TMDA/DMD/MDA/PROT/004
Revision: 00

THE UNITED REPUBLIC OF TANZANIA

MINISTRY OF HEALTH

TANZANIA MEDICINES AND MEDICAL DEVICES AUTHORITY

PROTOCOL FOR PERFORMANCE LABORATORY EVALUATION OF COMBINED HIV AND SYPHILIS SEROLOGY ASSAYS

FIRST EDITION

OCTOBER, 2025

TMDA Headquarters, Plot No. 56/1, Block E, Kisasa B Centre, Hombolo Road, P. O. Box 1253, Dodoma - Tanzania, Tel: +255 (26) 2961989/2061990/ +255(22) 2450512/2450751/2452108, Email: info@tmda.go.tz, Website: www.tmda.go.tz, Toll free: 08001100834

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

This first edition of the protocol intends to establish a well-documented procedure for conducting performance evaluation of in vitro diagnostic tests for detection of antibodies to HIV-1/2 and Treponema pallidum. This protocol serves as a guide and set a clear expectation on performance evaluation reports that are mandatory in marketing authorization applications and reduce round of request for additional data which used to prolong the evaluation time and hence delays in approval. It is my hope that the information included in this protocol will assist manufactures and/or applicants to conduct performance evaluation studies in Tanzania Mainland and subsequently expedite the verification process of the results by the Authority.

I would like to extend my heartfelt appreciation to the dedicated TMDA staff, particularly to Ms. Rehema Mariki, Dr. Goodluck Gotora, Mr. Christian Kapinga, Ms. Edina Zebedayo, Mr. Edinanth Gareba, Mr. Octavian Aron Ngoda, Mr. James Tanguye, Ms. Emmanuela Mkalawa and Ms. Adelina G. Gadiye from Muhimbili National Hospital (MNH) whose diligent efforts were instrumental in perfecting this document. Furthermore, it is with honor I appreciate the outstanding technical contribution offered by Prof. Willy Urassa, a World Health Organization (WHO) consultant, from the early stages of development of this protocol.

Special thanks are also owed to the International Medical Devices Regulators Forum (IMDRF), the World Health Organization (WHO) as well as Medical Device Coordination Group (MDCG) established by European Union (EU) for making their guidelines accessible for referencing and shaping this document.

The proper use and implementation of this protocol will ensure that the detection of antibodies to HIV-1/2 and Treponema pallidum which have been designated to be of public health importance will perform clinically as claimed by manufacturers.

Dr. Kissa W. Mwamwitwa
Director for Medical Devices and Diagnostics Control

FOREWORD

A combined HIV and syphilis serology assay, also known as a dual HIV-syphilis rapid diagnostic test (RDT), is a single blood test that simultaneously detects antibodies for both HIV and syphilis, offering a significant advantage by enabling efficient screening for both infections, particularly in high-risk populations, and facilitating early diagnosis and treatment to prevent further transmission, especially for mother-to-child transmission (MTCT) of both diseases; this approach is particularly important due to the high co-infection rates between HIV and syphilis, where one infection can increase the risk of acquiring the other.

It is through this assertion that the TMDA has established this protocol which outlines the requirements for the applicant to design and execute performance evaluation studies in Tanzania. Data on performance evaluation is needed by TMDA as part of marketing authorization of in-vitro diagnostics in Tanzania. This is one of the legal proceedings enriched in the Tanzania Medicines and Medical Devices Act, Cap 219.

Applicants and manufacturer are henceforth compelled to read the details as provided for in this protocol to streamline and expedite the approval process of diagnostics targeting Tanzania market. Together with the existing legal framework including TMDA Act and Regulations in force, applicants will also be required to refer the ISO 20916 standard when conducting performance evaluation studies. Studies done comprehensively and pursuant to this protocol will allow for generation of adequate data and subsequent authorization of in-vitro diagnostics for marketing in Tanzania.

It is envisaged that the protocol will be helpful to our stakeholders and harmonize the performance evaluation process and therefore any suggestions for improvement of the same are likewise welcomed.

Dr. Adam M. Fimbo
Director General

ABBREVIATIONS

EIA	-	Enzyme Immunoassay
HIV	-	Human Immunodeficiency Virus
IFU	-	Instructions for use
ISO	-	International Organisation for Standardization
MTCT	-	Mother-to-Child Transmission
NAT	-	Nucleic Acid Test
PCR	-	Polymerase Chain Reaction
PI	-	Principal Investigator
QC	-	Quality Control
RDT	-	Rapid Diagnostic Test
TMDA	-	Tanzania Medicines and Medical Devices Authority

DEFINITION OF TERMS

Authority
Means the Tanzania Medicines and Medical Devices Authority or the acronym "TMDA" established under section 4(1) of the Act.

Applicant
Means any person or institution or company that applies formally for performance evaluation of IVD in Tanzania.

Clinical Performance
Means ability of an IVD to yield results that are correlated with a particular clinical condition/ physiological state in accordance to target population and intended use.

Conformity Assessment
Means the systematic examination of evidence generated and procedures undertaken by the manufacturer, under requirements established by the Authority, to determine that an IVD is safe and performs as intended by the manufacturer and, therefore, conforms to the Essential Principles of Safety and Performance of Medical Devices.

Diagnostic Sensitivity
Means the ability of a device to identify the presence of a target marker associated with a particular disease or condition.

Diagnostic Specificity
Means the ability of a device to recognize the absence of a target marker associated with a particular disease or condition.

Ethics Committee (EC)
Means an independent body composed of members with expertise in both scientific and non-scientific arenas which functions to ensure the protection of human rights and the well-being of research subjects based on six basic principles of autonomy, justice, beneficence, no maleficence, confidentiality and honesty.

Error rate
Means a measure of the degree of prediction error of a model made with respect to the true model.

In Vitro Diagnostics Devices
Means a device whether used alone or in combination, intended by the manufacturer for the in vitro examination of specimens derived from the human body and animals principally to provide information for diagnostic, monitoring or compatibility purposes. IVD include reagents, calibrators, control materials, specimen receptacles, software, and related instruments or apparatus or other articles and are used for example for the following test purposes: diagnosis, aid to diagnosis, screening, monitoring, predisposition, prognosis, prediction and determination of physiological status.

Manufacturer
Means any person or a firm that is engaged in the manufacture of IVDs.

Negative likelihood ratio (LR-)
Means a probability that a person with the disease tested negative divided by the probability that a person without the disease tested negative.

Negative predictive value
Means the ratio of subjects truly diagnosed as negative to all those who had negative test results (including patients who were incorrectly diagnosed as healthy).

Performance evaluation
Means the assessment and analysis of data to establish or verify the performance (analytical performance and where applicable, clinical performance) of an IVD.

Positive predictive value
Means the ratio of patients truly diagnosed as positive to all those who had positive test results (including healthy subjects who were incorrectly diagnosed as patient).

Positive likelihood ratio (LR+)
Means probability that a positive test would be expected in a patient divided by the probability that a positive test would be expected in a patient without a disease.

1 Introduction

This protocol describes the procedures required to perform an evaluation of in vitro diagnostic tests antibodies towards HIV-1/2 and Treponema pallidum (syphilis) submitted for TMDA for market authorization. This protocol is not intended to replace validation and verification studies that need to be conducted by the manufacturer in order to fulfil TMDA product dossier requirements.

The performance evaluation determines the accuracy of all assays for the discriminatory detection of antibodies to HIV-1/2 and Treponema pallidum (hereafter referred to as HIV/TP assays) comparison with established performance criteria. These characteristics include sensitivity, specificity, negative and positive predictive values. In addition, several operational characteristics are assessed including the suitability for use in laboratories and/or testing settings with limited infrastructure. Assays for the detection of HIV and syphilis antibodies are covered in this protocol.

The TMDA HIV/TP specimen reference panel comprises approximately 800 serum/plasma specimens from Tanzania. Of 400 HIV seropositive specimens, 200 are TP seropositive and 200 are TP seronegative. Of the 400 HIV seronegative specimens, 200 are TP seropositive and 200 are TP seronegative. Evaluation will also include HIV and TP lot-to-lot variation panel.

2 Study objectives.

2.1 Overall objective

To evaluate and compare the accuracy of currently available HIV/TP RDTs for the discriminatory detection of HIV-1/2 antibodies and/or Treponema pallidum antibodies against established performance criteria.

2.2 Specific objectives

The specific objectives of the performance evaluation are:

- To determine the sensitivity and specificity of currently available HIV/TP RDTs for the detection of antibodies to HIV-1/2 and/or Treponema pallidum as compared to a reference result (For HIV; EIA and line immunoassay. For TP; treponemal EIA, and treponemal particle agglutination [TPPA],

- To evaluate the operational characteristics of HIV/TP assays, e.g. ease of performance, utility of specimen type, inter-reader variability, rate of invalid runs/devices, suitability for use in extreme climates (high/low temperatures, high humidity), and suitability for use in countries with limited infrastructure (no/limited electricity, no/limited clean water, and inadequate means of biosafety disposal).

3 Study implementation.

3.1 Performance evaluation laboratory

The laboratory shall hold one of the following certifications for quality management within the laboratory: ISO 17025 *General requirements for the competence of testing and calibration laboratories*, ISO 15189 *Medical laboratories: Requirements for quality and competence* or equivalent.

The Head of the Laboratory will act as the Principal Investigator (PI) for the work performed by the laboratory.

3.2 Training, performance evaluation and supervision

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

- The PI will be responsible for training the laboratory professionals on the details of the evaluation protocol and on the performance of each assay undergoing evaluation.
- Only those personnel who have received specific training for a particular assay evaluation will be employed.
- 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 on the agreed data collection sheets and are accurate and up to date.
- It is important to plan work in advance and follow standard operating procedures as prepared and controlled by the laboratory.

- To reduce the risk of adding an incorrect specimen to a test device/well, before starting the test run, the operator will prepare worksheets and label all tubes, dilution vessels, test devices or plates with the specimen's unique number.
- Because objective, machine-generated, permanent results for simple/rapid diagnostic tests are not feasible, it is essential that the PI emphasizes to the operator performing the tests the need for accurate recording of results and recordkeeping.
- To minimize the risk of error, it is recommended that the results are read and recorded independently by three trained staff members.
- To allow immediate correction of erroneous recording of results (rather than differences in visual interpretation), the PI or designee should assess the results as soon as possible to allow her/him to return to the original test device to investigate apparently discordant readings.
- For the performance evaluations performed at the laboratory, at least one representative result from both Hepatitis B antigen positive and negative specimens will also be recorded by taking electronic images. Unexpected test results will also be digitally recorded as well an image of the instructions for use.

3.3 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, 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 guidelines on laboratory safety should be carefully followed by the laboratory staff.

3.4 Storage of assays

All reagents must be stored as indicated in the instructions for use. Some assays may not need refrigeration. If refrigerated storage space is inadequate to store the entire test kit, they may be divided so that labile reagents can be refrigerated separately from the non-labile supplies. Calibrated thermometers are placed at each location where reagents and specimens are stored, i.e. ambient, refrigerator and freezer. Temperatures are recorded daily on the laboratory temperature logs. The lot numbers of the test kits received/used, and their expiry dates are recorded on the individual run worksheets.

Two separate production lots (with different lot numbers and different expiry dates) will be requested for evaluation, according to the following definition of a lot: *"The amount of material that is uniform in its properties and has been produced in one process or series of processes. The material can be either starting material, intermediate material or finished product.*" Each lot shall comprise different production (or manufacturing, purification, etc.) runs of critical reagents. Furthermore, lots must be sourced from a representative production run and not produced especially for the purpose of this evaluation. TMDA will verify this information before the product assessment has been finalized.

4 Specimens

4.1 Clinical performance specimen panel

The panel consists of approximately 800 serum/plasma specimens of Tanzania origin. There are 400 anti-HIV positive specimens, of which 200 are anti-TP positive and 200 are anti-TP negative. There are 400 anti-HIV negative specimens, of which 200 are anti-TP positive and 200 are anti-TP negative.

4.2 Specimen collection and storage

Collection of specimens for the HIV/TP specimen reference panel

Specimens are collected as serum or plasma (with EDTA as an anti-coagulant). Specimens will be sourced from Tanzania HIV and syphilis clinics.

Newly collected specimens are assigned a unique identification number at the collection site and then assigned with a specimen identification number upon arrival at the laboratory. The specimens are processed and aliquoted into working volumes of 250µl volumes and stored at -20 °C until testing commences. During the period of testing the specimens are stored at 2 to 8 °C and this time period does not exceed one week. After the completion of testing, they are again stored at -20 °C. Each aliquot does not undergo more than two freeze/thaw cycles. Table 1 - HIV/TP specimen reference panel

	Anti-HIV positive	**Anti-HIV negative**	**TOTAL**
Anti-Treponemal pallidum positive	200	200	400
Anti-Treponemal pallidum negative	200	200	400
TOTAL	400	400	800

4.3 Characterization of the HIV/TP specimen reference panel

The HIV/TP specimen reference panel is characterized for HIV using a standardized combination of assays i.e. a testing algorithm. These results are used to determine the HIV status of each specimen for the purpose of this evaluation see Figure 1.

Initially, each specimen is tested on the Vironostika® HIV Ag/Ab (bioMérieux) EIA and Enzgnost Anti-HIV 1/2 Plus (Siemens Healthcare Diagnostics) EIA in parallel.

Specimens that are non-reactive on both EIAs are not further tested and are assigned anti-HIV negative.

Specimens with discrepant EIA results AND with dually reactive results on both EIAs are tested on the INNO-LIA™ HIV I/II Score (Innogenetics) line immunoassay.

Specimens that are negative by line immunoassay are further tested on Innotest® HIV Antigen mAb (Innogenetics) EIA and if found non-reactive then are assigned anti-HIV

negative. If found to be neutralisable for HIV-1 antigen, the specimen is considered HIV-1 antigen positive and anti-HIV negative but are excluded from the panel.

Specimens that are indeterminate by line immunoassay are further tested on Innotest® HIV Antigen mAb (Innogenetics) EIA and if found non-reactive then are excluded from the panel. Specimens that are reactive for antigen (and neutralisable) are assigned as HIV-1 antigen positive and anti-HIV inconclusive but are excluded from the panel.

Specimens that are positive by line immunoassay are assigned as anti-HIV-1 positive or anti-HIV-2 positive. Those specimens that cannot be discriminated (i.e. anti-HIV positive) are further tested on the NEW LAV II Blot (BioRad Laboratories). Specimens that are indeterminate or negative by the NEW LAV II Blot are assigned as anti-HIV-1 positive. Specimens that are positive by the NEW LAV II Blot are assigned as anti-HIV positive.

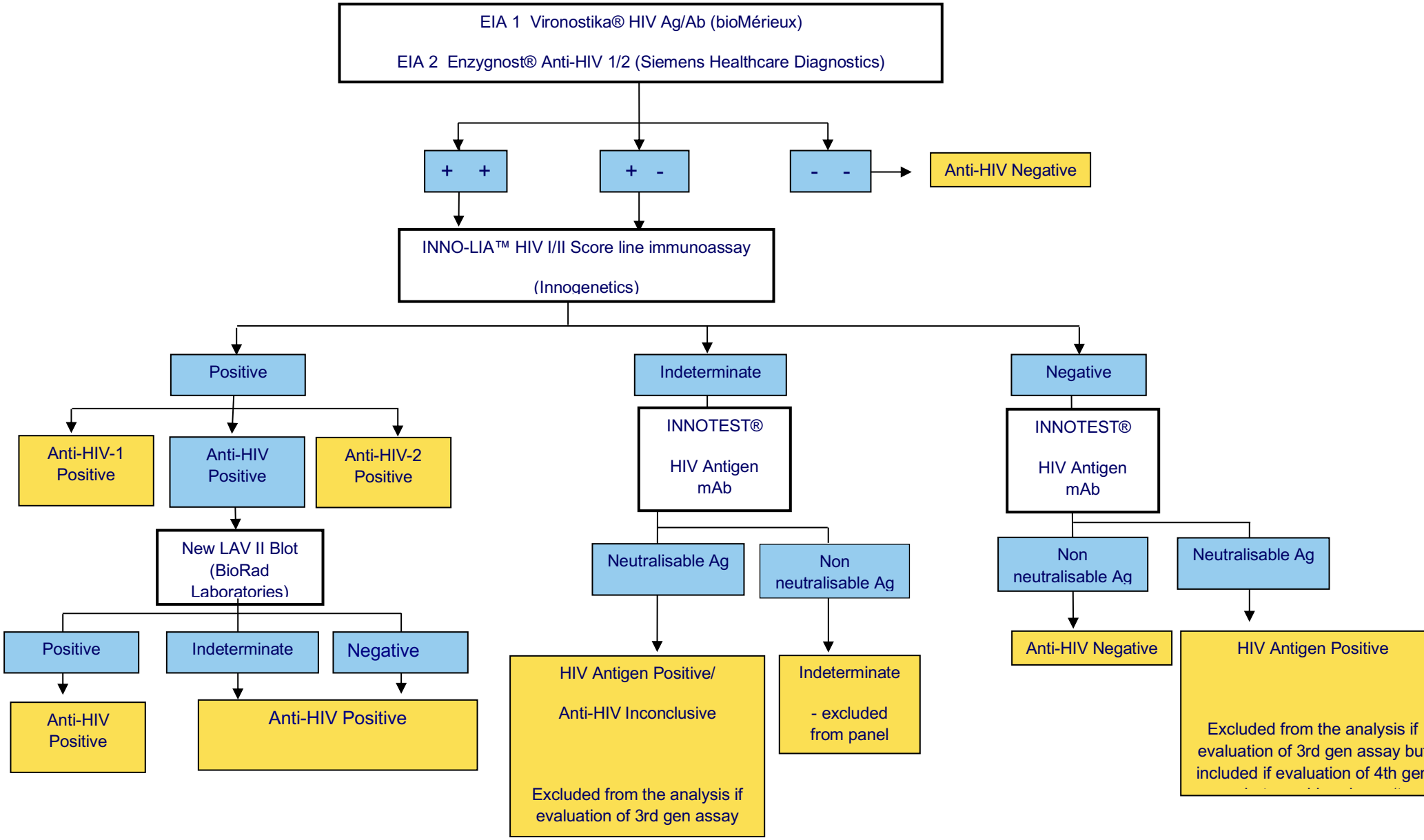

Figure 1 - Testing algorithm for HIV characterization of the HIV/plasma specimen reference panel

4.4 Characterization of the TP reference results

The HIV/TP specimen reference panel is characterized for TP using a standardized combination of assays i.e. a testing algorithm. These results are used to determine the TP status of each specimen for the purpose of this evaluation see Figure 2.

Initially, each specimen is tested on treponemal enzyme immunoassay (EIA) Vitros Syphilis TPA Assay (Ortho Clinical Diagnostics).

Specimens that are non-reactive on EIA are not further tested and are assigned anti-TP negative.

Specimens that are reactive on EIA are further tested by a Treponema pallidum particle agglutination (TPPA) (SERODIA-TP.PA, Fujirebio Inc). Specimens that are reactive on TPPA and EIA are assigned anti-TP positive.

Specimens that are reactive on EIA but non-reactive on TPPA are assigned anti-TP indeterminate but are excluded from the panel.

The use of rapid plasma regain (RPR) test which detects antibodies towards phospholipid antigens such as cardiolipin is useful to further characterize specimens with discrepant results between the assay under evaluation and the reference results. The non-treponemal assay; BD Macro-Vue™ RPR Card Tests (Becton Dickinson) will be utilized.

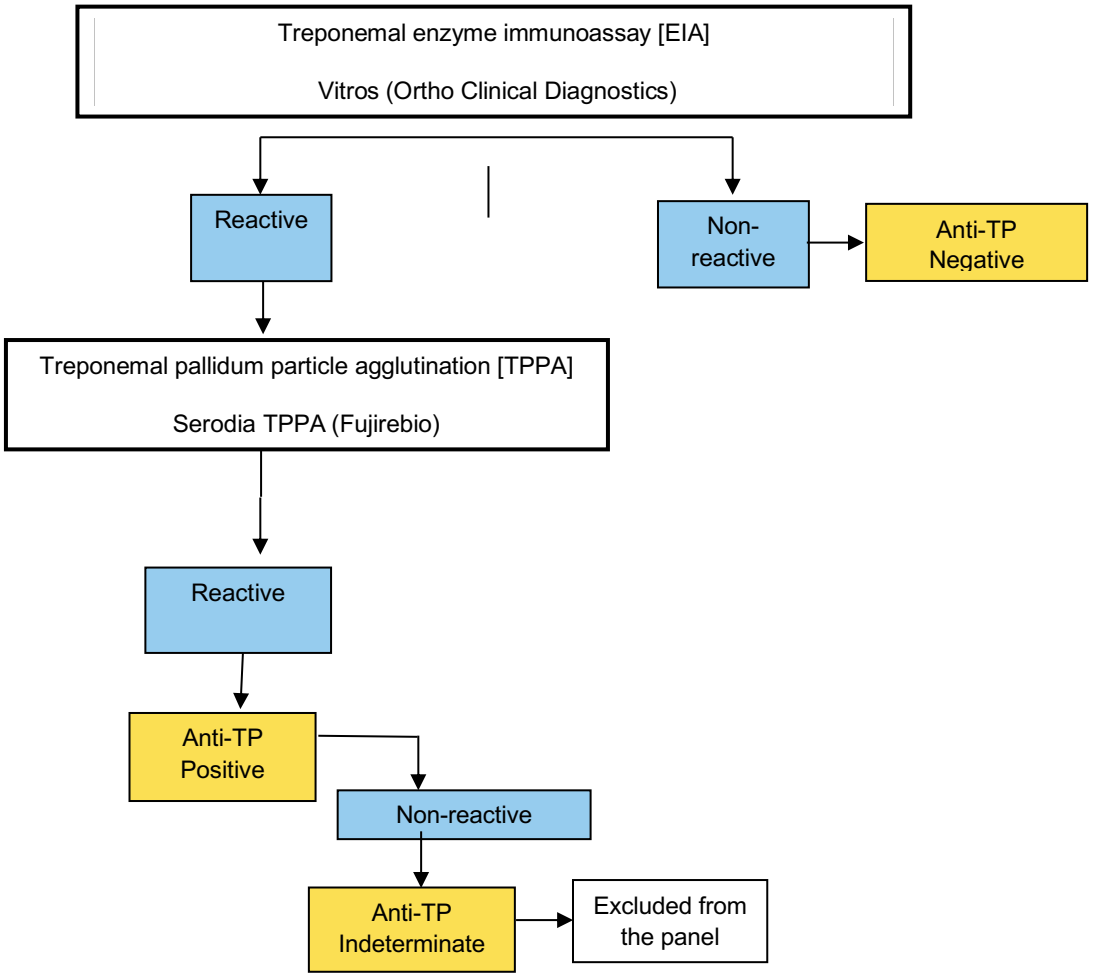

Figure 2 - Testing algorithm for TP characterization of the HIV/TP specimen reference panel

4.5 HIV lot-to-lot variation and TP lot-to-lot variation panels

Lot-to-lot variation is assessed by testing the same ten dilution series (comprised of 2-fold dilutions of 10 stock HIV positive specimens in commercially available normal human serum) on two separate production lots of the assay under evaluation in the same testing session.

A similar set of ten dilution series (comprised of 2-fold dilutions of 10 stock anti-TP positive specimens in commercially available normal human serum) on two separate production lots of the assay under evaluation in the same testing session.

4.6 External quality control specimen

See section 6 for further details.

Table 2 - Overview of specimen panels used in laboratory evaluation.

Panel name	Number of specimens
HIV/TP specimen reference panel	200 HIV positive, TP negative 200 HIV positive, TP positive 200 HIV negative, TP negative 200 HIV negative, TP positive
HIV and TP lot-to-lot variation panels	16-member dilution series of 10 HIV seropositive specimens, (160 in total) 16-member dilution series of 10 TP seropositive specimens, (160 in total)

5 Laboratory testing

With regarding to PCR a ‘test run’ for archived frozen sample is recommended for samples with only one thawing cycle.

Kit controls, if available and external quality controls will be conducted as stated in the quality manual of the specific laboratory.

5.1 Review of the instructions for use

Each assay under evaluation is used strictly in accordance with the instructions for use (IFU) issued by the manufacturer. The laboratory will send a copy of the IFU to TMDA upon delivery of the test kits and prior to the commencement of the laboratory evaluation.

The IFU must be reviewed against the IFU submitted to TMDA as part of the dossier assessment for market authorization. If the IFU has been updated since dossier submission, the technical officer-in-charge of laboratory evaluation will inform TMDA technical officer-in-

charge of dossier assessment of any ramifications for dossier assessment prior to the laboratory evaluation commencing.

Any specific procedural aspects of the IFU that should be reinforced or clarified, such as use of specimen transfer device included within the test kit, will be communicated by email to the laboratory, prior to commencement of the evaluation

5.2 Sequence of testing

The specimen reference panels are run in order that approximately one half of the specimen panel will be run with the one lot and the other half of the panel with the other lot. The specimens of the HIV/TP specimen reference panel should initially be tested in singular and in a blinded manner.

Lot-to-lot variation is assessed by testing the same set of dilution series (comprised 2-fold dilutions of 10 stock HIV and 10 stock TP positive specimens) on two separate production lots of the assay under evaluation.

For the purpose of evaluating the rapid diagnostic tests, a 'test run' is defined as a consecutive run of simple/rapid tests of the same production lot performed during the same 'session'. A 'testing session' might be considered to be a morning or afternoon.

5.3 Criteria to continue testing.

After testing the first 50 specimens (for rapid tests) an interim analysis is done, and results are communicated to TMDA. If less than 90% of the results are concordant with the reference results, then the evaluation is stopped, and troubleshooting should be done by TMDA and the manufacturer.

5.4 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 are recorded on the data collection sheets including where the control line does not appear or in any other way the test result is invalid as defined by the IFU. For test results that are indeterminate according to the IFU, the results are recorded on data collection sheets.

Testing was performed according to the test kit instructions for use (IFU) by one operator. The results will be read by that operator and one additional reader. The second reader will be blinded to the results reported by the first reader. The intensity of the test and control lines will be graded according to the reading scale in Table 2. When consensus for the sample reading is not met, a third independent reader will record their results and the eventual consensus (2 of 3 readings being the same) will be used as the final results. A grading of 0 will be considered a negative result. Any test result that was invalid according to the IFU was recorded. For RDTs invalid results were when the control line did not appear or invalid due to obviously defective test device or associated kit components. For PCR, invalid test runs will occur when the kit controls failed the manufacturer's validation criteria.

All test results are recorded on standardized test result worksheets and then entered in a Microsoft Excel spreadsheet for further data analysis.

Where applicable for PCR, test results will be recorded electronically directly from the plate reader and then entered in a Microsoft Excel spreadsheet for further data analysis. Printed records will also be generated from the PCR reader, and these will be used to cross-check a sample of the imported data.

For subjectively read assays such as rapid diagnostic tests or line immunoassay, the intensity of band/line/spot is additionally entered into the data collection sheet

5.5 Recording test results

All test results are recorded on standardized test result worksheets and then entered in a Microsoft Excel spreadsheet for further data analysis.

For subjectively read assays such as rapid diagnostic tests the intensity of band/line/spot is additionally entered into the data collection sheet. The intensity rating system reads as described in Table 1.

Table 1. Result legend for subjectively read assays.

Scoring index	**Intensity reading scale**
0	Non-reactive
1	Very Weak
2	Weak
3	Medium to Strong Reactivity
7	Invalid

6 Quality control and interpretation of test results

6.1 Test kit controls.

Manufacturer/Applicant supplied positive and negative test kit controls will be run as indicated in the IFU for all test formats included in each test run for EIAs and at the commencement of each testing session for rapid diagnostic tests and other formats. Where positive and negative test kit controls are not supplied by the manufacturer/applicant, as will be the case for many rapid diagnostic tests, the external quality control specimen will act at the control specimen.

6.2 Internal control lines for rapid diagnostic tests

Generally, rapid diagnostic tests contain a control band, line or spot to determine migration of the reagents, or the sample has occurred. Most control bands/lines/spots will become visible with the addition of reagent (i.e. buffer). However, some rapid diagnostic tests will contain a control band/line that becomes visible with addition of specimen (i.e. presence of IgG). It is imperative that the exact nature of the control band/line is ascertained and

included in the report. An experiment is performed to verify this point, if not explicitly mentioned in the IFU.

6.3 External quality control specimen

The laboratory will supply an external quality control (QC) specimen which is tested in singular at the beginning of each test session for rapid diagnostic tests.The QC specimen represents a weakly reactive Hepatitis B positive sample. The QC specimen is prepared by the laboratory or acquired commercially, depending on the assay under evaluation.

6.4 Competency panels

User competency must be established for each assay by each operator before the evaluation commences. This may be established at the time of assay demonstration by the manufacturer or for training purposes.

6.5 Limits of acceptability

All results on test kits controls and QC specimens are entered on the data collection sheets. Should the test kit controls or the QC specimen not give results within the expected ranges, evaluation testing on that assay is suspended until the cause has been identified and a satisfactory solution identified. Such problems must be communicated immediately to TMDA and recorded on the data sheets. The PI is responsible for carefully checking all data entry forms for legibility, accuracy and completeness.

7 Analysis of data

7.1 Invalid devices

The number of invalid test runs (if PCR) is recorded as the number of invalid runs as a percentage of the total number of runs performed for clinical specimens only (excluding the lot variation panels).

The number of invalid devices (if rapid diagnostic test or other format) is recorded as the number of invalid test devices as a percentage of the total number of devices used for the evaluation testing with clinical specimens (excluding, lot to lot variation panels).

Invalid results may mean test results as defined by the instructions for use such as where the control line/band/spot does not appear or invalid due to obviously defective test device or defective transfer pipette.

7.2 Inter-reader variability

The inter-reader variability is calculated when test results must be read without any objective reading instruments i.e. rapid diagnostic tests. Three persons independently interpret each test result. The inter-reader variability is expressed as the percentage of specimens for which initial test results are differently interpreted (i.e. reactive or non-reactive or indeterminate, if applicable) by the independent readers for the clinical specimens (excluding commercially acquired panels, TMDA reference preparations, culture supernatant panels, lot to lot variation panels).

8 Performance characteristics from specimen reference panel

The following methods are used to calculate the performance characteristics for each assay under evaluation and are closely linked to the reference testing results gained during specimen panel characterization.

Table 4 - 2 x 2 table for calculation of performance characteristics

		Results of reference testing		
		+	–	Total
Results of assay under evaluation	+	**a** **True positives**	b False positives	a+b
	–	c False negatives	**d** **True negatives**	c+d
	Total	a+c	b+d	

8.1 Sensitivity

Sensitivity is the ability of the assay under evaluation to detect correctly specimens that contain the analyte (reference results positive). Thus, sensitivity is the number of true positive specimens identified by the assay under evaluation as positive (a), divided by the number of specimens identified by the reference assays as positive (a+c), expressed as a percentage.

$$\textbf{Sensitivity} = \frac{a}{a+c} \; x100\%$$

8.2 Specificity

Specificity is the ability of the assay under evaluation to detect correctly specimens that do not contain the analyte (reference results negative). Thus, specificity is the number of true negative specimens identified by the assay under evaluation as negative (d), divided by the number of specimens identified by the reference assays as negative (b+d), expressed as a percentage.

$$\textbf{Specificity} = \frac{d}{b+d} \; x100\%$$

8.3 Confidence intervals

The exact 95% confidence intervals for binomial proportions are calculated for both sensitivity and specificity.

8.4 Confidence intervals

The 95% confidence intervals are calculated for both sensitivity and specificity in order to assess the level of uncertainty introduced by sample size, etc. Exact 95% confidence intervals for binomial proportions were calculated from the F-distribution. [Armitage, 2002; Kirkwood, 2003]

8.5 Positive predictive value (PPV)

The probability that when the test is reactive that the specimen does contain the analyte of interest (HIV-1/2 antibodies and/or TP antibodies). PPVs at various prevalence rates will be calculated for each analyte separately, according to the following formula.

$$PPV = \frac{(prevalence)(sensitivity)}{(prevalence)(sensitivity)+(1-prevalence)(1-specificity)}$$

Negative predictive value (NPV)

The probability that when the test is negative that a specimen does not contain the analyte of interest (HIV-1/2 antibodies and/or TP antibodies). NPVs at various prevalence rates will be calculate for each analyte separately, according to the following formula.

$$NPV = \frac{(1-prevalence)(specificity)}{(1-prevalence)(specificity)+(prevalence)(1-sensitivity)}$$

The probability that a test result will accurately determine the true infection status of a person being tested varies with the prevalence of HIV and/or syphilis infection in the population from which the person comes. In general, the higher the prevalence of HIV or syphilis infection in the population, the greater the probability that a person testing positive is truly infected (i.e., the greater the positive predictive value [PPV]). Thus, with increasing prevalence, the proportion of individuals testing false-positive decreases; conversely, the likelihood that a person whose test result is negative is truly uninfected (i.e., the negative predictive value [NPV]), decreases as prevalence increases. Therefore, as prevalence increases, so does the proportion of individuals testing false-negative.

The PPV and NPV are calculated at a prevalence of 0.1%, 1% and 5%.

8.6 Indeterminate results

For the specimen reference panel only: specimens which are found to be indeterminate (grey zone) by the criteria stated in the instructions for use should be retested in duplicate on the same lot number of assay. In the case that the testing result cannot be resolved after all testing, the specimen is to be called indeterminate and included in sensitivity/specificity calculations. A value for initial sensitivity and specificity are calculated based on the results obtained for the assay under evaluation on the first lot. The final sensitivity and specificity values are calculated taking into consideration the repeat testing performed on a same lot and further testing second lot of the assay under evaluation.

8.7 Discrepant results

Those specimens with results that are consistent with the reference results i.e. the characterized specimen results undergo no further testing. Those specimens with results discrepant from the reference result are retested in duplicate using the same lot number by the same operator. The results that occur two out of three times are recorded as the final result. If the result is again discrepant, the specimen is retested on a second lot number, if available. If the result on the second lot is concordant with the reference result, no further testing is required. If the result is still discrepant from the reference results, the result is recorded as is.

An initial sensitivity and specificity are calculated based on the initial results obtained for the assay under evaluation on the first lot. The final sensitivity and specificity are calculated taking into consideration the results of repeat testing performed on a second lot of the assay under evaluation i.e. if found to be concordant on the second lot will be recorded as such, and if found to be discrepant on the second lot will be recorded as such.

Specimens that are discrepant for HIV are further tested on Innotest® HIV Antigen mAb (Innogenetics) EIA. Specimens discrepant for TP are further tested on rapid plasma regain (RPR) test which detects antibodies towards phospholipid antigens such as cardiolipin (BD Macro-Vue™ RPR Card Tests, Becton Dickinson).

8.8 Interpretation of results from lot-to-lot variation panels

The results of both the HIV and TP lot-to-lot variation panels for the two production lots are compared.

9 Analytical performance characteristics

9.1 Results from lot-to-lot variation panel

The results of the lot-to-lot panel for the two production lots are compared and a variation of +/- 1 dilution series is considered acceptable. The number of series with acceptable and non-acceptable variation is reported.

9.2 Laboratory professional's appraisal

The technical aspects of the assay under evaluation are assessed by the Laboratory professional(s) who performed the testing. These assessments, along with other selected assay characteristics, contribute to an overall appraisal of each assay’s suitability for use in small laboratories. To enable comparison between assays, a scoring system is used to rate specified operational characteristics (Annex 1).

9.3 Report preparation and dissemination.

The preliminary data analysis and drafting of the report will be carried out by the evaluating laboratory according to pre-defined report templates.

The draft report will be shared simultaneously with TMDA and the manufacturer.

9.4 Acceptance criteria

The following criteria will be used to assess the assay under evaluation. Other parameters included in this evaluation are provided for information but are not used as pass/fail criteria for this assessment.

Table 2. Minimum acceptable performance for HIV/TP serology assays in the TMDA prequalification performance evaluation

Parameter	Rapid diagnostic tests
Initial sensitivity estimate	≥ 99%
Final specificity estimate	≥ 98%
Inter-reader variability	≤5%
Invalid rate	<5%

10 Roles and responsibilities

10.1 Responsibilities of the Evaluating Laboratory

i. Ensure availability of HIV/TP specimen reference panel, lot-to-lot variation, panel;
ii. Conducting the performance evaluation in accordance with internationally recognized best practice;
iii. Preparation of QC specimens and proficiency panels;
iv. Preparation of draft report on laboratory evaluation;
v. Advising TMDA 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.

10.2 Responsibilities of TMDA

i. Technical advice to the PI;
ii. Technical and administrative management of the laboratory evaluation;
iii. Verification of the draft report, seeking of comments from manufacturer in case of List one laboratories;
iv. Preparation and dissemination of the final report;
v. Formal contacts with authorized contacts of the manufacturers.

BIBLIOGRAPHY

Armitage P, Berry G, Matthews, JNS, Statistical Methods in Medical Research, 4th Edition. Blackwell Scientific Publications, Oxford, 2002

Kirkwood B, Stern J, Essential Medical Statistics 2nd edition, Blackwell Science Ltd. 2003

WHO. Guidance on regulations for the transport of infectious substances 2021-2022. https://iris.who.int/bitstream/handle/10665/339825/9789240019720-eng.pdf?sequence=1

WHO. Guidance for procurement of in vitro diagnostics and related laboratory items and equipment.https://iris.who.int/bitstream/handle/10665/255577/9789241512558-eng.pdf?sequence=1

EN 13612:2002 Performance evaluation of in vitro diagnostic medical devices
ISO 17025 (General requirements for the competence of testing and calibration laboratories)
ISO15189 (Medical laboratories — Particular requirements for quality and competence)

Annex 1: Operational characteristics and ease of use

This assay requires/does not require laboratory equipment and can/cannot be performed in laboratories with limited facilities or in non-laboratory settings. *If applicable, add specifics on why it cannot be used in laboratories with limited facilities: e.g.* the instrument requires a stable source of electricity and significant physical space. Furthermore, training and implementation of good laboratory practice is essential to obtaining accurate results. *If technical support was needed during evaluation:* Adequate technical support from manufacturer or representative is critical.

The assay was found easy to use / not easy to use by the operators performing the evaluation. *If applicable, add specific information provided in comments in the ease of use table.*

For RDTs (and ELISA)

Key operational characteristics	
Number of steps*	X steps in total X steps with precision pipetting *(only for serum/plasma)*
Time to result	X minutes
Endpoint stability (interval)	X minutes (the test can be read between xx and xx minutes after addition of specimen/diluent)
Internal QC	Yes/no, *insert brief description. [The test has an internal control line. The presence of the control line indicates that migration of liquid has occurred; however, it does not guarantee that the correct specimen type or volume was added or that the test procedure was followed correctly.]*

** Definition: each action required to obtain a result (excluding specimen collection, device preparation – opening the pouch), e.g. for RDTs: add specimen, add buffer (2 steps).*

For instrument-based assays

Key operational characteristics	
Number of steps for one specimen*	X steps in total X steps with precision pipetting
Number of steps for instrument management**	X steps per run/day
Time to result for one test/run	X minutes
Operator hands-on time for one test/run	X minutes
Level of automation	
Quality controls	QC are/are not provided by the manufacturer and should be purchased separately. *Add information*

	on type of QC (eg. high positive, low positive, negative)
Operating temperature	xx- xx °C, any comments on temperature stability of conducting the test.
Result display and connectivity	Results are displayed on the instrument / connected computer. They may be printed using a standard/specific printer. The results can be exported to the laboratory information system and other health information systems.
Power sources	Main power / Battery / Solar power The use of a UPS is recommended, as stable electricity is required
Biosafety *(outside of infectious specimen handling)*	Operators reported no biosafety concerns for the user. *Add information if applicable*
Waste	The volume of liquid was is approx. xx per test/run. The volume of solid waste is approx. xx per test/ run. Waste disposal requires / does not require specific measures in addition to usual laboratory biohazard waste disposal procedures. *Add information if applicable.*
Calibration	Calibrators are/are not provided by the manufacturer and should be purchased separately. *Add frequency of calibration recommended.*
Maintenance	Daily / Weekly / Monthly / Yearly / No maintenance is required.
Other specific requirements	*If applicable (eg. space requirements, weight to surface area ratio, installation by manufacturer)*