be measured using 10 replicates of four QC specimens (two subtypes, two different concentrations), on three to five instruments of the same brand.

8.3.1.2. Linearity
Linearity will be estimated using serial dilutions of a set of the stock specimens, described in Table 1 of Section 6.1. Linearity will be analysed using a dilution series of the five different QC specimens, each of different subtypes. Each dilution series will consist of seven members. Five replicates will be run of each
member of the dilution series. The dilutions will cover a clinically relevant range i.e. from 100 copies/ml to 100 000 copies/ml.

The statistical analysis for linearity will be performed by polynomial regression of the measured value against dilution. Within accepted limits of quantitation, this should be a straight line with a slope of 1.00 when using log units of viral load.

8.3.1.3. **Limit of detection**

The limit of detection (LoD) is the lowest concentration of analyte that can be consistently detected in \( \geq 95\% \) of specimens tested under routine laboratory conditions and in a given specimen matrix. It defines the analytical sensitivity (ISO/IEC, 2007). The LoD will be therefore defined as the lowest viral concentration detected with a positivity rate equal or higher than 95%.

In order to estimate the limits of detection for each assay under evaluation, at a minimum, 24 replicates of a five member dilution series concentrating on the lower end of the manufacturers’ claims for the dynamic range of the assay will be used. The five dilutions will be spread across the manufacturers’ claimed LoD and centred around the claimed LoD with two dilutions above and below (two-fold and four-fold higher and lower). The 24 replicates will be separated in a minimum of eight runs. The 3rd WHO International Standard for HIV-1 RNA preparation will be used for this purpose.

8.3.1.4. **Robustness**

If applicable, the robustness experiment will allow the determination of the well-to-well cross-contamination rate of high throughput platforms or potential carry over in low throughput instruments. The robustness of the different platforms will be assessed by running 20 positive subtype B specimens alternating with 20 negative specimens. The concentration of the positive specimens will be \( 10^6 \) copies/ml.

8.3.2. **Clinical Performance**

Specimens used for the evaluation of the clinical performance will be clinically-derived and collected by the testing laboratory. Details of the panel are described in Table 2.

8.3.2.1. **Trueness of measurement**

The trueness will be reported as bias or mean difference. The bias or mean difference is defined as the difference between the expected test results and an accepted reference value obtained from an instrument of methodology of a higher order. In this case, the reference value will be defined as the result obtained using the reference method. The bias or mean difference will reflect the average difference between the viral load results obtained with the reference method and the method under evaluation.

The level of agreement between the different specimens will be evaluated using the Bland-Altman analysis (M. Bland, 1986); i.e. through a graphical representation of the plot of the difference between the measurements using the two different methods (assay under evaluation and reference method) for each
data point against their mean. The limit of agreement is the 95% confidence interval of the difference between the methods which is bias ±1.96 SD (standard deviation).

Additionally, the level of agreement between the different platforms under evaluation will be estimated using the same method as described above.

8.3.2.2. **Rate of misclassification**

The rate of misclassification will be defined as the percentage of results that were incorrectly identified around the clinically relevant threshold of 1000 copies/ml, which relates to virological failure as defined by WHO (World Health Organization, 2016). The rate will be calculated in order to estimate the percentage of patients that would have been incorrectly classified as “failing treatment” (false positives) or incorrectly classified as “treatment successful” (false negatives) compared to results obtained with the reference method.

8.3.2.3. **Sensitivity and Specificity**

**Table 1 2x2 table for calculation of misclassification**

<table>
<thead>
<tr>
<th>Results of reference testing</th>
<th>+</th>
<th>−</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>True positives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>−</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>False negatives</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Results of assay under evaluation</th>
<th>+</th>
<th>−</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>a+b</td>
<td></td>
</tr>
<tr>
<td>True positives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>−</td>
<td>c+d</td>
<td></td>
</tr>
<tr>
<td>False negatives</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>a+c</th>
<th>b+d</th>
</tr>
</thead>
</table>

**Sensitivity**

Sensitivity will be calculated as the number of true positive results, i.e. “individuals in true virological failure” (viral load > 1000 copies/ml) compared to true positives by the reference method.

\[
Sensitivity = \frac{a}{a + c} \quad \text{(see Table 1)}
\]
Sensitivity will be expressed as a percentage.

**Specificity**
Specificity will be calculated as the number of true negative specimens, i.e. “individuals not failing treatment” (viral load < 1000 copies/ml) identified by the index method compared to true negatives by the reference method.

\[
Specificity = \frac{d}{b + d}
\]

(see Table 1)

Specificity will be expressed as a percentage.

**Confidence intervals**
The 95% confidence intervals are calculated in order to assess the level of uncertainty introduced by sample size. Exact 95% confidence intervals for binomial proportions will be calculated from the F-distribution (P. Armitage, 2002) (B. Kirkwood, 2003).

**Discrepant results**
Discrepant results are defined as results that vary by more than 0.5 \(\log_{10}\) from results obtained with the reference testing results.

Those specimens with results that are consistent with the reference testing results undergo no further testing. Specimens with test results discrepant from the reference testing will be retested by the same operator on the assay under evaluation if sufficient specimen is available.

**9. Operational Characteristics**
The operational characteristics of the assay will be assessed by the laboratory technicians performing the evaluation testing using a standard evaluation sheet (See annex 1), in order to give an appraisal of the assay under evaluation. Special attention should be paid to the IFU in order to evaluate whether these instructions are sufficient for WHO Member States end-users. Comments on the IFU must be made in the report if it does not meet an acceptable standard for any of the following criteria (clarity, presentation, content, safety instructions).

**10. Report preparation**
For evaluations conducted in List 1, WHO Prequalification Evaluating Laboratories, the preliminary data analysis and drafting of the report will be carried out by the laboratory according to a pre-defined report template and sent to WHO in a timely manner. WHO will verify the data and draft report and send to the authorized contact designated by the manufacturer for comment. Data generated and the report prepared
by a List 2 WHO Prequalification Evaluating Laboratory will be shared simultaneously with WHO and the manufacturer. Manufacturers will have one month right of reply. After one month has elapsed, the report will be accepted as final by WHO, regardless if comments are submitted. The final report will be prepared and disseminated by WHO. A copy of the final report will be sent to the authorized contact designated by the manufacturer and to the laboratory.

11. Materials and supplies
Manufacturers will provide the products and any equipment necessary for the evaluation free of charge.

12. Roles and responsibilities

12.1. Responsibilities of the WHO Prequalification Evaluating Laboratory
   i. Ensure availability and maintenance of all specimen panels;
   ii. Conducting the performance evaluation in accordance with this protocol and good laboratory practice;
   iii. Preparation of draft report of the laboratory evaluation;
   iv. Advising WHO on operational characteristics of assays evaluated.

All source data, data analysis records and all correspondence are retained and archived for a period of at least ten years.

12.2. Responsibilities of WHO
   i. Technical advice to the PI;
   ii. Technical and administrative management of the laboratory evaluation;
   iii. Verification of the draft report, seeking comments from the manufacturer;
   iv. Preparation and dissemination of the final report;
   v. Formal contacts with the manufacturers.

Any publication by WHO of the results of these evaluations and the WHO recommendations derived therefrom will, however, be accompanied by the following disclaimer:

The mention of specific companies or of certain manufacturers’ products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

WHO and the WHO Prequalification Evaluating Laboratory, do not warrant or represent that the evaluations conducted with the HIV test kits referred to in this document are accurate, complete and/or error-free. WHO and the WHO Prequalification Evaluating Laboratory disclaim all responsibility for any use made of the data contained herein, and shall not be liable for any damages incurred as a result of its use. This document must not be used in conjunction with commercial or promotional purposes.
13. Other documents and tools required

Standard Operating Procedures
SOP_PQDx_224_Overaching Procedure for Molecular Evaluations

Master Templates
PQDx_230 Report template VL technologies

Other Tools
PQDx_292_Technician’s appraisal of operational characteristics

References