Technical Guidance Series (TGS)
for WHO Prequalification – Diagnostic Assessment

Guidance on test method validation for in vitro diagnostic medical devices
WHO Prequalification of IVDs

Preface

WHO Prequalification – Diagnostic Assessment: Technical Guidance Series

WHO Prequalification is coordinated through the Department of Essential Medicines and Health Products. WHO prequalification of in vitro diagnostic medical devices (IVDs) is intended to promote and facilitate access to safe, appropriate and affordable IVDs of good quality in an equitable manner. The focus is on IVDs for priority diseases and their suitability for use in resource-limited settings. WHO Prequalification undertakes a comprehensive assessment of individual IVDs through a standardized procedure that is aligned with international best regulatory practice. It also undertakes post-qualification activities for IVDs to ensure their ongoing compliance with prequalification requirements.

Products that are prequalified by WHO are eligible for procurement by United Nations agencies. The products are then commonly purchased for use in low- and middle-income countries.

IVDs prequalified by WHO are expected to be accurate, reliable and able to perform as intended for the lifetime of the IVD under conditions likely to be experienced by a typical user in resource-limited settings. The countries where WHO-prequalified IVDs are procured often have minimal regulatory requirements, and the use of IVDs in these countries presents specific challenges. For instance, IVDs are often used by healthcare workers who do not have extensive training in laboratory techniques, in harsh environmental conditions, in the absence of extensive pre- and post-test quality assurance capacity, and for patients with a disease profile that differs from the profiles encountered in high-income countries. Therefore, the requirements of WHO Prequalification may differ from the requirements of high-income countries, or those of the regulatory authority in the country of manufacture.

The Technical Guidance Series (TGS) was developed following a consultation held on 10–13 March 2015 in Geneva, Switzerland. The consultation was attended by experts from national regulatory authorities, national reference laboratories and WHO prequalification dossier reviewers and inspectors. The guidance series is a result of the efforts of this and other international working groups.

This guidance is intended for manufacturers interested in WHO prequalification of their IVD. It applies in principle to all IVDs that are eligible for WHO prequalification for use in WHO Member States. This guidance should be read in conjunction with relevant international and national standards and guidance.

The TGS guidance documents are freely available on the WHO website.
## Contents

Acknowledgements ................................................................. 2
List of contributors ................................................................................................................. 2

1 Abbreviations ........................................................................ 3

2 Definitions ............................................................................ 3

3 Introduction ........................................................................... 5

4 Scope ................................................................................... 5

5 Terminology for test method validation ................................. 6

5.1 Explanation of the terms characterization, verification and validation........................................................................ 6

5.2 Explanation of the terms accuracy, trueness and precision ........................................................................ 7

6 Uses of test method validation in the lifecycle of the IVD ........................................................................ 8

7 Test methods .......................................................................... 8

7.1 Categories of test methods .................................................. 8

7.2 Statistics and test methods .................................................. 9

7.3 Quantitative and qualitative assays .................................... 9

7.4 Specimen panels and test methods .................................... 9

8 Variability in the test method ................................................. 10

9 Planning for test method validation ...................................... 11

10 Examples of test methods and their validation ................... 12

10.1 Validation of test methods related to cleaning processes ........................................................................ 12

10.2 Validation of test methods for raw materials ........................................................................ 14

11 References ........................................................................... 20
Acknowledgements

The draft document *Technical Guidance Series for WHO Prequalification – Diagnostic Assessment: Guidance on test method validation for in vitro diagnostic medical devices* was developed with support from the Bill & Melinda Gates Foundation and UNITAID. This document was prepared in collaboration with Dr J Duncan, London, United Kingdom; Ms D Healy; Ms R Meurant, WHO; and with input and expertise from Dr V Alcón and Ms D Lepine; IVDD Section, Medical Devices Bureau Health Canada, Ottawa, Canada; Dr S Hojvat, Maryland, United States of America (USA); and Dr S Norman, California, USA. This document was produced under the coordination and supervision of Ms D Healy, Ms R Meurant and Ms I Prat, Prequalification Team – Diagnostic assessment, WHO, Geneva, Switzerland.

List of contributors

The draft technical specifications document was posted on the WHO website for public consultation on 19 May 2016. Various stakeholders – including manufacturers submitting to WHO prequalification of IVDs, IVD manufacturing industry associations, various national and international regulatory bodies, and IVD standards organizations – were informed of the consultation in order to solicit feedback. A 2-month response period was provided.

Comments were received from the following: Dr A Colling, Australian Animal Health Laboratory, CSIRO, Victoria, Australia; Dr J Duncan, London, United Kingdom; Ms T Milic, University of Western Australia, Western Australia, Australia; and Dr D Obradovich, Abbott Laboratories, Abbott Park, Illinois, USA.
1 Abbreviations

COA certificate of analysis
HPLC high performance liquid chromatography
IVD in vitro diagnostic medical device
R&D research and development

2 Definitions

The section below provides definitions which apply to the terms used in this document.

Accuracy: The closeness of agreement between a test result and the accepted reference value. (1)

Lot: Defined amount of material that is uniform in its properties and has been produced in one process or series of processes.

Note 1. The material can be either starting material, intermediate material or finished product. (2)

Characteristic: Distinguishing feature

Note 1 to entry. A characteristic can be inherent or assigned.

Note 2 to entry. A characteristic can be qualitative or quantitative. (3)

Note 3 Characterization: a description of the distinctive nature or features of something. (3)

Control material: Substance, material or article used to verify the performance characteristics of an in vitro diagnostic medical device. (4)

Control procedure: Activities at the point of use to monitor the performance of an IVD medical device.

Note 1. In the IVD medical device industry and in many laboratories that use IVD medical devices, these activities are commonly referred to as quality control.

Note 2. Quality control may monitor all or part of the measurement procedure, from the collection of samples to reporting the result of the measurement. (4)

In vitro diagnostic medical device (IVD): A medical device, whether used alone or in combination, intended by the manufacturer for the in vitro examination of specimens derived from the human body solely or principally to provide information for diagnostic, monitoring or compatibility purposes.

Note 1. IVDs 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, determination of physiological status.
Note 2. In some jurisdictions, certain IVDs may be covered by other regulations. (5)

**In vitro diagnostic reagent/IVD reagent:** Chemical, biological or immunological components, solutions or preparations intended by the manufacturer to be used as an IVD. (2)

**Lifecycle:** All phases in the life of a medical device, from the initial conception to final decommissioning and disposal. (6)

**Limit of detection, detection limit:** Measured quantity value, obtained by a given measurement procedure, for which the probability of falsely claiming the absence of a component in a material is $\beta$, given a probability $\alpha$ of falsely claiming its presence.

  Note 1. IUPAC recommends default values for $\alpha$ and $\beta$ equal to 0.05.
  
  Note 2. The term analytical sensitivity is sometimes used to mean detection limit, but such usage is now discouraged. Modified from (7)

**Limit of quantitation, quantitation limit:** Lowest value of measurand in a sample which can be measured with specified measurement uncertainty, under stated measurement conditions. (7)

**Measurand:** Quantity intended to be measured. (7)

  Note 1. The term "measurand" and its definition encompass all quantities, while the commonly used term "analyte" refers to a tangible entity subject to measurement (i.e. the measurand describes what is causing the result of the measurement, including the nature of the specimen, and the analyte describes the particular component of interest)

**Objective evidence:** data supporting the existence or verity of something

  Note 1. Objective evidence can be obtained through observation, measurement, test, or by other means. (2)

**Performance claim:** Specification of a performance characteristic of an IVD medical device as documented in the information supplied by the manufacturer.

  Note 1. This can be based upon prospective performance studies, available performance data or studies published in the scientific literature. (2)

  WHO Note. “Information supplied by the manufacturer” includes but is not limited to: statements in the instructions for use, in the dossier supplied to WHO and / or other regulatory authorities, in advertising, on the internet referred to simply as “claim” or “claimed” in this document.

**Precision:** The closeness of agreement between independent test results obtained under stipulated conditions. (1)

**Quality:** Degree to which a set of inherent characteristics of an object fulfils requirements. (3)

  WHO Note. For the purpose of this document these requirements include fitness-for-use, safety and performance.
Quality assurance: Part of quality management focused on providing confidence that quality requirements will be fulfilled. (3)

Ruggedness (robustness): A measure of an analytical procedure’s capacity to remain unaffected by small but deliberate variations in method parameters and provides an indication of its reliability during normal usage. (8)

Standard method: A method that is (metrologically) traceable to a recognized, validated method.

Non-standard method: A method that is not taken from authoritative and validated sources. This includes methods from scientific journals and unpublished laboratory-developed methods. (8)

Trueness: The closeness of agreement between the average value obtained from a large series of test results and an accepted reference value. (1)

Validation: Confirmation by examination and provision of objective evidence that the requirements for a specific intended use have been fulfilled. (3)

Verification: Confirmation through the provision of objective evidence that specified requirements have been fulfilled. (3)

3 Introduction

The purpose of test method validation is to ensure that a method consistently produces results fit or appropriate for a specific purpose. Testing must have a useful purpose and the result from the test must be shown to be meaningful and to give the expected (and appropriate) information. In order to ensure meaningful results, the test method must be validated; otherwise the measurement has little purpose and no value. By using validated test methods, a manufacturer can have confidence that claims made in respect to the quality and performance of an IVD are supported by objective evidence.

4 Scope

This document is intended to provide guidance on the validation of the test methods used in manufacturing of an IVD. Sometimes test methods are referred to as analytical methods but in the context of establishing the design, the development and manufacture of an IVD, “test method” is the more commonly used and a more appropriate description since not all testing is analytical. Minimal specific guidance relating to test method validation is available for IVD manufacturers despite the abundance of guidance for the analytical chemistry or pharmaceutical industries (e.g. those from Eurachem (10), Eurolab (11), ICH (12), WHO (13) and FDA (14)) or for clinical laboratories compliant with ISO 15189 (15). This document provides information on validating the test methods used by manufacturers of IVDs in their research and development (R&D), quality control and quality assurance laboratories; it must be read as an adjunct to those formal guides mentioned previously.

This document is not intended to give guidance on validation of the IVD itself. For this, it is recommended to refer to “Principles of performance studies TGS-3” (16) in this series. Qualification of instrumentation is outside the scope of this document although the test
WHO Prequalification of IVDs

methods used in qualification must be validated (17). This document does not outline statistical methods for analysis of the required data.

5 Terminology for test method validation

5.1 Explanation of the terms characterization, verification and validation

Although internationally accepted definitions exist for the terms characterization, verification and validation, the following explanation is provided to give greater clarity with relation to test method validation.

Characterization, verification and validation are essential terms. For the purposes of this guidance document “characterization of a test method” refers to an experimental procedure and the documentation of its characteristics. It is undertaken in order to provide objective evidence of what a method is capable of consistently achieving under defined conditions. The characteristics of the assay are the numerical values proven and documented for each of the method attributes such as sensitivity, specificity, limit of detection etc. Each attribute should be evaluated using an appropriate, validated test method.

Verification is the documentary proof that particular specifications have been met. When designing and developing an IVD, relevant attributes such as cost, and those for performance such as precision, sensitivity and stability are identified and given numerical specifications in design input documentation. It is subsequently the role of the R&D department to design an IVD that will meet those specifications. The R&D department consequently identifies valid test methods to demonstrate that the specifications have been met (verified) in the new design. Once design has been established, further numerical specifications are produced by the R&D department to ensure that the specifications of each attribute will be met consistently in routine production and leading to quality manufacturing. These new specifications are assigned to control critical production points and may include those for acceptance of raw materials, in-process materials, cleanliness of equipment, qualification of instrumentation and for the finalized IVD to verify its manufacture. Again, it is also the role of the R&D department to identify appropriate test methods to monitor these specifications. An example of verification is related to incoming goods inspections; each time a raw material is purchased its properties will be verified against the specification using a validated test method.

Validation is the documentary proof that the particular requirements for a specific intended use can be consistently fulfilled (9). VIM edition 3 (7) defines validation as, “verification against needs for a specific use” (i.e. the specification for that use). Within this guide, consistency is essential: it is an expectation that every lot of an IVD will behave as all other lots and will continue to meet design inputs. To ensure this, it is necessary to have validated test methods for measuring and/or monitoring specifications that will consistently produce results fit for purpose. The test methods must be validated to ensure that the results of measuring and/or monitoring are meaningful. For example, the need for accurate measurement of a raw material weighed in micrograms will not be achieved by using a weighing device with tolerance measured in grams. A test method using such an instrument would not be valid for the
intended use. Thus, for the example provided, a test method should be specified that has the necessary accuracy and precision for measuring such weights, and an instrument and procedure identified that will consistently achieve this requirement during its use. The test method is then validated to produce results fit for purpose.

Validation of a test method is distinct from its characterization. Characterization is documentation of some or all of the features of the method; validation is ensuring that the relevant characteristics are appropriate for the specific intended use. Validation of a method to be used widely, and for standard methods, often begins with complete characterization. However, for each specific intended use it is likely that only a subset of the characteristics will be relevant and must be evaluated.

Published guides to test method validation provide the broad characteristics of assays as set out in Table 1.

Table 1: Examples of characteristics of assays

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trueness</td>
<td>bias, recovery, accuracy, matrix effects</td>
</tr>
<tr>
<td>Precision</td>
<td>repeatability, intermediate precision, reproducibility</td>
</tr>
<tr>
<td>Selectivity</td>
<td>specificity, interferences</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>traceability, working range, linear range, limit of detection, limit of quantitation, uncertainty at clinically significant threshold values</td>
</tr>
<tr>
<td>Reproducibility</td>
<td>precision under defined conditions: repeatability, ruggedness, robustness</td>
</tr>
<tr>
<td>Stability</td>
<td>Ability to maintain the characteristics throughout the period of intended use</td>
</tr>
<tr>
<td>Productivity†</td>
<td>speed, hazards, cost</td>
</tr>
</tbody>
</table>

† Productivity is not usually mentioned in test method validation texts but is important in manufacturing environments. Although cost is not a factor considered in risk minimization (6), it should be a consideration in choice and validation of test methods.

Usually only a small selection of the possible characteristics and attributes will ever be studied for a test method specifically developed for a single purpose.

5.2 Explanation of the terms accuracy, trueness and precision

The terms accuracy, trueness and precision have specific meaning for technical documentation. Trueness and accuracy are the values obtained for the method under investigation and relative to a value accepted as truth, being established through the use of an accepted traceable calibrator or derived by testing using an accepted reference measurement method on the same item as measured by the method under test. Without an accepted value neither trueness nor accuracy can be given for a test method, only a percent (positive and/or negative) agreement.

Trueness is a measure of the closeness of agreement between the accepted value and the average of a large [infinite] number of results from a test or assay method under review. It is expressed as a bias: “the result from this test method has a bias of ± x units”. Trueness is a characteristic of the method.
**WHO Prequalification of IVDs**

*Precision* is a measure of the closeness of agreement between independent test results obtained under stipulated conditions using the same test method and encompasses concepts such as repeatability and reproducibility but not trueness, depending on the specified conditions. It is expressed in terms of a standard deviation or related measures “the precision (or repeatability or reproducibility etc.) of this test method under these conditions is ±y-units”. Precision is a characteristic of the method.

*Accuracy* is a measure of the closeness of agreement between the accepted value as documented and the result of a measurement. It is a characteristic of that single measurement and has components from both trueness and precision of the test method. Each time the test method is performed, the accuracy of the measurement is likely to be different because of experimental error and the imprecision of the test method. If a test method requires the documented result to be the average of several individual measurements, the accuracy is related to that average; the fact of limited replication does not convert accuracy to trueness although it may improve the accuracy. Accuracy is expressed in terms of bias: “the accuracy of that measurement was -z-units” (1, 7).

Both precision and trueness for a particular method, and subsequently the accuracy of an assay by that method, can be influenced by the concentration of the measurand. As such, when characterizing a method, knowledge of the performance of the assay should be obtained over a range of foreseeable measurand concentrations, to ensure the validity of any assumptions regarding the performance of the method.

### 6 Uses of test method validation in the lifecycle of the IVD

Testing in the R&D phase of the lifecycle of a commercial IVD is often to ensure that the work in progress will meet the input requirements or to verify that the test development meets those requirements. Design requirements such as a claim of lack of interference from similar analytes will need to be supported by evidence generated using validated test methods.

During production, testing is usually employed to ensure that the material being tested meets its specifications. Test methods will use classical analytical chemistry or biochemistry to evaluate the quality of materials coming into, or synthesized by the factory e.g. commercial chemicals, enzymes, recombinant proteins, peptides or nucleic acids. Validation of the test methods will ensure that the correct attributes are measured appropriately.

### 7 Test methods

#### 7.1 Categories of test methods

Test methods can be categorized as standard or non-standard. Standard methods are metrologically traceable to a recognized, validated method and do not require additional characterization by IVD manufacturers. Pharmacopoeia and various national regulations document approved standard methods. In contrast, non-standard methods must be individually characterized and validated for the intended use. However, all
methods must be assessed as appropriate for the specific intended use, and must be verified as being used correctly \(\text{(6, 13, 19)}\).

### 7.2 Statistics and test methods

It is recommended to seek expert statistical advice during the planning stage of all experiments to ensure that sufficient numbers of specimens are tested to provide statistically powered results. These are required to justify any claim, and to provide reasonable estimates of uncertainty.

Frequently, statistical differences will be found that have no practical consequence. For that reason practical differences, or limits of confidence, should always be defined before experiments are performed.

### 7.3 Quantitative and qualitative assays

Most test methods will produce numeric, quantitative results, but some assays can only produce a qualitative output: the binary result of analyte present or analyte absent relative to a particular cut-off value. For qualitative assays some of the characteristics listed in section 7 Table 1 cannot be enumerated without applying advanced statistical methods. For guidance on this issue see Valcárcel et al. \(\text{(20)}\).

Some IVDs are intended only to produce qualitative results in users’ environments but it is usual (and in most cases is essential) that the test methods used in manufacture (quality assurance, quality control) will provide a quantitative result. In most cases qualitative assays can be adjusted to provide a quantitative result, either from an instrumental reading, e.g. for an enzyme immunoassay, or against a graduated reference scale (semiquantitative reporting of a present/absent result as is the case with many rapid diagnostic tests or photometric reading of a rapid diagnostic test as a measure of the amount of target analyte bound in the test zone). If a quantitative result cannot be obtained, then experiments and results must be designed so that they can be analysed by appropriate qualitative statistical methods. Documenting an outcome as merely positive or negative without giving an uncertainty estimate is rarely sufficient, particularly for test methods intended to characterize an IVD (e.g. for stability, precision, sensitivity) or for release-to-sale testing. For more information, refer to “Panels for quality assurance and quality control of in vitro diagnostic medical devices TGS-6” \(\text{(21)}\).

### 7.4 Specimen panels and test methods

Test methods used to verify design and consistent production will frequently involve the choice and use of panels of specimens in order to determine and/or monitor quality characteristics of an IVD. Panels must be designed and specimens selected to ensure that data generated usefully demonstrates that the specifications have been met. Designing a valid method to assess sensitivity of antibody detection for example, will need to take into account the fact that testing of dilutions of a strong positive specimen will not produce results that reflect the performance of the assay with respect to seroconversion sensitivity. Similarly, the panel composition for release-to-sale and
stability testing must employ panels/test samples demonstrated to represent the intended performance of the IVD relative to real, critical specimens. For example release and stability testing should use samples consisting of target sequences of the same nucleic acid type as the biomarker in a matrix. Due to availability and biosafety issues, it cannot reasonably be expected that patient specimens are used, but if surrogate materials are used, they must be demonstrated to behave similar to patient specimens in all critical aspects. It is useful to note that test methods used for an IVD in both the design and development phases as well as during production can have various applications, for example the experiments undertaken at release-to-sale can be also be used with adjusted criteria in demonstrating the stability of the IVD.

8 Variability in the test method

A critical attribute of all test methods is that they must be less variable than the parameter being evaluated (i.e. have a higher precision). The variability of the test method must not conceal variability in that which is tested. This requirement is usually studied as “gauge R&R” (Gauge Repeatability and Reproducibility, refer to Burdick et al. (22)) but the process and methodology applies to any measuring system, not just to gauges. As a rule of thumb, the gauge should have a variance of less than 20% of the variance of the “test-piece”. In this context, it is unreasonable to make claims based on one lot of an IVD evaluated in one or two similar laboratories, regardless of how many individual specimens are tested – the test method here being the overall study, e.g. of specificity verification, stability, interference. It is important to understand the variability between lots of IVD and the test method (e.g. the overall stability study) must be capable of revealing it. For instance, it is an accepted published practice to use three lots of an IVD to demonstrate stability (23), however, no guidance is provided on what actions are required when significant lot-to-lot variability is identified during stability testing. Shelf-life should be assigned statistically taking into consideration the variance between lots (see “Establishing stability of in vitro diagnostic medical devices TGS-2” in this series (24)). Due to the requirement of high precision of the test method, using a previous lot of the IVD as a simple comparator is unlikely to meet the requirements of being a validated test method unless there is sufficient knowledge and control of the variability associated with each lot. The concerns regarding variability apply equally to all claims, including specificity and sensitivity.
9 Planning for test method validation

The flow of the test method validation process is shown in Figure 1. The steps will be described in detail in the examples of test method validation to follow.

Understanding the real, intended purpose of the test method is critical. The US FDA (25) states that: “Design input is the starting point for product design. The requirements which form the design input establish a basis for performing subsequent design tasks and validating the design. Therefore, development of a solid foundation of requirements is the single most important design control activity”. Both the US FDA and ISO 13485 (25, 26) require that “design and development outputs shall meet the input requirements”. Compiling requirements and comparing outputs to inputs is essential for activities that require detailed planning and execution.

Once the input requirements, which define the exact purposes for the test method are documented and agreed upon (e.g. to ensure a particular claim will be met for the whole shelf-life of the IVD), the required characteristics of the test method can be specified and given measurable attributes, usually following thorough risk assessments and development work.

The following are examples of input requirements: ability to detect an increase of 0.5% in the invalid result rate of an IVD; ability to detect 10% loss of sensitivity for a particular epitope; evidence that infectivity of a positive control material is reduced by >100-fold. It is important that the assignment of design input specifications, and accordingly the assigned acceptance criteria, are clearly defined as the first stage in the process, rather than by undertaking experimental testing using a test method that is under consideration to determine minimum performance requirements.

The method can then be developed and validated against these predetermined needs. Without the predetermined needs the method cannot be validated, merely characterized.
WHO Prequalification of IVDs

Where the analysis of incoming raw materials is supported by well-established standard methods and/or published routine test method validation, (e.g. chemical analyses, protein, peptide or nucleic acid sequencing or terminal analyses, spectroscopy, chromatography, electrophoresis, electro-blotting) the link between specifications, predetermined test method characteristics, the capability of the method and the utility of results must still be documented.

For quality assurance and quality control of an IVD, the test method frequently requires decisions on the attributes to be measured and numbers of specimens for testing panels. The finalized test methods and specimens to be used are usually developed together during the R&D for an IVD. Given the important relationship between the chosen test method and the testing panel, it is almost always too late to try to find appropriate specimens for the panels after R&D is completed.

10 Examples of test methods and their validation

This section gives examples of test methods and their validation for some aspects of IVD manufacture. The examples were generated following WHO analysis of the deficiencies in evidence provided dossiers submitted to support prequalification. The analysis concluded poor understanding of the importance of test method validation, and its necessity in support of claims (for example on lot-to-lot reproducibility, stability, specimen types). A manufacturer will need to evaluate each phase of work, processes and materials and adapt the procedure outlined in section 9 to the particular IVD. The examples in this document are neither authoritative nor complete. However, if the test method is not valid and documented, the claim is not supported.

10.1 Validation of test methods related to cleaning processes

a) Introductory discussion

This example of test method validation is of validation of the methods used in verifying cleanliness, not validation of the cleaning process itself. The specifications of what constitutes "clean" must be ascertained on a case-by-case basis. This is usually from risk analysis based on chemical knowledge of the reasons the cleaning process is necessary and some experimental evidence: why cleaning is essential, the cleaning agents used and the probable subsequent uses of the cleaned item. Once the specifications for cleanliness are known, proven and documented, the requirements of the test methods can be defined.

As a simple example consider cleaning a vessel used for preparing conjugates, last used for a conjugation of a monoclonal antibody with an enzyme and now to be used for preparing other conjugates.

b) Define the purpose of the testing

The residual conjugation chemicals, antibody and enzyme and subsequently any cleaning agents must all be removed in order not to contaminate the next solutions in the vessel.

The vessel will be cleaned with pressurized hot water containing an organic anionic detergent followed by alkali and acid rinses and finally rinsing with distilled water and drying.
c) **Identify the required characteristics**

The characteristics required for any test method used to assess the cleanliness of a vessel are, trueness, sensitivity, selectivity and precision for each of the possible contaminating analytes.

The criteria for successful cleaning could be based in a standard operating procedure requiring vigorous extraction of the cleaned vessel with a defined volume of distilled water prior to any drying stage in order to avoid artefacts from an unclean vessel appearing clean because of difficulty in detecting dried-on contaminants. The methods must be able to detect any contamination of the water that could in principle affect subsequent use of the vessel.

d) **Assign numerical values to the attributes**

Typically the specification for the water after rinsing could be: less than 10 ppm of total organic carbon, less than 5 ppm of residual protein, less than 1 ppm of residual detergent, less than 1 µM in conjugation related reagents and a conductivity of less than 0.5µS.

These are typical specifications for equipment used for fermentation or protein handling and experience shows that if they are achieved, the vessel is likely to be acceptably clean for these purposes. However the utility of the cleaning process with these specifications would need to be validated (27) for the specific use before finalizing and validating the test methods.

e) **Select or develop the method(s)**

The extraction of the water from the vessel is part of the test method: it is required to demonstrate during validation by R&D that a second, similar extraction would contain unmeasurable amounts of the potential contaminants and, independently, that nothing of practical importance would leach into the next solution to be used in the vessel. This aspect of the work requires different methods to those used for the routine verification. These more sensitive methods are required to be validated in the R&D phase of the work (validated for use in the conjugation solution).

The validation of the (non-standard) test methods used in this type of work is thoroughly exemplified in the formal pharmaceutical test method validation guides. Validation of the total organic carbon measuring system originates from the manufacturer’s specification and the subsequent performance qualification.

As the measurements are made in almost pure water it would not be necessary to verify lack of interference from other constituents of the matrix. Similarly if the conductivity meter was specified as being capable of accurate readings superior than the requirement, further validation would not be necessary.

Residual detergent would be measured using an instrument, for example high performance liquid chromatography (HPLC). The method chosen would require characterization of the sensitivity, accuracy and precision for this purpose and the specific detergent involved. Functionality and conformation of proteins would not survive the acid and alkali washes so any contaminating protein would be measured by a chemical technique, defined in the standard operating procedure for the cleaning
process. The definition of the test method is essential as each method of protein quantitation (Lowry, biuret, binding of various dyes and HPLC) provides marginally different results for specific proteins. The sensitivity of the method would be demonstrated to be capable of meeting the requirement and the precision to show that the stated level of protein could be determined with sufficient accuracy. Chemical methods may be used to analyse cleanliness relating to the conjugation reagents. The test method must be defined and the precision and sensitivity in the matrix of distilled water demonstrated to be appropriate.

f) Compare performance with requirements and use the methods if adequate

Once the methods have been characterized and proven to meet the required numerical attributes they can be used in routine verification of the cleaning process.

Over time as the process is shown to consistently produce cleanliness within the required specification, testing would be minimized: i.e. the process itself would be validated (consistently providing cleanliness fit for purpose). However this can only be done by prior use of validated test methods.

10.2 Validation of test methods for raw materials

10.2.1 Routine commercial materials

Most commercial chemicals (salts, acids, alkalis, sugars) have standard analytical methods from pharmacopoeia, (needing no further evaluations except for verification of proper use and documentary evidence of the required level of quality in the materials). The scope of testing for routine commercial chemicals requires individual assessment. This should be easy with reputable suppliers however, it is important to note that contracts with all external suppliers for such raw materials contain a clause requiring notification of any changes made to their manufacture.

10.2.1.1 Components

Components to accompany the IVD such as sachets of drying agents, specimen collection devices and tubes, transfer pipettes or dropper bottles will require testing (and documentation) against the specific requirements of the IVD.

Example: validation of an incoming test for transfer pipettes used to drop specimen into an IVD.

a) Define the purpose of the testing

The purpose of the testing is to demonstrate that across the lot of transfer pipettes, the volume delivered meets the specification provided by the R&D department during the development of the IVD.

The specifications provided by the R&D department for volume transfer by pipette would have been validated to demonstrate that when the volume delivered is within specification, the product is able to meet the claims of the assay (sensitivity, specificity, precision, etc.) through the assigned life of the IVD.
b) **Identify the required characteristics**

The required characteristics are the trueness and the precision of the lot of transfer pipettes, for each specimen type claimed.

Further specifications that need to be evaluated for such pipettes may include: orifice diameter and overall length (measures of trueness and sensitivity required), ability to deliver discrete drops easily by untrained individuals (i.e. an in-use precision measure). Each of these requires a validated specification, numerical limits and consequently a test method validated as giving the required information. The following example below is only for the volume measurements.

c) **Assign numerical values to the attributes**

The specification for volume delivered into the IVD, validated by the R&D department, might be “not less than 30 µL and not more than 45 µL of specimen to be delivered in two drops from the pipette” which in the instructions for use would be translated to “add two drops of specimen using the dropper pipette provided”. The specification of the pipettes would be “to deliver 35-40 µL ± 2 µL in two drops and evaluated across the lot”. R&D would have validated this specification for each specimen type claimed.

From that specification, the requirement of the test method would be a bias of < 1 µL in the range 30 – 45 µL and a precision of < ± 0.8 µL (variance ≈20% of that allowed in the volume specified for the pipettes).

d) **Select or develop the method**

The test method (weighing drops of water) is unlikely to introduce bias or imprecision beyond that in the specification (based on the assumption that properly maintained and calibrated weighing instrumentation is accurate) so the most important validation aspect is the relationship between drops of water and drops of each specimen type from the pipettes. The number of randomly selected pipettes and the proportion of lots to be tested are calculated on the basis of acceptable risks (28) as confidence in the supplier increased.

It is unlikely that the quality assurance incoming goods inspection team have access to the specimens claimed for the IVD (e.g. fresh whole blood, fresh serum, cerebrospinal fluid). Hence any volume measurements on a substitute liquid (e.g. water) with volume estimated by weight must be validated. Drop volumes and variances of different liquids differ due to density and surface tension effects.

The exact method and required specifications would be documented in a standard operating procedure in addition to data recording, monitoring requirements and a reference to the validation of the test method.

e) **Compare performance with requirements**

The required characteristics of the test method can be measured and compared against the specification (the precision and the number of pipettes to be tested). Consequently the method may be used if found to be fit for purpose.
WHO Prequalification of IVDs

As can be seen, method validation requires planning, experimental work and documentation beyond merely defining the method (“weigh some drops”) and the specification for the component.

10.2.1.2 Package labels, instructions for use, vials, stoppers etc.

Ancillary materials (e.g. printed matter, packing materials, containers, stoppers) will at a minimum, need inspection prior to acceptance. Inspection is a test method. As usual, the requirements of the test method (the inspection) must be defined once the purpose of the testing is understood. The attributes of the test method can be evaluated (e.g. the pre-defined consumer risk and hence the proportion of the incoming delivery to examine, the capability of the inspectors to distinguish and record the attribute) and the method validated to consistently assure appropriate quality. Regardless of inspection method chosen, its capability to detect flaws at the required level of risk must be documented, as must be the standard operating procedure for performing the inspection.

10.2.2 Constituents critical to IVD performance

Critical constituents must be decided on a risk assessment basis. Nitrocellulose membranes, some detergents and all complex biological reagents (peptides, proteins, and oligonucleotides) are assumed to be critical (unless there is evidence to the contrary).

Testing critical constituents typically involves techniques such as HPLC, spectroscopy, mass spectrometry, sequencing and various forms of electrophoresis. All instrumentation is assumed to be correctly documented, qualified and operable and the instrumentation itself will not contribute to bias or uncertainty in the example below. The latter assumptions are usually true for the biological systems described here; it is the nature of the measurements made that requires validation. Assessment of critical constituents should be more detailed than for non-critical constituents (even if the critical materials are obtained commercially). A commercial supplier cannot know the exact use of the material and can only give a general certificate of analysis (COA). Where components are deemed critical, it is necessary to have in place a supplier quality agreement which includes the requirements for a supplier to notify the IVD manufacturer of any proposed changes to a supplied component, or to the process of manufacturing of specific components, This would allow for adequate time for production planning and validation activities.

10.2.2.1 Example of acceptance testing for a low molecular weight constituent

Some detergents (those containing a polyether bond e.g. Tween, Triton) easily and quickly generate peroxides (29).

a) Define the purpose of the testing

Peroxides can disrupt enzyme activity and the conformation of recombinant proteins and some peptides. For this reason it may be considered necessary to monitor the peroxide content of detergents used in an IVD, either at the incoming goods check or often just prior to use.
b) **Identify the required characteristics**

The characteristics required are sensitivity (range and uncertainty at specific concentrations), trueness (accuracy, bias) and precision.

c) **Assign numerical values to the attributes**

R&D should have proven the stability of the IVD in studies using various lots, some of these lots at the end of their shelf-lives (18, 24). Preliminary stability experiments should lead to knowledge of the maximum permissible concentration of peroxide, (specified as, for example, < 6.0 µM) in the detergent used routinely in the manufacture of the IVD.

An example of specifications for the test method derived from the R&D requirements could be “no bias, sensitivity of 5 µM ± 1 µM at a concentration near 55 µM” (i.e. ability to distinguish between 50 and 55 µM and to allow for 10-fold dilution of a stock solution), with a peroxide specification of <55 µM in the stock solution: to give an acceptable margin of safety relative to the permissible concentration and the test method variance.

d) **Select or develop the method**

Several suitable methods for measurement of peroxide in aqueous detergent solutions are available. However, as they are not standard methods, they necessitate characterization and validation of the required sensitivity, bias and precision near the permissible concentration of peroxide in solutions of the particular detergent.

e) **Compare performance with requirements**

Clear specifications and justification of both method and expected result are required.

10.2.2.2 Acceptance testing of molecules with defined structures

For short peptides and oligonucleotides, the COA from an established and reputable supplier usually gives sufficient structural detail (e.g. proof of sequence and terminal residues (usually by mass spectrometry) and freedom from synthetic artefacts and residues (usually by HPLC)). As a result, further acceptance measurements are usually not required. However, for peptides containing cysteine (or cystine) residues it might be necessary to monitor the state of the received material to ensure lack of oxidation (or reduction), requiring a validated method for measurement of sulphydryl content.

Monoclonal immunoglobulin G class antibodies (but not polyclonal antibodies) are normally robust in production and a COA of identity and purity is generally sufficient. Polyclonal antibodies, which vary in avidity and precise epitopic dependence from animal to animal, may require similar functionality testing to that suggested in the following for recombinant proteins. This is also true for immunoglobulin M class antibodies, whether monoclonal or polyclonal.

10.2.2.3 Acceptance testing of recombinant proteins and polynucleotides

Recombinant proteins and polynucleotides require more complex testing than for molecules with a simple sequence. The functionality of recombinant proteins and polynucleotides is frequently dependent on the conformation of the macromolecules
and specificity depends on the precise impurities present (among other things). The material’s specification must include requirements for functionality, for purity based on similarity of contaminants between lots, measures of sequence integrity and molecular conformation.

The test methods involved in preparing satisfactory COAs, or of providing evidence of satisfactory in-house preparations, are much more complex than those in the elementary examples given above. However, validation of the methods follows exactly the same principles. Usually the methods are well known analytical procedures. It may be the case that they are not “standard” methods but adequately known and characterized so that if used appropriately and with documented justification, they do not require further validation.

A problem observed in many submissions to WHO prequalification is that either the methods are not used at all, or the output is not appropriate for the task. The following section discusses these major issues. It does not provide detail of the process of validation, but an expectation of how the well-known test methods will be used.

Both purity and conformation are critically lot dependent and are not usually documented in sufficient detail in a commercial COA to provide objective evidence of inter-lot reproducibility. A standard commercial COA for a recombinant protein usually provides a result for purity from a gel after electrophoresis (e.g. “>95 %”) without specifying the exact concentrations and molecular weights of the impurities (so allowing lot-to-lot variation within the specification and hence potential for specificity and stability issues). A COA will also usually provide a molecular weight, which is determined from a gel or a Western blot. However, neither technique are adequately sensitive to demonstrate minor post-translational modifications, nor capable of providing any information about conformation (allowing potential sensitivity and selectivity issues). Quantitation of results from both stained and blotted gels is not reproducible without special techniques and gives only approximate values (30). A COA should always give some measure of uncertainty in the stated values of both molecular weight and quantity. A competent COA should also include amino- and carboxy-terminal amino acid analyses to ensure absence of minor proteolysis during purification.

Choice of correct methods, and knowledge of their limitations, is a major deficiency in most WHO prequalification submissions. There is insufficient proof that there is no lot-to-lot variability, neither in those critical materials nor in the final IVD made from them.

Before committing to purchase or process a substantial amount of a new lot of recombinant protein, a careful manufacturer will need to check that the new lot will detect the most difficult specimens for a particular IVD (examples: seroconversion, latent stage, unusual serotype specimens) with the sensitivity and specificity claimed for the IVD and with approximately the same utilization (devices per milligram) as the lots used to validate the IVD itself. Proficient manufacturers develop testing of identity, integrity and functionality of polynucleotides to be used in IVDs.

These tests for complex critical reagents can only be specified for commercial or for in-house reagents by the manufacturer of the IVD, since the requirements are unique to the IVD and its validated claims. Nevertheless, the methods used must be validated as suitable for use.
The sensitivity, specificity and utilization measurements on new lots of recombinant proteins are usually made by preparing the IVD on a small scale and testing against defined panels of specimens proven to monitor the stated parameters with satisfactory efficiency.

Test method validation is to ensure that the panels do indeed monitor the expected parameters.
WHO Prequalification of IVDs

11 References


8. US FDA Volume II: Methods, Method Verification and Validation ORA-LAB.5.4.5 October 2003, Revised August 2014. [http://www.fda.gov/ScienceResearch/FieldScience/ucm171877.htm](http://www.fda.gov/ScienceResearch/FieldScience/ucm171877.htm)


11. Eurolab: Validation of Test methods. General principles and concepts (In cooperation with EAL, European cooperation for Accreditation of Laboratories); EL1545/96; 1996 [http://www.eurolab.org](http://www.eurolab.org)


WHO Prequalification of IVDs


The Technical Guidance Series for submission to WHO Prequalification – Diagnostic Assessment is developed to assist manufacturers in meeting prequalification requirements for their IVD. Further information on this guidance and other Technical Guidance series documents email diagnostics@who.int