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

Panels for quality assurance and quality control of in vitro diagnostic medical devices

TGS–6

Draft for comment 22 May 2017
The WHO Prequalification Programme 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. The WHO Prequalification Programme 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 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 be 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 health care workers without extensive training in laboratory techniques, in harsh environmental conditions, without extensive pre- and post-test quality assurance (QA) capacity, and for patients with a disease profile different to those encountered in high-income countries. Therefore, the requirements of the WHO Prequalification Programme may be different to the requirements of high-income countries, or of the regulatory authority in the country of manufacture.

The Technical Guidance Series was developed following a consultation – held on 10–13 March 2015 in Geneva, Switzerland – which 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. It should be read in conjunction with relevant international and national standards and guidance.

The TGS guidance documents are freely available on the WHO website.
Panels for quality assurance and quality control of in vitro diagnostic medical devices

TGS–5

Draft for comment 22 May 2017

Contents

Acknowledgements .................................................................................................................................................. 6

1 Abbreviations and definitions ........................................................................................................................... 7

1.1 Abbreviations.................................................................................................................................................. 7

1.2 Definitions.................................................................................................................................................... 7

2 Introduction ....................................................................................................................................................... 12

2.1 Key concept .................................................................................................................................................. 12

2.2 Rationale for the use of QA and QC panels ................................................................................................. 12

2.3 Purpose of this document ............................................................................................................................ 13

2.4 Limitations of this guidance .......................................................................................................................... 13

3 WHO prequalification requirements ................................................................................................................ 15

3.1 Manufacturer responsibility .......................................................................................................................... 15

3.2 Standards ..................................................................................................................................................... 15

3.3 Suitability for use in Member States ........................................................................................................... 16

4 Basic principles for developing panels ............................................................................................................ 17

4.1 Core principle ............................................................................................................................................... 17

4.2 Quantitation is essential ............................................................................................................................... 17

4.3 Verification and validation ............................................................................................................................. 18

4.4 Use of surrogate specimens in panels ........................................................................................................ 18

5 Use of panels ..................................................................................................................................................... 19

5.1 Release-to-sale panels ................................................................................................................................. 19

5.2 In-process panels ....................................................................................................................................... 19

5.3 Stability panels ............................................................................................................................................ 19

5.4 Reproducibility during evaluations ............................................................................................................ 20

5.5 Analytical sensitivity and range: quantitative assays .................................................................................. 21

5.6 Control materials provided to users ........................................................................................................... 21
Panels for quality assurance and quality control of in vitro diagnostic medical devices

6 Selection of specimens ........................................................................................................... 22
  6.1 General comments ........................................................................................................... 22
  6.2 In-house panels ............................................................................................................... 23
  6.3 Certified reference materials .......................................................................................... 29
  6.4 International conventional calibration materials .............................................................. 29
  6.5 Measurement variation .................................................................................................... 30
  6.6 External regulatory panels ............................................................................................... 30

7 Relationship between panel members and claims .............................................................. 33
  7.1 Process control charts ..................................................................................................... 34

8 Maintenance of panels ........................................................................................................ 35
  8.1 Storage ............................................................................................................................. 35
  8.2 Replacement ....................................................................................................................... 35

9 References ............................................................................................................................ 38

10 Annex 1 – Examples ............................................................................................................ 41
  10.1 Example 1: Correlation of a QA / QC panel member with a critical specimen ............. 41
  10.2 Example 2: Correlation under stress conditions .............................................................. 44
  10.3 Example 3: Validation of a release-to-sale panel for anti-HCV ........................................ 47
  10.4 Example 4: A release panel for a flow cytometer for the enumeration of CD4 T-cells... 49
  10.5 Example 5: Nucleic acid testing .................................................................................... 51
  10.6 Example 6: Use of imposed specimens as QA / QC panel members ............................. 53
Acknowledgements

The document **Panels for quality assurance and quality control of in vitro diagnostic medical devices** was developed as part of the Bill & Melinda Gates Foundation Umbrella Grant and the UNITAID grant for “Increased access to appropriate, quality-assured diagnostics, medical devices and medicines for prevention, initiation and treatment of HIV/AIDS, TB and malaria”. The draft was prepared in collaboration with Dr J Duncan, London, United Kingdom; Ms K Richards, WHO Geneva, Switzerland and Ms R Meurant, WHO Geneva, Switzerland and with input and expertise Dr S Hojvat, Maryland, United States of America; Dr E Cowan, Maryland, United States of America and Dr D Milic, WHO Geneva, Switzerland. This document was produced under the coordination and supervision of Kim Richards and Josée Hansen WHO/HIS/EMP, Geneva, Switzerland.

The draft guidance has been posted on the WHO website for public consultation on 22 May 2017.
1 Abbreviations and definitions

1.1 Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CE</td>
<td>Conformité Européenne (European Conformity)</td>
</tr>
<tr>
<td>CV</td>
<td>Coefficient of variation</td>
</tr>
<tr>
<td>CLSI</td>
<td>Clinical and Laboratory Standards Institute</td>
</tr>
<tr>
<td>CRM</td>
<td>Certified Reference Material</td>
</tr>
<tr>
<td>EIA</td>
<td>Enzyme-linked immunoassay</td>
</tr>
<tr>
<td>GHTF</td>
<td>Global Harmonization Task Force</td>
</tr>
<tr>
<td>HBsAg</td>
<td>Hepatitis B surface antigen</td>
</tr>
<tr>
<td>HBV</td>
<td>Hepatitis B virus</td>
</tr>
<tr>
<td>HCV</td>
<td>Hepatitis C virus</td>
</tr>
<tr>
<td>HIV</td>
<td>Human immunodeficiency virus</td>
</tr>
<tr>
<td>IFU</td>
<td>Instructions for Use</td>
</tr>
<tr>
<td>IS</td>
<td>International Standard</td>
</tr>
<tr>
<td>ISO</td>
<td>International Organization for Standardization</td>
</tr>
<tr>
<td>IU</td>
<td>International Unit</td>
</tr>
<tr>
<td>IVD</td>
<td>In vitro diagnostic or in vitro diagnostic device</td>
</tr>
<tr>
<td>NAT</td>
<td>Nucleic Acid Test, Nucleic Acid Testing</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>PEI</td>
<td>Paul Ehrlich Institute</td>
</tr>
<tr>
<td>QA</td>
<td>Quality assurance</td>
</tr>
<tr>
<td>QC</td>
<td>Quality control</td>
</tr>
<tr>
<td>QMS</td>
<td>Quality management system</td>
</tr>
<tr>
<td>RDT</td>
<td>Rapid diagnostic test</td>
</tr>
<tr>
<td>R&amp;D</td>
<td>Research and development</td>
</tr>
<tr>
<td>SI</td>
<td>International System of Units/Système International d’Unités</td>
</tr>
</tbody>
</table>

1.2 Definitions

The definitions given below apply to the terms used in this document. They may have different meanings in other contexts.
Acceptance criteria: A defined set of conditions that must be met to establish the performance of a system.

Source: (1)

Numerical limits, ranges, or other suitable measures for acceptance of the results of analytical procedures.

Source: (2)

Batch/lot: A defined amount of material that is uniform in its properties and has been produced in one process or series of processes.

Note: The material can be either starting material, intermediate material or finished product.

Source: (1), definition 3.5

Component: Part of a finished, packaged and labelled in vitro diagnostic device (IVD).

Note: Typical kit components include antibody solutions, buffer solutions, calibrators or control materials.

Source: (1), definition 3.12

Constituent: Raw materials used to make a component.

Source: (1), definition 3.57

Control material: A substance, material or article intended by its manufacturer to be used to verify the performance characteristics of a medical IVD.

Source: (3), definition 3.4 and (1), definition 3.13

Certified reference material (CRM): Reference material, accompanied by a certificate, one or more of whose property values are certified by a procedure that establishes metrological traceability to an accurate realization of the unit in which the property values are expressed, and for which each certified value is accompanied by an uncertainty at a stated level of confidence.

Source: (4), definition 3.8

Design input: The physical and performance requirements of an IVD that are used as a basis for IVD design.

Source: (5), definition f

Diagnostic sensitivity: The proportion of patients with a well-defined clinical disorder (or condition of interest) whose test values are positive or exceed a defined decision limit (i.e. a positive result and identification of the patients who have a disease).

Note 1: The clinical disorder must be defined by criteria independent of the test under consideration.

Note 2: The term "diagnostic sensitivity" (Europe) is equivalent to "clinical sensitivity" (United States)

Source: [http://htd.clsi.org]
Panels for quality assurance and quality control of in vitro diagnostic medical devices

Evidence: Information that can be proved true, based on facts obtained through observation, measurement, test or other means.

Source: Modified from (6), definition 3.8.3

Instructions for use (IFU): Information supplied by the manufacturer to enable the safe and proper use of an IVD.

Note: Includes the directions supplied by the manufacturer for the use, maintenance, troubleshooting and disposal of an IVD, as well as warnings and precautions.

Source: (1), definition 3.30

WHO comment: In the United States, the acronym IFU occasionally stands for “indications for use”, and the acronym IU stands for “intended use” or “indications for use”. The International Organization for Standardization (ISO) definition and requirements (1) for IFU cover the intended use and the precise method of use.

International conventional calibrator: A calibrator whose value of a quantity is not metrologically traceable to the International System of Units (SI), but is assigned by international agreement.

Note: The quantity is defined with respect to the intended clinical application.

Source: (4) definition 3.11

In vitro diagnostic medical device: 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; they are used, for example, for diagnosis or to aid diagnosis, screening, monitoring, predisposition, prognosis, prediction and determination of physiological status.

Note 2: In some jurisdictions, certain IVDs may be covered by other regulations.

Source: (7)

International unit: An arbitrary unit assigned to a WHO International Standard by the WHO Expert Committee of Biological Standardisation.

Source: (4), Section 4.2.6

IVD reagent: Chemical, biological or immunological components, solutions or preparations intended by the manufacturer to be used as an IVD.

Source: (1), definition 3.28

WHO comment: This document uses the terms “IVD” and “IVD” reagent interchangeably.
Panels for quality assurance and quality control of in vitro diagnostic medical devices

**Life cycle:** All phases in the life of a medical device, from the initial conception to final decommissioning and disposal.

*Source: (8), definition 2.7*

**Metrological traceability:** Property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons, all of which have stated uncertainties.

Note 1: Each comparison is affected by a (reference) measurement procedure defined in a calibration transfer protocol.

*Source: (4)*

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

Note 1: The specification can be based on prospective performance studies, available performance data or studies published in the scientific literature.

*WHO comment:* “Information supplied by the manufacturer” includes but is not limited to: statements in the IFU, 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 guide.

*Source: (1), definition 3.51*

**Quality assurance (QA):** Part of the quality management focused on providing confidence that quality requirements will be fulfilled.

*Source: (6), definition 3.3.6*

**Quality control (QC):** Part of quality management focused on fulfilling quality requirements.

*Source: (6), definition 3.3.7*

**Risk management:** The systematic application of management policies, procedures and practices to the tasks of analysing, evaluating, controlling and monitoring risk.

*Source: (8)*

**Statistical process control:** Activities focused on the use of statistical techniques to reduce variation, increase knowledge about the process and steer the process in the desired way.

*Source: (9), definition 2.1.8*

**Trueness of measurement:** Closeness of agreement between the average values obtained from a large series of results of measurements and a true value.

*Source: (4), definition 3.33*
Validation: Confirmation by examination and provision of objective evidence that the particular requirements for a specific intended use can be consistently fulfilled.

Source: (5), definition z; (6), definition 3.8.13.

Verification: Confirmation by examination and provision of objective evidence that specified requirements have been fulfilled.

Source: (5), definition aa; (6), definition 3.8.12.
Panels for quality assurance and quality control of in vitro diagnostic medical devices

TGS–5

2 Introduction

2.1 Key concept

A “panel” is a collection of well-characterized materials and specimens that is used to monitor aspects of the function of an IVD or its components. Probably the most important use of a panel is to verify, before a lot of an IVD can be released-to-sale, that the lot will consistently meet all its quality critical metrics until the end of its assigned shelf life (not just at the time of its release-to-sale, also referred to as batch release).

Panels are also used for in-process control, during reproducibility and stability studies and for some aspects of design validation. The same materials might be used for each of these purposes, but would be assigned different acceptance criteria for the different functions.

The manufacturer should be able to justify their rationale for assigning specific acceptance limits when panel samples are being tested at any point throughout the product lifecycle. The rationale can be documented as part of the risk management process; alternatively, it can be included in the design control documentation when statistical techniques are used as part of the process for establishing performance characteristics.

2.2 Rationale for the use of QA and QC panels

There is a regulatory requirement for CE-marked products for the detection of infectious diseases listed in Annex II of the In Vitro Diagnostic Medical Devices Directive 98/79/EC (IVDD) that, “The manufacturer’s release testing criteria shall ensure that every batch consistently identifies the relevant antigens, epitopes, and antibodies” (10: Section 3.4.1).

It is expected that this will be shown for all of the relevant specimen types claimed for the IVD (e.g. serum, whole blood and urine), even for rest-of-world products that fall into the high-risk categories C and D of the Global Harmonization Task Force (GHTF) classification (10). However, subject to documented risk evaluation by the manufacturer, this requirement could be relaxed to testing of only the most searching specimen type(s). It is also a regulatory requirement that the manufacturer provides objective, scientifically sound evidence to support all claims made regarding the performance of an IVD (e.g. stability, reproducibility and sensitivity). It is not reasonable to verify all performance
metrics for every lot manufactured, but use of a well-designed lot release panel will – subject to appropriate risk assessment and validation work – provide a high degree of assurance that each lot of the IVD will consistently perform as claimed for its assigned life.

In addition to testing of the IVD at release, conventional practice is to evaluate intermediates in the of manufacture (referred to as “in-process testing”). Use of carefully chosen panels will provide evidence that the intermediates meet their specifications and that the manufacturing process is in statistical control. In-process panels are also used to provide evidence that a lot is homogeneous in performance metrics from the beginning to the end of each production process.

Panels are also needed for stability and reproducibility studies used during IVD development, and for aspects of verification at the end of the product’s assigned shelf life. Results over time – as dictated by the developer’s quality management system (QMS) – must be shown to be within predetermined and validated specifications.

2.3 Purpose of this document

The purpose of this document is to provide IVD manufacturers with guidance on possible approaches to preparing validated panels for QA and QC; for example, choosing the materials, assigning meaningful criteria to them, storing them and replacing them when necessary. It describes the expectations for WHO prequalification in terms of the QA and QC information to be provided in dossiers submitted according to PQDx_018 (11: Section 6.2.1). It also provides guidance on information that might be requested during QMS inspections according to PQDx_014 (12: Section 7.2.2), following requirements in ISO 13485:2016 (13: Clauses 7.3.4 and 8.2.6).

2.4 Limitations of this guidance

This document should not be taken as a prescriptive checklist of what must be performed, but as a guide on how to improve processes and generate the evidence needed to ensure a comprehensive, systematic procedure with an appropriate risk management plan. As explained in Section 3.3, the expectations of the WHO prequalification might be more stringent than the requirements of the users and
regulatory authority in the country of manufacture. Wherever possible the guidance attempts to explain the reasons for these additional expectations. Other approaches to accommodating these further expectations can be provided in dossiers submitted for WHO prequalification, if supported by rigorous risk assessment or other evidence.

The examples included in this document apply to the principles outlined here only. Manufacturers must perform their own product-specific risk assessment for each of their IVDs. The risk evaluation must be related to the specific product, in its specific format (e.g. antigen sandwich, next-generation sequence methodology, etc.), and to the specific target analyte and the specific, claimed intended use and users.

Depending on the particular categorization of the product, additional requirements may apply in particular jurisdictions. Such regulatory and legal requirements are specific for each regulatory authority; they are beyond the scope of this document but should be documented in the design input requirements and their effects should be evaluated by risk analyses.
3 WHO prequalification requirements

WHO requires the following details for prequalification:

...any in-process and final product testing ...

... an overview of verification, validation and quality control activities for all stages of design and manufacture (including purchased components, in-process products, and finished products).

Provide the batch release criteria ...

*Source: (11: Section 6.2.1)*

The extent of the information provided in the dossier will vary. For in-process, the control points and test methods would probably be noted on process flow diagrams, with a link to the risk assessments that describe the necessity for that control. For stability or for reproducibility work, and especially for release-to-sale (i.e. batch release), a full description of the panel would be expected, including the reason for the inclusion of each panel member, its characterization, the criteria assigned and the validation of the test method. If required during any on-site QMS inspection (12), the full information about each QA and QC control point and test method should be available in the design history file.

The information provided must demonstrate the link to the *predetermined user requirements* and to product development.

3.1 Manufacturer responsibility

It is a manufacturer’s responsibility to ensure that all claims made regarding the performance of the IVD are supported by evidence that is objective and scientifically sound.

3.2 Standards

WHO recommends that manufacturers be familiar with the standards and guidance documents listed in the *Standards applicable to the WHO prequalification of in-vitro diagnostics (14)*, and take them into account when planning, assessing risk and developing QA and QC procedures.
3.3 Suitability for use in Member States

Information on the use and value of panels for QA and QC in the dossiers of a product submitted to WHO must reflect the expected environmental conditions and the normal usage conditions and methods encountered by the users in WHO Member States. The environmental conditions might differ from, and be more extreme than, those in the country of manufacture; for example, more extreme temperature and humidity, and different contaminating microorganisms. Each of these factors must be considered not only in the design input documentation (in the risk management section) but also in validating the IVD and in developing QA procedures to verify adequate performance through the life cycle of the IVD.
4 Basic principles for developing panels

4.1 Core principle

The core principle underlying any testing is that the results of the test are relevant to the investigation; that is, the test methods must be validated (15). For development of QA and QC panels, the materials and panel members chosen must be applicable to the task in hand, and their utility must be validated and documented. A panel member should be chosen for the purpose of showing stability (or sensitivity, reproducibility, etc.) of some aspect of an IVD, so that if the IVD becomes unstable (or insensitive, irreproducible, etc.) in that respect, the test result from that panel member will reflect the potential for generation of an incorrect result, with resulting incorrect patient management. Often, the data provided in WHO prequalification submissions do not adequately support the conclusions drawn because the panel members have not been properly characterized and validated for their stated use. Examples of appropriate (and inappropriate) choice of specimens are given in Annex 1 of this document, and in the guides and sample dossiers available online (16).

4.2 Quantitation is essential

Panels are used in quantitative, semiquantitative and qualitative processes. It is important to be able to show whether a parameter is different from what is expected and, if so, by how much. That change can then be related to the predetermined limits within which the IVD will function, as shown by the panel validation work (see Section 0 of this document). Finding a panel member to be merely positive or negative is often uninformative. Many IVDs are not intended to produce quantitative results; however, for certain situations (e.g. release-to-sale, stability work, robustness studies and process control) it is essential to be able to assign some form of quantitative values to the results from panel members. The methods for doing this should be determined through research and development (R&D).

For most rapid diagnostic tests (RDTs), the intensity of the colour of the signal from the IVD can be compared with a calibration scale. Degradation of signal intensity may be a sign that the IVD is degenerating. It is sometimes argued that the relative intensity of the
signal from a qualitative IVD antibody test is not related to the antibody content of the specimen. Although that is true when comparing different specimens it is not true for the relative signal generated from an IVD that has degenerated or been manufactured incorrectly (nor for dilutions of a single specimen); in such cases, the signal changes and that change can be used to monitor potential changes in the device.

Some IVDs based on a nucleic acid test (NAT) cannot be forced to give a quantitative signal, and some qualitative IVDs need to be assigned a sensitivity (with confidence limits) relative to an international conventional calibrator. In such cases, the parameter that can most easily be used to monitor stability, and verification at release, is the detection limit of each claimed analyte, which is almost always required to be stated either in the IFU or in regulatory submissions. The detection limit (or any other threshold value) is found by probit or logistic analysis of replicates of dilutions of the target analytes in each claimed specimen type, but it is beyond the scope of this document to discuss the use of these techniques in panels. For an introductory text see (17: Section 6.2.6) and (18: Appendix C). The data should be analysed statistically rather than by eye, and a statistical plan should be developed before starting the experiments.

4.3 Verification and validation

The following discussion often uses the terms “verification” (or “verify”) and “validation” (or “validate”). These terms are defined in Section 1.1; however, to reiterate, verification means providing proof of meeting predetermined specifications, whereas validation means providing proof of consistently meeting predetermined needs, one of which is always that the validated factor is robust and reliable.

4.4 Use of surrogate specimens in panels

Frequently, IVD lots must be verified as being able to detect particular specimens that are often available in only limited quantities, because of either rarity or restriction in supply from an authority. In such cases, surrogate panel members must be found that, when tested in the assay of interest, mimic the rare specimens but are more readily available (see Section 6 and Example 1 in Annex 1).
5 Use of panels

5.1 Release-to-sale panels

As noted in Section 2.2, for many years it has been accepted as best practice that the release criteria for IVD will ensure that the device consistently performs as claimed over the assigned shelf life of the lot. It is not necessary to verify every aspect of performance at release-to-sale with every specimen type claimed (provided that R&D analytical matrix studies have shown performance equivalence). Nevertheless, the documentation of claims that are not verified at release must include a stringent risk assessment and evidence of validation during the development of the device.

5.2 In-process panels

QA panels used to verify individual process stages (including testing of incoming materials) and the performance of manufactured intermediates will be varied, and will require critical evaluation and validation. Materials used in these in-process panels might be similar to those used in other panels, particularly the release-to-sale panel, but the criteria assigned will usually be different. For example, the in-process criteria assigned to a panel member used for release of a coated microplate to the next stage of manufacture might be wider than the criteria assigned to the same panel member when used at release-to-sale, because some variability in the coating is accommodated by adjustment of a conjugate.

5.3 Stability panels

Stability of each critical aspect of the IVD must be evaluated, subject to risk assessment where the necessity to evaluate is determined. Some of the performance claims that are not verified lot by lot at release-to-sale will probably be validated as part of the stability programme. Hence, the QA panels used for stability studies, although similar to those used for releasing lots to sale, are likely to be more comprehensive.

It is good practice to establish an ongoing stability monitoring programme where necessary (19: Section 4.1), to verify that lots released-to-sale maintain their claimed performance at the end of their assigned shelf lives. This verification is achieved using
Panels for quality assurance and quality control of in vitro diagnostic medical devices

IVDs retained by the manufacturer and evaluated at some predetermined time at or after their expiration date, employing a panel designed specifically for the purpose (20). The panel will probably be the same as that used for release-to-sale but it is likely to include additional material to monitor some of the quality metrics validated for stability and therefore not verified at release (e.g. in-use stability, open-pack stability and specificity). The criteria might be different from those of the same panel members in the release-to-sale panel. For example, if there is a known and allowable degeneration of signal over the life of the device the requirement at release will be higher than that at end of life so that the device will continue to meet the claim over its assigned life.

The manufacturer can extend a product’s shelf-life claim by using IVD lots kept under ongoing stability conditions and evaluated at appropriate intervals. The panel used for this, as with the panel used for validation, will probably be more complex in its composition than the release-to-sale panel.

5.4 Reproducibility during evaluations

The QA panel used for repeatability and reproducibility studies need not be complex – unless a risk assessment shows otherwise. It is generally sufficient to use dilutions of selected specimens or calibrants with concentrations near decision points (e.g. cut-off values or clinically important concentrations) in each of the claimed specimen types. A significant aspect of precision studies that is frequently absent from submissions to WHO prequalification is the relationship between the repeatability and reproducibility as measured by the developers of the assay and by intended users in their own environment. If there is an important difference between the findings of the developers and of the intended users, the assay needs further development in some respect. Best practice is to give the users involved in external performance evaluation / clinical validation a reproducibility panel that is identical to the one used by the R&D department during development work, and then to ensure that the panel is tested, along with the specimens, each time the assay is performed. The reproducibility data collected during external performance evaluation should cover the variability between the testing sites, the testers and the lots of IVD used. Comparison of these data with similar in-house data,
and the development of an experimental design that will enable the data to be collected in the first place, requires statistical advice that is beyond the scope of this guide.

**5.5 Analytical sensitivity and range: quantitative assays**

Panel members used in development work for verifying the limits of detection, quantitation and sensitivity at clinically important thresholds and range of quantitative assays are likely to be a subset of the materials eventually used in the release-to-sale panel. For most quantitative assays, certified reference materials or international conventional calibrators will be available. In the worst case, only a manufacturer’s calibrator will be available – ISO 17511 *Section 4.2.2g*) provides information about assigning values to such material. These panels are of fundamental importance in verifying the performance of an IVD.

**5.6 Control materials provided to users**

Control materials provided to users are intended to assure users that IVD performance is consistent with its intended use and the manufacturer’s claims – a function related to that of QA/QC panels. Such materials are sometimes called “run controls”, and they differ from materials provided as part of external QA programmes. These controls are not part of the manufacturer’s own QA system; nevertheless, the controls chosen and the criteria they must meet are similar to those of the manufacturer’s internal QA panels. Controls may be supplied along with the IVD as individual components or as a line on a membrane (e.g. in the case of an RDT), or they may be available to purchase separately; however, in all cases, ISO 15198 *(3)* is applicable. The claims given for run controls in IFU need careful validation, and it must be shown that these controls do in fact provide evidence that the IVD has or will function as claimed. Submissions to WHO prequalification rarely document that the control materials meet the relevant claim in the IFU; for example, for serological assays, most run controls merely show addition of a reagent, or flow, but not that the IVD would meet its claimed quality critical performance.
6 Selection of specimens

6.1 General comments

Some IVD require use of fresh specimens; for example, assays for CD4, some NIVDD.

AT IVDs and some assays requiring capillary whole blood. If stable surrogate specimens cannot be generated and validated then all the required panels must be generated and evaluated before each use. This will require use of a predicate test method, validated for both the purpose and the specific specimen type to be used. The variance of the test method must be proven not to conceal variance in the IVD being verified. This requirement is usually studied as gauge repeatability and reproducibility (R&R), but the process and methodology applies to any measuring system, not just to gauges. Statistical analysis of gauge R&R is well documented (21).

6.1.1 Inactivation

Many specimens used in QA panels are reactive for infectious organisms and must therefore be handled with appropriate caution. All control materials should undergo universal screening; for example, for human immunodeficiency virus (HIV), hepatitis B virus (HBV), hepatitis C virus (HCV), human T-lymphotrophic virus (HTLV- I/II) and syphilis. They should also be shown to be negative for transmissible (i.e. infectious) agents, unless a particular analyte is essential for demonstrating a performance characteristic of an IVD or suitable negative specimens are not available.

It is often convenient to inactivate the organisms specific for the assay so that the specimens can be handled with lower risk; for control materials provided to users (see Section 5.6), inactivation is generally essential. Thermal inactivation is commonly used for HIV-positive specimens but it is often not realized that the routine conditions for such inactivation (≥56 °C for ~30–60 minutes) do not adequately inactivate virus dried down in blood fractions, as might be found round stoppers or seals (22). Even the conditions commonly used for HBV inactivation (≥65 °C for ~16–20 hours) might not completely inactivate dried specimens containing HIV. Where thermal inactivation is used, it must be performed and documented correctly, using properly calibrated thermal measurements and taking care to ensure that no dried specimen resides on the tops of containers.
Panels for quality assurance and quality control of in vitro diagnostic medical devices

Chemical inactivation with tri(n-butyl)phosphate–detergent mixtures has been shown to efficiently inactivate enveloped viruses (23, 24), and for some purposes chemical inactivation might be more appropriate than thermal inactivation.

Whatever method of inactivation is used, it is important to verify that the treatment does not affect the analyte of interest. For antibodies there is rarely a direct problem; however, heat-treated serum or plasma are well known to cause false reactivity in some assays. For antigen and nucleic acid detection, the risk of affecting the analyte by inactivation is relatively high. Care must be taken to show that the measurement is not affected by the treatment. It must at least be shown that signal and end-point titre from the specimen are not affected, and that both newly manufactured and aged IVD product detect the treated specimen in exactly the same way as they detect the untreated specimen. If the treated specimen is to be used as a control material for more than one IVD, the proof of validity must be documented for each IVD.

6.1.2 Dilution

Most specimens used in QA panels will probably be dilutions of stronger specimens. In such cases, it is important to document that the diluent is appropriate. The diluent must not interfere in the assay in any way and must give a negative signal. In addition, if the diluent is a specimen negative for the analyte of interest, it must be shown not to contain any related antibodies, antigens, polynucleotides or inhibitors, even if it appears not to interact with the IVD under test. When the QA / QC process requires a titration curve (e.g. to monitor sensitivity relative to an international conventional calibrator), the result from the diluent must always be included in the documented measurements. A titration curve in the absence of documentation of the signal from the diluent is invalid.

6.2 In-house panels

6.2.1 Characterization

Panel members must be well characterized. Annex 1 of this document provides some examples of characterization for release and stability panels. Normally, specimens are chosen that are non-reactive for infectious agents other than the agent of specific interest, but sometimes this is not possible. If the reagent is reactive for an agent other
than that of interest, then it must be proven and documented that the other agents will not affect the primary assay. Where specimens are to be used in routine panels, the process documentation must contain warnings alerting operating staff to the presence of other infectious agents.

6.2.2 Verification or validation

It is incumbent on the manufacturer to develop appropriate QA and QC panels. Risk assessment of each key performance and functionality aspect, and experience with development of the IVD, should indicate which features must be verified at release-to-sale, during stability work, in the process of manufacture and for end-of-life security, and which aspects can be validated as routinely fit for purpose and only rarely need to be verified. Efficient development work will minimize the need for verification by providing comprehensive documentary evidence of validation for as many aspects as possible. However, according to regulation or current practice, some functionalities must always be verified at release-to-sale, at least to some extent. The correct, consistent functionality of each critical constituent (enzyme, primer, antibody, antigen, other active biological substances such as protein A and streptavidin, and agents to suppress or monitor false reactivity or immune complex disruption etc.) is normally expected to be verified at release-to-sale. However, the ability to function with each specimen type claimed should be determined unless other strong evidence is available.

6.2.3 Sensitivity

As noted previously, although verification panels for QA at release-to-sale might contain a restricted set of material, the panels used to validate key metrics must be comprehensive and provide the evidence required to allow documentation of the validity. The example that follows illustrates what might be expected for validation of epitope functionality safer. In antibody detection assays, recombinant fusion proteins are frequently used. For example, for syphilis antibody detection, the fusion protein might comprise epitopes from the three treponemal proteins (TpN15, TpN17 and TpN47); for anti-HIV-1 it might comprise epitopes of gp41 and gp120; and for anti-HCV it might comprise epitopes of NS3 and core. It must be validated that each of those epitopes will function consistently to the
end of the assigned life of the IVD. The development panel will therefore prove that for all lots of the recombinant, in all lots of the complete device, each epitope functions as expected and remains consistently functional throughout stability studies. Testing one lot of recombinant, even in several lots of IVD, is not sufficient to validate the system, because it is known that different purification runs can subtly modify the epitopic efficiency, particularly that of fusion proteins and the various epitopes within them. The development panel used in these studies will therefore specifically and independently monitor the epitopes involved, and show that the activities remain consistent and stable. The R&D department faces the difficult challenge of selecting specimens to provide such proof. Usually, specimens can be found that react strongly by Western blot (or another method) with just one of the epitopes required, and that by dilution can be made essentially specific for that epitope, at least so far as a development panel is concerned (see also Section 6.2.5). Once the panel members have been identified (usually throughout the R&D phase of an IVD life cycle) they can be used each time the recombinant is prepared and used. The results, collected and analysed over time, can either support validation of only one of those specimens in the release panel, or show the need for all of them as part of the in-process panels.

6.2.4 Prozones

Prozones, also known as “hook effects” occur when dilution series of specimens with a high concentration of analyte give a maximum signal stronger than the signal from the undiluted specimen. The following relates to assays for which the immunological response profile for a target analyte is not clearly established or in which the analyte is heterogeneous; for example, assays for antibodies or for antigens occurring in multiple forms, such as hepatitis B surface antigen (HBsAg) and histamine rich protein 2 (HRP2) for malaria. Assays for well-characterized, homogeneous analytes (e.g. small metabolites and some hormones) can usually be verified not to prozone with a single artificial specimen containing well above the highest concentration of the analyte found in a clinical setting.

The possibility of a prozone depends largely on the format of the assay. If the format chosen is known to prozone, it is imperative to show that the balance of reagents will make prozones rare and prozoning to negative even more rare. Prozones are commonly
associated with particle agglutination assays, but can also occur with immunometric
assays and with lateral flow devices. Two-step assays — that is, add specimen to capture
system, separate (wash), add label, wash — are not susceptible to prozones. However, the
normal form of lateral flow device — that is, add specimen to conjugate pad, cause
specimen and conjugate together to flow over capture line — will always have the
potential to prozone with sufficiently strong specimens. Prozones are very sensitive to the
exact concentrations of reagent in the IVD; hence, studies of lot-to-lot variation are
necessary.

Selection of specimens for the panel to validate consistent lot-to-lot absence of prozone
is not difficult, but there are two issues that need to be taken into account.

The first issue, as noted, is that of the number of specimens to be chosen. Because
prozones are relatively rare, a reasonable number of specimens must be evaluated. The
number would need to be obtained from knowledge of the concentration range of
analyte likely to be found and a stringent risk analysis — at least 10 specimens would be a
suitable minimum.

The second issue is the selection of the specimens. It is necessary to choose specimens
that have a strong signal and a very high titre in an assay format that cannot prozone.

Once the specimens and their numbers have been selected it is a simple matter to test all
of them on several lots of the device, to show that there is no increase in signal after
dilution of the specimens in a diluent validated not to interfere with the test. Any increase
whatever in signal strength for any of the specimens implies a potential prozone and
needs further analysis.

6.2.5 Rare specimens

Sometimes it is necessary to validate detection of critical specimens (e.g. seroconversion,
rare subtypes or antigen or nucleic acid in immune complexes); however, the genuine
specimens might be too valuable to use in lot release-to-sale panels or in stability panels.
In such cases, the reactivity or the classes of antibody and epitopic specificity are unlikely
to be found in commonly available specimens. Instead, it is necessary to validate
surrogate specimens. For immune complex and seroconversion work, near-
seroconversion specimens can sometimes be shown to mimic the critical individual seroconversion specimens claimed. Such specimens might be available in existing seroconversion series where the later members are not of particular interest in themselves. Immune complexes at later stages of an infection can be important – for example, in HBsAg testing (25) – in that case, specimens can be prepared by dilution of analyte into individual specimens to cover an appropriate range of antibodies (at least ad and ay related, in the HBsAg testing example). For rare subtypes, it might be necessary to prepare surrogates synthetically either biochemically or by evaluating numbers of available specimens until one that mimics the target is found. Further information is given in Example 1 of Annex 1.

For instance, a specimen with a rare marker might be tested once with a device and the result recorded. Such a single evaluation does not provide evidence of consistency and it gives no indication of the extent of any lot-to-lot variability that might be present. Therefore, a properly constituted panel is required that can be used on a sufficient number of occasions to prove – using lots that are as varied as possible, and lots that are at the end of their assigned lives – that any variability is within a predetermined acceptable range.

6.2.6 Specificity

Specificity of IVD is largely controlled by the additives in wash solutions and diluents (e.g. detergents, chaotropic agents, cell constituents and masking proteins); monitoring the functioning of such additives is an important role of QA / QC panels. Specificity must be monitored during stability work both for components and for complete devices.

It is usual to collect false reactive specimens and interfering specimen types (26) during the R&D phase of device design, and then to monitor and control lot-to-lot variation by using these specimens in release-to-sale panels. When an unusual type of false reactive or interfering substance is identified post-market, either submitted as a complaint or identified in published literature, the specimen (if available in sufficient volume) should be added to release-to-sale panels, and monitored during the statistical process control to ensure that lots do not show excessive reactivity with that specimen.
6.2.7 Specimen types

The claims to usability of different specimen types are usually validated (and documented) in the R&D phase of the IVD life cycle, and verification at lot release is restricted to one or two specific specimen types. This is a particular issue for whole blood specimens with some flow IVD where there is variability between lots and between devices in clearance of the red cell debris from the results window. In general, the whole blood to be tested must be fresh unless otherwise validated, and a sufficient number of individual specimens must be used to verify that the device will operate as expected.

6.2.8 Associated materials

Some materials used in an IVD are not directly regarded as functionality related; for example, antimicrobials, stabilizers for enzymes, agents to quench false reactants and desiccants. Each of these materials must be validated in some way, and a panel of materials, or a chemical method, prepared with which to document acceptable performance. The antimicrobials used in an IVD submitted to WHO prequalification need special consideration, both in terms of the material used (which must be capable of protecting the device against a range of aggressive microorganisms) and stability under harsh conditions. The panels of microorganisms used to validate antimicrobial efficacy can be derived from various pharmacopoeia – for example (27) – and should be used to show efficacy at the end of the assigned shelf life of the IVD.

6.2.9 Acceptance criteria

Acceptance criteria are specific indicators or measures employed in assessing the ability of a component, structure or system to perform its intended function. The manufacturer should be able to justify the rationale for assigning specific acceptance limits when panel samples are being tested at any point throughout the product lifecycle. The rationale can be documented either as part of the risk management process, or as part of the design control documentation when statistical techniques are used as part of the process for establishing performance characteristics.
6.2.10 Conclusion

Selection of specimens for preparation of appropriate panels to verify or support validation of claims for all manufactured lots requires an in-depth knowledge of the particular analyte, the mechanism of the particular assay and the clinical intent of the IVD. Each claim (which will originate from a requirement in the input documentation) and each statement in any documentation of an IVD needs careful consideration as to whether the claim or statement is to be validated or verified for lot release. This consideration needs start early in R&D so that appropriate specimens and procedures can be identified and proper validation performed.

6.3 Certified reference materials

Certified reference materials (CRM) are available for many IVDs intended to measure analytes that are homogeneous at a molecular level. CRM can frequently be traced to calibrants with an accurate concentration and uncertainty in SI units. Whenever they are available, materials traceable to SI units must be used to validate and verify quantitative IVD. For this type of analyte the production of panels is usually a matter of diluting the pure material into each specimen type required and proving commutability (4). CRM not traceable to SI units used in biological assays are usually international conventional calibrators, as discussed in the sections below.

6.4 International conventional calibration materials

Many standards from WHO and the Paul Ehrlich Institute (PEI) fall within the category of international calibration materials. Examples include the 3rd International Standard for HBsAg (28), and the 1st International Reference Panel for HBV genotypes for HBsAg-based assays (29). Such standards are used for calibration of quantitative assays and for evaluation of analytical sensitivity of qualitative assays. These materials are in too limited supply to use routinely in panels, so it is usual for to prepare manufacturer`s working standards from more readily available specimens, calibrated against the reference materials by methods described in ISO 17511 (4).

The 2nd HBsAg WHO international standard is serotype adw2, genotype A, while the 3rd international standard is a mixture of serotypes ayw1 and adw2. Relative reactivity of
these standards in HBsAg assays from different manufacturers is different, and so when verifying claims related to this type of standard, it is important to define the version used. The variation can be a result of molecular heterogeneity, which is likely to be detected differently by different antibodies. The same applies for IVDs intended to detect antibodies – if reference materials are changed it is likely that the apparent sensitivity of devices from different manufacturers will also change.

6.5 Measurement variation

It is expected that any quantitative result will be accompanied by an uncertainty statement (4); therefore, any sensitivity claimed in IFU or assigned to panel members should always be associated with the uncertainty. The uncertainty should take into consideration variation both within and between lots.

6.6 External regulatory panels

Further details on external regulatory panels are given in Example 6 of Annex 1. Some national agencies and some regulatory authorities provide panels of specimens that the IVD must detect correctly before it can be released-to-sale. These panels cannot control all aspects and all claims about the IVD from all the manufacturers that they regulate because of the variety of reagents in those IVDs and the different specimen types involved. A properly designed regulatory panel might be able to verify some key metrics, but not all that a manufacturer has claimed. For example, some HIV-1 antibody detection systems include only gp41, whereas others include a variety of epitopes from gp41, gp120, gp160 and p24. It is unlikely that a panel provided by a regulator could monitor each of these epitopes in a critical way so as to ensure that the claims related to each could be verified.

In the example of the previous paragraph, depending on how the regulator designed and measured the balance of antibodies in a changed panel member, as assay that did not detect anti-p24 could, after a change to a panel member that resulted in reduced anti-gp41 but augmented anti-p24, suddenly appear insensitive although it had not changes at all and would continue to perform as claimed. An IVD that detected anti-p24 might appear more sensitive or remain that same on the changed panel member.
A different problem might be encountered if correct detection of a regulatory panel were used as the manufacturer’s sole criterion for release-to-sale. Some large national regulatory panels – intended to set a baseline of performance, although each manufacturer’s IVD is unique – set an accuracy requirement of, for example, correct detection of nine out of 10 specimens. Such a criterion is perfectly acceptable if used appropriately; however, if it were used as a sole criterion, a manufacturer could release a lot with an acknowledged 10% false detection rate.

Manufacturers must not rely exclusively on external panels to prove compliance but must control products according to their claims while also meeting and using regulatory requirements intelligently. Despite these cautionary comments, from a manufacturer’s point of view, regulatory panels might be useful in maintaining lot-to-lot consistency.

For certain high-risk IVD in GHTF class D (10), the European Commission Common Technical Specifications (30: Section 3.4.2) states that, “The manufacturer’s batch release testing for screening assays shall include at least 100 specimens negative for the relevant analyte”, but gives no further comment on choice of specimens or the required result. Presumably, the intent is to detect specificity problems but the (two-sided) 95% confidence interval around zero false reactives in 100 specimens is 0–3.6% false reactive. For the manufacturer of an IVD claiming a typical 0.05% or fewer false reactives, these 100 specimens superficially add little value, particularly if the 100 specimens have been pre-selected as negative on that IVD and then stored. Certainly, the results with these 100 specimens could not be used as the sole criterion for verification of specificity at release.

Again, for a manufacturer, the importance of these types of panels is as part of a performance consistency monitoring programme.

For assays with a quantitative output, the 100 specimens might give information if the background were monitored and corrected for minor changes in blank readings. A change in apparent signal for the negative specimens or an increased skewness towards the cut-off value might signal that specificity problems might be found with larger numbers of specimens. However, experience also shows that false reactions are likely to be from sporadic high signals, unrelated to the background as a whole but related to impurities in the system, changes in conformations of proteins (leading to new possibility of incorrect
epitopic reactions) or changes in the population or individuals tested, such as a vaccination programme (e.g. vaccination against influenza in the autumn).
7 Relationship between panel members and claims

The particular panel members must fail their criteria if the IVD will not attain the relevant claims at the end of its labelled life. For example, if there is a known change in activity or a drift in signal, the release panel criteria must be set so that each lot will still meet claims at the end of life, despite the drift. Similarly, the end-of-life panel must show that the device has, or has not, met its claims. The signals from the panel members must therefore be correlated with the signals from the critical specimens, providing information on how signals from both the panel member and the critical specimens will change over the life of the IVD. For panel members controlling sensitivity and specificity, the routine R&D work on stability and robustness should have shown the correlation and the change of signal that can be tolerated for as wide a range of conditions as possible; for example, different constituent purification or syntheses, different balances of constituents and various forms of stress (e.g. temperature and humidity) of routine devices. This gives assurance of correlation under conditions that might occur in manufacturing; it also emphasizes the necessity of preparing and validating potential panel members throughout the R&D phase of the IVD life cycle, when these parameters are all accessible. For some components that appear completely stable under routine conditions, extremely stressful conditions must be used to force change of signal in the critical specimens and the panel members to obtain points for the correlation studies (see Examples 1 and 2 in Annex 1). Correlation is the key: the signals from the panel member and the related critical specimen or antigen must follow each other. It is not necessary that the signals be the same, just that the correlation be understood and applied. For example, in many microplate enzyme immunoassays (EIAs) the positive control material is manufactured with an optical density (OD) of around 1.5–2.0, and the criterion in the IFU for a successful assay is >0.8 OD. Thus, the device could have lost more than 50% of its activity and still be declared as efficacious. These same EIA seroconversion series are often claimed have a specimen to cut-off ratio of 1:1. Even if the control material is a dilution of a specimen from a late stage of the infection and bears no relationship to seroconversion, it is unlikely that those seroconversion claims would be upheld if the device had lost more than 50% of its activity on any kind of specimen. It is irrelevant that the device is normally stable and does not
lose activity – if the control material is claimed to function at 0.8 OD, then so must the IVD, otherwise the method of validation of the control material would have failed. The same principle applies to all QA / QC panel members.

### 7.1 Process control charts

As noted in Section 4.2, data from panel members should be made quantitative. Best practice is to plot successive values lot by lot for both release-to-sale and in-process data, generating “process control” charts. There is a large body of literature on how to design and interpret charts to monitor incipient or actual changes in the process, or in the state of the product at release-to-sale, as shown in (31). Such charts for release-to-sale data are expected to be available for inspection (12), given that this is a method of verifying lot-to-lot consistency in manufacture, assuming the QA panel used is valid.
8 Maintenance of panels

8.1 Storage

Fully characterized and validated specimens are valuable; therefore, appropriate storage is wise. To minimize the possibility of degradation, the stock from which the specimens are to be prepared by dilution is usually split between two locations, and stored frozen at a temperature below its eutectic point (usually between $-20$ °C and $-40$ °C), so that liquid water does not occur. Larger stocks of prepared panel members are also usually stored at these temperatures, as are the working size aliquots. Storage of some of the working aliquots at $\leq -70$ °C acts as insurance in case of doubt over the stability of those stored at $-40$ °C. It is unlikely that material stored at both $-40$ °C and $-70$ °C would degrade at the same rate; hence, if differences are found, instability could be suspected. Although it is probably impractical to undertake full stability studies on frozen stored specimens, the number of freeze–thaw cycles permissible and the time allowable for storage unfrozen must be documented unless specimens are thawed, used within a short space of time and then discarded.

The volume in the stored working aliquots needs to be determined from the rate of use, the stability as a liquid and the validated number of freeze–thaw cycles permissible. Storing in too small a volume in too large a vial should be avoided to minimize any effects from potential freeze drying.

The temperature in freezers used for storing panel members must be monitored, ideally electronically, and the temperature record must be retained. Normally, there is an alarm system to alert staff remotely if a freezer shows signs of failure.

8.2 Replacement

Panel members are key in monitoring and verifying the performance of the IVD and its components; thus, any change must be controlled rigidly. The pre-launch risk evaluation of an IVD should consider the situation that the parent stock of a panel member might not satisfy requirements for the expected commercial life of the product. Factors to be taken into consideration are the likelihood of being able to obtain an equivalent specimen and the extent of characterization already available for the existing
specimen. When fresh specimens must be used in QA panels, the risk assessment must also evaluate the validation of the acceptance criteria. As pointed out previously, a thorough, documented understanding of the role of each panel member and of the characteristics required to achieve the role is vital and is an essential part of the development effort of the IVD. This documentation and standard operating procedure (SOP) to control and direct replacement of panel members should be in place from the time the finalized manufacturing documentation of the IVD is prepared (i.e. before verification and validation of the device).

8.2.1 Replacement from an existing stock or with well-defined analytes

Replacement of working vials of panel members is simple if the parent material (e.g. serum pack or manufacturer’s working calibrator (4)) is available. Replacement is simply a matter of dilution into a validated diluent (see Section 6.1.2) and assignment of new criteria to the diluted material. To be sure of conformity, it is best to verify that the signal from the newly made panel member is the same as that from a stock previous version, by assaying the new and previous members side by side in sufficient replication and on sufficient lots of the IVD (or component if it is an in-process control), taking into account the issues discussed in the next paragraph. If there are slight differences, a new criterion can usually be applied by proportion of the difference. The new panel members must be prepared under the usual change control system, and it must be possible to trace the change from the identification number of the new working stocks and the records of use of the panels.

Changes by comparison are known to lead to drift. It is therefore important to reserve some of the working panel members validated during device development and stored at ≤−70°C for use in the comparisons when panel members are changed, and not merely to compare with the previous version of that panel member.

8.2.2 Replacement from a new stock

Paragraph 8.2.1 could apply to replacement of panels of analytes with a well-defined molecular structure (e.g. polynucleotides and some viral proteins), although the parent stock might be exhausted. Even for relatively well-defined entities such as HBV it is clear
that the serotype (for HBsAg testing) or genotype (for NAT) must be taken into
consideration for replacement panel members, the precise level of identity required must
be defined (e.g. ad or ay, ayw or ayr, and ayw1 or ayw3) and the replacement must be
shown to correlate with the original in several distinct lots of the IVD.

Replacement of panel members from a new parent material is more difficult for
heterogeneous analytes. In antibody testing, the spectrum of antibodies (e.g. exact
epitopes detected, exact conformation of those epitopes, affinities and composition in Ig
classes) varies between specimens, making replacement of like for like tedious. Ideally, in
this situation a complete re-validation of the new specimen should be performed,
showing correlation with claimed critical specimens for different lots and for accelerated
stability. Subject to stringent and documented risk assessment related to product claims
and consistency, this might not be necessary; however, it does require detailed
consideration of exactly what is required, particularly for specimens intended to control
critical epitopes such as in seroconversion to HIV, HCV NS3 or syphilis TpN15.
9 References


Panels for quality assurance and quality control of in vitro diagnostic medical devices


Panels for quality assurance and quality control of in vitro diagnostic medical devices


10 Annex 1 – Examples

The anecdotal examples that follow are not comprehensive but are intended to be illustrative. Manufacturers and their R&D groups know their own assays better than anyone else, so they are in the best position to develop and validate their panels. The suggestions in the examples are based on selection of the panels for confirmation of claims but the same principles apply to all method validations; that is, assurance that the test method is providing a meaningful result (e.g. if it is a stability test, that the output really does prove stability relative to claims and not just stability for a particular, possibly irrelevant, spectrum of reactivity).

It is regarded as good practice, subject to the evaluation of residual risk before launch of the product, to verify the validated claims on occasion over the life cycle of the IVD. This will be achieved using the QA panels and test methods initially devised. For some attributes this is a requirement (e.g. whole device stability, ISO 14971 (1) and ISO 23640 (2)) but for others it is more flexible (e.g. specificity, by active review of data from users).

10.1 Example 1: Correlation of a QA / QC panel member with a critical specimen

10.1.1 Background

Critical specimens are likely to be rare and expensive, and our unlikely to be available over the commercial life of an IVD. It might be possible to use critical specimens at all stages of device characterization and at release-to-sale, but in view of the expense and restricted availability of such specimens, validation and documentation of the methods involved in their replacement becomes crucial. Therefore, the critical specimens themselves are not particularly suited for inclusion in QA / QC panels and substitutes must be found. Dilutions of random specimens from late stages of the infection are unlikely to substitute for seroconversion, because the class, the affinity and even the epitopes involved are likely to be different. Similarly, dilutions of specimens from one stage of an infection or specimen type may well not monitor the behaviour of the IVD for a different stage or specimen type. For NAT, the presence of different amounts of analyte in organisms in immune complexes or in perinatal whole blood could present problems hence, throughout the development work, R&D should be searching for readily available
specimens that mimic the rare, critical ones, and proving that the specimens chosen do indeed monitor the expected characteristic. Obtaining the correct panel is not an insignificant task, the main thing is that throughout the development of the device, the potential panel member and the critical specimen type should correlate, and should have been tested together whenever possible. Frequently, several specimens must be used and tested every time the critical specimen is tested – this allows the correlation to be obtained and non-correlating specimens to be rejected. Once the correlation is understood, it is possible to establish the criteria for the signal generated by the IVD for that panel member and hence to have evidence of meeting the corresponding claim.

The following example is drawn from the development of an EIA for antibodies to Treponema pallidum when it was known (from the input requirements) that a particular specimen in a particular regulator’s collection had to be detected in order to allow that IVD to be sold in various jurisdictions. The cut-off of the EIA had been fixed at 0.15 OD above the mean value of the control negatives and the regulator’s specimen gave a signal of ~0.3 OD above the mean value negatives on routine lots. The specimen was only available to IVD manufacturers in limited quantities but it was in the regulator’s own panel (along with other specimens that presented no challenges). After risk evaluation, it was decided that it was necessary to supplement the QA release-to-sale panel with a mimic of this specimen to ensure consistent compliance with the regulator’s requirement because no other specimen in the panel monitored behaviour with this type of specimen.

10.1.2 Development of a valid panel member

R&D characterization of the specimen showed it to have predominantly IgM antibodies to TpN17 (found by in-house Western blots against a panel of recombinant proteins made for the assay’s development). A search through the in-house collection of specimens found only a limited number that had this characteristic. To be sure that the chosen specimen would mimic the regulator’s specimen, it was tested alongside the regulator’s specimen when the IVD was in development, with several different experimental batches of the recombinant fusion protein and with the lots also severely stressed by accelerated stability-like studies. That led to a Passing & Bablok regression (3), as shown in Fig. A1.
Panel members for quality assurance and quality control of in vitro diagnostic medical devices

Fig. A1 The lower values were from stressed IVD, the higher from R&D IVD with varying concentrations of recombinant from different purifications.

10.1.3 Establishing the criteria for the panel member

a) The regression appeared reasonable and the putative panel member gave about twice the signal from the regulator’s specimen.

b) The regulator’s specimen had a signal of about 0.3 OD on routine lots so it was decided to dilute the in-house specimen twofold in negative serum (in order to have a greater volume available).

c) Having shown that the dilution did indeed have half the signal of the undiluted material on several lots of the device, a criterion of 0.45 OD was assigned to the panel member in the release specification. That value was calculated from the observation that the panel member as diluted gave a signal of about 0.35 OD with routine lots (which gave about 0.3 OD with the regulator’s specimen) with a standard deviation of about ± 0.02 OD, and to allow for a known loss of activity of about 10% over the assigned life of the IVD.

d) For end-of-life testing criteria, this panel member was assigned a value of 0.4 OD.
10.4 Comment

This was an exceptional case – it will not always be possible to perform such complete regression studies – but it serves to illustrate the principle. For qualitative IVD or IVD with discrete steps of signal (e.g. when read from a scale of colour with integral values), regression is much more difficult to show. However, important factors are the proof that the panel member varies with the critical specimens it is to monitor, and that the criterion assigned will ensure the claims are met at end of life.

10.2 Example 2: Correlation under stress conditions

10.2.1 Background

In Example 1, correlation between a rare, critical specimen and potential panel member was proven using data obtained from both accelerated stability-like stress and changes in recombinant preparation and concentration. Sometimes, routine accelerated stability conditions are insufficient to effect change in the IVD, and it appears to be completely stable. It is still necessary to show that the panel member and specimen will correlate as the IVD is stressed, thus helping to ensure that unforeseen issues in manufacture are likely to be detected (see Section 0).

This example is of a stable RDT for anti-HIV based on lateral flow and with a recombinant polymeric gp41 as the constituent for HIV-1 detection. The RDT provided a qualitative result but an in-house step-wise graduated colour chart was developed for QA/QC purposes to quantify line strength. The chart was validated both by serial dilution of specimens and by changes to the concentration balances of recombinants used. Independent readers were in agreement with the assigned classification of results of test devices about 90% of the time, and were never more than one grade-step in disagreement, over the whole range of gradation. On this basis it was concluded the colour chart was valid for routine use, with both ratios and differences of grade being meaningful.

Detection of critical specimens in a number of commercial seroconversion series was claimed. Between different recombinant purification lots, the appropriate panel members were shown to correlate well with both IgG-first and IgM-first seroconversions (as
reflected by class-specific Western blots, relative activity on second generation (IgG only) and third generation (IgG and IgM) commercial assays, and by protein A-based and protein L-based research assay. There appeared to be no loss of activity either with panel members or the critical seroconversion specimens over the claimed life of the assay (24 months claimed, 27 months stability data, 4–40 °C, humid or dry storage). Risk assessment found that because activity of the gp41 component of the assay depended critically on the conformation of the recombinant (4, 5), further correlation to include stress conditions would give more assurance that the panel members could detect subtle changes in the recombinant that might not be detected in R&D conditions, but might affect the manufacturing, perhaps in the long term.

### 10.2.2 Extended correlation work

Knowledge of the IVD suggested that stressful conditions could include storing the IVD out of its protective pouch for various lengths of time under humid conditions at elevated temperatures or by freezing and thawing it several times, again out of its protective pouch. Trial experiments with the putative panel members showed that both these sets of conditions caused loss of sensitivity: the more the stress, the more the loss. Neither of these conditions were likely to occur either with users or in manufacture; nevertheless, they were judged to give additional, different and useful information on the state of the IVD beyond simply using different lots and concentrations of the recombinant.

It was not practical to test the critical specimens more than once under each condition because of rarity value, but the panel members could be tested several times at each condition. The work was done with three lots of the IVD made to final documentation on routine equipment. Lot A was at the end of its assigned life, whereas Lot B and Lot C were recently manufactured but from different lots of critical constituents.

Some of the correlation data for the QA panel member (QA 123) and critical seroconversion specimens 1 and 2 is shown in Fig. A2. The work was done under isochronous conditions (6: Section 4.3.2), with RDT being taken from routine storage and placed under stress at different times, starting at 10 hours before testing, then 7 hours and so on for the data shown, in order that all specimens could be tested in random order.
at about the same time. Each number represents the result from an individual test of an RDT kept under the stated conditions before use. The result was read by two readers independently and the score relative to the colour scale was recorded for each.

**Fig. A2. “Correlation” data**

A value such as 3.5 indicates that the two independent readers gave different values, but these were never more than 1 grade different (e.g. 3 and 4 gives 3.5, whereas 1 and 0 gives 0.5).

**Comments**

Although the analysis of these data are not statistically rigorous, the general trends can easily be seen; for example, the panel member was inactivated at about the same rate as this pair of critical specimens (which happened to be IgG-first seroconversions). Data from IgM-first seroconversions and the relevant QA / QC panel member and from the corresponding freeze–thaw experiments were equivalent to these results.
This study was approved by the risk assessment group as sufficient to show satisfactory relationships between the QA panel members and the critical specimens, when taken together with the similar data obtained with different lots of the recombinant, as stated above. Approval to manufacture and release the IVD was given in relation to these panel members, which were considered to have a low risk of failing to detect non-compliant product.

The panel member QA 123 was assigned a release-to-sales criterion of ≥3 but ≤5 on the basis of this and other evidence, and of ≥3 for end-of-life verification testing (for which the IVD is kept at a constant 40 °C for more than its assigned life).

### 10.3 Example 3: Validation of a release-to-sale panel for anti-HCV

This example applies both to RDTs with line intensity related to a graduated scale and to EIAs.

#### 10.3.1 Background

As noted previously (Section 5.1), a release-to-sale panel must be validated to provide evidence that the tested lot will meet all the claims for the device for at least the entirety of its labelled shelf life. The claims will be derived from user input requirements and will generally include specificity, sensitivity to seroconversion, sensitivity to genotypes, capability to work on whole blood (for RDTs) and on serum and various plasma types, a functional internal control and so on. A review of the input documentation and the risk analyses is important, and is likely to produce a number of other attributes that will be claimed for the IVD but are not directly performance related (e.g. microbiological stability and operating temperature ranges). R&D should have developed the device with these and the other input requirements in mind. During the development, R&D will probably have noticed that at least NS3 and core antigens from HCV are required, and that some lots of antigen (especially NS3\(^1\)) detect some seroconversion series later than other lots. They will probably have noticed also that some lots of antigen are less specific than

---

\(^1\) NS3 (proteases and RNA helicase) is a known to be a difficult antigen to use in product development; it seems to include both conformational and linear epitopes, and the exact conditions of recombinant culture and purification are critical to functionality – see, for example, European patent EP 1 471 074 (7) and Mondelli et al. (1994) (8).
others, especially NS5, if that is included in the IVD. Depending on the format of the assay there might be separate components to detect IgG and IgM. Whole blood is known to cause flow problems in many IVD, and there should be an input requirement related to invalid rates; the risk analysis should have reviewed this and R&D should have developed the device to overcome the problem. A competent design input risk assessment should have identified a range of other factors related to safe usage and other factors that will bring “added value” to the users of the device and so bring commercial advantage.

1054 10.3.2 Development of a valid panel

For release-to-sale panels, not all claims associated with the IVD need to be addressed, depending on the risk assessment (e.g. microbiological stability, which can be validated during R&D from antibiotic efficacy and stability studies). For HCV, genotype detection can probably be validated in R&D and not verified at release. Risk evaluation must occur continuously in R&D to minimize the amount of work at release by proven and documented validation of as many factors as possible.

Characteristics of IVD that are known from R&D studies to vary between lots must be verified at release (e.g. sensitivity, specificity or invalid result rate), and lot variance must be shown to be within acceptable limits. Also, it is generally accepted that concentrations and functions of critical constituents (e.g. antigens, antibodies, and biologicals such as protein A and streptavidin) must be verified at release. The number of specimens to be tested can be minimized by making appropriate choices, so that each specimen can monitor the condition of more than one thing (e.g. anti-NS3 first seroconversion and the NS3 antigen in the system).

Once the claims to be verified have been decided, the specimens used to monitor them can be selected and evaluated as in Example 1 of Annex 1. A comparison of the published commercial line-assay results for critical specimens with those from [dilutions of] potential panel members might help in deciding on the initial choices for an anti-HCV IVD.

For specificity monitoring, it is usual to include known falsely reactive specimens found during R&D; some falsely reactive specimens are available from the commercial suppliers for inclusion in QA / QC panels. Methods for monitoring invalid result rates must be
devised and validated on a device-by-device basis because fresh specimens must normally be used for these studies, particularly if there is a whole blood claim.

Suggested anti-HCV QA/QC panel members with these considerations in mind are given in the TGS2 guidance document for establishing stability of in vitro diagnostics assays and components (9).

### 10.3.3 Establishing the criteria for panel members

The particular panel members must fail their criteria if the device will not attain any of the claims at the end of its labelled life. The signals from the panel members must therefore be correlated with the signals from the critical specimens, as noted in Example 1. For panel members controlling sensitivity and specificity, the routine development, stability and robustness work should have shown the correlation and the loss (or gain) of signal that can be tolerated; it should also have shown that the test method is satisfactorily less variable than the variability between lots of the device itself (see gauge R&R in Section 6.1).

### 10.4 Example 4: A release panel for a flow cytometer for the enumeration of CD4 T-cells

#### 10.4.1 Background

The user input requirements for a quantitative assay generally include:

- a specified limit of detection (LoD), with or without a specified limit of quantitation (LoQ);
- a measure of accuracy at clinically relevant thresholds;
- (possibly) a specified number of calibrator points, and hence provision of a calibrator solution) and
- (probably) a specified linear range or, more scientifically, a range with a specified accuracy and precision.

For a CD4+ T-cell enumerating assay, there will almost certainly be a specificity requirement for CD4+ T-cells but not for other cells that might carry the CD4 determinant (e.g. monocytes). The design input risk analysis should have identified other well-known
CD4-related problems, including potential analyte-specific interference (e.g. tuberculosis with its effects on CD4-bearing monocytes and, for some RDTs, malaria with its effects on red blood cell haemolysis). A competent design input risk assessment should have identified a range of other factors related to safe usage, and others that will bring “added value” to the users of the device and so bring commercial advantage. Many of these factors might lead to claims that require verification at release in addition to evaluation during device characterization; see, for example, the WHO sample product dossier for CD4 IVD (10).

10.4.2 Development of a valid panel

For release-to-sale panels, not all claims associated with the IVD need to be addressed, depending on the risk assessment. For example, monoclonal antibodies from a reputable source are unlikely to change in avidity. However, device specificity could conceivably change from lot to lot, depending on the precise purity of the monoclonal antibodies; also, there could easily be variation between lots in the exact proportions of the monoclonal antibodies used in manufacture, which could lead to differences in ratios of cell counts and effects from analyte-specific interference. As noted previously, risk evaluation must occur continuously in R&D to minimize the amount of work at release, by proven and documented validation of as much as possible; in particular, any characteristics of the IVD that are known by R&D studies to be variable between lots must be verified at release.

Selection of QA / QC panels for an analysis such as CD4+ T-cell enumeration presents problems because the analyte cannot easily be stored – even short periods of time at 2–8 °C can affect the apparent cell count. It might be possible to design the device so that blood could be stored, or some form of stabilized or artificial “cells” might be developed. In each of those cases, the relationship between the behaviour of real blood and the substitute must be proven rigorously throughout device development. As in the HCV example, it must be proven and documented that there is satisfactory correlation between the real specimens and the substitutes.
Where fresh blood is used at release-to-sale, the relationship between the device being tested and the device used as the calibration method must be validated for all lots of the tested device. A previous batch of the device is unlikely to adequately meet the requirements as a validated test method because it could lead to a drift in device performance.

### 10.4.3 Establishing the criteria for panel members

The general comments from the previous examples are applicable; however, for the quantitative aspects of testing the criteria must be established directly from the numerical claims (e.g. LoQ: 20 CD4+ T-cells/µL, accuracy: ±25 at 350 CD4+ T-cells/µL). Of course, the criteria must be set so that the claims will be met at end of life, so the criteria might well be different from the claim. The criteria for any substitutes for fresh blood must be shown to correlate with the same attributes in the blood in the same way as the substitutes for critical specimens in Section 10.3.3.

Criteria must be established and documented during R&D for panel members monitoring claims that are not related directly to quantitation of the analyte (e.g. specificity on specimens from patients with tuberculosis and invalidity rate on specimens from patients with malaria). Similarly, criteria must be established and documented for panel members monitoring appropriate reactivity of the critical, potentially variable, components of the device.

### 10.5 Example 5: Nucleic acid testing

#### 10.5.1 Background

There is a wide variety of NAT methodology, qualitative and quantitative, based on DNA and RNA, and with intended use ranging from blood donor screening to infant diagnosis. The main issues relating to the assay concern specimen collection and preparation, inhibitory substances varying by specimen, stability of the enzyme systems involved and contamination during manufacture or in use. Occasionally, sub-genotypes present problems with some systems. Claimed specimen types usually include whole blood, often in the form of dried blood spots, in addition to buccal fluid, various types of respiratory samples, cerebrospinal fluid, stool, serum and plasma. Control materials, including
calibrators for quantitative assays, are normally supplied with the IVD. As for all devices, each of the issues found during design input and process risk management processes must be dealt with and the methodology developed to minimize any effects. The nature and type of specimens to control all the hazards must be defined during the R&D phase of the work.

**Development of a valid panel**

As usual, each of the claims that cannot be validated during R&D for all lots over the commercial life of the device must be verified lot by lot at release-to-sale. In particular, the panels must demonstrate maintenance of sensitivity or LoD, or accuracy and precision for all claimed specimen types and all claimed genotypes if the latter cannot be validated in R&D. There are international conventional calibrators against which in-house QA / QC panel members can be standardized for most organisms of concern to WHO\(^1\) (see also Section 6.4), and the derived secondary standards should be enough to monitor most claims for both qualitative and quantitative assays. ISO 17511 (14) describes the types of calibrators and methods for tracing secondary, in-house, standards to internationally accepted calibrators that are expensive and in restricted supply. For genotype claims, it might be necessary to refer to collections such as the 1st WHO International Reference Panel for HBV Genotypes for NAT-Based Assays (15), or the 1st International Reference Panel HIV-1 RNA Genotypes (16). However, it is likely that the manufacturer will be more aware of the difficult subtypes from in-house and user testing than might be revealed by these panels. In that case, critical specimens will probably need to be included in the panels, at least for in-house validation of attributes such as stability and inhibition. Whenever in-house panels are made, the correlation under varying conditions with the internationally accepted materials must be demonstrated, as for any substitute specimen; proof of ISO 17511 (14) metrological traceability alone is not sufficient.

---

\(^1\) For example, the 3rd HIV 1 International Standard (11), the 4th International Standard for Hepatitis C Virus for Nucleic Acid Amplification Techniques (12) and the Third International Standard for HBsAg (13).
Proof of functionality of extraction, especially from dried blood spots, and removal of inhibitory substances will almost certainly need to be verified lot by lot. This might require specimens in the panel or simply chemical analysis of the reagents.

10.5.3 Establishing the criteria for panel members

Refer to previous examples.

10.6 Example 6: Use of imposed specimens as QA / QC panel members

10.6.1 Background

Some regulatory authorities require that specimens supplied by them (with criteria also supplied by them) be detected appropriately before each lot of an IVD can be released-to-sale. Regulators cannot know the claims of every IVD within their purview, so such specimens cannot act as a valid release-to-sale panel. They are a requirement and, as such, will be in the design input documentation, but they should not be used as a sole basis for the manufacturer to verify release-to-sale.

10.6.2 Reasons for not using imposed panels for QA / QC as the main panel members

See also Section 6.6.

a) A regulator’s panel might not control all components. The following are actual examples. A regulator provided a panel for HCV release-to-sale that contained virtually no NS3 antibodies, so the lot could still be released even though the NS3 component of a device could be lacking. Another regulator changed an HIV release panel member from one that contained virtually no anti-p24 to one that contained high levels of anti-p24, both dilutions of positive specimens. The signals on some EIAs appeared to be equivalent, but not those on others, and the specimen did not correlate with seroconversion sensitivity – in fact it contributed little and certainly did not monitor critical aspects of devices.

b) A panel member might be changed by the regulator, as in point (a), so that although still being detected appropriately as “positive”, it could allow a change in assay sensitivity to occur without being detected; alternatively, the device manufacturer could change his manufacturing to meet the new imposed panel
member acceptance criteria with no control over the influence of the manufacturing change on specificity or sensitivity to meet his established claims.

c) The requirements on the panel might be such that the device is not controlled within its claims. For example, if 19 of 20 panel members were required to be detected as negative, then a lot with about 5% false reactives could be released despite a claim of a much better specificity.

d) If a lot meets the regulator’s requirements but the manufacturer decides the lot would not meet its claims and so does not release it, the manufacturer would be expected to document why the lot was not released as part of its QMS, to establish an auditable trail.

e) Even international calibrators are not necessarily valid for release of a device against a sensitivity claim. When the first international HBsAg standard was replaced by the second, this changed the apparent sensitivity of a number of assays (17). Claims must always be cited against particular versions of international conventional calibrators.

f) A manufacturer is responsible for confirming and maintaining his claims over the commercial life cycle of the IVD, despite external regulatory changes.

10.6.3 Appropriate uses of regulators’ panels by manufacturers

It might be possible to validate the regulator’s panel as in Section 10.3.3, and use it appropriately. However, the correlation work would be the same as the manufacturer choosing the specimens and there would be no security against the regulator changing the panel without notice.

Using the regulator’s panel alongside the manufacturer’s own panel to monitor changes by routine trend analysis might occasionally be informative – usually about changes by the regulator. However, panel testing is done independently of the manufacturer by a regulatory authority. For example, for NAT devices, the panels are designed to test for a lot’s ability to detect a minimally acceptable number of copies/mL, defined by a regulatory authority. Any adverse comments on the composition of a regulator’s panel should always be discussed with that authority.
Panels for quality assurance and quality control of in vitro diagnostic medical devices


Panels for quality assurance and quality control of in vitro diagnostic medical devices (TGS–5)


1299  17 Ferguson M, Health A, Lelie N, Nubling M, Nick S, Gerlich W et al. WHO Working Group on Hepatitis and HIV Diagnostic Kits: report of a collaborative study to 1) assess the suitability of candidate replacement international standard for HBsAg and a reference panel for HBsAG and 2) to calibrate the candidate standard in IU. 2003; (http://apps.who.int/iris/handle/10665/68581, accessed April 2017).