A Practical Approach to Biological Assay Validation

Joke Ederveen

Sponsored by the Dutch Ministry of Housing, Spatial Planning and the Environment (VROM)

2010

Progress report number 08090

PROGRESS
project management & engineering

A Practical Approach to Biological Assay Validation

Author:

J.C. Ederveen

Contact:

Joke Ederveen
Progress, Project Management & Engineering
P.O. Box 60; 2130 AB Hoofddorp, the Netherlands
Tel. +31 23 563 5016 / +31 6 4637 7226
j.ederveen@progress-pme.nl
www.progress-pme.nl

Steering committee:

Dr. D.A. Bleijs, RIVM, GMO Office, Bilthoven
Dr. N.A. Kootstra, Academic Medical Center, University of Amsterdam
Drs. D. Louz, RIVM, GMO Office, Bilthoven
Drs. P.C. van Mourik, Sanquin Pharmaceutical Services, Amsterdam
Dr. B.P.H. Peeters, Central Veterinary Institute, Lelystad
Ing. A.L.M. Wassenaar, RIVM, GMO Office, Bilthoven
Dr. G. van Willigen, Leiden University Medical Center

Sponsor:

This report has been ordered and sponsored by the Dutch Ministry of Housing, Spatial Planning and the Environment (VROM).

Acknowledgement:

The assistance of Dr. M.A.N. Rits in the preparation of this manuscript is highly appreciated.

Hoofddorp, May 2010

Summary

This report is written from the perspective that 'validation' is a familiar concept, but in practice often interpreted and applied in different ways amongst scientists, government officials involved in authorizing experimental laboratory work, and inspectors. The document is meant as practical guide for the execution and implementation of biological assay validation in laboratory research.

Extensive literature on the subject of assay validation is available, but its content is often too abstract or not specific enough or not useful for the aimed type of research. Therefore, in this report, the basic concepts of 'validation' and 'validation performance characteristics' are explained in simple terms and elucidated by examples.

In the report the possible types of assay categories are illustrated and it is shown how each laboratory experiment can be assigned to one of these categories. Depending on the purpose of the assay, the relevant characteristics of an assay are determined that have to be validated.

Assay validation is in fact a continuous process. The validation process starts with a description of the purpose of the method, followed by the development of the assay and the definition of the performance characteristics. It continues with documentation of the methodology and the validation results. During the in-use phase of the assay, there is continuous monitoring to assure that the assay still generates results in accordance with the performance characteristics as originally determined. Revalidation is applicable if the method is changed, or has been out of use for a while, when it is applied for another material or for a new purpose.

For clarity, in this report the assay validation process is systematically subdivided in five different validation phases, which is suitable for validation of biological assays in general. The validation process is demonstrated by detailed, concrete examples of frequently used biological assays.

Samenvatting

Dit rapport is geschreven vanuit het perspectief dat het begrip ‘validatie’ breed bekend is, maar in de praktijk op verschillende manieren geïnterpreteerd en gehanteerd wordt door onderzoekers, overheidsfunctionarissen betrokken bij het verlenen van toestemming voor experimentele laboratoriumwerkzaamheden en inspecteurs. Het rapport is bedoeld als een concrete handleiding voor de uitvoering en implementatie van validatie van biologische assays in laboratoriumonderzoek.

Er bestaat ruimschoots literatuur over dit onderwerp, maar vaak is de inhoud te abstract of niet specifiek voor de betreffende methode of niet bruikbaar voor de beoogde toepassing van het onderzoek. Daarom worden in dit rapport de begrippen 'validatie‘ en 'validatie prestatiekenmerken‘ in eenvoudige termen uitgelegd en zijn illustratieve voorbeelden opgenomen.

Het rapport behandelt vier mogelijke soorten assay categorieën waarin elk laboratoriumexperiment afhankelijk van de vraagstelling ingedeeld kan worden. Het type assay en het beoogde doel bepalen welke relevante eigenschappen van de assay gevalideerd moeten worden.

Assay validatie is in feite een continu proces: het validatieproces begint met de omschrijving van het doel van de methode, gevolgd door het ontwikkelen van de test en het bepalen van de prestatie-karakteristieken. Het proces wordt voortgezet met het documenteren van de methode en het vastleggen van de validatieresultaten. Tijdens de gebruiksfase van de assay wordt voortdurend gemonitord om vast te stellen of de assay nog steeds aan de prestatiekenmerken, zoals deze oorspronkelijk zijn bepaald en vastgelegd, wordt voldaan. Hervalidatie vindt plaats als de methode wordt aangepast, als de assay een tijdlang niet is gebruikt of als de assay gebruikt gaat worden met een ander type materiaal of voor een andere beoogde toepassing.

Voor de duidelijkheid is het validatieproces in dit rapport systematisch onderverdeeld in vijf verschillende validatiefases – een concept dat bruikbaar is voor de validatie van biologische assays in het algemeen. Het validatieproces wordt gedemonstreerd aan de hand van concrete voorbeelden van veel toegepaste biologische assays.

Purpose and scope

The purpose of this document is to provide general understanding about biological assay validation. Illustrative practical examples are shown of how the concept of validation can be applied to commonly used assays: virus infectivity assays, immunoassays (ELISA) and polymerase chain reaction (PCR) assays. This document focuses on the use of these assays for the detection, identification or quantitation of viruses in the context of activities with genetically modified organisms (GMOs) in the laboratory.

This report is meant as a guidance document for scientists, laboratory personnel, (biological) safety officers, auditors and regulatory bodies for developing a validation plan. This report should not be regarded as a prescription for the way specific assays should be validated, because this is context-dependent and mainly driven by the purpose of the assay, the scientific questions asked and the risks associated with the outcome of the assay. Although examples are provided in the report, the selection of the assay or test method that is best suited for a particular application, as well as to which extent feasibility or validation studies have to be done, is beyond the scope of this report.

Table of contents

Summary 2

Samenvatting 3

Purpose and scope 4

Table of contents 5

PART I: VALIDATION PRINCIPLES AND VALIDATION PARAMETERS 8

1 Introduction 9

2 The validation process 11

2.1 Introduction on assay validation 11

2.2 To what extent is validation required? 11

2.3 Assay categories 14

2.4 Assay development, performance and maintenance 16

2.5 Method verification 21

2.6 Concluding remarks 22

3 Assay validation parameters 23

3.1 Accuracy 23

3.2 Precision; repeatability, inter-assay variation, reproducibility 26

3.3 Linearity and range of the measurement 33

3.4 Specificity 34

3.5 Limit of detection (detection limit) 35

3.6 Limit of quantitation 37

3.7 Robustness 38

3.8 Application of validation parameters 40

4 Documentation 42

4.1 Raw data 42

4.2 Specifications 42

4.3 Procedures 43

4.4 Reports 44

PART II: EXAMPLES OF BIOLOGICAL ASSAY VALIDATION 46

5 Test for detection of virus infectivity ($TCID_{50}$ assay) 47

5.1 Introduction 47

5.2 Validation parameters 48

5.3 Stage 1. Selection and feasibility 49

5.4 Stage 2. Development and standardization 53

5.5 Stage 3. Performance qualification 55

5.6 Stage 4. Performance validation 59

5.7 Stage 5. Maintenance and improvement 60

5.8 Method verification 60

5.9 Points to consider 60

6 ELISA for HIV-p24 64

6.1 Introduction 64

6.2 Validation parameters 64

6.3 Stage 1. Selection and feasibility 65

6.4 Stage 2. Development and standardization 66

6.5 Stage 3. Performance qualification 71

6.6 Stage 4. Performance validation 72

6.7 Stage 5. Maintenance and improvement 72

6.8 Method verification 73

6.9 Points to consider 75

7 Polymerase Chain Reaction: PERT assay 78

7.1 Introduction 78

7.2 Validation parameters 78

7.3 Stage 1. Selection and feasibility 79

7.4 Stage 2. Development and standardization 81

7.5 Stage 3. Performance qualification 86

7.6 Stage 4. Performance validation 88

7.7 Stage 5. Maintenance and improvement 89
7.8 Method verification 89
7.9 Points to consider 90
PART III: GENERAL ISSUES RELATED TO ASSAY VALIDATION 93
8 General remarks and recommendations 94
8.1 Reference standards 94
8.2 Training 94
8.3 Contract Research Organisations (CROs) 94
9 Glossary 96
9.1 Terms and definitions 96
9.2 Abbreviations 99
10 References 102

PART I: VALIDATION PRINCIPLES AND VALIDATION PARAMETERS

In this part, the general validation process and principles are described in order to help the reader in gaining understanding of what assay validation is and what parameters are necessary for certain applications for the different assay categories.

Chapter 1 is a general introduction.

In Chapter 2 the different stages in the assay validation process are explained and the differences in requirements for both an entire new method and an existing method are elucidated.

Chapter 3 is an outline of validation parameters. It explains the parameters in basic terms for those who are not yet familiar with the concept of validation or want to refresh their knowledge of this subject.

Chapter 4 focuses on the advantages of good documentation. It covers documentation aspects related to validation and contains recommendations for the use of standard procedures, validation protocols, and reports. It shows how an assay fit for purpose should be documented.

1 INTRODUCTION

Good science requires well-planned, well-executed and well-documented experiments followed by a meaningful interpretation of data. Experiments should be based on generally accepted scientific principles, and appropriate controls should be included to demonstrate that the experimental setup is working as expected. Besides a sound study design, it is important that the results of experiments are accurate, reliable and reproducible. In order to demonstrate these characteristics, assay validation comes into place.

Assay validation is the evaluation of a test method to determine its fitness for a particular use. In a validation process, the performance parameters of an assay are studied to verify that they are sufficient for providing the data to answer a particular problem or question for which the assay is intended to be used. In addition, the validation process may provide understanding of the limitations of the assay.

Validation can be seen as an ongoing process of the assay in use, in which important elements and concepts such as its design, selection, feasibility, development, establishment of the method, performance characteristics, monitoring and trending, optimization, changes and revalidation are embodied. This is called a 'validation life cycle approach'. For many applications, however, assay validation is constrained by time and resources and only factors inherent to the optimization and standardization of an assay are considered relevant. Therefore, the extent of the required validation depends on the purpose of the assay result.

In the pharmaceutical community, risks related to safety of human and animal health care products are addressed by a detailed set of guidelines and regulations covering the development, production and testing of these products, including procedures related to the use of analytical methods. Assay validation is an integral part of the underlying quality system. The importance assigned to it, is illustrated by the existence of several regulatory and guidance documents covering this subject that have been published by official agencies or professional bodies like the Food and Drug Administration (FDA) [2], The European Medicines Agency (EMEA) [8, 27], and the International Committee on Harmonisation (ICH) [1]. These documents provide a framework for correct and reliable testing and describe the fundamental principles, boundaries and specifications that are important in producing accurate and reproducible results during the manufacturing and testing of medicines.

In addition to the risks related to medicine use and patient safety, there is an increased demand to perform risk assessment dealing with risks and consequences of exposure to man and environment when working with potentially hazardous biological agents −including GMOs− in the laboratory. The employment of validated assays may be a necessary regulatory requirement in the process of minimizing possible risks and facilitating activities at appropriate containment levels. Currently no official guidelines in this field are available.

Despite the availability of extensive documentation related to assay validation, often questions related to the implementation of method validation remain, especially to specify which assay parameters are required to enable a valid interpretation of the results. Assay validation is not an easy or straightforward task and may be difficult to learn from a text book. This document intends to provide some practical guidance and advice on this specific topic.

The structure of this report is as follows: in Part I general principles of assay validation are explained. Simple illustrative examples are provided with emphasis on biological assays. The process of validation itself is described and general terms of validation parameters are explained in order to help the reader in gaining a better understanding of the meaning of these parameters and the context that they are used in. References to literature for obtaining further information are provided.

In Part II practical examples of three types of biological assays are provided, namely a cell culture based infectivity assay, an immunoassay and a PCR assay. For each assay, an example of a validation process is provided and known and possible pitfalls and points to consider are presented as well.

Part III contains general aspects that have not been addressed in Part I and part II.

2 THE VALIDATION PROCESS

2.1 Introduction on assay validation

Validation of an assay is the process to determine its fitness for a particular use. Appropriate validation is solely the responsibility of the laboratory that uses the method to produce the results.

Much literature on the subject of validation has been published and there is general consensus on the definition and application of assay validation principles: 'Assay validation involves documenting, through the use of specific laboratory investigations, that the performance characteristics of the method are suitable and reliable for the intended analytical applications. The acceptability of analytical data corresponds directly to the criteria used to validate the method.' [1].

The principles and practices of validation of analytical procedures are covered by the International Conference on Harmonisation (ICH) and they are published in guideline Q2(R1), 'Validation of Analytical Procedures' [1]. This document is used as a basis for other useful publications and guidelines, e.g. the EURACHEM Guide on method validation [3] and the FDA Guideline on Biological Method Validation [2].

Typical words used in the definition of validation provided above are 'documenting', 'suitable', 'reliable' and 'intended application'. This report will address these aspects in a practical way.

A validation study can be seen as series of experiments establishing performance characteristics of the assay under investigation such as accuracy, specificity and variation. However, taking into account that an assay suitable for its purpose should provide meaningful results, the concept of assay validation is generally seen in a broader perspective. The assay design, the type and nature of samples to be tested and the usefulness of the test results should also be taken into account. Therefore, the next sections of part I of this document address the general validation principles covering all aspects of the entire process. In Part II of this document, a practical approach of these principles is applied to examples of an infectivity assay, an immunoassay and a polymerase chain reaction assay.

2.2 To what extent is validation required?

Before going into a detailed description of the validation process itself, it is useful to know that the validation requirements depend on the purpose of the method.

Purpose	Assay Category	Example of Test	Relevant aspects	Validation parameters
To confirm the identity of a virus	Identification	PCR	Cross reactivity Matrix effect Specificity of PCR amplicon	Specificity
No more than 1000 *E.coli* bacteria are present per gram of tablet	Quantitative test for impurity	Bacterial culture assay	Can the test accurately quantify *E.coli* in the tablets?	Accuracy
			Are the titre measurements reproducible and precise?	Precision
			Are only *E.coli* or also other bacteria detected?	Specificity
			Can the assay demonstrate titres as low as 1000 bacteria/gram with acceptable accuracy and precision?	Quantitation limit
Absence of Replication Competent Lentivirus (RCL) in a batch of vector material	Limit test for impurities	Cell culture assay	Does the matrix influence the replication of RCL e.g. by inhibition and/or competition?	Specificity
			To what concentration can the assay detect virus and what sample volume can be handled in the test?	Detection limit
To determine the titre of the vector stock	Assay (concentration)	ELISA	Can the assay accurately determine the titre?	Accuracy
			To what range is the assay linear and in what dilutions can the titres accurately and precisely be measured?	Linearity Range Limit of Quantification
			Is the signal specific for the vector?	Specificity
			Is the measurement precise and accurate between days and with different technicians?	Repeatability Precision

Table 1. Four examples of biological tests, their purpose, relevant aspects and the validation parameters to demonstrate the required extent of validation. Note that only one example of a test is given, but that more assays can meet the intended application.

In order to get a first understanding of the subject, Table 1 shows practical examples of analytical tests. Depending on the intended application of the assay, relevant aspects of the assay can be identified, leading to validation parameters that need to be addressed. Further down in this Chapter, the different assay categories mentioned in Table 1 are explained. In Chapter 3, the validation parameters will be discussed. But first three cases are presented to further elucidate the possible scope of assay validation.

Case 1 describes the use of an entire new, quantitative, method. Without having extensive knowledge of assay validation parameters, it is clear that many aspects

regarding the validity of these measurements must be taken into account. The measurements should produce consistent and reliable quantitative results. Both assay development en assay validation are applicable. The differences in validation requirements between Case 1 and Case 2 are substantial.

Case 1: Measurement IgG antibodies against a viral vaccine (*'New method'*).

Consider the conduct of a clinical trial with a new viral vaccine in a group of human participants. An important aspect of the study is to monitor the development of IgG antibodies after immunization. Some important aspects that need to be addressed with respect to the measurement of the antibody titres are:

- Is the measured titre correct?
- Can the antibody titre be measured in sera from the entire panel of human individuals?
- Does the assay measure specifically IgG antibodies or also other types of immunoglobulins?
- Is the antibody response specific for the viral antigen in the vaccine or does it react with other virus strains?
- In what range can the titre be measured?
- When is a sample positive and when negative? How well can the day to day results be compared?

In Case 2, a previously validated method is changed. The method variation is expected to be minor and can be addressed in a simple validation experiment.

Case 2 : New bacterial growth medium (*'Method change'*).

A new medium is developed for the isolation of a specific recombinant *E. coli* strain. A validation study is required to determine if this medium is suitable.

The question: is the new medium as efficient as the old medium?

In this case a simple experiment may be sufficient to establish the new medium's validity: determination of the %-recovery on the new medium compared with the old medium.

Case 3 is an example in which a previously validated assay is transferred to another laboratory. The performance characteristics of the assay are known, but transferring the assay to another laboratory implicates many changes, e.g. other equipment, source of materials, environmental conditions, standards, samples, personnel. A method

verification process is applied to demonstrate that the method is still performing according to its original (validated) performance characteristics.

Case 3: Virus assay (*'Method transfer'*).

A virus plaque assay, used to determine infectious titres in samples, has been developed and validated in diagnostic laboratory X. A report containing the details on the method as well as validated performance characteristics (accuracy, robustness, repeatability, limit of detection etc.) is available.

Laboratory Y is implementing the same method, using the same protocol and materials. The assay is transferred and set up in Laboratory Y.

A 'Method verification' study is performed, to demonstrate that the same accuracy and precision are obtained as in the previously validation study in Laboratory X. If the criteria are met, the assay at Laboratory Y is considered validated.

To be able to ascertain the requirements, assays can be divided into assay categories depending on their purpose (see below).

2.3 Assay categories

In general, the purpose of measurements can be divided into one of the following four assay categories:

- Identification ('Can the identity of a certain compound be confirmed?')
- Quantitative test for impurities ('How much of a certain contaminant is present?')
- Limit test for impurities (Demonstration of 'Not more than' or 'Absence' of a certain contaminant)
- Assay ('What is the concentration and/or potency of the active component?')

Examples of each of the assay categories are given in Table 1, see column 'Assay category'.

When assessing the validation requirements, one should realise that the relevant validation parameters (Dutch: *'prestatiekenmerken')* depend on the specific purpose of the measurement, which can be either qualitative or quantitative.

Qualitative measurements include assessment of the presence of a particular substance (e.g. 'is melanin present in the milk powder?') or have identification purposes ('can the active component present be unmistakably identified?'). This type of assay is generically indicated as '*Identification*'.

Quantitative measurements analyse the amount of a substance. This is either to determine the concentration of an active component (the actual product), or testing for impurities that are usually present at a lower concentration or -preferably- should be absent.

The three different types of quantitative assays with examples are:

- *Quantitative impurity test*:
  e.g. 'it is acceptable that the vector batch may contain a low titre of infectious virus, but the titre may not be higher than 100 infectious virus particles per ml'.
- *Limit test for impurities*:
  e.g. 'no infectious virus particles may be present in a 100 ml aliquot of inactivated viral vaccine.'
- *Assays*:
  e.g. 'measure the titre of viral vector in a gene therapy product'.

In the ICHQ2(R1) guideline on assay validation [1], and many other related guidelines documents, e.g. the United States Pharmacopeia [10], tables with the different assay categories and the required validation parameters are present, which are in fact all based on the same concept laid down in the ICHQ2(R1). In Figure 1 the type of measurements and the specific parameters to be assessed are summarized schematically.

Understanding the differences of the four assay categories helps the laboratory in choosing the relevant validation parameters for a particular method.

Useful background information about to what level of validation is required, can be found in the EURACHEM Guide 'The fitness for Purpose of Analytical Methods. A Laboratory Guide to Method Validation and Related Topics'. Although not specifically written for biological assays, the general principles are well explained [3]. Practical examples related to the validation of alternative methods for control of microbial quality are provided in the European Pharmacopoeia, general text 5.1.6. [28].

Now that the different assay categories have been explained, it may be clear that the assay performance characteristics differ upon the purpose of the method. The next step is to look into detail at the validation process itself.

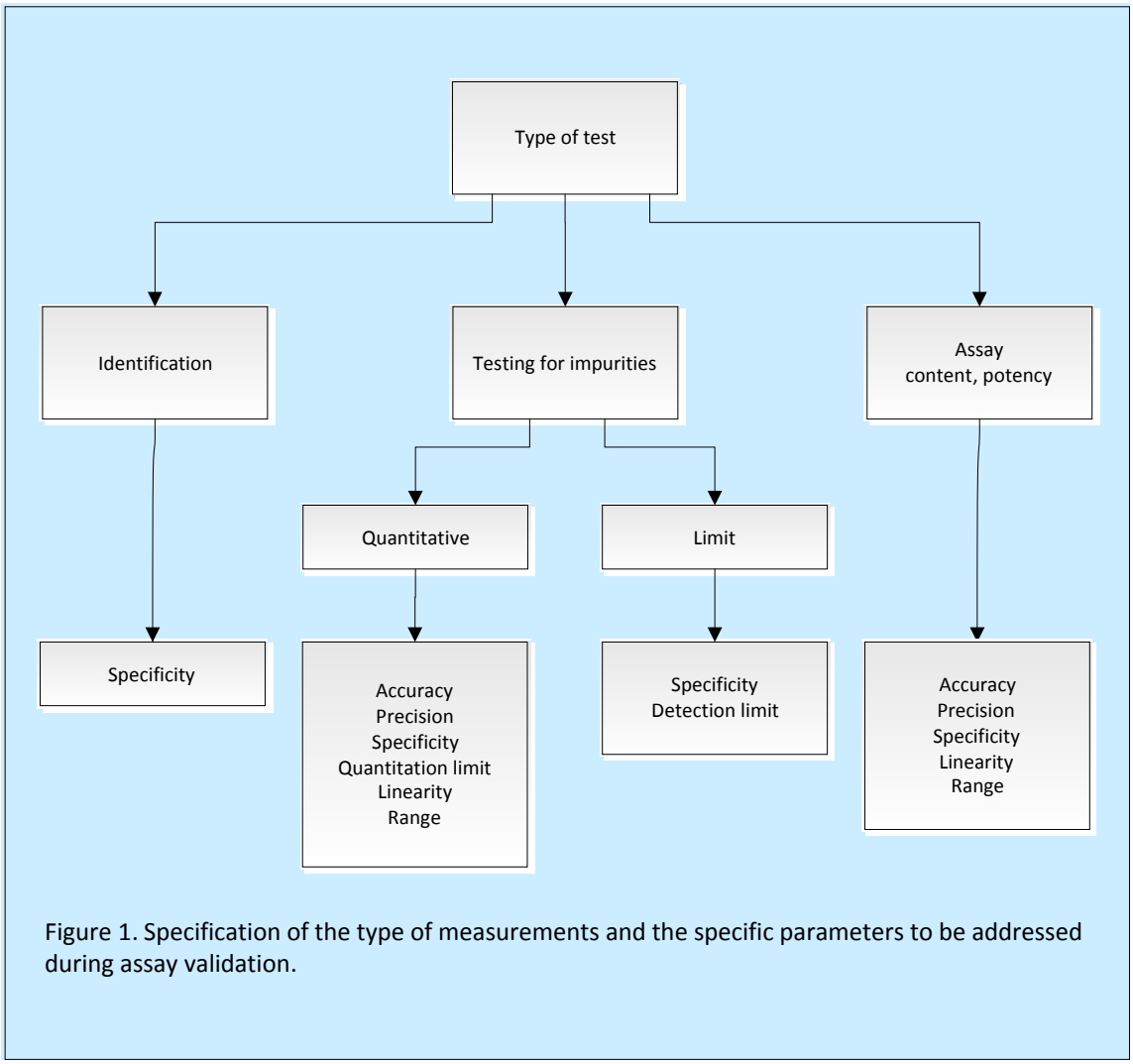

Figure 1. Specification of the type of measurements and the specific parameters to be addressed during assay validation.

2.4 Assay development, performance and maintenance

2.4.1 The validation process stepwise

In Sections 2.2 and 2.3 it is shown that the extent of validation very much depends on both the type of assay and on its purpose. A validation study actually goes further than the performance of validation experiments alone and there are several steps taken in a validation process in order to qualify an assay as suitable for its use.

Method selection, feasibility, development, suitability testing, performance characterization, performance validation, documentation, maintenance and improvement, are important steps in the overall life cycle of a 'fit-for-purpose assay'. Some laboratories often perform these activities without realising that they are actually carrying out steps of the validation process itself. In addition, sometimes the activities as mentioned above are carried out, but just not appropriately documented.

There is no (official) general consensus on how the assay validation process should be divided into stages, although in some literature practical guidance for defining steps is provided. The World Organization for Animal Health in Chapters 1.1.3 and 1.1.14 of the OIE Manual of Diagnostic Tests and Vaccines for Terrestrial Animals [14, 15] defines five practical stages. In analogy, for practical purposes these five stages are also suitable to define the biological assay validation process.

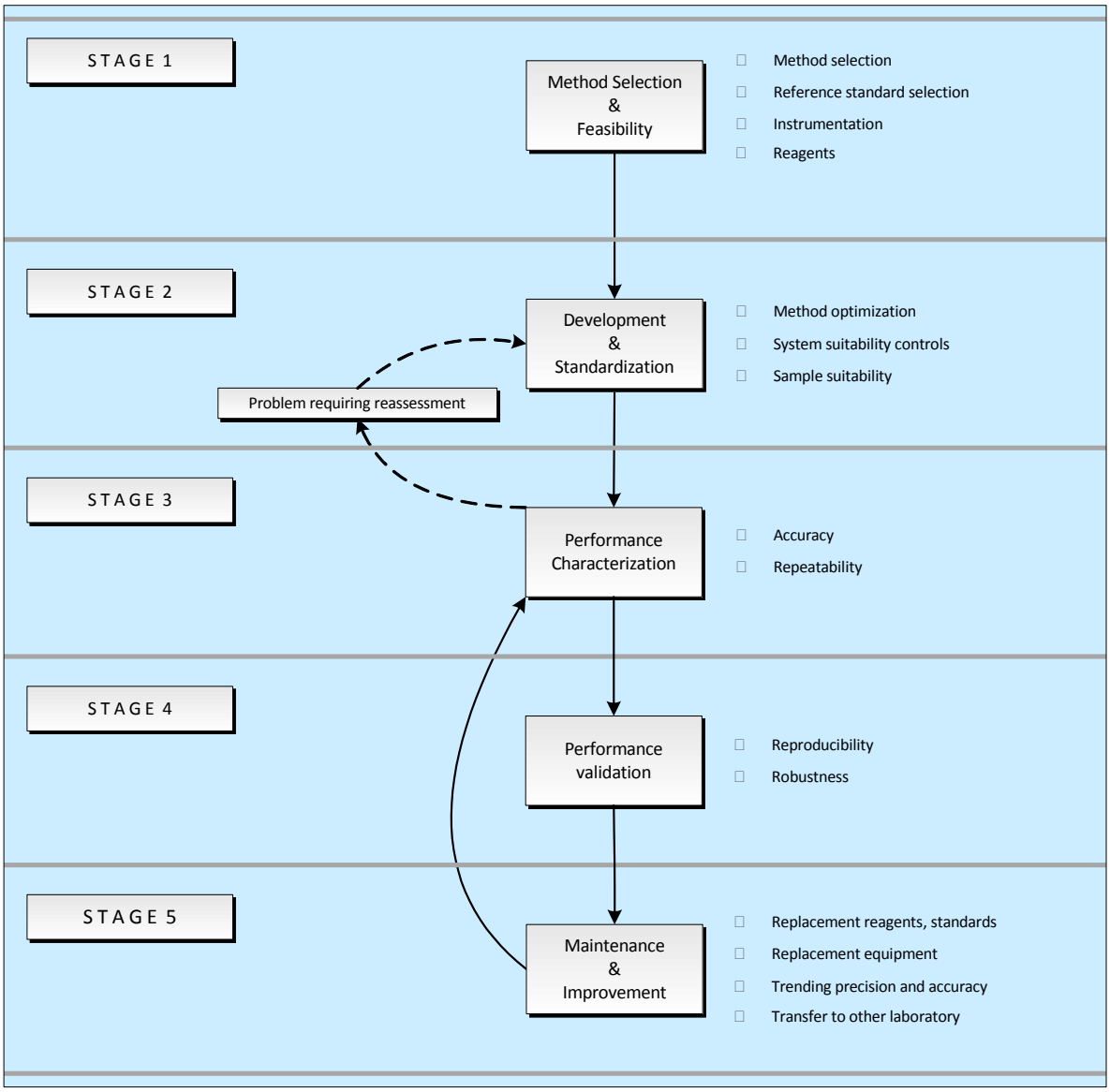

Figure 2: Five typical stages of an assay validation process.

In Figure 2, the five stages are depicted to which will be referred to for the remaining text in this report.

Stage 1: Method selection and feasibility

Stage 2: Development and standardization

Stage 3: Performance characterization

Stage 4: Performance validation

Stage 5: Maintenance and improvement

2.4.2 Stage 1. Method selection and feasibility

a) Method selection

The choice of the method is related to the purpose and the required output of the experiment and is therefore an important aspect of the validation process. Three examples below elucidate this concept.

Besides that the assay should be fit for its purpose from a scientific perspective, other factors can influence the suitability of a method as well. Logistical and operational limitations like running costs, equipment availability, reagent availability, and biosecurity, are examples of such factors. Often the suitability is limited by the lead time of an assay, e.g. a virus culture method taking 14 days may not provide the results in time, while the ELISA test, providing results in one day, can be appropriate.

Example 1: If the absence of infectious (wild-type/replication-competent) virus particles in a cell culture used for the production of non-replicating (defective) recombinant virus must be demonstrated, an infectivity assay is a relevant assay, whereas Polymerase Chain Reaction (PCR) and Western blot assay are not suitable. Both these latter methods may also detect defective viral particles and hence are likely to produce false positive results. The PCR and Western blotting assays are therefore not specific enough for this purpose.

Example 2: To demonstrate the absence of 2 or more virus particles/ml in a cell culture fluid, an assay should be used that reliably can demonstrate at least 2 virus particles/ml in this particular type of cell culture fluid. A robust assay that can only detect 100 or more virus particles/ml, is not suitable due to lack of sensitivity and consequently may give false negative results.

Example 3: If the presence of less than 1000 viruses/ml in a particular product is acceptable, the assay that can reliably detect 2 viruses/ml and the assay that is validated to detect 100 viruses/ml in this product, could both be suitable assays. This does not necessarily mean that an assay detecting 2 viruses/ml is a *better* method than the method that can 'only' detect 100 viruses/ml. If the method with a higher limit of detection (100 viruses/ml) has a better reproducibility and is more robust than the assay with the lower limit of detection (2 viruses/ml), the former assay (100 viruses/ml) may be preferred.

b) Feasibility

In the feasibility phase the required resources and the assay format are investigated and selected. Reagents like antibodies, primers, cell banks, virus seeds and reagents are collected, properly prepared and aliquoted. Control materials –in the best case reference materials– are obtained, aliquoted and certified. Instrumentation is installed and the methodology is evaluated. Feasibility studies are performed to investigate if the assay (system) is actually working.

2.4.3 Stage 2. Assay development and standardization.

Experiments are conducted to find the optimum conditions for parameters like temperature, duration, equipment settings, reagents, etc. For instance 'checkerboard' tests are performed to test different concentrations against a feasible choice of other parameters to find the optimal concentration. The specificity of the method is demonstrated and the limit of detection is assessed.

In this stage also the *system suitability* controls are developed. All measurements susceptible to variations in analytical conditions should be suitably controlled. A list of assay controls should therefore be established in combination with the criteria they should fulfil. An assay is considered valid –and may only be used to generate data of an unknown sample– if the criteria for the specific assay run are met. System suitability does not only cover the actual test, but may also include the preparation of the samples as well as control samples.

The verification process, whereby it is established that the conditions are suitable for a specific type of sample, is called *sample suitability*. Whereas in the system suitability process the functionality of the method or equipment is evaluated by means of assay controls ('does the test/equipment itself properly work?'), in the sample suitability process the functionality of the method is evaluated in relation to the product or sample to be tested ('is the test suitable for this particular type of sample?').

An example of a verification process is given by e.g. testing for infectious virus in faeces samples. Although an assay may be suitable to detect a particular virus, the stool sample can be too toxic for the cells used in the infectivity assay to serve as an appropriate readout system resulting in consistently non-valid assay results. Or, in case of a PCR assay, a sample can not be processed properly to isolate genetic material of sufficient quantity and quality. The test and/or sample preparation may therefore require modifications to overcome such limitations.

2.4.4 Stage 3. Assay performance characterization

Once the method is standardized and system suitability tests are in place, the assay is ready for the first qualification experiments, in which validation parameters are established, using standard or reference samples of virus, antigen, plasmid etc. (see Chapter 3 for the definition of validation parameters). Replicate assay runs are performed and the results provide an indication of the precision and accuracy of the assay. The results should meet the preset criteria that are chosen based on the preliminary data from the method development and –taking in account the purpose of the assay– in accordance with the requirements. If the results comply with the requirements, the assay is ready for the next validation stage. If not, the assay goes back to the previous stage for additional development and standardization.

2.4.5 Stage 4. Assay performance validation

Once the method has been successfully developed, no further modifications are required and assay characteristics have been established, the assay is ready for the performance validation. In this stage the performance of the assay is demonstrated, showing that the assay is reproducible when it is performed e.g. on different days, by different technicians, with the range of samples for which the assay is intended, etc. Also small fluctuations in assay conditions, that could occur under normal situations, are mimicked and the results should demonstrate that the assay is robust.

Guidance on selection of the parameters to be validated can be found in relevant literature e.g. the EURACHEM guide on method validation [3], Annex 5 of the WHO technical guide TRS823 [7], Chapter 15 of the WHO GMP Validation guide [6], the laboratory procedure of the Office of Regulatory Affairs (ORA-labs) of the FDA [12], The World Organization for Animal Health (OIE standards) [14, 15], the FDA Guidance on Bioanalytical Method Validation [2], or the draft EMEA Guideline on Validation of Bioanalytical Methods [27]. Chapter 3 provides a comprehensive explanation of all possible validation parameters.

If the assay was developed as an alternative to an existing, already validated method, the performance may be compared with the old method. The existing method can then be used as a formal reference method for verification of the results found with the new assay (equivalence study). An example is a PCR assay that is developed to replace a conventional culture technique. The accuracy and suitability of both methods may be compared using a panel of identical field samples.

2.4.6 Stage 5. Maintenance and improvement

The validation process does not end once the assay has been introduced. The assay is to be consistently monitored for precision and accuracy, e.g. by trending of the results obtained with the reference standard over time. Proficiency testing is another tool to monitor the performance of an assay. An example is a ring test, where a panel of samples is tested simultaneously by different laboratories using the same method.

If the results show a trend towards either better or poorer performance, an investigation to find the reason of the event should be performed, and implementation of corrective measures are required. After implementation of the appropriate corrective measures, several validation tests may be required to demonstrate the original test performance level is obtained again.

Typically, after modification, the assay goes back to Stage 3 of the process and reconsideration of the assay performance characteristics testing should be done. In the next section 'method verification', more is described about the validation requirements after trouble-shooting and/or improvement of a method.

2.5 Method verification

Once a test has been validated, small changes in the procedure may occur over time. For example reagents or reference standards are depleted and need replacement by other batches. Equipment may brake down or new software is installed. New technicians are trained for performance of the method or a method that has been developed and validated, may be transferred to another laboratory.

When a method developed in Laboratory A is transferred to Laboratory B, after implementation a demonstration will be required that the test is running as good as it was validated in Laboratory A (even if both laboratories are located within the same institute).

If problems arise during the verification process, these problems need to be assessed, and the assay may even need to go back to Stage 2 of the validation process for additional development and standardization. More extensive validation testing may then be required and qualification tests that are normally done only in Stages 2 to 4, may need to be repeated.

Another situation can occur when a method that has not been in use for a while is started up again for a series of experiments. In all these aforementioned examples and situations, it may not be necessary to perform or repeat the entire assay validation, but to perform merely a method verification using the criteria that were previously defined in Stage 3 of the validation process (see the flow diagram in Figure 2 for the five stages of validation).

Method verification is also typically applicable for commercial assays ('kits') that already have been validated by the manufacturer. Obviously not all validation experiments need to be repeated, but the user will need to verify that the assay in the laboratory is running according to the manufacturer's specifications. In such cases, a method verification process, usually existing of a limited number of experiments, is performed to demonstrate that previously determined parameters can be repeated.

Another example of method verification is where new sample types are applied to a validated standard method. E.g. an assay is in use for the detection of impurities in human serum samples and is now to be used for the detection of impurities in bovine serum. Experiments need to be conducted with bovine sera to demonstrate that the positive controls –when spiked in bovine serum– show the expected recovery and that negative controls do not react. If the results meet the pre-set criteria, the assay is considered suitable for use.

2.6 Concluding remarks

In this Chapter it has become clear that (the process of) validation is a series of activities rather than a single occurring event. The validation process starts with method selection and its development, and once the assay is validated, the performance of the assay is to be continuously monitored (Figure 2).

For new methods, more extensive studies are required than for already existing (validated) assays. Once the test is in use and modifications are deemed necessary, a limited number of experiments may be sufficient to demonstrate that the assay's original test performance level is achieved and functional parameters meet the acceptance criteria.

The extent of validation required is determined by the type of assay, while the type of assay as such is determined by the purpose of its intended use. Quantitative measurements are required when the actual concentration of an analyte needs to be determined. In case the assay needs to demonstrate 'absence' or 'identity', testing of fewer validation parameters is usually sufficient.

3 ASSAY VALIDATION PARAMETERS

In literature, validation parameters have been identified to characterize the analytical performance of an assay. These definitions are all based on the concepts established in the ICH Guideline on Validation of Analytical Procedures Q2(R1) [1]. This guideline is widely accepted and applied by the scientific community. Other useful literature on general aspects of assay validation is -for instance- the EURACHEM guide [3], The FDA Guidance on Bioanalytical Method Validation (BMV) [2], the WHO guidelines [6, 10] and the draft EMEA Guideline on Validation on Bioanalytical Methods [27]. All these documents are available on internet, see Chapter 10 for references.

In this Chapter 3, the general definitions of assay validation parameters are outlined and, further explanation is provided in order to make them easier accessible to the reader who is less experienced or not familiar at all with the concept of validation. For this purpose, random examples and figures are introduced that do not originate from guidelines, but are solely meant to assist the reader in getting a more comprehensive picture of the topic.

3.1 Accuracy

The accuracy of the procedure is defined as the closeness of the results obtained by the procedure to the true value. It expresses the closeness of agreement between the value which is accepted either as a conventional true value (in house standard) or an accepted reference value (international standard) and/or the found value (mean value) obtained by applying the test procedure a number of times. Accuracy is sometimes also called 'trueness'. In Example 1 the accuracy of three 'counting machines' is calculated.

Example 1. Accuracy of counting machines

The accuracy of three counting machines is determined. On each machine three experiments are performed using a reference box containing an accepted reference value of exactly 1000 units.

- Machine 1

  counted 993, 994 and 998 units respectively.

  The mean value is 995.

  The accuracy is 99.5%, namely (995/1000) X 100%.

- Machine 2

  counted 995, 995 and 995 units respectively.

  The mean value is 995.

  The accuracy is 99.5%.

- Machine 3

  counted 1003, 954 and 1028 units respectively.

  The mean value is 995.

  The accuracy, again, is 99.5%.

Is any of the machines acceptable? This will be addressed further below.

In Dutch language 'accuracy' ('*juistheid*') is often confused with precision ('*precisie*' or '*nauwkeurigheid*'). Many people unfamiliar with assay validation confuse the terms precision and accuracy. Also, Dutch language dictionaries do not clearly discriminate between these terms. The difference between accuracy and precision is visualized in Figure 3.

The accuracy of a testing method can be determined by applying the procedure to samples of material that have been prepared with quantitative accuracy. Samples can be 'spiked' with an exact amount of the analyte and the recovery is subsequently determined in the assay ('spike recovery'), see Example 2.

It is recommended to test quantities throughout the range of the assay, including samples containing 10% above and below the expected range of values. It is also possible to determine accuracy by replicate analysis of samples containing known amounts of the analyte (a control panel). The BMV guidance <2>, for instance, recommends including a minimum of three concentrations in the range of the expected concentrations. The mean value should be within 15% of the actual value expected, except at the Lower Limit of Quantitation (LLOQ), where it should not deviate by more than 20%. The deviation from the mean from the true value serves as the measure of accuracy.

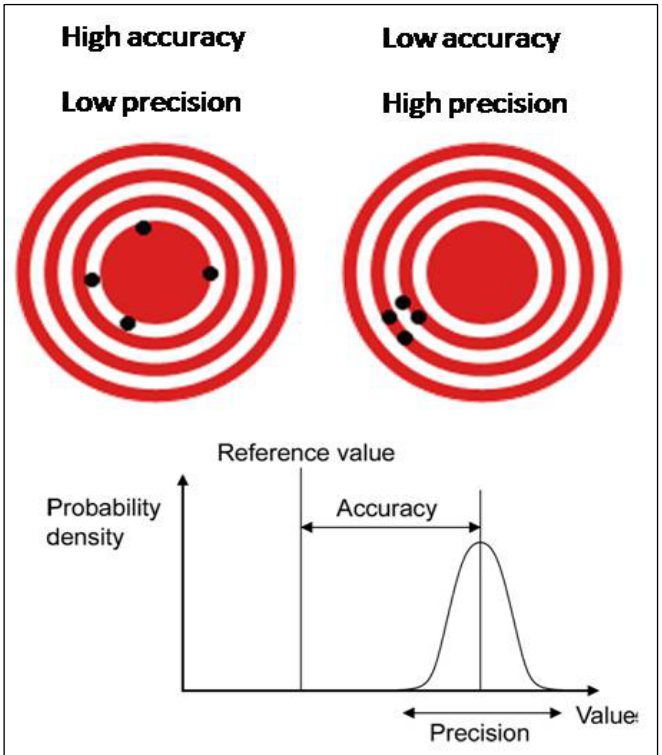

Figure 3. Relationship between accuracy and precision depicted as bullets and depicted in a hypothetical graph

In bioassays, accuracy is often determined in spike and recovery studies: a known sample amount is added to the excipients (=the product without the active ingredient) and the actual drug value is compared to the value found by the assay. The accuracy is expressed as the bias or the % error between the observed value and the true value: (assay value/actual value) x 100%). For biological products, accuracy determination is often not possible, because pure standards –like international reference standards– may not be available. For such products, a comparison can be made to a reference product

which is run in parallel in the same assay. Acceptable results are based on specifications for the ratio of the sample value to the reference value [6].

If no reference standards are available, accuracy may also be determined by comparing the results with those obtained using an alternative procedure that has already been validated [7]. For example the virus titre determined with a Q-PCR assay (alternative assay) is compared with the titre obtained by virus titration in cell culture (conventional assay). It should be mentioned that conventional (bio)analytical assay experts are very reluctant to adopt this approach of performing accuracy determination, because of lack of a direct link to the absolute value of a standard. Sometimes however, in biological assay validation, this is the only available option.

Although accuracy assessment is an important parameter in the validation of an assay, depending on the purpose of the method, determination of this parameter may not be required. If the method has shown satisfactory sensitivity and specificity, in some cases this can already be sufficient.

The determination of the accuracy in an analytical method can only be established when other relevant parameters of the assay like precision specificity, and linearity, have already been determined.

Example 2. Accuracy determination in an immunoassay.

Procedure:

Use 3 spiking concentrations.

Prepare 2 samples of each concentration of a reference in the excipient solution.

Test these 6 samples in triplicate in one run (18 samples).

Measure the expected vs. de average measured value (based on these 18 values) and calculate the % recovery.

The % recovery is the accuracy.

3.2 Precision; repeatability, inter-assay variation, reproducibility

3.2.1 Precision

Intuitively, one may not be satisfied with any of the counting machines in Example 1, but if one had to choose, most probably machine 2 would be selected. This is because people tend to prefer the precision of this machine.

The precision of an analytical procedure is the degree of agreement among individual test results. It describes the closeness of individual measures of an analyte when the procedure is applied repeatedly to multiple aliquots of a single, homogenous volume of the biological matrix. Precision is measured by the scatter of individual results from the mean when the complete procedure is applied repeatedly to separate, identical samples drawn from the same homogenous batch of material.

Precision is usually expressed as the standard deviation (SD) or as the coefficient of variation (CV) of a series of measurements [1]. The CV is the standard deviation of the assay values, divided by the concentration of the analyte. It is also called 'relative standard deviation' (RSD). The standard deviation, relative standards deviation, and confidence intervals are typically parameters to report when determining the precision.

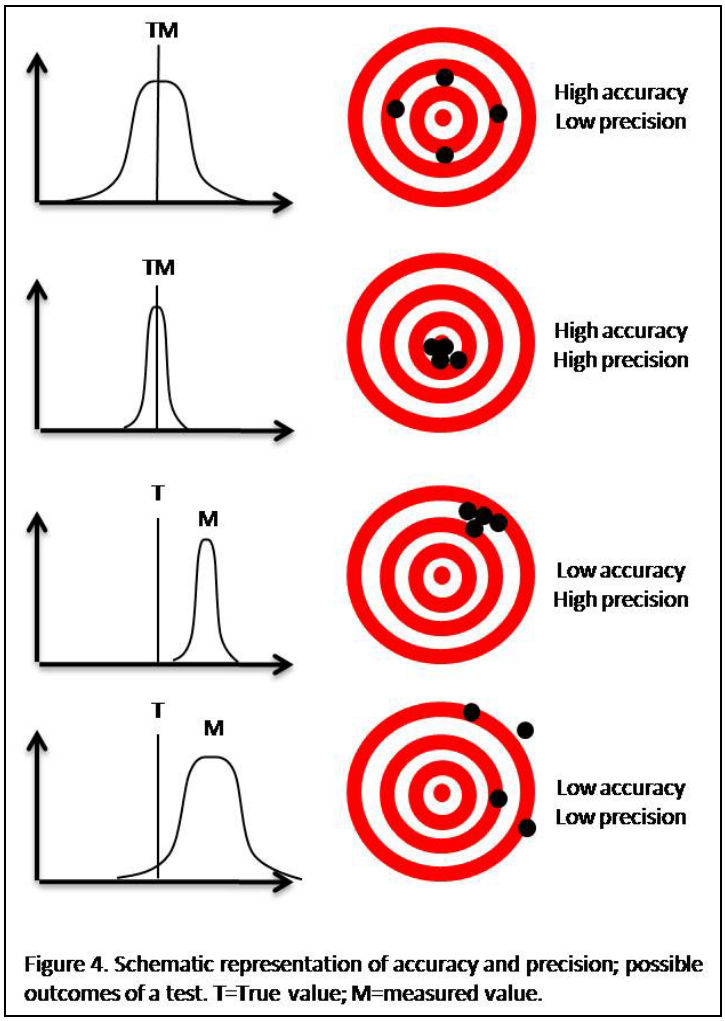

Figure 4. Schematic representation of accuracy and precision; possible outcomes of a test. T=True value; M=measured value.

In bioanalytical method validation, precision is measured as a minimum of five determinations per concentration and a minimum of three concentrations in the range of the expected concentration is recommended. The precision is determined for each concentration level and should not be more than 15% of the coefficient of variation (CV), except for the lower limit of Quantitation (LLOQ) where it should not exceed more than 20% of the CV [2].

When the counting machines from Example 1 -in the previous section- would be applied to microbial assay validation, the acceptation criteria may be different than for other applications. You would probably be very satisfied if you could buy a machine that is able to count 100 bacteria with an accuracy of 99.5% and a precision of <2%. All three machines match these criteria.

In Example 3 the counting machines are used again for demonstrating how to find a suitable machine for its purpose.

Several types of precision can be measured:

- Repeatability (Dutch: '*herhaalbaarheid*')

  The precision of multiple determinations of a single sample in a single test run.

- Inter-assay variation (intermediate precision).

  The precision for multiple determinations of a single sample, controls and reagents analysed in several assay runs in the same laboratory.

- Reproducibility.
  The precision between laboratories, usually determined in collaborative studies [6].

Definitions and aspects of these three different types of precision are discussed in the next sections.

Example 3. Is the counting machine acceptable?

Consider the same counting machines and their three measurements with '1000 units' as in Example 1.

Machine 1: (993, 994, 998) Mean =995, SD[(1)]= 2.65 CV[(2)]= 0.27

Machine 2 (995, 995, 995) Mean =995, SD= 0.00 CV= 0.00

Machine 3 (1003, 954, 1024) Mean =995, SD= 15.39 CV= 1.53

Is this acceptable?

It all depends on the purpose of the machine!

	Mean	Accuracy[(3)]	Precision(CV)[(2)]
Machine 1	995	99.5%	0.27
Machine 2	995	99.5%	0.00
Machine 3	995	99.5%	1.53

[(1)]SD = standard deviation of the three measurements

[(2)]CV = SD / mean

[(3)]Accuracy = (mean value / actual value) x 100%. The actual value = 1000

If the machine is a coin counting machine, used by a saving bank with the purpose of counting your money, of course this will not be acceptable. You would expect an accuracy of 100.0 % and a precision of 0.00. Alternatively, the validation criteria could be formulated as: 'All three tests should provide exactly a count of 1000'.

In microbial assay validation, the criteria may be different. You would probably be very excited if you could buy a machine that is able to count 1000 bacteria with an accuracy of 99.5% and a precision of < 2%. All three machines comply with these criteria.

3.2.2 Repeatability

The repeatability is defined as the precision of the procedure when repeated by the same technician under the same set of conditions (identical reagents, same equipment, experimental settings, laboratory, etc) and within a short interval of time –usually the same day. The repeatability of the procedure is assessed by carrying out complete, separate determinations on separate, identical samples of the same homogenous batch of material, and thus provides measure of the precision of the procedure under normal operating conditions.

Repeatability can be assessed, for instance, by testing a minimum of 9 determinations over a minimum of 3 concentration levels covering the specified range; e.g. 3

concentrations with 3 replicates each of the total analytical procedure. Alternatively, a minimum of 6 determinations at 100% of the test concentration can be performed. See Example 4 for a typical experiment on the measurement of repeatability.

Repeatability is also referred to as 'intra-assay precision', 'intra-assay variation' or 'within-assay variation' [1, 6].

Example 4. Repeatability. Determination of the precision of a homogenous sample at various points of the curve in a single assay.

Take the homogenous sample and prepare three dilutions of it (high, mid, low concentrations within the range)

Test 10 replicates of each dilution of the sample (a total of 30 measurements) using the assay with the standard curve in parallel.

Calculate the average and standard deviation for each point on the standard curve

Calculate the CV = (SD:mean)X100% for each point on the curve.

Calculate the CV for all points on the curve.

Calculate the 95% confidence level.

The validation is valid if the precision meets the preset specification.

3.2.3 Inter-assay variation

The inter-assay variation is also called 'inter-assay precision', 'intermediate precision' or 'intermediate variation'. It considers the random error introduced by factors like specific equipment, technicians, laboratories, days, and so forth. This is in contrast to systemic or systematic error (=bias), which is determined in the accuracy of the method.

The extent to which the inter-assay variation should be established, depends on the circumstances. Often, not all factors need to be investigated separately. To assess the extent of inter-assay variation the use of an experimental design matrix is encouraged, meaning that different parameters can be addressed in combined experiments (see Example 5). In addition to the more common parameters (days, technicians, materials), typical examples of intermediate precision are the establishment of virus titres in early and end-of life cell culture stages, the effects of minimum-maximum duration of incubation periods and testing at minimum-maximum temperatures (in case a broad temperature range is given).

N.B. There is some dissimilarity on the term intermediate precision. In some literature, intermediate precision is considered as part of repeatability, while other authors consider intermediate precision tests as part of reproducibility (see 3.2.4). Occasionally, even both interpretations are applied in one article [7]. Although semantics are sometimes confusing, when establishing the validation policy, it should be verified whether intermediate precision is required. Intermediate precision pertains to the establishment of the effects of random events on the precision of the analytical procedure, considering random error introduced by factors such as specific equipment, technicians, laboratories, and days. It is not meant to include the systematic error (bias).

Precision in biological assays

The 95% confidence limits for results of within-assay variation (intra-assay precision) normally should be on the order of ±0.5 ^{10}log of the mean. Between-assay variation (inter-assay precision) can be monitored by the inclusion of a reference preparation. To be acceptable, the estimate of whose potency should be within approximately 0.5 ^{10}log of the established mean estimate, in the laboratory. Assays with lower precision may be acceptable with appropriate justification.

Example 5. Determination of inter-assay precision (CV) of a homologous sample at various points on the standard curve.

Take the homogenous sample and prepare three dilutions of it within the range (high, mid, low concentrations: C_h C_m C_l respectively).

Test triplicate aliquots of each dilution of the sample in three different assay runs (run 1, run 2, run 3). Overall there are 9 measurements per assay and 27 measurements in total.

Include the following variations:

Day-to-day variations (Monday, Tuesday, Wednesday)

Technician-to-technician variation (David and Mary)

Lot-to-lot variations of assay materials (kit 1, kit 2, kit 3)

The test matrix is provided in the table below:

Assay run	1		2		3		
Day	Monday		Tuesday		Wednesday		
Technician	David		Mary		David		
Assay material	Kit 1		Kit 2		Kit 2		
	Acc.	**Prec.**	**Acc.**	**Prec.**	**Acc.**	**Prec.**	**Inter-assay precision**
Result for C_h	$A_{Ch,1}$	$CV_{Ch,1}$	$A_{Ch,2}$	$CV_{Ch,2}$	$A_{Ch,3}$	$CV_{Ch,3}$	$CV_{Ch1,2,3}$
Result for C_m	$A_{Cm,1}$	$CV_{Cm,1}$	$A_{Cm,2}$	$CV_{Cm,2}$	$A_{Cm,3}$	$CV_{Cm,3}$	$CV_{Cm1,2,3}$
Result for C_l	$A_{Cl,1}$	$CV_{Cl,1}$	$A_{Cl,2}$	$CV_{Cl,2}$	$A_{Cl,3}$	$CV_{Cl,3}$	$CV_{Cl1,2,3}$

Calculate the mean and standard deviation (SD) for each point on the standard curve for each individual test and determine the accuracy (A) and precision (CV) for each of the 9 points. The accuracy and precision results from all 9 test results should meet the preset specifications.

Provided all validation assay runs are valid, the inter-assay precision for each of the three dilutions (C_h C_m C_l) can be calculated for each point on the curve between the assay runs as follows.

CV = (SD/mean) X 100%.

Remark:

If an assay run is not valid, the result is skipped completely, and the run is repeated. If more than one run is invalid, this indicates something is not in control. The validation should be postponed and the reason for the failure should be investigated. The assay may not be robust enough and may need further development or optimisation. After the method is optimised, the validation is repeated including three new runs.

3.2.4 Reproducibility

Reproducibility is defined as the inter-laboratory variation, which stands for how well an assay can be reproduced elsewhere without loosing its performance characteristics. Reproducibility is assessed when e.g. a laboratory implements a new technique that has been published in literature or when commercial tests are used for the first time in a laboratory. It may also be assessed in collaborative studies between laboratories −usually applied to standardization of methodology [1, 6, 7].

N.B. Reproducibility is not the same as the comparison of results obtained from two different methods (regardless of whether the results are derived from the same laboratory or from different laboratories). Determining reproducibility with different methods is scientifically not sound.

3.3 Linearity and range of the measurement

The linearity of an analytical procedure is the ability (within a given range) to produce results that are directly proportional to the concentration (amount) of analyte in the sample [1]. Linearity can be measured as the slope of the regression line and its variance (see Figure 5 for a graphical visualization). Alternatively, linearity can be measured as the coefficient of determination and correlation coefficient [6]. In some cases, to obtain linearity between assays and sample concentrations, the test data need to be subjected to a mathematical transformation (e.g. logarithm) prior to the regression analysis. For the establishment of linearity, a minimum of five concentrations is usually recommended.

In some assays, e.g. immunoassays, the relationship between response and concentration is not linear, not even after mathematical transformation. In this case, standardization may be provided by means of a calibration curve [7]. The calibration curve is the relationship between response and known concentrations of the analyte. A calibration curve should be prepared in the same biological matrix as the sample or separately controlled by means of spike samples. The number of standards used in constructing the calibrate curve depends on the anticipated range of analytical values and the nature of the analyte/response relationship.

Example 6. Determination of the range.

Prepare 8 sample dilutions across the claimed range.

Test each dilution in triplicate in three runs.

Calculate for every run, for each of the 8 sample dilutions, the average value, standard deviation and CV (=precision per sample dilution).

Calculate for every run the %recovery for each of the 8 sample dilutions, based on the found average concentration and the actual concentration (=accuracy).

The range is determined by the highest and lowest concentration with satisfactory accuracy and precision in all three runs.

Alternative calculation method:

Analyse each set of dilutions within the assay run as a linear curve and calculate the coefficient of correlation. The range is defined by the highest and lowest concentration with an acceptable coefficient of correlation in all three runs.

The determination of the linearity will identify the range of the analytical assay. The range of the procedure is the interval between the upper and the lower levels (the highest and lowest concentrations) of analyte that can be determined, with acceptable precision, accuracy and linearity [1, 6]. These characteristics are determined by application of the procedure to a series of samples having analyte concentrations spanning the claimed range of the procedure [7]. Example 6 shows how the range can be calculated.

3.4 Specificity

The specificity of a method is its ability to measure accurately and specifically the analyte in the presence of components that are expected to be present in the sample matrix. A method can be 'specific' for one or more components of a mixture, but 'non-specific' for another component. Selectivity and recovery are parameters that are covered by the term specificity.

Selectivity is a measure of the degree of interference (or absence thereof) in the analysis of complex sample mixtures. Selectivity may be expressed in terms of the bias of the assay results obtained when the procedure is applied to the analyte in the presence of expected levels of other components, compared to the results obtained on the same analyte without added substances. Selectivity, like accuracy, is expressed as the bias or the % error between the measured and the known value [6].

The recovery of an analyte in an assay is the detector response obtained from an amount of the analyte added to and extracted from the biological matrix, compared to the detector response obtained for the true concentration of the pure authentic standard. Recovery pertains to the extraction efficiency of an analytical method within the limits of variability. In quantitative assays, recovery experiments are performed by comparing the analytical results for extracted samples at three concentrations (low, medium and high) with non-extracted standards that represent a 100% recovery [2].

For a suitable assay, recovery of the analyte need not to be 100%, but the extent of recovery of an analyte and of the internal standard should be consistent, precise and reproducible. Lack of specificity of an individual analytical procedure may be compensated by other supporting analytical procedures.

3.5 Limit of detection (detection limit)

The detection limit (DL) or limit of detection (LOD) is the lowest level of analyte that can be detected, but not necessarily quantitated as an exact value [1]. The DL is usually expressed in terms of a concentration of analyte in the sample. Several approaches for establishing the DL are possible, depending on whether the procedure is non-instrumental or instrumental.

In case of non-instrumental procedures, like visual evaluation (e.g. scoring for CPE or evaluation of an SDS-PAGE), the DL is determined by analysis of samples with known, diminishing concentrations of analyte and by establishing the minimum level at which the analyte can be reliably detected. When the final measurement is based on instrumental reading (e.g. ELISA), the background response will have to be taken in account (the signal-to-noise characteristics of the responses observed). For procedures which exhibit baseline noise, the detection limit can be based on the signal-to-noise ratio. Determination of the signal-to-noise ratio is performed by comparing measured signals from samples with known low concentration of analyte with those of blank samples and by establishing the minimum concentration at which the analyte can be reliably detected. A signal-to-noise ratio of 3:1 is generally considered acceptable for estimating the detection limit.

In a virus titration assay, a DL of 2 Plaque Forming Units (PFU)/ml means that the assay can demonstrate the presence of the virus, when the concentration is at least 2 PFU/ml. It does not mean that the titre can be accurately measured at this concentration: the LOQ (limit of quantitation) is much higher, for example it can be e.g. up to 20 PFU/ml for the particular assay.

Another method to determine the DL is based on the standard deviation of the response and the slope of the calibration curve.

DL= (3.3 x σ) / Slope of the calibration curve.

The slope may be estimated from the calibration curve of the analyte. The estimate of the standard deviation (σ) can be carried out in a variety of ways, e.g. based on an appropriate number blank samples (e.g. 10, [3]), based on standard deviation of the regression line, or the Y-intercept of control samples in the range of the detection limit [1].

Example 7. Determination of the detection limit (DL) of an immunoassay.

For a bioassay the DL is the minimum concentration of a substance that generates a consistent response greater than the background of the test. Responses of 2 tot 3 times the standard deviation, are usually satisfactory limits. For an immunoassay, when Optical Density (OD) values are measured, often the DL is three standard deviations above the background. The DL can be determined as follows:

Prepare the 'sample': a standard concentration (reference standard) in the appropriate solution. The concentration should fall within the established range.

Prepare the 'blank solution': a zero concentration in the solution.

Perform the immunoassay according to the standard procedure three times (=three runs). Test the blank solution and the sample in duplicate in every test.

Confirm for each run that the criteria for a valid assay run are met (the precision and the calibration curve are in accordance with the criteria).

Calculate the average and standard deviation of the OD-value of the blank solution.

Calculate the average of the OD-value of the sample.

Calculate the 'OD per concentration' by dividing the average OD of the sample by the concentration of the reference standard.

Calculate the DL as: (3 x σ of the blank solution) / (OD per concentration).

N.B. In case the estimated value for the DL is obtained by calculation or extrapolation, it is expected that these estimates are actually confirmed by independent analysis of a suitable number of samples near the DL.

3.6 Limit of quantitation

The 'Limit of Quantitation' (LQ) is provided by the lowest and highest concentrations of analyte in a sample that can be determined with an acceptable degree of accuracy and precision. Like the DL, the LQ is usually expressed in terms of a concentration of analyte per volume. It can further be specified as 'Lower Limit of Quantitation' (LLOQ) and 'Higher Limit of Quantitation' (HLOQ) and in fact these values determine the range of the assay. Often, when the LQ is given, actually the LLOQ is meant.

Dutch synonyms of LQ are '*kwantificeringslimiet'* and '*bepalingslimiet'*. The latter term is often confused with the Dutch word '*detectielimiet*'.

When the final measurement is based on an instrumental reading (e.g. ELISA), the background response will have to be taken in account (the signal-to-noise characteristics of the responses observed). For procedures that exhibit baseline noise, the LLOQ can be based on signal-to-noise ratio. Determination of the signal-to-noise ratio is performed by comparing measured signals from samples with known, low concentration of analyte with those of blank samples and by establishing the minimum concentration at which the analyte can be quantified to an acceptable degree of certainty. A typical signal-to-noise ratio in determination of the LLOQ is 10:1.

LLOQ= (10 x σ) / Slope of the calibration curve.

The slope can be calculated as described for the determination of the DL above. Figure 5 depicts the relation between DL, LLOQ, HLOQ as well as the slope and the linear range.

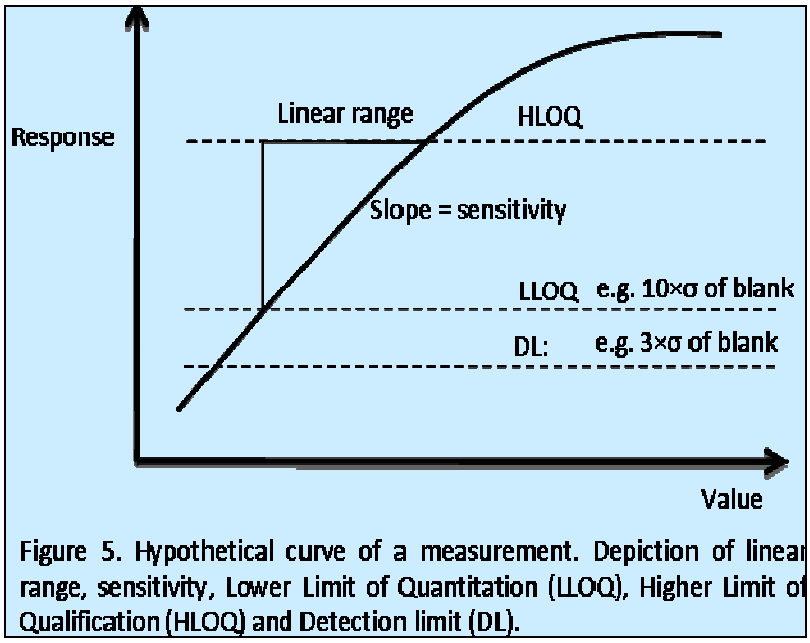

Figure 5. Hypothetical curve of a measurement. Depiction of linear range, sensitivity, Lower Limit of Quantitation (LLOQ), Higher Limit of Qualification (HLOQ) and Detection limit (DL).

3.7 Robustness

The robustness of an analytical procedure is a measure of its capacity to remain unaffected by small, but deliberate variations in method parameters. It provides an indication of its reliability during normal usage [1, 5, 7, 10]. Variation in incubation times, minor variations in pH of a reagent, varying humidity, assay temperatures, mobile phase compositions, flow rates, injection volumes and primer concentrations are typical examples that can influence the result of an assay. Under each of these conditions, the accuracy and precision of other assay parameters can be measured to see what variations in the assay conditions can be tolerated and the extent to which the test is reproducible [6].

Where there is a need to further improve the method, improvements can probably be achieved by concentrating on those parts of the method known to be critical. Robustness is normally evaluated during the method development phase (Stage 2), typically by the originating laboratory, before collaborating with other laboratories. Robustness tests are normally applied to investigate the effect on either precision or accuracy [3]. They generally result in a set of system suitability parameters that should be used to ensure the validity of the analytical method during regular use. Some additional robustness tests are performed in Stage 4 (assay performance validation, see Figure 2) to demonstrate that small fluctuations of assay conditions do not affect the validity of the assay. Robustness tests are not standard included during method verification. In a properly

validated commercial assay the robustness aspects should have been addressed by the manufacturer.

A note on the term 'Sensitivity'

The term sensitivity is not used unambiguously in relation to assay validation. In official guidelines, sensitivity is defined as 'the capacity of the test procedure to record small variations in concentration'—think for instance the slope of the calibration curve. Although it is recommended to avoid a more general use of the term sensitivity (see references [3, 7], frequently the term sensitivity is confused with other validation characteristics like precision, limit of quantitation, and —more often— limit of detection. In scientific papers the term 'sensitivity' is often used to refer to detection limit. In this respect an assay that can detect 100 viruses/ml is considered more sensitive than an assay with a detection limit of 10^5 viruses/ml.

3.8 Application of validation parameters

In Chapter 2 the method validation process is outlined and the different assay types are described: Identity tests, Quantitative limit tests, Limit tests and Assays (see Section 2.3, Figure 1). Now that also the different validation parameters have been explained in this Chapter, the different assay types and the requirements of assay validation for the specific types of tests are listed again. In Table 2 the recommended validation parameters versus the type of analytical procedure as defined in the ICHQ2(R1) guideline [1] are presented. For convenience and clarity, in this Section of the report the ciphers I, II, III and IV are introduced to identify the different types of analytical assays. In Part II of this report practical examples of assay categories II, III and IV are described.

Parameter	**Type of Analytical Procedure**			
	I*	**II**	**III**	**IV**
	Identification	**Testing for Impurities**		**Assay**
		Quantitative	**Limit**	
Accuracy *('Juistheid')*	-	+	-	+
Precision *('Herhaalbaarheid')*				
Repeatability	-	+	-	+
Intermediate	-	+(1)	-	+(1)
Specificity(2)	+	+	+	+
Detection Limit	-	-(3)	+	-
Quantitation Limit	-	+	-	-
Linearity	-	+	-	+
Range	-	+	-	+

-: not normally evaluated +: normally evaluated

(1): In case where reproducibility test has been performed, no intermediate precision is needed.

(2): lack of specificity of one analytical procedure could be compensated by other supporting analytical procedures(s).

(3): may be needed in some cases.

*Note: The ciphers I, II, III and IV are not derived from ICHQ2(R1), but introduced by the author.

Table 2. Recommended Validation parameters versus the analytical procedure as recommended by ICH Q2(R1).

Robustness is not specifically listed as a parameter, it is however expected that this is included in the appropriate stages of the method validation process [1].

Most pharmaceutical companies write a policy procedure in which they refer to this table and in addition specify the required levels of validation for the different stages of pharmaceutical research and development (e.g. for pre-clinical, clinical and post-licence use). Three different levels of validation are often specified. For assays that are used during the initial development phase, less extensive data for precision and accuracy are required than for assays used to support testing of products for intended clinical trials. For marketed products, extensive validation data are required and more effort is placed on robustness testing to demonstrate the consistent performance of the assay over time, taking into consideration the use of different batches of reagents, equipment variables and storage effects of materials as well as other possible variations.

4 DOCUMENTATION

An integral part of the validation process is appropriate documentation of all results obtained. Formally, if an assay is validated, but the process is not (fully) documented, there is no trace that the assay is performed the way it is validated.

There are four types of documents typically related to assay validation:

- Raw data
- Specifications
- Procedures
- Reports

There exists no consensus for specific requirements for the layout of standard documentation in the laboratory. Every laboratory can establish its own formats and provide recommendations to laboratory personnel and scientists on how to use them. The use of standard templates is highly recommended, as they have shown to be both effective and efficient.

4.1 Raw data

It is obvious that careful documentation of laboratory activities in laboratory log books or standard forms are key elements of good validation practise. In all phases of the validation process, like during routine performance of an assay, all operators should write down all relevant data of experiments accurately and promptly.

Not only assay runs that are valid, but also experiments that were unsuccessful should be documented. Failed experiments should be evaluated and this may help to improve the method and anticipate necessary modifications of the method. Ignoring failed experiments may also mask a potential problem with an assay, which is inefficient and can be frustrating.

The raw data are the original registrations of the activities demonstrating 'who-when-what-with-where' and they are vital for the reconstruction of experiments. They should not be replaced by e.g. summary reports from processed data.

4.2 Specifications

Once the assay is developed and becomes ready for validation, it is important that criteria are specified before the experiments are performed, otherwise the acceptability

of an assay may be driven by other forces than good science. Only validation experiments should be performed for which clear acceptance criteria are specified and documented.

A validation protocol can be used to describe the validation procedures and to document the specifications, so it is a description of the experiments and the criteria that need to be fulfilled. The validation protocol can refer to the standard procedure as described in the next section. Alternatively, a draft standard procedure can be written and included, or attached to, the validation protocol.

Once the validation is completed, the requirements of the system suitability tests, assay controls, materials, equipment, etc. are also fixed. It is useful to include these specifications in the written standard procedure.

4.3 Procedures

Once the assay has been validated, it is crucial that the method is consistently performed as it was validated. The same materials are used, and suitability controls and assay controls are applied with the same specifications as during the validation process.

For this purpose, the validated method is carefully documented in a procedure or a standard protocol. In the industrial setting this is often called a Standard Operating Procedure or 'SOP'. The documentation is not limited to a general description of the assay, but preferably includes all relevant aspects that contribute to the performance of the assay.

An SOP should cover the relevant aspects of the analysis, from the time the sample is collected until the results of the analysis are reported, including data recording and data processing, sample preparation and description of analytical tools such as methods, reagents, equipment, instrumentation, use of positive and negative controls, reference strains, acceptance criteria, calculation, statistical methods, and procedures for verification of results.

A standard procedure may contain visual aids, like photographs, flow charts, figures and may also refer to other procedures.

If the method changes, e.g. after improvement, or when new types of samples are included in the scope of the method, the SOP will be updated and a new version number and issue date are assigned.

The SOP is a useful communication tool when training laboratory personnel. Accurately following the standard procedure will help to prevent drifting away from the validated procedure over time and may prevent time consuming troubleshooting.

The SOP may contain the following (non-exhaustive) list of elements:

- The scope of the assay, in order to specify what the purpose and limitations of the method are, for which sample (matrix) the assay is suitable, which concentration range is applicable, etc.
- A list of reagents and materials, including blanks, and control/reference standards. Specifications on the material grade, and catalogue numbers for reagents
- Details on storage conditions, preparation instructions, expiration of raw materials and prepared reagents
- Description of apparatus and equipment, performance requirements where relevant
- Sample preparation methods and sample storage conditions, as well as expiration periods of prepared samples and controls
- Performance of the method with sufficient details on the sequence of operations, critical parameters (time, temperature), the use of samples and controls, precautionary statements, etc.
- Criteria of assay validity and system suitability tests
- Safety aspects
- Evaluation and calculation of the assay result
- The format of reporting the results.

4.4 Reports

In addition to writing an SOP, it is recommended to document the test conditions together with the results of the validation study in a laboratory journal and to write a validation summary report. The validation report will contain at least a summary of the performance characteristics and the conditions they were obtained under. The raw data (the primary laboratory results) can be included as attachments to the validation report. Alternatively, reference can be made to the laboratory note books and/or a location where the raw data are stored. In addition to, or as part of the validation report, a development report can be generated, to summarize the process of development and standardization. Chapter 7 of the draft Guideline on the Validation of Bioanalytical

Methods from November 2009 [27] contains recommendations for the content of validation study report for the validation of analytical methods used for pharmacokinetic sample analysis. The same recommendations may be useful for analytical methods for other purposes.

Both the development report and the validation report can be very helpful documents in case of trouble shooting, method verification or technology transfer to other laboratories. These documents are particularly useful in case modifications to the method are required. Furthermore, both the development and validation report function as communication tools for quality officers and regulatory bodies and may also be valuable for training purposes.

The validation report is the ultimate documented evidence that an assay is fit for use.

PART II: EXAMPLES OF BIOLOGICAL ASSAY VALIDATION

Three examples of validation of biological assays are worked out in detail in three consecutive Chapters. For each assay, the process is described with the same lay-out, following the five stages of validation as defined in Part I.

Chapter 5	Chapter 6	Chapter 7
TCID50	ELISA	PCR
1. Introduction to the method 2. Validation parameters 3. Stage 1. Selection and feasibility 4. Stage 2. Development and standardization 5. Stage 3. Performance qualification 6. Stage 4. Validation 7. Stage 5. Maintenance and improvement 8. Method verification 9. Points to consider		

The Chapters 5, 6 and 7 each start with an introduction of the method describing the assumed applicability of the assay, followed by the definition of the validation parameters that are relevant for the intended use. Also a section on method verification is included, describing the procedure for a previously validated or commercial assay. The Chapters 5, 6 and 7 are each completed by a section 'Points to consider', containing some general remarks and recommendations.

5 TEST FOR DETECTION OF VIRUS INFECTIVITY ($TCID_{50}$ ASSAY)

5.1 Introduction

5.1.1 General

In this Chapter the Tissue Culture Infectious Dose 50% ($TCID_{50}$) assay, a conventional virus titration method, is used as an example to demonstrate the validation of a method that is used to quantify the infectious virus content of a sample.

In the assay, the starting material containing infectious virus is divided into samples that are increasingly diluted. The diluted samples eventually contain such a small amount of the original solution that the last samples will not contain infectious virus particles anymore. Dilutions are applied in replicates to small cell cultures (usually a microtiter plate or a multi-well culture dish) and incubated during a certain period in which infectious virus, if present, will induce cytopathic effect (CPE) on the cells. Based on the number of positive/negative wells in correlation with the applied dilutions, an estimation of the virus titre is given.

5.1.2 The $TCID_{50}$ determination in viral clearance studies

Before continuing with a description of the method and the validation strategy, it is relevant to understand the purpose and applicability of the $TCID_{50}$ assay and to provide some background information on viral clearance studies for which these assays frequently are applied.

In many situations there is a requirement to investigate the capacity of a procedure to remove or to inactivate viruses. This is for instance relevant during vaccine production or during production of medicines derived from human blood or plasma. Also in disinfectant efficacy studies the reduction of virus titres is studied. In all these situations, virus reduction experiments are usually performed in the laboratory after downscaling the original (production) process or by mimicking the actual situation as closely as possible.

In virus reduction studies, the starting material is deliberately 'spiked' with a known virus that is either the same virus or a virus with similar characteristics as the virus of interest. In any case it is a virus of which the infectivity can be assayed. The virus reduction capacity of the process can be evaluated by comparing the difference of the $TCID_{50}$ titres estimated before and after reduction.

Viral reduction studies (also called viral clearance studies) are frequently referred to as 'virus validation studies', but should not be confused with studies for the validation of

virus testing methods. More background information on the design of viral reduction studies can be found in guidance documents like the ICH Q5A 'Viral safety evaluation of biotechnology products derived from cell lines of human or animal origin' [31] or the EMEA guidance Virus validation studies: The design, contribution and interpretation of studies validating the inactivation and removal of viruses' [8].

For the calculation of virus infectivity, virus titrations in cell culture are performed and frequently the 'Tissue Culture Infective Dose 50% assay' ($TCID_{50}$) is used for this purpose (See Table 3 for an example of the outcome of a $TCID_{50}$ assay). In this Chapter this method is explained and an indication of how such an assay can be validated is provided.

This Chapter will not further detail on virus reduction studies, but will be confined to the purpose of the method for the $TCID_{50}$ determination in general.

5.2 Validation parameters

For the determination of the titre of a virus stock, from which a considerable titre is expected to be present, the type of analytical procedure is 'Assay', a Type IV assay, see Figure 1 in section 2.3 and Table 2 in section 3.8. As can be seen from the table, the relevant validation parameters are **accuracy, precision, specificity, linearity** and **range.**

If the purpose of the assay is to provide accurate and precise titres of both low and high concentrations of virus (which is applicable to virus reduction studies), then in addition to the previous mentioned parameters, the **quantitation limit** (Type II analytical test, Quantitative test for impurity) would also be relevant. This is for instance applicable if $TCID_{50}$ titres are established during virus clearance studies in which log reduction values are calculated.

In this Chapter, for the purpose of virus reduction studies, the $TCID_{50}$ will be validated for both Type II and Type IV assays. The determination of the validation parameters will be addressed in the Stages 2 and 3 of the validation process.

Only when the infectivity assay is also meant to provide a 'yes or no' answer to either the presence or the absence of virus, **limit of detection** would be a relevant validation parameter. However, the small sample volume of the $TCID_{50}$ assay is not suitable for such a 'limit test'. Some aspects related to limit testing of infectivity assays are addressed in the Section 'Points to consider' at the end of this Chapter (Section 5.9).

The **robustness** aspects are addressed in the development phase (Stage 2) and the validation phase (Stage 4) of the assay.

5.3 Stage 1. Selection and feasibility

5.3.1 Purpose and scope

The purpose of the method is to determine the titre of infectious HSV-1 in samples with both high titres and low titres of HSV-1. The assay is used to estimate the infectious virus titre in samples before and after virus reduction treatment.

5.3.2 Method selection

Justification of the $TCID_{50}$ assay

An infectivity test has one important advantage over methods like molecular biological assays or immunobinding assays: it can discriminate between infectious and non-infectious virus. For studies in which the reduction of infectivity should be measured, infectivity tests are suitable and molecular and immunobinding assays are not suitable as the latter assays may also demonstrate non-infectious virus and are therefore not specific.

The $TCID_{50}$ assay is a quantal assay (samples are being diluted to an end-point and the titre is estimated) whereas in Plaque assays actual viruses numbers (plaques) are counted and titres usually are more precise and accurate than in a quantal method. Also, the assay duration of the plaque assays is shorter than the $TCID_{50}$ assays (e.g. 3 days versus 7 days respectively). In favour of the $TCID_{50}$ are the larger numbers of samples that can be handled, the lower costs and the fact that relatively less cells are required. Also, in general being less sensitive to variations in cell density, the assay can be more robust than the plaque assay. Furthermore, the longer duration of the test is not a limitation for the purpose of the assay.

Choice of virus and cells.

HSV-1 is a model virus for virus reduction studies. Important factors for the selection of the cell line to be used for virus titration are:

-The cell line should be susceptible for the virus

-The virus should produce a clear cytopathic effect

-The virus, can be grown to relative high titres, in order to demonstrate a high reduction potential

-The cells should be capable to grow to a confluent monolayer in microtiter wells

-The monolayers should remain intact over the entire incubation period of the assay

Vero cells are selected as the indicator cells because they are susceptible for HSV-1 and in addition they have been widely used in $TCID_{50}$ assays. These cells can be easily grown and remain well attached to the surface, even when they grow to complete confluence.

$TCID_{50}$: Method outline

Samples from which the virus titres are to be determined, are serially diluted in cell culture medium: 1 ml of sample is diluted with 9 ml culture medium and 1 ml of the 1:10 dilution is further diluted with 9 ml culture medium, etc.

Vero cells are seeded one day before inoculation with 10^4 cells per well in a microtiter plate. From each virus dilution, eight wells with Vero cells are inoculated with 100 µl/well and the plates are placed in an incubator at 37ºC. After 7 days –during which no refreshment of culture medium took place– the culture medium is removed and the remaining cells are fixed with formalin and stained with a crystal violet dye solution. After the final rinse, the plates are ready for evaluation. From the wells that stain blue/purple the Vero monolayer is still intact and these wells are scored negative (-). From wells that are empty/colourless, the Vero cells have vanished due to virus infection and these wells are scored positive (+).

5.3.3 Calculation of the virus titre and the standard deviation in a $TCID_{50}$ assay

A typical outcome of an assay with 12 dilution steps, performed in 8-fold, is shown in Table 3. 'Total' is representing the total number of positive responses (X_i in the formula below) for each dilution step and p_i is representing the reaction rate. These values are used to estimate the virus amount in the original solution. The 95% confidence interval for the log titre of the virus stock, as determined by the method of Spearman-Kärber (SK) [8, 21, 22, 23], is 6.250 ± 2x0.292= 5.667 - 6.834 $TCID_{50}$/100µl, see for calculations the formula below.

Replicate	[d]Log Dilution											
	1	2	3	4	5	6	7	8	9	10	11	12
1	+	+	+	+	+	+	-	-	-	-	-	-
2	+	+	+	+	+	-	+	-	-	-	-	-
3	+	+	+	+	+	-	-	-	-	-	-	-
4	+	+	+	+	+	+	-	-	-	-	-	-
5	+	+	+	+	+	-	-	-	-	-	-	-
6	+	+	+	+	-	+	+	-	-	-	-	-
7	+	+	+	+	+	+	+	-	-	-	-	-
8	+	+	+	+	+	-	-	-	-	-	-	-
Total (X_i)	8	8	8	8	7	4	3	0	0	0	0	0
p_i	8/8= 1	8/8= 1	8/8= 1	8/8= 1	7/8= 0.875	4/8= 0.500	3/8= 0.375	0/8= 0	0/8= 0	0/8= 0	0/8= 0	0/8 0

Table 3. An example of a read out of a serial dilution assay to determine the $TCID_{50}$ titre. '+' means that virus was present, '-' means no virus was present. The assay was performed with 12 dilutions and 8 replicates per dilution; 'Total (X_i)' is the number of positive wells per dilution, 'p_i' is the fraction positive wells/negative wells, '[d]Log' is de logarithm of the dilution factor between every dilution step.

The SK method estimates the titre using the following formula:

$$m = X_k - d/2 + d \Sigma (X_i / n)$$

or

$$m = X_k - d/2 + d \Sigma p_i$$

Where:

m = the negative logarithm of the titre, considering the inoculum volume per well

X_k = the logarithm from the reciprocal value of the lowest dilution where all replicates are positive

d = the [10]logarithm of the dilution factor

(d= 1.00 when dilution factor is 1:10,
d=0.30 when dilution factor is 1:2,
d= 0.48 when dilution factor is 1:3, etc.)

n = the number of replicates per dilution

X_i = the number of positive wells per dilution, starting with dilution X_k

p_i = reaction rate (= X_i /n)

Example of a $TCID_{50}$ titre calculation

In the example from the titration in Table 3, the titre is calculated as follows (assuming a dilution step of 1:10 (d=1)

$X_k = 4$

$d = 1$

$n = 8$

X_i see row 'Total'

p_i see bottom row

$m = 4 - 1/2 + 1\ (8/8 + 7/8 + 4/8 + 3/8)$

$= 4 - 0.5 + 2.75$

$= 6.25$

With an inoculum volume of 100 µl/well, the log end point dilution is 10^{-m}, and the log titre is therefore 6.25/100µl or 7.25/ml ($10^{6.25}$ $TCID_{50}$/100 µl or $10^{7.25}$ $TCID_{50}$/ml).

Example of a **standard deviation** (SD) calculation <35> of the log titre determined in the titration

$SD = \sqrt{[d^2 \Sigma \{p_i(1-p_i)/(n-1)\}]}$.

$SD = \sqrt{[1^2 \Sigma \{p_i(1-p_i)/(8-1)\}]}$.

As:

p_1 till $p_4 = 1.000 => 1\ (1-p_i)/7 = 0$

$p_5 = 0.875 => 0.875\ (1-0.875)/7 = 0.016$

$p_6 = 0.500 => 0.500\ (1-0.500)/7 = 0.036$

$p_7 = 0.375 => 0.375\ (1-0.375)/7 = 0.033$

p_8 till $p_{12} = 0.000 => 0\ (1-0)/7 = 0$

$SD = \sqrt{\{1^2\ (0.016 + 0.036 + 0.0.033)\}}$

$SD = \sqrt{0.085} = 0.292$

5.3.4 Materials

Typically there is no official reference standard for HSV-1 that can serve as a reference stock. Usually a HSV-1 virus bank is prepared and frozen at -70ºC in single use aliquots to serve as the reference stock during the validation of the assay and to serve as reference strain during routine use of the test.

For the Vero cell bank it is recommended to define the maximum split rate and maximum passage number of the cells as well as the standard cell culture procedure and cell

culture medium, including the type of medium and the foetal bovine serum (FBS) concentration.

The performance of the $TCID_{50}$ assay does not require specific equipment other than present in a normal cell culture laboratory, taking into account the biocontainment or biosafety level required for activities with the virus strain. The incubator used to incubate the cell cultures should be monitored (either continuously or periodically) to verify that the specifications of temperature, humidity and CO_2-concentration are met.

5.4 Stage 2. Development and standardization

5.4.1 Development and optimization experiments

In this stage first the optimum test format, e.g. the cell density, the inoculum volume, the culture medium and the incubation period, is to be found. This is determined by a number of development runs or checkerboard experiments. The conditions with optimal results are chosen. The method is described in detail, e.g. the type of cell culture plates (96-wells or 24-well plates), dilution with serological pipettes in sterile tube or dilution in 96-wells using multichannel automated pipettes, the cell culture medium, serum concentration etc. In this phase of the validation process some preliminary robustness tests are performed, e.g. the effect of the seeding cell density or the incubation duration is studied.

5.4.2 System suitability controls

System suitability controls are defined during the standardization phase of the assay and will become part of the routine assay.

The system suitability controls will consist of:

-The positive control: the titration of the positive HSV-1 stock

-The cell control: cells seeded, but not inoculated

-The negative control: wells inoculated with virus diluent.

As a positive control, the HSV-1 stock with the established titre is titrated to verify that the titre and the precision within the test run are according to the specification, demonstrating that the assay conditions were suitable and that the cells were susceptible to the virus. The cell- and negative controls are used to demonstrate absence of cytotoxic properties and absence of accidental contamination of the media and diluent.

5.4.3 Preliminary specificity

During this phase of the assay, experiments are conducted to verify that the diluent used to prepare the virus dilutions does not produce cell death or cytotoxicity and that the Vero cell cultures stay healthy during the entire incubation period. Further studies for specificity are performed in Stage 3 of the validation process, once other parameters of the assay have been determined. For the purpose of demonstrating viral clearance in a reduction process, the sample suitability is part of the qualification of the assay for its particular use.

The effect of the matrix of the samples should be addressed during sample suitability testing that is conducted prior to use of the $TCID_{50}$ assay for a specific purpose.

5.4.4 Preliminary precision experiments

Virus titrations suffer from the problem of variation common to all biological assay systems.

Both systematic and random variation can arise within an assay as a result of dilution errors, statistical variability and differences within the assay system which are either unknown or difficult to control. Random effects are likely to be greater when different assay runs are compared ('between-assay variation‘ or ‘inter-assay variation') than when results within a single assay run are compared (‘within-assay variation' or ‘intra-assay variation'). In the standardization phase of the assay, it is verified that the assay format is suitable with respect to the number of replicates and the dilution factor between the replicates.

Initial precision experiments are performed to verify that the assay is capable to determine titres within the required confidence level, using the chosen number of replicates and the dilution factor. These experiments may be conducted using an arbitrary chosen number of replicates (e.g. 4, 8, 16) and smaller or larger dilution steps (e.g. dilution factor 1:3, 1:5, 1:10 dilution). However, the more replicates tested and the smaller the dilution factor, the more precise the assay will be.

A general accepted 95% confidence limit for biological assays is the mean ± 0.5 log <31><8>. The variation within the assay is determined by calculating the standard deviation and based on this the confidence level is set as mean ± 2SD (in log values).

In section 5.3.3., in the example in Table 3, the standard deviation of 0.292 resulted in a confidence level of 0.584 log (0.6 log) and this exceeds the maximum value of ± 0.5 log.

The SD of this assay is higher than acceptable if used for the described method in this Chapter.

For a more precise titre determination, the number of replicates can be increased (e.g. n=16 instead of n=8) and/or the dilutions steps can be smaller, e.g. dilutions 1:3 (d=0.48) instead of 1:10 (d=1.00). Assuming that the same titration results would be considered as in Table 3, but the dilution steps were reduced from 1:10 (d=1) to 1:3 (d=0.48), the standard deviation would be:

$SD = \sqrt{\{0.48^2 (0.016 + 0.036 + 0.033)\}}$

SD= 0.140.

An SD of 0.140 provides a confidence interval of mean ± 0.280 log −rounded off ± 0.3 log−, and this is acceptable. When the number of replicates is reduced from n=8 to e.g. n=6 or n=4, it is obvious that the SD will be higher. In many cases, for the calculation of titres and reduction values, a minimum of 8 replicates is considered appropriate.

In some situations the precision of the assay is of less importance. For instance when the assay is used to screen samples for infectivity to determine the peak titre in fractions of a column eluate, a $TCID_{50}$ assay performed with less replicates and larger dilution steps can be more practical and may already be suitable.

5.4.5 End of Stage 2

Once the assay conditions have been defined and described in the (concept) procedure, the materials are fixed, and the results from several development experiments provide confidence that the precision can be obtained, the required assay is ready for the assay qualification (Stage 3). The titre of the HSV-1 reference stock will be established by at least 6 independent assays and when all individual measurements differ by no more than 0.5 log from each other.

5.5 Stage 3. Performance qualification

5.5.1 Precision

In Part 1 (Section 3.2) is explained that in general the precision of an assay can be classified into repeatability (within-assay variation), intermediate precision (between-assay variation) and reproducibility (between laboratories variation).

In this section the determination of repeatability and the intermediate precision of the $TCID_{50}$ assay are described. Reproducibility of the assay is applicable only when the

assay is transferred to another laboratory, and is shortly addressed in Stage 4 of the validation process (Section 5.6).

The repeatability is calculated from more than one assay run in a short period of time (e.g. one day), using multiple samples with multiple concentrations and performed by one operator. In Table 4 an example is provided for such an experimental setup. The assay is conducted in two runs on one day, three samples, each in triplicate, to provide 18 samples. For each individual assay, an SD value is obtained. Each SD from the individual assay (usually consisting of one plate) meets the preset criteria of <0.25 and the pooled SD provides the repeatability of the TCID50 assay.

Sample	**Run 1**		**Run 2**	
	Titre	**SD**	**Titre**	**SD**
S1 (low)	3.74	0.19	3.56	0.19
	3.39	0.21	3.82	0.18
	3.74	0.18	3.48	0.22
S2 (medium)	5.65	0.19	5.39	0.19
	5.91	0.21	5.91	0.23
	5.74	0.18	5.65	0.24
S3 (high)	7.48	0.20	7.39	0.22
	7.82	0.21	7.74	0.24
	7.56	0.20	7.39	0.20
Pooled SD	0.205			

Table 4. Repeatability testing of $TCID_{50}$ assay. Three samples S1, S2, S3 are titrated in triplicate in two independent runs performed by one operator on the same day, resulting in log titres and SD values from 18 assays. Assays consists of 8 replicates and the titre and SD are calculated using the SK-formula (see section 5.3.3). The pooled SD is the repeatability of the $TCID_{50}$ assay, as is calculated based on the individual SD's from the 18 assays.

The repeatability can be calculated based on the mean SD of the 18 individual SD values. Alternatively a pooled SD can be calculated:

$$\text{Pooled SD} = \sqrt{\{[(n_1-1)SD_1^2 + \ldots.(n_i-1)SD_i^2] / [n_1 + \ldots.n_i - i]\}}$$

Where:

SD= standard deviation for each assay
n= number of replicates within each assay
i = number of assays

Intermediate repeatability

The intermediate precision or inter assay precision is determined by using the assay of the high titre positive control stock in at least six different variants with variation in days, operators, and pipettes.

From each of the assay runs the titre and the SD are calculated. The results should fall within the established confidence level of the assay as determined for the positive control stock used as the reference stock (mean log titre ± 2SD).

The inter assay precision is determined by calculation of the pooled SD as described above, or by the mean SD of the six individual assays.

Example of an assay matrix for intermediate precision

Assay	Day	Technician	Multichannel pipette	Batch cells
1	1	A	P	S
2	1	B	Q	S
3	2	A	Q	T
4	2	B	P	T
5	3	A	Q	U
6	3	A	P	V

5.5.2 Linearity

The linearity of the assay is determined using 5 concentrations of virus and subsequent testing in triplicate assays. The experiments may be combined with those for the establishment of the precision of the assay (section 5.5.1). The results are analyzed by common linear regression analysis (y=ax + b) using all data points (15 values) with the expected log titre on the X-axis and the found log titre on the Y-axis. Acceptance criteria should be defined prior to the regression analysis. A correlation coefficient of ≥0.99 is characteristic for a $TCID_{50}$ assay.

5.5.3 Quantitation limit

The LLOQ is the lowest concentration in which the titre can be accurately and precisely measured and should fall within the linear range. For $TCID_{50}$ assays, the titre calculation can only be measured accurately if the lowest dilution provides CPE in all replicates (for Table 3, the p_i =1 for at least the lowest dilution step). The quantitation limit is relevant

in case the exact titre of virus samples should be established in samples with low titres. The LLOQ is dependent on the initial dilution of the sample prior to performing the $TCID_{50}$ assay: the higher the pre-dilution, the higher the LLOQ is. A pre-dilution might be necessary to overcome cytotoxicity, as will be determined in the sample suitability test of the assay.

5.5.4 Range

The range in which a precise and accurate titre measurement can be performed, is covered by the titres values lying between the LLOQ and the maximum titre for which the assay has shown linearity.

In viral clearance studies often results are obtained where even the lowest dilution does not lead to positivity in all wells, or even where all wells are negative for all the tested dilutions. Titres are then presented as 'less than', whereby the titre is calculated assuming that all wells from the lower –but not actually tested –dilutions would have been positive. These 'less than titres' are always lower than the LLOQ.

If no positive wells are found in a titration, the titre is calculated as if one well from the lowest tested dilutions would have been positive and is expressed as '<xx $TCID_{50}$/volume'. This 'less than' titre is sometimes referred to as the LOD, but this is not correct. For the establishment of the LOD, it is expected to evaluate the minimum amount of virus that can be reliably detected in 95% of the experiments. See section 5.9 for some more information on the LOD in infectivity assays

5.5.5 Accuracy

As for most viruses, no formal reference standard exists for HSV-1. As a result the accuracy of the assay can not be measured. Instead, the system suitability controls can verify that the test is able to provide the results with sufficient accuracy by testing the provisional reference strain with established titre in each assay run.

5.5.6 Specificity (sample suitability)

The initial test for specificity has been completed in Stage 2 of the validation process. Once the assay has been qualified, it may be used for other types of samples then it was originally validated for. The sample matrix may interfere with the assay, either by cytotoxicity, causing false positive reactions, or by affecting the susceptibility or replication of the virus, resulting in a false negative result. The specificity of the $TCID_{50}$ assay should therefore be determined for every type of sample that is applied to the assay.

If the assay is used to quantify HSV-1 in fractions from a sucrose gradient, the samples consisting of the highest sucrose concentration should be tested for suitability. If the assay is used to titrate HSV-1 in serum and the sample would consist of pure serum, the sample suitability test is performed using pure serum. Sample suitability testing has two different stages.

The first stage is the cytotoxicity test, to verify the performance of the test with a sample in the absence of the virus. For example the sucrose sample is diluted and the Vero cells are inoculated. Cell cultures are observed for the absence of cytopathic effect. In case the assay is used for the efficacy testing of disinfectants, 'cytotoxicity controls' are prepared by dilution of the disinfectant (after neutralization of the disinfectant if applicable) and the effect of every dilution on the cells is observed after the entire incubation period. The dilution that does not cause cytotoxicity is selected as the starting dilution in the assay. Only the starting dilutions that do not produce CPE may be suitable for use.

The second stage concerns the interference of the matrix with the virus titre. Some matrixes do not inactivate the virus, but prevent the cells from replicating virus, resulting in false negative results or titres that are lower than the actual value. Thus, if a dilution does not produce CPE, it should be verified that the dilution does not inhibit the growth of low virus titres.

In viral clearance studies it is expected that sample suitability runs are performed by dilution of the reference strain in the lowest dilution(s) of the sample matrix not showing cytotoxicity. The titre of the reference strain should be the same as the reference strain in the normal dilution, taking into account the normal variation of the assay (same confidence interval). The presence of neutralizing antibodies in the matrix may also result in found titres lower than the actual values. Proper sample suitability controls will reveal such effects.

5.6 Stage 4. Performance validation

This stage will be relevant in case the assay will be used over extended periods of time, and when the assay is providing critical information such as required for potency assays of manufactured drug products or when used in stability studies of medicines.

Experiments include additional intermediate precision testing with multiple runs, applying deliberate variation of assay conditions. They are used to gain more information about the robustness of the assay, and applied for e.g. cell passages at the minimum and maximum passage levels (e.g. Vero cells at passage 160 and passage 195 are tested), testing of two different batches of FBS or use of different equipment (multichannel pipettes, incubators).

A typical reproducibility study, to establish the variation between laboratories, is exemplified by a determination of three representative samples at low, medium and high concentration in three individual experiments conducted by each laboratory.

For an assay that is qualified for the use in viral clearance studies, this validation stage is not applicable, as the validity of the assay is demonstrated per protocol and the experiments are usually performed within a confined range of time. In addition, in the viral clearance study the sample suitability and system suitability controls are included to demonstrate the validity of the assay.

5.7 Stage 5. Maintenance and improvement

For a $TCID_{50}$ assay once developed, usually little improvement is required. A spreadsheet is maintained with the date of test, titre of the reference stock and SD of the assay. In particular, the replacement of cell banks and positive control stocks may require attention. If there is a trend in the titre of the stock, the stability of the reference stock may be investigated periodically by checking in another type of assay (e.g. plaque assay).

Often, like with other assays, the precision of the assay improves over time, especially when run by experienced technicians. On the other hand, when the assay shows impaired precision over time, it can be necessary to modify the assay e.g. by adapting seeding densities, or to choose smaller dilution steps or a higher number of replicates per dilution.

5.8 Method verification

The $TCID_{50}$ is not an example of a test that is commercially available. A verification of the method would typically be applicable if the assay is not used for a while, if a new cell bank is used, or if new operators are qualified for the assay. In such case an appropriate qualification would consist of three assay runs using the positive control stock. For each assay the titre is determined within the confidence limit of the established titre, and the SD of each assay is less than 0.25.

5.9 Points to consider

5.9.1 Probability of detection of viruses at low concentrations

A similar principle as described for the $TCID_{50}$ determination can be applied when testing large volumes of samples containing low virus titers. An example is the investigation of absence of replication competent viruses in large volumes of vector batches, where there

is a small chance that they are present [32]. This is relevant in testing for replication competent virus in gene therapy medicinal products, but the same principles apply when screening products for absence of production of infectious virus for biosafety reasons, or when larger volumes are tested to demonstrate absence of infectious virus after a virus inactivation step. Such an assay is a typical 'Limit test' (Type III, see Table 2 in Section 3.8) for which the LOD of the assay should be determined. In order to determine the LOD of the method, the following points should be considered.

At low virus concentrations (e.g., in the range of 10 to 1000 infectious particles per litre) a sample of a few millilitres may either contain infectious particles or not.

The probability *p* that this sample does not contain infectious viruses is:

$p = ((V-v)/V)^n$

where V (litre) is the overall volume of the material to be tested, v (litre) is the volume of the sample and n is the absolute number of infectious particles statistically distributed in V.

If V >> v, this equation can be approximated by the Poisson distribution:

$$p = e^{-cv}$$

or, c = ln *p* /-v, where c is the concentration of infectious particles per litre.

For example, if a sample volume of 1 ml is tested, the probabilities *p* at virus concentrations ranging from 10 to 1000 infectious particles per litre are:

c	**10**	**100**	**1000**
p	**0.99**	**0.90**	**0.37**

For a product with a concentration of 1000 viruses per litre (1 virus/ml), in 37% of sampling, 1 ml will not contain a virus particle. This is in fact similar to a situation if all 96 wells from a microtiter plate would be inoculated with an exact amount of 1 infectious virus particle/well: Statistically, 37% of the wells would be negative (purple) and 63% would be positive (blank). See Figure 6 for a visualization of such a result.

In accordance with the Poisson distribution, in order to detect −with 95% probability− at least 1 infectious virus, a sample from 3 ml should be inoculated [19]. The required test volume = (-1/c) ln (1-p).

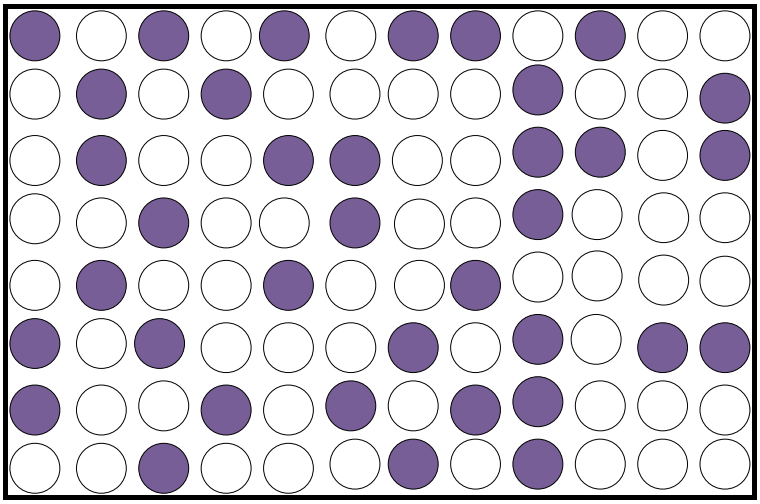

Figure 6. Visual representation of the Poisson distribution resulting in a theoretical number of positive wells of 63%.

5.9.2 Dilution

The manner in which dilutions are made (type of pipettes, dilution tubes, diluent), contributes to the precision of the assay. More precise measurements are obtained when a dilution is made in large volume dilution tubes (e.g. 5 or 10 ml volumes, from each tube multiple replicates are inoculated) compared to dilutions made in microtiter plates, for which often multichannel pipettes are used. The manner in which dilutions are made, should be part of the assay procedure, and be similar during both validation and actual use.

5.9.3 Carry-over

To avoid carry-over effects, a new pipette (or pipette tip) should be used for every dilution step. Neglecting this practice may result in artificial high titres, especially when high concentration virus suspensions are assayed. This will become obvious during linearity testing in Stage 3 of the validation process. It is one of the most frequently occurring errors in the $TCID_{50}$ assay –caused by growing nonchalance over time– and this issue should be addressed during training of technicians. It can be useful to include an illustrative experiment in the training program, like titration of a high-titre sample with and without changing pipette tips in parallel.

5.9.4 Media changes

No media changes should be performed after the first round of virus reproduction is completed (usually a few hours), as this will inevitably result in cross contamination leading to false positive results or an incorrect titre being higher than the actual value.

5.9.5 Use of alternative cell lines

Due to differences in susceptibility of different cell lines, the titre of a reference standard can differ from one cell line to another, even when the method is validated. Virus titres obtained on Vero cells will be different from titres obtained on other cells types (e.g. human MRC-5 or HEK293 cells). If virus titres are provided, it is necessary to refer to both the method and the cell type, e.g. 'The titre is $10^{7.20}$ $TCID_{50}$/ml when titrated on Vero cells'.

6 ELISA FOR HIV-P24

6.1 Introduction

An Enzyme-Linked Immunoassay (ELISA) [25, 26] is a widely used method for the detection of antigens and antibodies. The HIV-1 p24 ELISA is an immunoassay for the detection of the *gag* structural protein present in HIV-1 and HIV-1 derived lentiviral vectors. Such an assay can be used for –amongst others– the following purposes:

- Detection of HIV-1 in patients by testing serum or plasma samples
- Detection or quantitation of lentivirus (LV) particles in cell cultures or vector batches
- Demonstration of the absence of Replication Competent Lentivirus (RCL).

Due to cost savings, laboratories often develop their own ELISA instead of using commercial kits. Irrespective whether a system is bought or developed in house, the laboratory that uses the assay is responsible to properly validate the method for its intended use.

In this Chapter, the validation of a p24 assay (used to detect and quantify HIV-1 related p24 protein) is addressed for its purpose to detect and quantify the expression of LV particles in cell samples. An example of a validation process for an entire new ELISA (sections 6.2 till 6.7) is described, as well as the verification process for an already validated commercial assay (section 6.8). In section 6.9 points to consider with respect to the validation of immunoassays in general as well as points to consider specifically for the p24 ELISA for the detection of LV are provided.

6.2 Validation parameters

6.2.1 Selection of parameters

As the assay is used to quantify p24 at low concentrations, it is a Type II analytical procedure (Quantitative test for impurities, see section 3.8) and the relevant validation parameters, according to the ICH Q2(R1) [1], are: **accuracy**, **precision**, **specificity**, **quantitation limit, linearity** and **range**.

As the assay will also be used to provide a 'yes or no' answer for the presence of LV particles, in addition the **limit of detection** is a relevant parameter. Also, as for all assays, **robustness** testing is applicable.

The validation parameters will be assessed in a logical order during the different Stages of the validation process as described in sections 6.3 and further. For instance, accuracy can only be determined once the linear range and precision of the assay have been established.

In section 6.9.1 some other examples of immunoassays and the validation requirements in relation to their purpose are provided.

6.3 Stage 1. Selection and feasibility

6.3.1 Purpose and scope

For the purpose of this validation exercise, a method is used to demonstrate the absence of RCL. A second validation exercise concerns the quantitative estimation of LV production.

6.3.2 Method selection

The p24 ELISA is suitable for detection of replication competent virus as well as defective and non-replicating particles. The assay detects p24 gag protein, a structural protein of LV particles, enabling measuring of p24 protein in test samples as an indication for virus production or titre determination. Performance of a p24 ELISA only in order to demonstrate the presence of replicating competent viruses is not a suitable approach: for this specific purpose it is essential that the p24 ELISA is preceded by cell culture with a sensitive cell line.

6.3.3 Method outline

The method consists of a double antibody sandwich ELISA of which the principle is outlined in Figure 7. The 24-kilodalton protein (p24) is a major structural core protein from HIV-1. Purified monoclonal antibodies with high specificity and affinity for this viral protein are immobilized onto the surface of the well of an ELISA microtiter plate, and subsequently able to selectively capture p24 protein that is present in the sample. A second antibody to which an enzyme is coupled will react to the p24 antigen in the next step. The bound enzyme will enable a substrate reaction that results in the colour development that can be measured as an OD signal at the wavelength optimal for the type of substrate.

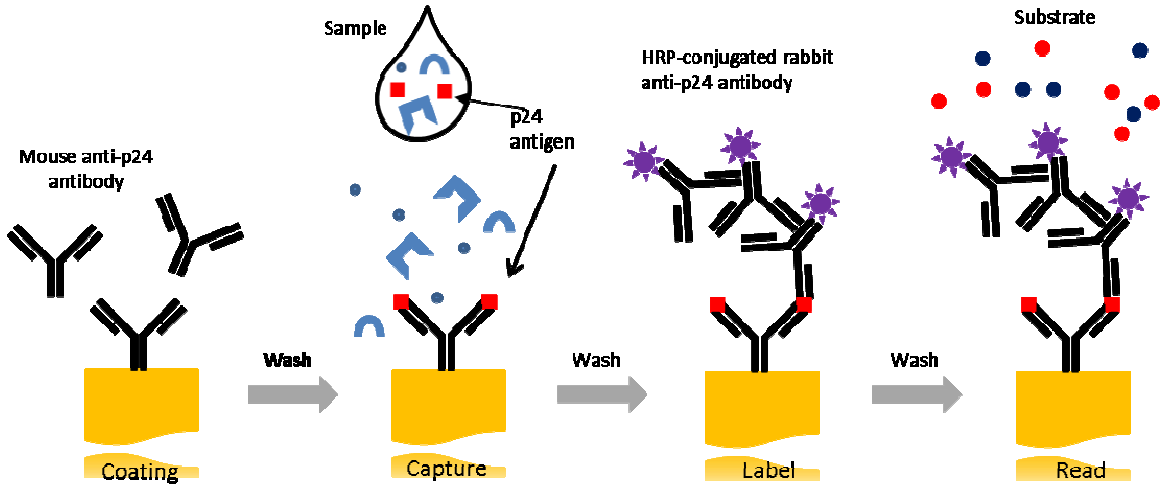

Figure 7. The principle of a HIV-1 p24 ELISA. The wells from a microtiter plate are coated with Mouse anti-p24 antibody. Following incubation of the plate with the sample to capture the target analyte (p24 antigen), residual components are washed away. Then, incubation with a Horse Radish Peroxide (HRP) conjugated second antibody takes place, followed by a third wash step. After addition of substrate, the HRP produces a color signal, from which the intensity can be measured in an ELISA reader being an indication for the amount of p24 originally present in the sample.

6.3.4 Materials

A detailed list of reagents, equipment and accessory materials should be defined. The antibodies belong to the most critical components and are carefully selected. Equipment is installed and calibrated. A p24 reference standard is prepared, stored in multiple, single use aliquots and subsequently calibrated for p24 content by sending out to a reference laboratory.

6.4 Stage 2. Development and standardization

Checkerboard experiments are performed to find the suitable assay conditions like antibody concentrations, incubation times, temperatures and optimization of washing steps. Once satisfactory back-ground and signal to noise ratios are established, the method is standardized. The cut-off value and system suitability controls are defined.

6.4.1 Cut-off value

The cut-off value, being the level above which samples are considered positive, is calculated as the mean negative control value of the OD-value plus two times or three times the standard deviation of the blank samples. The cut-off value may differ slightly between assay runs. However, a maximum value should be defined in the standard procedure. Instead of the standard deviation in the particular assay, a fixed value representing the typical SD of the assay can be taken –based on the SD results in the

assay qualification. In that case also requirements to the maximum variation in the blanks should be included in the procedure.

6.4.2 Preliminary specificity

A set of different culture fluids harvested from cell cultures is tested in the ELISA and should be non-reactive (no cross-reactivity is observed). These culture fluids are from different cell types and may include different types of cell culture media and serum sources. The culture fluids should not react above the cut-off value, while culture fluids spiked with LV particles near the presumed LLOQ, should react.

6.4.3 Preliminary robustness

Using the reference control material, some preliminary robustness experiments are performed. For instance, the influence of minor changes in antibody concentrations and incubation duration on the signal to noise ratio as well as randomization of samples on the plates and the performance of assays in small experiments and larger experiments (with longer waiting times between washing steps etc.) are investigated. For a satisfactory assay minor changes should not influence the result of the test. Nevertheless additional robustness testing may be required depending on the use of the assay. This is usually done in Stage 4 of the validation process.

6.4.4 Limit of detection

For a quantitative assay the LOD may not be relevant. However, for a limit test (Type III assay), to provide a 'yes or no' answer, the LOD is a relevant parameter. The LOD is the minimum concentration of p24 that generates a consistent response greater than the background of the test. Responses of 2 to 3 times the standard deviation of the background are reported as satisfactory limits [2, 3].

The LOD can be determined as follows [6]:

Prepare a standard concentration of p24 and a zero concentration (=blank solution) in the solution. Perform the ELISA assay three times in duplicate according to the standard procedure and measure the OD values in the six assays. Calculate for the six assays the mean and the SD of the OD_{blank} values. Calculate for the p24 sample the mean OD_{sample} value.

The LOD is calculated from

LOD (in g p24/ml) = (mean OD_{blank} + 3xSD_{blank}) / (mean OD_{sample}/ p24 concentration in sample).

The cut-off value is entered into the calibration curve and a p24 concentration is obtained from that.

An alternative way is to serially dilute 6 different samples with a spiked known titre, and to test each dilution series in three independent experiments. The concentration as calculated from the dilutions at which each of the 6 samples are still positive, is considered the LOD of the assay [34].

The LOD is a concentration lower than the Lower Limit of Quantification, and therefore the LOD is an indication, but not an accurate and precise value. The difference between cut-off value and LOD is that the cut-off value is a OD-value, which is calculated for each assay, while the LOD is determined only once during the qualification of the assay. The LOD is a concentration (p24/ml) and is also called the analytical sensitivity of the assay.

6.4.5 Linearity and range

In order for an ELISA to give accurate results, there must be an excess of antibody (both captured and conjugated) relative to the analyte (p24) being detected. Only if there is excess of antibody, the dose response curve is positively sloped and the assay can provide an accurate quantitation of p24 antigen. When the concentration of p24 begins to exceed the amount of antibody, the dose response curve will flatten and with further increase it may even become negatively sloped. This phenomenon is termed *'Hook Effect*'. For samples with high p24 concentrations the possibility exists that some samples may have analyte concentrations in excess of the antibody, which results in erroneously reduced titres. During the validation this effect should be addressed, when establishing the linear working range. Before carrying out the experiments to determine the precision and accuracy of the method, the expected concentration range should be defined.

The linear range is determined as the interval between the lower and upper concentration of p24 for which a suitable level of precision, accuracy and linearity is obtained. The LLOQ and the HLOQ are the lowest and highest concentrations of the range

respectively, which are repeatedly found in the linear range with sufficient precision and accuracy.

To establish the working range, tests with serial dilutions from low to (very) high concentrations are determined. The OD values are plotted in a graph against the p24 antigen concentrations and the visible (plotted linear range) is determined.

To establish linearity, the closeness of observations to a straight line is measured. This is done by calculation the correlation coefficient for dilutions of p24 over the claimed range.

Six to eight dilutions (excluding the zero concentration) spiked with known p24 concentrations within the claimed range are prepared. Each dilution is tested in triplicate, in three independent runs (9 tests per dilution).

a) For each sample in each assay run the actual results versus the expected results (the results are in concentration of p24/ml) are plotted. Each set of dilutions are analyzed in a linear curve and the coefficient of regression is calculated.

b) The accuracy (%recovery) and the precision for each dilution are calculated. The range is determined by the highest and lowest concentration with satisfactory accuracy and precision for each of the three assays.

6.4.6 Assay controls, system suitability

An example of how assay acceptance criteria, system suitability and test sample evaluation for the ELISA can be presented is described below. After the qualification of the assay, the requirements of the assay controls and system suitability may be adjusted from preliminary to definite values and as such incorporated in the standard assay procedure.

Example: System suitability and assay controls

Note: the figures are for indicative purposes only and should not be considered as real values!

- Substrate blanks

The OD values (blanked against water) should be below 0.050, otherwise the assay should be rejected for all samples.

- Negative control

The OD values of the negative control (0 ng/ml p24) should be below 0.150 for two of the three wells and the mean NC based on the three values must be below 0.150, otherwise the assay must be rejected for all samples.

- Positive control

All three replicates must show reactivity by an OD value of > 0.600 and the mean OD-value should be >0.800, otherwise the assay must be rejected for all samples.

- The standard curve

All samples above the established LLOQ of the assay must be reactive (OD above the cut-off value). The standard curve is plotted by quadratic regression and should show a linear regression coefficient of at least 0.980.

- System suitability

System suitability samples are sometimes called 'QC samples'. These are matrix samples spiked with a known concentration of p24 and the recovery should be ±20% in the middle range and ±25% for concentrations near the LLOQ or HLOQ. E.g. for a valid assay a sample spiked with the p24 standard to 100 ng/ml, should demonstrate a concentration of 100 ± 20 ng/ml as derived from the standard curve.

- Evaluation of test samples

  -Samples that are reactive (react above the cut-off value) and have titres below the LLOQ, are retested in duplicate in order to confirm the results. If one or both retests from the sample is reacting above the cut-off value again, the sample is considered positive with a concentration below the LLOQ. If the retest with one or both duplicate samples is negative, the sample is reported as negative.

  -Samples that are reactive and have an extrapolated titre at the LLOQ or higher, are considered positive. If the titre is higher than the HLOQ, the test should be repeated after pre-dilution of the sample in order to find a results established within the linear range of the test.

6.5 Stage 3. Performance qualification

Once the conditions of the assay have been set and preliminary qualification experiments have shown that the assay is consistently meeting the assay criteria, the assay is ready for the qualification experiments.

6.5.1 Precision

Intra-assay precision

The reference standard is diluted to provide three spiking concentrations near the LLOQ, the middle of the range, and near the HLOQ respectively. Ten replicates of each sample are tested (30 samples in one test). The average and standard deviation for each sample are calculated. The CV is calculated for each concentration within the assay run.

Inter-assay precision

The reference standard is diluted to provide three spiking concentrations near the LLOQ, the middle of the range, and near the HLOQ respectively. Test triplicates of each sample (9 samples) in three different assays (27 sample results). The assays should be performed on different days, by more than one technician and using different batches of critical materials, if possible. The average and standard deviation for each sample are calculated. The CV is calculated for each concentration between the assay runs.

For an ELISA the precision is typically 10-20% (CV).

6.5.2 Accuracy

The accuracy is determined in the matrix for which the assay is applicable. The reference standard is diluted to provide three spiking concentrations near the LLOQ, the middle of the range and near the HLOQ respectively. Two individual samples for each concentration are prepared. These six samples are tested in triplicate in one run. The expected versus the average measured value is calculated for each sample by calculating the % recovery (assay value / expected value x 100%). The recovery provides the bias of the assay. The criterion for accuracy determination is typically 80-120% [6, 27].

6.5.3 End of Stage 3

Once the qualification runs are performed, the criteria for assay control and system suitability are evaluated and adapted if necessary. The LLOQ as well as the LOD of the assay are established and sufficient data are obtained to finally set the criteria for assay controls and system suitability controls.

6.6 Stage 4. Performance validation

Stage 4 of the assay may be applicable when the assay is used over a longer time and when used to produce critical data, e.g. for potency assays in pharmaceutical manufacturing. Testing may include −but is not limited to− additional robustness testing, freeze thaw stability of samples, lot-to-lot variation and stability of reagents. Also accuracy determination by parallelism may be relevant e.g. to determine that the recovery in a particular matrix is consistent with the lower and higher concentrations. Proficiency testing and reproducibility experiments can be part of this Stage of assay validation. In the context of this report, these aspects will not be further elaborated.

6.7 Stage 5. Maintenance and improvement

6.7.1 Assay trending

In order to monitor the performance of the assay over time, and to ensure maintenance of the validated status of an assay, it is useful to plot the titre of the p24 reference stock and/or QC samples, as well as the standard deviation from every ELISA run, into a database in relation over time. If the titre is changing significantly, an investigation should provide the cause of the assay drift. Some possible factors are: environmental conditions, equipment failure or maintenance, slight changes in procedures, increase in capacity, deterioration of washing buffers, new reagent lot, expired reagents, changes in the matrix of the sample, storage conditions, inadequate pipette calibration, systematic calculation errors or human (operator) errors.

6.7.2 Titre reference material

Prolonged storage may affect the titre of reference stocks. Also, the titre can be affected due to incidents like temporarily placing the box with reference material outside the freezers during freezer cleaning actions. It is recommended to periodically send out samples of the reference strain to monitor the titre as an indication of the stability of the reference material.

6.7.3 New materials

Over time reagents get depleted and new batches of materials are obtained. It is useful to include in the standard procedure how new batches of reagents and microtiter plates will be qualified for use. Significant changes in materials, e.g. a new batch of coating antibodies or a new source of conjugate require more extensive validation, similar as described for Stage 3 in the section on method verification (section 6.8).

6.7.4 Improvement

Improvements of the method that may affect its performance should be evaluated for their possible impact and may require validation as described in the section on method verification.

6.8 Method verification

6.8.1 Use of commercial assays

Often, reagents necessary for an ELISA are available in commercial assay kits, consisting of a set reagents as well as instructions for their proper use. The kits are to be used in accordance with the manufacturers' instructions. In addition, it is important to ascertain that the kits are suitable for the analysis of the substance to be examined, with particular reference to selectivity and limit of detection [24]. General guidance concerning immunoassay kits is provided by the WHO, TRS 658, part 2 [6]. In the next section an example is provided for a p24 antigen detection ELISA.

6.8.2 Validation of a commercial p24 antigen ELISA

Commercial assays for p24 antigen detection are widely available and often provided with detailed manuals in which system suitability tests as well as performance characteristics of the assay are provided (e.g. Perkin Elmer, [20], Cell Biolabs, [24]). If a validated commercial assay is available, there is no need to repeat the entire validation. It is common practise that the relevant parameters like accuracy and precision are repeated with the use of the reference standard that is included in the kit. In this section the process of method verification is shortly described.

When the use of a commercial validated assay is considered, first of all an assessment should be made if the kit was validated for the same purpose as its intended use in the laboratory. For example, if a kit was validated for use of detection of HIV-1 in human serum (where virus can be captured in immune complexes and need to be disrupted), the performance of the assay may be different then if used for the detection of antigen in cell culture samples with FBS.

Stage 1 consists of literature investigation and evaluation of the performance characteristics of the test kit. Equipment (reader, washer, pipettes) are installed and qualified and the required materials not included in the test kit are obtained. A p24 reference control material is obtained when not provided with the commercial kit.

Stage 2 consists of some trial runs to investigate if the system suitability controls and the standard curve actually work.

In Stage 3 the reproducibility of the commercial assay in the laboratory is assessed as well as a confirmation of the accuracy. The repeatability of the assay is determined within the range of the assay. There is no need to repeat the validation with respect to specificity, linearity, detection limit and range of the assay, as long as the assay remains unchanged and the correct system suitability controls are included. For quantitative assays a standard curve should be included in each assay run as well.

Example: A Stage 3 qualification study to assess reproducibility and accuracy of a commercial assay.

Three p24 positive samples are prepared in the diluent by dilution of the reference control to known concentrations within the range: near the LLOQ, in the middle range and near the HLOQ. The same samples are tested in three experiments in duplicate plates (=6 assays per sample). The mean p24 concentrations found for each of the 3 samples for each of the 6 assays tests are within 80-120% of the expected concentration, and the within assay (intra assay) CV is <10% and the between assay (inter assay) CV is <20% (assuming these were the performance characteristics as provided by the kit manufacturer).

In Stage 4 the applicability of the assay for the specific sample matrix is evaluated. The ability of the assay method to correctly quantify a known concentration of p24 in the sample matrix, is evaluated by spike and recovery studies. These 'matrix interference' experiments should demonstrate that the sample matrix does not interfere with the test. The matrix intervention should be evaluated for all types of matrices for which the assay will be used.

Example: A Stage 4 qualification study to assess matrix interference.

Samples are spiked with a known amount of p24 antigen and tested in the ELISA along with p24 antigen in the standard diluent. For a quantitative assay a typical acceptable recovery is 80-120%. For non-quantitative assays the spiked samples should provide a positive assay when spiked with low (near the LOD) and high concentrations. If interference is found, samples may be tested in higher dilutions in order to overcome the interfering properties. Of course this affects the LOD of the assay, but as long as the impact is understood, for many applications this can still be suitable.

Stage 5 of the validation process for a commercial assay kit is similar to that for a non-commercial 'in-house' assay as described in section 6.7 above.

6.9 Points to consider

6.9.1 Examples of validation of other immunoassays

Depending on the purpose of the immunoassay, the validation requirements differ. In addition to the Type II assay used as an example in this Chapter, other types of analytical procedures are possible. From each of the three other types of analytical methods as described in section 3.8, an example for an immunoassay is given below.

Type I (Identification) An immunofluorescence assay to identify a pathogenic agent. The only relevant validation parameter is 'specificity'.

Type III: (Limit test for impurities) A classical pregnancy test in order to detect the human chorionic gonadotropin hormone in urine to provide a 'yes or no' answer on pregnancy. If hormone levels need to be assayed with the purpose to have an indication on the pregnancy stage, the assay would need to be validated as a type II assay.

Type IV: To quantify an antigen content of a vaccine, from which the titre is expected to fall in a given range. All validation parameters are relevant, except detection limit and quantitation limit, as the requirement is to determine the concentration with maximal accuracy and precision and thus the assay should not be conducted in the lower concentration range.

6.9.2 Virus particle calculation

One should be reluctant to use the p24 ELISA results to calculate LV particles based on an estimation of number of molecules of p24 per LV [24]. Due to fact that *gag* protein can be over-expressed, or empty LV particles devoid of transducing activity may exist, the calculation of LV particle concentration may not be correct [13]. These limitations can not be overcome by proper validation of the assay. Due to the fact that incomplete or empty particles can be present, the ratio between (infectious) particles and ELISA activity may vary considerably between virus batches.

6.9.3 Recovery in samples containing initial p24

When samples already contain p24 antigen and are subjected to matrix validation, the recovery is based on both the initial and the added (spiked) antigen [34], see the example values below.

Sample	p24 initial ng/ml	p24 added ng/ml	p24 found ng/ml	%recovery
1	<4	50	42	84
2	<4	50	53	106
3	<4	50	52	104
4	5	50	57	104
5	25	50	73	97

6.9.4 Matrix intervention

Matrix intervention, meaning the influence of the test sample components other than the analyte, is actually a frequent reason for incorrect assay results. Both false negative and false positive results may occur and therefore matrix validation should not be underestimated. Ion strength, pH, overall protein concentration, antibodies and cross reactions are factors that are known to contribute to incorrect assay results. Examples of other factors that contribute to the performance of an ELISA are effects of coagulants and the differences between plasma and serum, the sample collection (tube materials), sample storage conditions and the (unexpected) interference of unknown substances on the assay [33].

6.9.5 Replacement of critical reagents

In case critical reagents need to be replaced, requalification experiments should be performed, similar to the method verification Stage 3 (see section 6.8). However, when critical changes to the method are made, e.g. another source of antibody is used instead of using a different batch of antibodies, revalidation should also include specificity testing and verification of the range of the assay.

6.9.6 Recommended criteria and methods by the European Pharmacopoeia

For the quantitative immunoassay, specific validation criteria and validation methods are specified in the European Pharmacopoeia (EP) section 2.7.1. [29]. These requirements are strictly applicable for the industry producing medicines for human and veterinary use, however the concept is useful for a broader range of applications as well and therefore some attention is given to this. The essentials of the EP recommendations are schematically provided on the left side in the scheme below (column 'EP 2.7.1 requirement'). On the right side an interpretation of the validation parameters used throughout this report is provided (column 'validation parameter').

EP 2.7.1. requirement	**Validation parameter**
The antibody or antigen significantly discriminates between the test and standard. For a labelled reactant, the corresponding reactant significantly discriminates between the labelled and unlabelled compound.	Specificity
The method is not affected by the assay matrix, that is, any component of the test sample or its excipients, which can vary between samples. These may include high concentrations of other proteins, salts, preservatives or contaminating proteolytic activity.	Accuracy (spike recovery)
The limit of quantitation is below the acceptance criteria stated in the individual monograph.	Range, LLOQ
The variance of the results meets the requirements stated in the individual monographs.	Precision
The order in which the assay is performed does not give rise to systematic errors.	Accuracy, precision, robustness

7 POLYMERASE CHAIN REACTION: PERT ASSAY

7.1 Introduction

In this Chapter the 'Product enhanced reverse transcriptase assay' (PERT) assay is used as an example of how a reverse transcriptase quantitative real time Polymerase Chain Reaction (Q-PCR) can be validated for the detection of retrovirus in cell culture samples as part of the assay to detect replication competent retrovirus (RCL). Although some laboratories use the PERT assay to quantify retrovirus particles, in this Chapter only the validation process for use of the assay for its purpose to detect if – 'yes or no' – retrovirus is present in the sample, is described. The same assay is also known as 'F-PERT assay', where the F refers to 'fluorescence' being the signal measurement function [17].

A description of the assay is provided in the text box in section 7.2. Briefly, the assay starts with concentration and isolation of the retrovirus from cell culture fluid by ultracentrifugation. The virus pellet is lysed and a fraction is added to a reverse transcriptase reaction mixture containing a generic RNA template and primer. If RT is present, complementary-DNA (cDNA) is produced. In the final stage of the method, the cDNA is brought into a PCR reaction mixture and by repetitive cycling a fluorescent marker is incorporated.

The PERT assay is an example of a method in which the validity of the method is determined for a great deal by system suitability and assay acceptance criteria and this principle is largely explained in the section on development and standardisation.

It should be clear that all experimental facts and figures are for demonstrative purposes only and are not based on actual experiments. The experimental and validation set-up however is characteristic for a PCR method to detect reverse transcriptase in cell culture samples. Furthermore, this Chapter does not conclude whether or not the PERT is suitable for the detection of RT activity in cell culture samples.

It is assumed that the reader of this Chapter has basic knowledge of the PCR methodology and therefore the principles of the PCR assay will not be elaborated.

7.2 Validation parameters

Although the activity of the reverse transcriptase enzyme will be evaluated in the test by real time Polymerase Chain Reaction, for its intended purpose the PERT assay is not quantitative. The assay provides a 'yes or no' answer. The type of assay is a 'Limit test

for impurity‘ or a type III assay and for this purpose the relevant validation parameters are **selectivity (specificity)** and **limit of detection**, see Part I, Section 2.3, Figure 1 and Section 3.8, Table 2.

The selectivity of the assay is initially established during the development phase in Stage 2 of the validation process (Section 7.4.1), after the assay is standardized and also in Stage 3 of the process during qualification of the assay (See Section 7.5.2). The DL of the assay is determined in Stage 3 of assay, see Section 7.5.1.

Other validation parameters like accuracy and precision, as well as linearity and range, are not relevant, as the assay does not produce quantitative results.

Note: The purpose of the assay is not to determine the titre of RCL in the culture assay. Such an assay would be a Type II or Type IV assay (see Table 2) and requires further development and standardization as well as other validation parameters.

7.3 Stage 1. Selection and feasibility

7.3.1 Purpose and scope

The purpose of the method is to detect if a sample contains RCL. Typically instead of detecting RCL directly in the sample, first HEK293 cells are inoculated with the sample to achieve RCL amplification by serial passaging of supernatants. If RCL is present, the HEK293 cells will multiply RCL and produce virus particles in the culture medium. It should be emphasized that only infectious RCL will be amplified on the HEK293 cells. The entire method consists of the culture part followed by detection of amplified retrovirus.

The assay that is considered in the example in this Chapter is meant to demonstrate whether or not retroviral particles are excreted by the HEK293 cells. The PERT assay is therefore used as a detector assay. Because the samples are tested after amplification on HEK293 cells, the detection limit and specificity of the method will be primarily determined by the infectivity assay and not by the detection assay. Therefore elements of specificity and DL of the infectivity assay are not part of the scope of this Chapter.

The method in this example is applicable for cell culture fluid samples.

7.3.2 Method selection

All retroviruses contain the enzyme 'reverse transcriptase‘ (RT) and a 'reverse transcriptase quantitative polymerase chain reaction‘ (RT-Q-PCR) can be used to detect RT activity in the supernatant of HEK293 cells. This RT-Q-PCR method is also referred to

as 'Product enhanced reverse transcriptase assay with measure of fluorescence' abbreviated as PERT or F-PERT [17, 19].

HEK293 cells do not produce retrovirus particles. Furthermore, the RT-Q-PCR is an established technique for the detection of RT in culture samples, as is demonstrated in several scientific papers [17]. The assay is specific for viruses containing RT (mainly retroviruses) but also versatile, because it is independent of the genetic configuration in which retrovirus particles may appear. The test may detect both infectious and non-infectious particles which is acceptable for its purpose, because the selectivity is made by the HEK293 cells that will only replicate infectious virus. The method is capable to detect RT from a wide range of retroviruses. Based on this information, from a scientific point of view, the selected assay is suitable.

7.3.3 Laboratory, instrumentation and materials

A validation process is a well-structured process and requires adequate planning ahead. Before continuing the stage of method development and standardization, the instrumentation and supplies need to be in place. Therefore all the requirements are identified and assembled. Standardizing an assay without having all the critical materials in place is ineffective, as the standardization may need to be repeated.

Laboratory: Separate areas for preparation of reagents, sample treatment, PCR reactions and post-PCR treatment are required, in order to avoid the generation of false positive results. These area's should be equipped with dedicated pipettes and freezers and/or fridges for storage of the reagents that are handled in the specific area.

While the list of materials is prepared up front, during the development and standardization phase some modifications may be required and changes to the list may be applicable. However, when the development phase is completed, the list is not to be changed anymore.

Equipment: A list of critical equipment is made. All equipment is checked to be in proper condition and well maintained. It is essential that the correct functioning of equipment is verified before an assay is standardized and validated. This will avoid wasting time and resources, as the equipment surely contributes to the actual test results. The ultra-centrifuge with suitable rotor is installed. The PCR thermocycler is installed and its optical system is calibrated. Operational tests are performed to verify that the thermocycler and the software are working according to the manufacturer's specification. Heating blocks and pipettes are calibrated.

Materials:

HEK293 cells: The cells are revived from a cell bank with known history. The cells are verified not to produce reverse transcriptase produced by endogenous retroviruses or by contamination with other retroviruses.

RCL control virus: A stock with known titre, calibrated against a reference standard, e.g. MLV 4070A ATCC VR-1450. This stock will also be referred to as the RCL reference stock. For this purpose, the culture fluid from a HEK293 cell stock infected with a laboratory RCL strain can be taken. The material is clarified by low speed centrifugation to remove cell debris and aliquoted in 1.2 ml volumes and frozen at -70ºC. The infectivity titre of the stock is determined in an assay against a formal reference standard, e.g. MLV 4070A ATCC VR-1450 [40].

Reverse transcriptase standard: Purified RT-enzyme with known activity will be used as a positive control to monitor the RT reaction. For this purpose RT-MLV is used. It is commercially available and the activity (U/ml) is provided on the certificate of analysis.

Reagents: Primer sets, buffers and RNA template are selected based on both functionality and quality grade, and purchased. Quality certificates are evaluated and stored. Where commercial kits are used, the individual components are described and the quality is recorded. Certificates of analysis can often be obtained when the lot number is entered in the 'technical' or 'support' sections at the website of the manufacturer. Some reagents may require qualification experiments in order to verify the quality, e.g. it can be necessary to verify that the RNAse inhibitor is free of RT activity.

Disposables: A list of critical disposables is made. A material is considered critical when it comes (or is likely to come) into direct contact with samples, controls or reagents. The list should contain the relevant specifications with requirements like 'sterile', 'RNAse free', 'powder free', etc. Also the type of the material is relevant: e.g. glass, polystyrene, vinyl. It is recommended to include both the specifications and the manufacturer name and catalogue number on the materials list. This is helpful when the materials get depleted and need re-ordered during routine use, but also in case the supply is discontinued and alternative products or suppliers need to be found.

7.4 Stage 2. Development and standardization

7.4.1 Development and optimization experiments

The first experiments have the purpose to test the method and to verify and/or optimize the conditions. Concentrations, volumes, temperatures, incubation times, cycle program etc. as well as the sample preparation technique may require modification.

The specificity of the method is established by testing a panel of different samples, e.g. supernatants from cells known not to produce retrovirus particles as well as samples from cells known to produce retrovirus particles. High cell densities and contaminating cell debris can be sources of RNA, DNA, polymerases and other contaminants that may interfere with the assay. Therefore different sample types, or samples with different preparation protocols, will be tested as well to see how they possibly interfere with the assay.

To determine the optimum concentrations of the cDNA, PCR reagents and the optimal annealing temperature, different reagent concentrations are tested in combination with a temperature gradient and various incubation times. Taking into account that specificity of the assay is of utmost importance (both false positive and false negative reactions may not occur), much effort should be put into obtaining a low back ground value.

It is not recommended for negative controls to have a Ct value (Cycle threshold value) under the maximum cycle number, as this is likely to generate false positive results (for instance, if the cut-off value is 40, it is not recommended that the maximum number of cycles is 45). On the other hand, sufficient signal to noise ratio should be obtained, in order to generate sufficient signal.

Much emphasis is therefore put on the conditions of the positive control and on the suitability of the matrix in each specific test. Positive controls should include multiple dilutions to demonstrate that both a higher concentration of RT and a higher concentration of cDNA will lead to lower Ct values. Therefore a pure RT standard (see PC1 and PC2 in Table 5) as well as RCL reference virus, both in the absence and the presence of the sample (PC3, PC4, PC5 and HX1, HX2 in Table 5), should serve as positive controls

The PERT assay when used as an assay to demonstrate the absence of RCL in a sample, is an example of a 'limit test for impurities' with specificity and DL as relevant validation parameters. The validation process of the PERT assay described in this Chapter demonstrates the importance of assay feasibility and development as well as the importance of the system suitability controls in every run being part of the entire validation process.

PERT-assay: Test method.

Samples consisting of HEK293 cell culture fluid are clarified by low speed centrifugation, then passed through a 0.2 µm filter, and subsequently 1 ml samples are ultracentrifuged to sediment the retrovirus particles. The pellets are resuspended in 50 µl of lysis buffer to provide the samples for the reverse transcription (RT).

Subsequently an RNA template (HCV RNA) that can be converted to cDNA by RT, is provided in the reaction mixture. The cDNA is used for the PCR reaction. In the presence of target sequence DNA, PCR yields a fluorescent signal. The fluorescence is increased cycle after cycle and reaches a threshold at a specific cycle number, called Ct. For each sample a Ct value is obtained. The more DNA is present, the lower the Ct value.

For reverse transcription, an RNA template is used with primer HCVRev1. The primer/template is prepared by mixing HCVRNA with primer HCVRev1, heating at 95ºC and chilling the mixture on ice. An RNAse inhibitor is added to prevent degradation of the template by RNAse that can be present in the sample. The primer/template is added to a buffer with ingredients and 3 µl from the 50 µl of sample or control (=lysed pellet) is added. Reverse transcription is allowed to proceed (e.g. 37ºC/5 hours). After template digestion by RNAse, samples or controls are diluted at appropriate dilutions and amplified in 40 cycles by Q-PCR using primers HCVFor1 and HCVRev2 in a PCR-Thermocycler with fluorescence readout system. Software calculates the Ct values for every sample. A sample with a Ct value of less than the cut-off value (Ct=40) is considered positive.

7.4.2 Application of appropriate controls

The assay is controlled by two different sets of criteria, namely:

- Assay acceptance criteria, to demonstrate that the conditions of the test are valid. These controls are also called ’system suitability controls' or ’run controls'. Once established, these assay criteria need to be applied and fulfilled in every assay run [11].
- Assay analysis criteria, providing the criteria to conclude if the test sample is positive or negative.

See Table 5 for a list of controls in the assay.

7.4.3 Assay acceptance criteria (system suitability controls)

The assay consists of separate phases and therefore the validity of each of these three stages (i.e. virus pelleting and lysis, cDNA generation by RT and amplification, respectively) must be controlled by using appropriate controls that are specific and representative for each of the phases.

To demonstrate the absence in the reagents of contaminating DNA, which could result in false positive reactions, negative controls are included. Also, it has been reported that both RNAse inhibitor and Taq DNA polymerase may contain endogenous RT activity that can result in interfering background signals [17]. For that purpose, to verify the absence of interfering RT activity, additional negative controls should be included.

In Table 5, controls N1, N2 and N3 pertain to DNA transcription and amplification reactions, while negative controls N4, N5 and N6 are related to cell culture fluid, centrifugation and lysis of the samples. The N4 and N5 controls are dilutions of the cDNA prepared from the HEK293 negative control cell cultures. Control N6 is the undiluted cDNA sample.

Positive controls are included to demonstrate that the reverse transcriptase reaction conditions are suitable for providing cDNA template that can be used as DNA template for the PCR amplification. The use of RT-enzyme with known activity ensures that the RT reaction and the subsequent PCR reaction is performed correctly (PC1, PC2). The 'whole virus RCL` controls (PC3, PC4, PC5) demonstrate that the sample preparation of the virus-containing cell culture fluid (centrifugation, lysis) has functioned correctly The PC3 and PC4 controls are dilutions from the cDNA from the MLV stock, PC6 is the undiluted cDNA sample. The function of controls HX1 and HX2 is to demonstrate sample suitability.

Code	Description	Purpose	Criteria
N1	Primer/template without RNAse inhibitor	Negative control to demonstrate the absence of RT activity or DNA contamination in the Taq DNA polymerase (if present: N1 and N2 will be positive while N3 is negative).	Ct=40
N2	Primer/template with RNAse inhibitor	Negative control to demonstrate the absence of RT activity or DNA contamination in the RNAse inhibitor (if present: N2 will be positive, while N1 and N3 are negative).	Ct=40
N3	RNAse inhibitor without primer/template	Negative control to demonstrate absence of DNA contamination in the reagents other than RNA inhibitor (if present: N1, N2 and N3 will all be positive).	Ct=40
N4	HEK293 cells, cDNA dilution 1:50	Negative control to verify the absence of back ground value by the sample matrix and to demonstrate the absence of cross contamination introduced from the environment or the reagents during the preparation of the virus lysate.	Ct=40
N5	HEK293 cells, cDNA dilution 1:10	Negative control. Purpose as for N4.	Ct=40
N6	HEK293 cells, cDNA undiluted	Negative control. Purpose as for N4. The undiluted matrix occasionally provides some background noise. This is considered acceptable as long as the 1:10 and 1:50 diluted negative controls (N4, N5) are negative.	Ct≤40
PC1	RT MLV, 5 x 10^{-6} Units	Positive control. A concentration of purified reverse transcriptase (RT) near the detection limit of the assay. For purpose see PC2.	Ct≤40
PC2	RT MLV, 5 x 10^{-5} Units	Positive control. A concentration of RT above the detection limit of the assay. Demonstrates that the buffers, template, primers and reagents as well as the thermocycling program and read out is functioning correctly.	Ct<40
PC3	RCL (virus) stock, cDNA dilution 1:50	Positive control. The virus pelletting procedure has been efficient and the materials, and reagents during this process are free from inhibitors.	Ct<40
PC4	RCL (virus) stock, cDNA dilution 1:10	Positive control, as for PC3.	Ct<40
PC5	RCL stock, cDNA undiluted	Positive control, as for PC3.	Ct<40
HX1	Unknown test sample (9 volumes) diluted with 1 volume undiluted RCL (virus) stock. cDNA undiluted	Sample control. To demonstrate that the sample is capable to produce a positive result when RCL is present. Absence of inhibition.	Ct<40
HX2	cDNA from HX1. cDNA dilution 1:5	Sample control. As above for HX1.	Ct<40

Table 5. Assay controls, their function and the acceptance criteria for the PERT assay. Note: Controls NC and PC are system suitability controls. The inhibition reactions HX1 and HX2 reactions are sample suitability reactions and are performed for every unknown test sample included in the assay run.

7.4.4 Assay analysis criteria and sample suitability

Unknown test samples may possess interfering properties or background noise in the RT or PCR reaction in variable amounts. For this reason, inhibitory controls HX1 and HX2 are included and the requirement is that the Ct<40. Only samples that do not show inhibition are suitable to be evaluated. See the points to consider, Section 7.9.7, how to handle the situation when samples fail to meet the criteria.

Furthermore, the cDNA products from the unknown test samples are tested in three dilutions, namely undiluted ('neat' sample), 1:10 diluted and 1:50 diluted. If the PCR reaction is positive in a higher dilution of cDNA (=lower DNA concentration), but negative with a lower cDNA dilution (=higher DNA concentration), the sample is considered to interfere with the PCR reaction.

Clear criteria should be defined for different assays to conclude when results of the unknown sample are considered positive or negative. Only conclusions towards a positive/negative result may be drawn if both the assay meets the system suitability controls and the sample suitability controls are valid.

An example of clear assay analysis criteria for test samples in the PERT assay:

- A sample will be considered negative when Ct=40 with the 1:50, 1:10 and undiluted ('neat' sample).
- A sample will be considered positive when Ct<40 with at least one dilution and:
    - If Ct <40 with sample diluted 1:10, the Ct has to be <40 for undiluted sample
    - If Ct<40 with sample diluted 1:50, the Ct has to be <40 for sample diluted 1:10 and with undiluted sample.

7.5 Stage 3. Performance qualification

Stage 3 starts when the assay conditions are fixed and the suitability controls and assay acceptance controls provide consistent results.

For the purpose of the RT assay, which is a limit test to detect impurities, two parameters are important, namely **specificity** and **detection limit.** (Note: as described in Part I of this document, the detection limit is also called 'sensitivity' or 'analytical sensitivity').

7.5.1 Determination of the detection limit

The establishment of the LOD is required to provide a reference value for the revalidation and method validation purpose of this qualification experiment. Also, the LOD is required

to define the concentration of RT-enzyme used in the suitability control (PC1 and PC2 in Table 5). In order to measure that the routine PERT assay is running according to the validated conditions, the RT-enzyme is used in known concentrations near and above the LOD.

Based on the assay results in the development runs, already an expectation of the limit of detection of the assay is obtained. The LOD is determined by testing concentrations of RT-enzyme in a standard curve in order to provide the lowest concentration that repeatedly provides positive results. A sample panel is prepared by accurately diluting the RCL reference stock with known titre in HEK293 culture medium. The dilutions made are: 0 - 10 - 50 - 100 - 500 RCL/ml. Multiple aliquots are frozen at -70ºC until testing.

Each sample taken in duplicate from the panel is also tested in duplicate in three independent assay runs. In this context duplicate means that duplicate aliquots of each sample are processed independently, starting with the centrifugation step. The LOD is the lowest concentration at which all samples are positive in each of the three assay runs. A summary of a series of validation experiments with possible results is provided in Table 6.

RCL/ml		Test 1		Test 2		Test 3	
	cDNA dil.	Ct1, Ct2	Result	Ct1, Ct2	Result	Ct1, Ct2	Result
0	Neat						
	1:10	All 40	negative	All 40	negative	All 40	negative
	1:50						
10	Neat			40, 35		35, 40	
	1:10	All 40	negative	40, 31	positive	40, 40	positive
	1:50			40, 40		40, 40	
50	Neat	28, 40					
	1:10	26, 40	positive	All <40	positive	All <40	positive
	1:50	35, 40					
100	Neat						
	1:10	All <40	positive	All <40	positive	All <40	positive
	1:50						
500	Neat						
	1:10	All <40	positive	All <40	positive	All <40	positive
	1:50						

Table 6. The results of three validation experiments in which a panel with 0-10-50-100-500 RCL/ml was tested in duplicate. The results Ct1 and Ct2 are from duplicate samples respectively, tested in three dilutions of the cDNA obtained after the RT reaction. The LOD is 50 RCL/ml, based on the fact that at this concentration all 3 experiments were positive. (System suitability controls were valid, results from negative and positive control are excluded in this table). Note that a sample is considered positive if one of both replicates is positive.

Assuming that the preliminary experiments during the standardization phase showed that dilutions containing 50 RCL/ml or higher were positive in the PERT assay and that dilutions containing 10 RCL/ml sometimes were positive and sometimes were negative, the LOD will probably be lower than 50 RCL/ml and higher than 10 RCL/ml. This is equivalent to 3 RCL and 0.6 of RCL per reaction respectively, because virus present in 1 ml of culture medium is concentrated to 50 µl in the lysate and from that lysate, 3 µl is tested in the RT reaction (see test method, Section 7.4.1).

7.5.2 Specificity

For the PERT assay, the specificity of the reaction is directed by the action of reverse transcriptase that is specific for retroviruses. For PCR assays, often the specificity of an assay is already demonstrated in the development phase of the assay (see Stage 2). It is generally known that reverse transcriptase is present in all retroviruses. Furthermore, the system suitability controls and the sample suitability controls are demonstrating specificity of the assay. If the PERT assay is used only for the detection of a known and specific strain-derived RCL, and this strain is similar to the positive control strain, the system suitability controls and sample suitability controls together may be sufficient to demonstrate specificity of the assay.

Many laboratories prefer additional specificity data to demonstrate that not only related RCL strains will be detected, but also other retroviruses. To demonstrate this, a panel of cells known to produce retrovirus may be selected, as well as a panel of cells known to be negative for retrovirus. For example, HIV-1, HIV-2 and Equine Infectious Anaemia Virus strains as well as C-type retrovirus like amphotropic Murine Leukaemia Virus can be included as examples of RCL strains. The panel is tested in the PERT assay and expected to test positive for retrovirus. It is not necessary to determine the LOD with each virus strain from the panel.

7.6 Stage 4. Performance validation

In this particular assay, where extensive system suitability and sample suitability controls are included in a each assay run, further validation experiments are not required. Assay validation in terms of robustness would come into consideration if the assay was used for 'release assays' of clinical materials, or if the PERT assay would be used to quantify retroviruses. In the example used for this validation exercise, this is not the case. Therefore further experiments in this stage are not performed.

7.7 Stage 5. Maintenance and improvement

Once the assay is routinely performed, a database is designed in which the dates of testing, and the Ct values from controls PC1 till PC5 and the negative controls N1 till N6 (see Table 5) from every test are stored. The performance of the assay is monitored over time. In order to trend the failure rate of the assay, also tests that were invalid may be relevant, however care should be taken to interpret values from invalid assays in relation to trending. If new reagents are used, or new technicians have been qualified, this may be recorded in a 'remarks' section of the database as well.

For practical reasons the entry of data into the data base preferably is done directly after completion of the assay and it can even be part of the SOP for the assay.

Incidental failures of an assay may happen. However, when repeated failures are observed that do not have a clear reason, an investigation is recommended in order to find the reason for the failures. Modification or improvement of the method could be required. If significant changes are implemented, the LOD of the assay may need to be confirmed using the RT-enzyme standard.

7.8 Method verification

There are several reasons why method verification is relevant. Below are some examples of situations when method verification is applicable and what parameters should be tested.

- New RCL reference stock. In case the RCL (virus) reference stock gets depleted, the titre of the a new RCL stock should be qualified. The new stock is tested in triplicate along with the old stock.
- New equipment. If new equipment or software is installed, a study is performed to confirm the LOD of the assay with the RT-enzyme stock.
- Operator qualification. Operators that are getting trained to perform the assay without supervision, must demonstrate that they are capable to perform the entire assay, including the sample preparation. To confirm that the training was successful, a panel with samples, including samples prepared with RCL titres at the LOD should provide valid results.
- Transfer to another laboratory. The cut-off value should be demonstrated and the LOD, as determined in Stage 3, should be confirmed.

7.9 Points to consider

7.9.1 Equipment

It is essential that the same critical equipment that is used for the validation of the test is also used during the actual test. Critical equipment is e.g. the thermocycler and related software. Most laboratories use dedicated pipettes.

7.9.2 Ultracentrifuge tubes

Ultracentrifuge tubes may affect the recovery of the virus pellets. Also, components derived from the tubes can interfere with the enzymatic reactions or tubes can contain DNA, RNA, digesting enzymes or inhibitors. It is relevant to specify the types of tubes in the standard procedure.

7.9.3 Aliquoting reagents

Reagents should be aliquoted and stored for single use as much as possible, in order to avoid multiple thaw and freeze cycles and to prevent them from getting contaminated during use.

7.9.4 Source of reagents

The source of the critical reagents and the RT-enzyme (manufacturer, specifications) is carefully recorded and when the stock gets depleted, the new stock from the same quality should be obtained. It should be calibrated against the old stock in a verification assay (see section 7.8).

7.9.5 LOD cannot be based on theoretical values

For an assay considered validated, a claim like 'The PERT assay is a very sensitive assay. It can detect as little as 10^{-9} units of MLV-RT, corresponding to 2.1 X 10^{2} molecules, a number present in 3-11 virions' –as was made in one of the first articles describing the PERT assay [17] – should only be made if this is actually verified in validation experiments. In this particular example, the sensitivity was calculated based on theoretical values; however, no evidence was provided that these criteria were actually valid.

7.9.6 A low LOD is not always required

It can be questioned whether an assay with a very low detection limit is useful as a read-out assay after virus amplification, where considerably high virus titres are expected. In general, assays with low detection limits are more likely to produce false positive results,

which can be very annoying and costly, and therefore require emphasis on robustness during the validation. This in particular is the case for laboratories that perform testing of materials from different origin.

7.9.7 What if samples are inhibitory?

In case a sample is inhibitory in the PERT assay while all other controls are satisfactory (Ct <40 for values of HX1 and/or HX2, see Table 5), the cell culture sample can be diluted e.g. 1: 5 or 1:10 prior to centrifugation. This affects the LOD, however this can be acceptable as long as the impact is understood.

7.9.8 Use of the PERT assay for quantitation of LV particles

The suitability of the PERT assay to quantify LV particles has its limitations. Due to the fact that incomplete or empty particles can be present, the ratio between (infectious) particles and RT activity may vary considerably between virus batches. Furthermore, the infectious virus titre on one cell line may be significantly (several logs) higher than on another cell line [4], either by difference in susceptibility of the cells, or by the capability of cells to produce virus particles in the cell culture fluid. In the section 'Lentiviral vector particle quantification' in the EMEA guideline on development and manufacture of lentiviral vectors [13], some methods are provided to quantify RT activity and how to relate those to the number of particles.

The PERT assay is a generic assay for detection of a broad range of retrovirus particles. A quantitative polymerase chain reaction performed on the vector RNA molecules is an alternative method to estimate the number of LV particles. Unlike the PERT assay, in such an assay the target is the RNA molecule instead of the RT-enzyme. An example is a validation study described by Ebeling et al. for the measurement of retrovirus particles by a quantitative *env*-RT-PCR [4]. Like for the PERT system, proper system assay controls are extremely important to minimize interference in the assay caused by contaminants.

7.9.9 Validation of other PCR assays

Some literature references on the validation of PCR assays for other viruses are included in Chapter 10.

The European Pharmacopoeia (EP) in section 2.7.1. [11] provides general guidance for the validation of PCR assays with a practical example for a HCV PCR assay. In the EP section 2.6.7. [30], a guidance with concrete recommendations for the validation of a mycoplasma PCR method is provided. Specificity, DL and robustness are parameters that are advised to be included for the method for the detection of mycoplasma, being a

possible contaminant in cell banks and virus seeds. This section also provides useful guidelines for a comparability study for the replacement of a culture method by an alternative technique (nucleic acid amplification).

Another example of relevant literature is a publication by Pang et al. [16] about the validation of a PCR assay for the detection of Norovirus –a virus that does not grow in cell lines– in stool samples. In this article the validation of the method is described and the performance of the PCR analysis is compared with results obtained by electron microscopy.

A validation of a PCR test for the detection of replication competent adenovirus in adenovirus vector product is described by Watabe et al., exemplifying a combination of a culture assay and readout by a PCR test [9].

PART III: GENERAL ISSUES RELATED TO ASSAY VALIDATION

Chapter 8 contains general remarks and recommendations related to assay validation. Lists with terms and definitions as well as abbreviations used in both Part I and Part II are included in Chapter 9. Finally, Chapter 10 contains a list of references to literature and websites.

8 GENERAL REMARKS AND RECOMMENDATIONS

8.1 Reference standards

It is important that the quality of the reference standard is ensured, as the quality affects the outcome of the analysis. Therefore, reference standards are preferably obtained from an authentic and traceable source. Certified standards such as compendial standards (WHO, EDQM), commercially available standards, or fully characterised standards prepared in-house or by an external non-commercial organisation can be suitable. Irrespective of the supplier, it is recommended to obtain a certificate of analysis or other relevant information to ensure quality, stability, storage conditions, expiration date, batch number and purity of the reference standards [27].

8.2 Training

Considering the fact that personnel is an integral part of the analytical method, only qualified personnel should perform routine testing. When new technicians are trained, it is recommended to complete their training process by repetition of precision and accuracy experiments as performed in Stage 3 of the validation process. Technicians can be considered qualified by demonstration that the assays they performed met the validation criteria. If the criteria are not met, additional training followed by requalification is necessary.

8.3 Contract Research Organisations (CROs)

In case a decision is made not to perform the assay 'in-house' but to send samples out for testing by contract laboratories, an assessment should be made whether or not these laboratories are using a validated assay.

A test performed at a CRO is not considered the same as using a commercial assay.

Even if a laboratory is licensed to perform studies under Good Laboratory Practice (GLP), it is important to verify that the intended assay is validated for its particular use and to check if the protocol contains the correct suitability controls. That said, GLP compliance of a contract laboratory is more strongly related to organizational aspects rather than to scientific aspects [18].

For example, in an article about the validation of RCL detection in which the PERT assay was used as a readout system, false positive results were obtained from a GLP testing laboratory [20]. Validation data and system suitability controls related to the PERT assay were unfortunately not reported, which could have been interesting and relevant information for this report. It would have been interesting to see whether inadequate validation and/or system suitability may have contributed to or caused the false positive results.

Upon request of a sponsor, CROs should be able to show documented evidence (=validation reports) of validation studies. Unless part of the business contract, it can not be expected that CROs provide comprehensive validation reports, because such report are often considered as intellectual ownership or property and kept confidential. Nevertheless, validation reports should be available for auditing on site for sponsors.

9 GLOSSARY

9.1 Terms and definitions

ACCURACY	The closeness of agreement between a test result and the accepted reference value. Synonym of 'trueness'. Dutch: '*juistheid*'.
BIAS	The difference between the expectation of the test results and the accepted reference value.
CALIBRATION CURVE	Graphical representation of measuring signal as a function of quantity of analyte.
CROSS REACTIVITY	Response (of a method) to analogous, metabolites, or other non-target components that may be present in the matrix(es).
DETECTION LIMIT	Same as limit of detection. The lowest amount of analyte that can be detected and identified, but not necessarily quantitated as an exact value. Or: the lowest analyte content, if actually present, that will be detected and identified.
ERROR OF MEASUREMENT	The result of a measurement minus the true value of the measured quantity.
FITNESS FOR PURPOSE	Degree to which data produced by a measurement process enables a user to make technically and administratively correct decisions for a stated purpose.
GLP	Good Laboratory Practice, meaning the principles as defined by the Organization for Economic Co-operation and Development (OECD) applicable for non-clinical safety testing.
LIMIT OF DETECTION	See Detection Limit.
LIMIT OF QUANTITATION	The highest and lowest concentrations of an analyte that can be determined with acceptable precision (repeatability) and accuracy under the stated conditions of the test.

LINEARITY	The ability of the method to obtain test results proportional to the concentration of the analyte.
METHOD VALIDATION	The process of establishing the performance characteristics and limitations of a method and the identification of the influences which may change these characteristics and to what extent. Verification process of a method to ensure it is fit for purpose, i.e. suitable for use in solving a particular analytical problem.
PRECISION	The closeness of agreement between independent test results obtained under particular conditions (degree of scatter). Dutch: '*herhaalbaarheid'*.
QUALIFICATION	Documented experiments demonstrating that a specific process will meet its pre-determined acceptance criteria.
RANGE	The interval between thee upper and the lower concentration (amounts) of analyte in the sample for which the analytical procedure has a suitable level of precision, accuracy and linearity.
RECOVERY	The amount of analyte that was added deliberately to a test sample (spike sample) prior to analysis compared to the fraction of the added amount that was measured in the spike sample.
REFERENCE MATERIAL	Material or substance whose property values are sufficiently homogenous and well established to be used for the calibration of an apparatus or used as a control sample during a measurement.
REPEATABILITY	Closeness of agreement between the results of successive measurements, carried out in the same conditions of measurement.
REPRODUCIBILITY	Conditions where test results are obtained with the same method on identical test items in different laboratories with different technicians using different equipment.
RESULT OF A MEASUREMENT	Value attributed to a measurement, obtained by the measurement.

ROBUSTNESS A measure of the capacity of an assay to remain unaffected by small, but deliberate alterations in method parameters, providing an indication of the reliability of the measurement.

SELECTIVITY The ability of a method to determine accurately and specifically determine the analyte of interest in the presence of other components in the sample matrix.

SENSITIVITY The change in the response of a measuring instrument divided by the corresponding change in stimulus. Not to be confused with Detection Limit (Limit of Detection).

SPECIFICITY The ability of a method to measure only what is intended to measure.

STANDARD DEVIATION A measure of how values are dispersed about a mean in a distribution of values.

TRUENESS The closeness of agreement between the average value obtained from a large set of test results and an accepted value (see Accuracy).

VALIDATION Confirmation by examination and provision of objective evidence that the particular requirements of a process (or method) for a specified intended use are fulfilled. Documented testing which demonstrates that a process consistently operates according to predetermined acceptance criteria.

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

9.2 Abbreviations

ATCC	American Type Culture Collection
cDNA	Complementary-DNA
BMV	Bioanalytical Method Validation
CPE	Cytopathic Effect
CRO	Contract Research Organization
Ct	Cycle Threshold
CV	Coefficient of Variation
DL	Detection Limit
DNA	Deoxyribonucleic Acid
EDQM	European Directorate for the Quality of Medicines
OECD	Organisation for Economic Co-operation and Development
ELISA	Enzyme-Linked Immunoassay or Enzyme-Linked Immunosorbent Assay
EMA	European Medicines Agency
EMEA	European Medicines Agency *(currently named EMA)*
EP	European Pharmacopoeia
FBS	Foetal Bovine Serum
FDA	Food and Drug Administration (United States)
F-PERT	Product Enhanced Reverse Transcriptase with Measure of Fluorescence *(=same as 'PERT')*
GLP	Good Laboratory Practice. *(*Note *For the explicit meaning of GLP see previous section, 'Terms and definitions')*
GMO	Genetically Modified Organism
GMP	Good Manufacturing Practices
HCV	Hepatitis C virus

HIV	Human Immunodeficiency Virus
HLOQ	Higher Limit of Quantitation
HSV	Herpes Simplex Virus
ICH	International Conference on Harmonisation
LOD	Limit of Detection
LOQ	Limit of Quantitation
LLOQ	Lower Limit of Quantitation
LV	Lentivirus
MLV	Murine Leukaemia Virus
NC	Negative Control
OD	Optical Density
OIE	The World Organization for Animal Health (*started as Office International des Epizooties*)
PC	Positive Control
PCR	Polymerase Chain Reaction
PERT	Product Enhanced Reverse Transcriptase
PFU	Plaque Forming Unit
Q-PCR	Quantitative Polymerase Chain Reaction
RCL	Replication Competent Lentivirus
RNA	Ribonucleic Acid
RSD	Relative Standard Deviation
RT	Reverse Transcriptase
SD	Standard Deviation
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SK	Spaerman-Kärber
SOP	Standard Operation Procedure
Taq	*Thermus aquaticus*

$TCID_{50}$	Tissue Culture Infectious Dose 50%
WHO	World Health Organization

10 REFERENCES

1 ICH Harmonised Tripartite Guideline; Validation Of Analytical Procedures: Text And Methodology; ICH Q2(R1), November 2005. (http://www.ich.org)

2 Guidance for Industry; Bioanalytical Method Validation; FDA/CDER/CVM; May 2001. (http://www.fda.gov)

3 The Fitness for Purpose of Analytical Methods; A Laboratory Guide to Method Validation and Related Topics, EURACHEM Guide, 1998. (http://www.eurachem.org)

4 Ebeling et al. Human primary T lymphocytes have a low capacity to amplify MLV-based amphotropic RCR and the virions produced are largely non-infectious. Gene Therapy 10: 1800-1806 (2003).

5 Accreditation for Microbiological Laboratories. EA/Eurachem Working Group, June 2002. (http://www.eurachem.org)

6 A WHO guide to good manufacturing practice (GMP) requirements. Part 2: Validation. WHO Geneva, 1997. (http://www.who.int/vaccines-documents)

7 WHO Expert Committee on Specifications for pharmaceutical preparations, Validation of analytical procedures used in the examination of pharmaceutical materials. Technical report series No. 823 Annex 5, WHO Geneva, 1992. (http://www.who.int/vaccines-documents)

8 Note for guidance on virus validation studies: The design, contribution and interpretation of studies validating the inactivation and removal of viruses. EMEA, London, 1996. CPMP/BWP/268/95. (http://www.ema.europa.eu/)

9 Ishii-Watabe et al. Detection of Replication-Competent Adenovirus Spiked into Recombinant Adenovirus Vector Products by Infectivity PCR. Molecular Therapy Vol. 8: 1009-1016 (2003).

10 United States Pharmacopoeia (USP); Design Analysis of Biological Assays <11>; Validation of Compendial Methods <1225>.[a]

11 European Pharmacopeia, 6th Edition, EDQM Strasbourg, Section 2.6.21. (Nucleic Acid Amplification Techniques, section 7, Quality Assurance).[a]

12 Methods, Method verification and Validation. Ora Laboratory procedure, FDA. ORA-LAB.5.4.5, revision 2005. (http://www.fda.gov)

13 Guideline on Development and Manufacture of Lentiviral vectors. EMEA CHMP/BWP/2458/03 (2005). (http://www.ema.europa.eu/)

14 The World Organization for Animal Health (OIE) Manual of Diagnostic Tests and Vaccines for Terrestrial Animals, Chapter 1.1.3; Principles of validation of diagnostic assays for infectious diseases. (http://www.oie.int/fr/normes/mmanual/A_00013.htm)

15 The World Organization for Animal Health (OIE) Manual of Diagnostic Tests and Vaccines for Terrestrial Animals, Chapter 1.1.4; Validation and quality control of polymerase chain reaction methods for the diagnosis of infectious diseases. (http://www.oie.int/fr/normes/mmanual/A_00014.htm)

16 Pang et al. Evaluation and validation of Real-time reverse transcription-PCR assay using the lightcycler system for detection and quantitation of Norovirus; Journal of clinical microbiology, (2004) p.4679-4685.

17 Pyra et al. Ultrasensitive retrovirus detection by a reverse transcriptase assay based on product enhancement; PNAS (1994), p1544-1548.

18 OECD Principles of Good Laboratory Practice, Paris (1998). (http://www.oecd.org)

19 Wilson et al. Evaluation of recommendations for Replication Competent Retrovirus Testing Associated with Use of Retrovirus Vectors. Human Gene Therapy 8:869-874 (1997).

20 Miskin et al. A replication competent lentivirus (RCL) assay for equine infectious anaemia virus (EIAV)-based lentiviral vectors. Gene Therapy 13, 196–205 (2006).

21 C. Spearman. The method of 'right and wrong cases' ('constant stimuli') without Gauss's formulae. British Journal of Psychology 2:227-242 (1908).

22 G. Kärber. Beitrag zur kollektiven Behandlung pharmacologischer Reihenversuche. Archiv für experimentelle Pathologie und Pharmakologie. 162:480-487 (1931).

23 Van Geloven et al. Regression analysis on serial dilution data from virus validation robustness studies. Statistica Neerlandica. 63:245-257 (2009).

24 QuickTitre™ Lentivirus Titre Kit, Cell Biolabs Product Manual VPK-107.

25 Van Weemen and Schuurs. Immunoassay using antigen-enzyme conjugates.

FEBS Letters 15(3): 232–236.9 (1971).

26 Engvall and Perlmann. Enzyme-linked immunosorbent assay (ELISA) quantitative assay of immunoglobulin G. Immunochemistry 8(9):871-874 (1971).

27 Draft Guideline on Validation of Bioanalytical Methods. EMEA/CHMP/EWP/192217/2009 (2009). (http://www.ema.europa.eu/)

28 European Pharmacopeia, 6th Edition, EDQM Strasbourg, Section 5.1.6. (Alternative methods for control of microbial quality).[a]

29 European Pharmacopeia, 6th Edition, EDQM Strasbourg, Section 2.7.1. (Immunochemical methods).[a]

30 European Pharmacopeia, 6th Edition, EDQM Strasbourg, Section 2.6.7 (Mycoplasmas).[a]

31 ICH Harmonised Tripartite Guideline; Viral Safety Evaluation of Biotechnology Products Derived From Cell Lines of Human or Animal Origin. ICH Q5A(R1); September 1999. (http://www.ich.org)

32 Guidance for Industry: Guidance for Human Somatic Cell Therapy and Gene Therapy. CBER (FDA) 1998 (http://www.fda.gov)

33 WHO Technical report series No. 658 Annex 5, Requirements for biological substances No.7, Addendum 1980. (http://whqlibdoc.who.int/trs/WHO_TRS_822.pdf)

34 Manual Perkin Elmer HIV-1 p24 ELISA; kit For ICD and Detection of HIV-1 p24 Antigen in Serum/Plasma and Cell Culture Supernatant. (https://las.perkinelmer.com/content/Manuals/MAN_HIV1p24ELISA.pdf)

35 Gesundheitsschutz 2008 51:937 – 945. DOI 10.1007/s00103-008-0615-5. Springer Medizin Verlag 2008.

[a] The contents of the United States Pharmacopeia and the European Pharmacopoeia are subjected to copyright and the text is not available on internet. Pharmaceutical companies usually have subscriptions, and the recent editions can be consulted in the library, at the Quality Control department and/or at the office of the 'Qualified Person' or pharmacist.

Note: Most references can be found on internet by entering the entire title of the document in a search engine like Google.

About the content

A practical manual about assay validation, primarily intended for scientists and government officials involved in either authorizing or inspecting laboratory experimental work with genetically modified organisms. The content is also useful for laboratory personnel performing biological assays in Quality Control and Research & Development laboratories. A description of the different stages of the validation process is given, followed by an explanation of validation parameters. It shows that validation actually is a continuous process and explains why assay controls and documentation are essential in this process. From three types of assays practical examples are given for both new and already existing assays. This document has been originated with close monitoring of a steering committee consisting of scientists, biosafety officers and governmental officers.

About the author

Joke Ederveen (1958) worked for nine years at the virology department of veterinary faculty of the Utrecht University before switching to the pharmaceutical industry in 1990. As study director she was responsible for microbiological and virus assays in a pharmaceutical company and later at a contract research organization. In 2003 she joined Progress-PME to work as project manager and advisor in multiple, diverse projects related to biopharmaceutical commercial production, development, quality and biosafety.

About Progress-PME

A company with experts offering consultancy, project management and interim-management in the (bio)pharmaceutical industry. www.progress-pme.nl.