



ISTA Method Validation for Seed Testing

Any hardcopy of this document is not controlled copy

CREATED	REVIEWED	APPROVED
2006	2006	30.11.2006
TCORD	TCOM	ECOM

Acknowledgements:

The ISTA Method Validation Working Group is grateful to the ISTA Seed Health Committee for permission to use their Handbook of Method Validation as the framework for this publication.

CONTENTS

1. SCOPE

- 1.1 Introduction
- 1.2 What is Method Validation?
- 1.3 Why is Method Validation Necessary?
- 1.4 Professional Responsibilities
- 1.5 How Should Methods be Validated?
- 1.6 Method Validation in ISTA

2. OBJECTIVES OF THE ISTA METHOD VALIDATION PROGRAMME

- 2.1 Develop a Validated Test Method for a New Species
- 2.2 Compare Different Methods to Choose a Test Method to Validate
- 2.3 Improve a Validated Test Method(s)
- 2.4 Compare the Present Validated Test Method with a New Method(s)

3. COLLABORATIVELY VALIDATED METHODS

- 3.1 Scope
- 3.2 Purpose
- 3.3 Parties Involved in Collaborative Method Validation and Their Responsibilities
- 3.4 Procedure
- 3.5 Related Procedures

4. PERFORMANCE VALIDATED METHODS

- 4.1 Scope
- 4.2 Purpose
- 4.3 Parties Involved in Performance Validation and Their Responsibilities
- 4.4 Procedure – Test Kit Methods
- 4.5 Procedure – Non-Test Kit Methods
- 4.6 Related Procedures

5. MODIFICATIONS TO ISTA RULES

- 5.1 Method Modification Categories
- 5.2 Validation Requirements for Method Modifications
- 5.3 Examples of Editorial Changes that Correct or Clarify Language

6. VALIDATION OF METHODS FOR TESTING FOR THE PRESENCE OF SPECIFIED TRAITS

- 6.1 Introduction
- 6.2 Previously Validated Methods
- 6.3 Inter-Laboratory Validation for a New or Modified Method
- 6.4 Single Laboratory Validation for a New or Modified Method

1. SCOPE

1.1 Introduction

ISTA, as an international organisation, has a goal of uniformity in seed testing. The need for seed testing methods that are reliable and reproducible among member laboratories is therefore a basic need for ISTA; this is achieved through The International Rules for Seed Testing (ISTA Rules). Before being accepted into the ISTA Rules, most test methods have gone through collaborative study among laboratories to ensure that the test procedure gives reliable and reproducible results in accordance with the given specifications of the test method. However, in the past, ISTA has not always had a consistent system for the introduction of seed test methods into the ISTA Rules, the process varying depending on which of ISTA's Technical Committees put forward the Rules Proposal. This problem was first addressed by the ISTA Seed Health Committee, who in 2000 produced their 'Handbook of Method Validation for the Detection of Seed-Borne Pathogens'. In 2002 the ISTA Executive Committee decided that method validation should apply to all seed quality testing, not just tests for seed health.

Much of the information for this Scope section has been sourced from The Eurachem Guide (The Fitness for Purpose of Analytical Methods).

1.1.1 Goal of the ISTA method validation programme

The goal of ISTA's method validation programme is to provide a system which allows the validation of new test methods, the comparison of equivalent test methods, and the maintenance (i.e. review) of existing test methods for the evaluation of seed quality.

1.1.2 Good laboratory practice

The concept of good laboratory practice is based on six principles of analytical practice which, taken together, are considered to constitute best practice. As applied to seed testing the six principles are:

- Seed quality testing should be carried out to satisfy an agreed requirement (i.e. to a defined objective).
- Seed quality testing should be performed using test methods and equipment that have been tested to ensure they are fit for purpose (i.e. validated).
- Staff performing seed quality testing should be both qualified and competent to undertake the task.
- There should be a regular independent assessment of the technical performance of a laboratory.
- Organisations performing seed quality testing should have well defined quality control and quality assurance procedures.

1.1.3 Purpose

The purpose of this publication is to:

- Increase readers' understanding of what is involved with method validation and why it is important.
- Provide details of how method validation can be achieved.

This publication is expected to be essential reading for members of ISTA's Technical Committees and for anyone else involved in developing new seed test methods (both ISTA members and non-ISTA members). Others may find it useful as a source of background information. Method Validation involves the collection and analysis of

data. An overview of statistical procedures to be followed in analysis of method validation trials is given in Appendix 1.

The most important terms used in this book are defined in the Glossary. ISO and IUPAC definitions have been provided wherever possible.

The next sections of this Scope provide answers to the following questions:

- What is method validation?
- Why is method validation necessary?
- How should methods be validated?

and outline where method validation fits in the process of inclusion of new or improved test methods into the ISTA Rules.

1.2 What is Method Validation?

Validation is defined by ISO as '*Confirmation by examination and provision of objective evidence that the particular requirements for a specified intended use are fulfilled*'. This can be interpreted for method validation as being the process of defining an analytical requirement, and then confirming that the method under consideration can measure what is required with suitable accuracy. Implicit in this is that it will be necessary to evaluate the method's performance capabilities i.e. accuracy, reproducibility, repeatability.

Method validation is usually considered to be very closely tied to method development. Indeed it is often not possible to determine exactly where method development finishes and validation begins. Many of the method performance parameters that are associated with method validation are in fact usually evaluated, at least approximately, as part of method development.

The ISTA Method Validation Programme therefore:

- Is a critical examination of a seed quality test to ensure that the description of the method is clear and complete, and that the procedure gives accurate, reproducible and repeatable results in accordance with the given specifications of the test method.
- Where appropriate, confirms the relationship between the results of a quality test and a practical expression of seed quality.

It is implicit in the method validation process that the studies to determine method performance parameters are carried out using equipment that is within specification, working correctly, and adequately calibrated.

1.3 Why is Method Validation Necessary?

Seed quality tests are made every day in laboratories around the world. There are different reasons for making these measurements. For example, as a way of valuing goods for trade purposes; supporting regulatory programmes; checking the quality of seed for planting and/or storage; in-house measure of quality assurance for the seed trade.

There is a cost in carrying out any testing, and important decisions will be made based on the results. For example, tests showing seed to be unfit for purpose may result in legal action and compensation claims. Clearly it is important to determine the correct result and be able to show that it is correct within the achievable and/or acceptable limits of variation.

1.4 Professional Responsibilities

If the result of a test cannot be trusted, then it has little value and the test might as well have not been carried out. When a customer commissions seed quality tests from a laboratory, it is assumed that the laboratory has a degree of expert knowledge that the customer does not have. The customer expects to be able to trust results reported and usually only challenge them when a dispute arises. Thus the laboratory and its staff have a clear responsibility to justify the customer's trust by providing the right answer to the analytical part of the problem, in other words, results that have demonstrable 'fitness for purpose'. Implicit in this is that the tests carried out are appropriate for the analytical part of the problem that the customer wishes solved, and that the final report (which for ISTA is the International Seed Analysis Certificate) presents the analytical data in such a way that the customer can readily understand it and draw appropriate conclusions. Method validation enables analysts completing quality tests to demonstrate that a method is 'fit for purpose'.

For a seed quality test result to be 'fit for purpose' it must be sufficiently reliable and reproducible so that any decision based on it can be taken with confidence. Thus the method performance must be validated and the uncertainty of the result, at a given level of confidence, estimated. Most of the information required to evaluate uncertainty can be obtained during validation of the method. For many of its existing test methods (e.g. germination, purity, other seed determination, vigour), ISTA uses tolerance tables rather than quoting levels of uncertainty. Whenever possible, new ISTA test methods should also use the tolerance table approach.

Regardless of how good a test method is and how skilfully it is used, a seed quality problem can be solved by the analysis of samples only if those samples are appropriate to the problem. Taking appropriate samples is a skilled job. For reporting results on the ISTA Orange International Seed Lot Certificate the ISTA Accreditation Standard requires sampling of seed lots to be under the control of the Accredited Laboratory. When the sampling is not under the control of the laboratory, results of analysis will need to be reported on the basis of the samples as received, and the report should make this distinction clear. ISTA achieves this by requiring such results to be reported on the ISTA Blue International Seed Sample Certificate.

1.5 How Should Methods be Validated?

1.5.1 Who carries out method validation?

The ISTA Rules provide laboratories with a large number of already validated methods which have been developed for wide-ranging use as a published standard procedure. Many of these have been through multi-laboratory collaborative studies, and **this is still the preferred way of carrying out the validation within ISTA** (Collaboratively Validated Methods, Table 1). These validation studies are usually organised by an ISTA Technical Committee (see Table 1).

However, the traditional ISTA comparative test can be slow, time consuming and therefore costly, and in some situations may not be required (for example: the addition of a new species to an existing method). In recognition of this, ISTA has introduced its Peer-Validation Programme (Table 1) for the validation of methods by laboratories working with only one or two others in order to improve efficiency and reduce costs, but still maintain an appropriate level of confidence in the validation process. Again these validation studies are usually organised by an ISTA Technical Committee (see Table 1).

1.5.2 What degree of validation is required?

Method validation can be an expensive process and inevitably it may be constrained by time and cost considerations. Starting with a carefully considered analytical specification provides the base on which to plan the validation process. The test organiser (see Table 1) should also take into account any constraints imposed; for example customer requirements, existing experience of the method, and the need for compatibility with other similar methods already in use within a laboratory or used by other laboratories. Some of the parameters may have been determined approximately during the method development stage. Often a particular set of experiments will yield information on several parameters, so with careful planning the effort required to get the necessary information can be minimized. Assistance with deciding what degree of validation is required for a test method will be provided by the relevant ISTA Technical Committee.

1.5.3 The analytical specification

The test organiser should define the performance requirements that a method must have to be suitable for solving the analytical problem. In response to this requirement, the test organiser can arrange the evaluation of existing methods for suitability and, if necessary, develop a new method.

1.5.4 Method development

Method development can take a number of forms. At one extreme, it may simply involve adapting an existing method, making minor changes so that it is suitable for a new application. For example, a method required to detect quality differences in *Glycine max* (soybean) might be adapted from an established method for *Pisum sativum* (peas). These two crops are closely related, their seed structure is similar and the causes of differences in quality are also similar. Thus, it is likely that the same principles of seed quality evaluation can be applied to both species. If, on the other hand, a method is required to determine quality differences in a grass or small-seeded vegetable species, an adaptation of the method for the evaluation of quality in peas may not be the best option and an alternative method may need to be considered/ developed.

At the other extreme, the evaluation of seed quality may start out with a few sketchy ideas and require the expertise and experience of analysts/researchers to devise a suitable method. This may involve significant innovation based on novel exploitation of known characteristics of the species and the causes of quality differences. This clearly involves a great deal more work, and initially at least, a degree of doubt as to whether the final method will be successful. It is not infrequent for method development to involve working on a number of different ideas simultaneously and eventually choosing one which best meets the requirements.

For both situations the iterative process of development and evaluation continues until the method is deemed capable of meeting the analytical requirement and can start to provide test data. The process then becomes one of evaluating the test performance and confirming that the method is suitable, i.e. method validation.

1.6 Method Validation in ISTA

The outcomes of the ISTA Method Validation Programme are new or improved test methods accepted for inclusion in the ISTA Rules. The Programme has one major category of method validation, Collaboratively Validated Methods (which may be either Multi-Laboratory Validated Test Methods or Peer Validated Test Methods), and a minor category, Performance Validated Test Methods, for use in special circumstances. Descriptions of each category can be found in Table 1.

Regardless of category there is a five-step process before publication of a validated method in the ISTA Rules:

1. Method selection and development.
2. Validation through comparative testing.
3. Review of comparative test results and preparation of a Method Validation Report.
4. Approval of validation status by the ISTA Technical Committee and preparation of a Rules Proposal for the method.
5. Final acceptance by the ISTA Membership which will allow publication of the validated method in the ISTA Rules.

Table 1. Summary of requirements for collaborative and performance validation of seed quality test methods.

Programme Guidelines	Collaboratively Validated Methods		Performance Validated Test Methods ¹
	Multi-laboratory Validated Test Methods	Peer Validated Test Methods	
Objective	Multi-laboratory characterisation of test method performance	Verification of test method performance by peer laboratories	Verification of performance claims of test method
Laboratory Requirements²	6-8 collaborating laboratories chosen by the test organiser	Minimum of two collaborating laboratories chosen by the test organiser ³	Minimum of one laboratory approved by the Technical Committee
Sample Requirements	Preferably six samples representing three levels of the quality component being assessed ⁴	Sample number determined by performance claims but should include samples of known characteristics chosen by test organiser. Could be similar to <i>Multi-laboratory Validated Methods</i>	Sample number determined by performance claims. Could be similar to <i>Collaboratively Validated Methods</i>
Sponsor⁵	Method developer/promoter	Any method developer	Manufacturer / Method developer
Test Organiser⁶	Approved or appointed by the Technical Committee	Approved or appointed by the Technical Committee	Approved or appointed by the Technical Committee
Reviewers⁷	Chosen and appointed by the relevant Technical Committee and Statistics Committee.	Chosen and appointed by the relevant Technical Committee and Statistics Committee.	Chosen and appointed by the Technical Committee.

Main Steps	(1) Method development and in-house data development (may include pre-collaborative studies). (2) Multi-laboratory study design (test plan) (3) Multi-laboratory collaborative study. (4) Study analysis and method validation report (5) Adoption (6) Publication	(1) Method development and in-house data development. (2) Peer-laboratory test protocol (3) Peer-laboratory verification (4) Study analysis and method validation report (5) Adoption (6) Publication	(1) Method development and in-house data development (2) Independent-laboratory study design (test plan) (3) Independent-laboratory collaborative study (4) Study analysis and method validation report. (5) Package insert review (for proprietary test kits) (6) Adoption (7) Publication
Information Reviewed	Information must include intra-laboratory (within) repeatability; inter-laboratory (between) reproducibility; and comparison to existing methods where such methods exist.	Information must include intra-laboratory (within) repeatability; inter-laboratory (between) reproducibility; and comparison to existing methods where such methods exist.	Information must include intra-laboratory (within) repeatability; may include inter-laboratory (between) reproducibility; and comparison to existing methods where such methods exist.
Status⁸	(1) Accepted by the relevant Technical Committee: Validated Method. (2) Accepted by ISTA Membership vote: ISTA Official Method.	(1) Accepted by the relevant Technical Committee: Validated Method. (2) Accepted by ISTA Membership vote: ISTA Official Method.	(1) Accepted by the relevant Technical Committee ⁹ : Validated Method. (2) Accepted by ISTA Membership vote: ISTA Official Method.
Publication	(1) Publication of test report in Seed Testing Validation Reports (Online and hard copy) (2) ISTA Rules	(1) Publication of test report in Seed Testing Validation Reports (Online and hard copy) (2) ISTA Rules	(1) Publication of test report in Seed Testing Validation Reports (Online and hard copy) (2) ISTA Rules
Review Period¹⁰	Review period determined by Technical Committee	Review period determined by Technical Committee ⁷	Review period determined by Technical Committee (test kit manufacturers are required to inform ISTA if changes are made to the test kit).

Table 1 notes:

1. Performance Validated Test Methods
 - These test methods include proprietary test kits (e.g. for seed health testing).
2. Criteria for Participating Laboratories
 - Only laboratories experienced in the related seed testing area, or applying specific techniques to be evaluated, should be invited to participate in the validation programme.
 - Participating laboratories must have the skilled personnel, appropriate facilities and equipment for performing the test under evaluation.
 - Participating laboratories must understand and accept their obligations in participating in the validation programme
 - In cases where there is some doubt as to the ability of a laboratory to perform the test a *pre-test* may be used. Based on the results of the pre-test the organiser will decide which laboratories can be involved in the validation programme.
3. Peer Validation
 - Peer validation provides a means of including test methods which the relevant Technical Committee has decided do not require a full multi-laboratory study.

4. Samples
 - Whenever possible seed of commonly traded quality with respect to the attribute under study should be used in method validation tests. Samples must have a stable level of the quality attribute detectable by the test under evaluation after transport to the participants.
 - Under some circumstances seed may be amended to produce the desired quality attributes for study.
5. Sponsor
 - It is anticipated that most method sponsors will be ISTA Technical Committees or staff of ISTA Member Laboratories. However the ISTA Method Validation Programme is also open to non-ISTA members who may also submit a method for validation.
6. Test Organiser
 - A test organiser is any person approved or appointed by the Technical Committee to develop and organise the validation. This person may also be a member of the Technical Committee, or the test sponsor.
7. Reviewers
 - Reviewers are appointed by the relevant ISTA Technical Committee which may seek advice on suitable technical reviewers from the sponsor or Test Organiser. One of the three reviewers is nominated by the ISTA Statistics Committee.
 - Technical reviewers may be:
 - persons known to be working in the subject area
 - persons who have published on a similar topic
 - authors cited in the submission bibliography
 - collaborators acknowledged in studies of similar topics
 - **not** a member of a laboratory participating in the test evaluation or of the organization the laboratory is a part of.
8. Status
 - Upon acceptance by the relevant ISTA Technical Committee, a method is considered Validated. Any validated method may then be proposed for inclusion in the ISTA Rules and must be accepted by a vote of the ISTA Membership before being published in the ISTA Rules.
9. Performance Validated Test Methods for 'in-house' use
 - Not all test methods validated in this category may be intended for publication in the ISTA Rules. (refer to Section 6 of this book).
10. Review
 - Once included in the ISTA Rules, each validated method will be subject to review after a time period determined by the relevant Technical committee. A reviewer will be appointed by the ISTA Technical Committee. The reviewer will advise the Technical Committee of any modifications required to the published test method, addressed under the following categories:
 - editorial
 - minor changes
 - method extension
 - species addition
 - procedure modification
 - substantive revision

2. OBJECTIVES OF THE ISTA METHOD VALIDATION PROGRAMME

As outlined in Table 1, test methods can be validated by collaborative multi-laboratory validation, peer validation or test performance validation, depending on the individual method and its intended use. The appropriate procedure can be used to:

2.1 Develop a Validated Test Method for a Species

Situation: No validated test method is available. A new test method is developed, or there is one in common use but not well defined or described.

Objective: To document that the method is reproducible and repeatable. A limited number of elements of the test method may be varied to determine the optimum procedure.

Result: A validated seed test method.

2.2 Compare Different Methods to Choose a Test Method to Validate

Situation: No validated test method is available but more than one test method is in common use. These test methods are not uniformly described or may not be fully developed, have not been subjected to peer review or published in an international scientific journal.

Objective: To compare the different test methods and to decide, on the basis of results of the comparative test, which test methods may be adopted as validated test methods. In these comparative tests a limited number of elements of the test methods can be varied to determine the optimum procedure.

Result: A validated seed test method.

2.3 Improve a Validated Test Method(s)

Situation: One or more test methods are already included in the ISTA Rules. As a result of new developments, elements of these methods should be changed to improve the test method.

Objective: To test whether new developments can be incorporated into the existing test method(s) in such a way that it/they are still reproducible and repeatable.

Result: A validated seed test method that utilizes the latest developments leading to improved performance and usability of the test method.

2.4 Compare the Present Validated Test Method with a New Method(s)

Situation: A new technique has been developed for the identification of differences in seed quality. The introduction of the new technique can improve the identification of quality differences. This may include not only the reliability of the identification of quality differences, but also aspects of the test such as cost and time to perform the test.

Objective: To compare the new technique with the present test method and, when it is equivalent or better than the present test method, to introduce the new technique as the validated test method.

Result: The new seed test method as the validated test method because it outperformed the previous test method in terms of one or more of reliability, reproducibility, performance and test costs.

3. COLLABORATIVELY VALIDATED METHODS

3.1 Scope

3.1.1 There are two stages in the validation of seed quality testing methods:

- Establishment of acceptable performance parameters within a single laboratory.
- Demonstration of acceptable performance in an inter-laboratory collaborative study (multi-laboratory or peer validation).

Completion of the second stage confers confidence in the performance of the test method and the results produced in different laboratories.

While ISTA prefers multi-laboratory validation, there is sometimes a need for peer validation (see 1.5.1 and Table 1) whereby the collaborative study is conducted with a minimum of two other laboratories.

3.1.2 The scope of seed test methods suitable for submission includes:

- New methods for validation.
- Revisions of validated methods to extend their applicability or improve their performance.

3.1.3 The ISTA Method Validation Programme is fully open. Any interested party may submit a test method for validation, and may submit comments on Validation Test Reports. Submissions, queries, and comments should be sent to the appropriate Technical Committee c/o the ISTA Secretariat.

3.1.4 All test methods that are submitted for validation studies are subjected to technical and statistical review organised by the Technical Committee.

3.2 Purpose

The purpose of this section is to give clear guidelines on the procedures to be followed for collaborative validation of seed testing methods.

3.3 Parties Involved in Collaborative Method Validation and Their Responsibilities

3.3.1 ISTA Secretariat

The ISTA Secretariat may be notified directly that a test method is to be submitted for validation, and if so will forward this information to the relevant ISTA Technical Committee. Once the ISTA Technical Committee has accepted a method (i.e. it has become validated), the ISTA Secretariat will arrange for the publication of the "Seed Testing Validation Report" both online and in hard copy.

3.3.2 ISTA Technical Committee

The ISTA Technical Committee receives notice, either directly or from the ISTA Secretariat, that a test method is to be submitted for validation. The Technical Committee is responsible for approving/appointing a test organiser, appointing two reviewers, maintaining a record of the review process, approval of the method, and preparation of a rules proposal for consideration by the ISTA membership. The ISTA Statistics Committee is responsible for appointing a statistical reviewer.

3.3.3 Sponsor

A sponsor is the person who originally developed the test method and notified ISTA that the method was being submitted for validation. For many of ISTA's Technical Committees, the sponsor may be the chair, or a member of the Technical Committee. However, a sponsor may also be a person outside of ISTA. He/she may be appointed by the Technical Committee as the Test Organiser.

3.3.4 Test Organiser

The Test Organiser is approved or appointed by the Technical Committee to prepare the draft test plan, prepare samples/materials for the comparative testing, conduct an 'in-house' test run of the entire test method, submit the draft test plan to the Technical Committee for review, prepare the final test plan, secure the participants, distribute the seed samples/materials/ information, receive results and analyse data, prepare a draft report, and after feedback, prepare the final report and submit it to the Technical Committee.

3.3.5 Test Participants

The test participants are individuals involved in seed quality testing who have been identified by the test organiser or Technical Committee as having the expertise and equipment required to carry out the testing required to provide data for method validation.

3.3.6 Reviewers

Reviewers are two individuals appointed by the Technical Committee to review the draft test plan and the final test report. In addition, the ISTA Statistics Committee appoints one individual to review the statistical methods to be used, and the analysis of results.

3.4 Procedure

3.4.1 Notification

3.4.1.1 The test sponsor notifies ISTA (via the Secretariat) that a test method is to be submitted for validation. The Secretariat forwards this information to the relevant Technical Committee. If the sponsor is a member of the Technical Committee, notification via the Secretariat is not required.

3.4.2 Draft Plan

3.4.2.1 Once approved/appointed by the Technical Committee, the Test Organiser prepares a detailed draft test plan with clear objectives, a well-defined time schedule for the species and test concerned, and following the criteria outlined in Section 1 (Table 1) and Appendix 2 and 3. Essential materials must be identified in the test plan together with any costs for which the participant may be charged.

3.4.2.2 The test organiser prepares the samples and other essential materials for the comparative test (see Appendix 2) and conducts an 'in-house' test-run of the entire test method using the actual samples to be used in the comparative test.

3.4.2.3 The test organiser submits the draft test plan to the Technical Committee along with an application for Test Method Validation (Appendix 4).

3.4.3 Review of Draft Test Plan/Preparation of Final Test Plan

3.4.3.1 The three reviewers (two technical, one statistics) review the draft plan (see Appendix 5) and return their reviews to the test organiser.

3.4.3.2 The test organiser incorporates the comments of the reviewers on the draft plan into the final test plan.

3.4.4 Participants

- 3.3.4.1 Conforming to the time schedule of the test plan, the test organiser prepares the list of 6-8 (multi-laboratory) or 1 – 3 (peer) participants for the comparative test. Participants can be staff of ISTA accredited laboratories, or member laboratories, or non-member laboratories with a proven record of experience in the quality test under study. The number of participants may be limited by the amount of suitable seed available for use in the test and should not be less than the minimum number of participants needed for the statistical analysis.
- 3.4.4.2 When necessary the ISTA Technical Committee may assist the test organiser to identify qualified participants. The test organiser makes the final selection of participants.

3.4.5 Comparative Testing

- 3.4.5.1 The test organiser distributes seed samples, registration forms, time schedule, data record sheets and the final test plan to all participants.
- 3.4.5.2 The participants perform the comparative test according to the test plan and test method. The participants return the raw data to the test organiser within two weeks of completion of the tests, and not later than the date shown in the time schedule.
- 3.4.5.3 The test organiser checks and analyses the results using the statistical procedures indicated in the test plan.

3.4.6 Test Reports

- 3.4.6.1 The test organiser prepares a draft report containing conclusions and recommendations.
- 3.4.6.2 The draft report is sent to the participating laboratories for comments. Comments must be sent back to the test organiser within one month. The test organiser decides which comments are to be incorporated into the conclusions and which remarks should be added separately. The final report is prepared within two months of the distribution of the first draft report.
- 3.4.6.3 The final test report (an ISTA Method Validation Report) which includes all results, statistical data and a working method (Appendix 6) is submitted to the Technical Committee.
- 3.4.6.4 Where documents are not presented in the required format, they will be returned to the author without review.

3.4.7 Technical Review of Test Data

- 3.4.7.1 The Technical Committee arranges for the two expert reviewers (one technical, one statistics) to review the data generated and the final test report. Reviewers must return their reports (Appendix 7) to the Technical Committee within 4 weeks.
- 3.4.7.2 The Technical Committee maintains a record of the review process and endeavours to ensure that it proceeds in a timely fashion.

3.4.8 Post-Technical Review

- 3.4.8.1 Following review, and based on the advice of the reviewers, the Technical Committee will:
- a) approve the method and report without revision
 - b) approve the method and report following minor revisions to the method and/or report
 - c) defer a decision pending major revisions to the method and/or report
 - d) reject the method.

- 3.4.8.2 Where reviews are considered inadequate or reviewers do not agree, the Technical Committee will, at its discretion, make a final decision or seek the advice of additional reviewers.
- 3.4.8.3 The Technical Committee will notify the test organiser and reviewers of its decision with reasons and, if appropriate, request revisions.
- 3.4.8.4 Methods in category (b) will only be considered approved following completion of the required revisions to the satisfaction of the Technical Committee.
- 3.4.8.5 Methods in category (c) will be re-considered by the Technical Committee following completion of the required revisions and may be subjected to further independent review. Completion of revisions will not guarantee approval, and a further validation round may be requested.

3.4.9 ISTA Rules Proposals

- 3.4.9.1 Following approval of the method by the Technical Committee, the method receives Validated status. The Technical Committee will then prepare a proposal for the ISTA Rules and submit it to the ISTA Rules Committee.
- 3.4.9.2 The Method Validation Report and the Rules proposal will be listed on the ISTA Online website prior to a vote by the membership. Copies of both will be available from the ISTA Secretariat. Any interested party may submit comments and data in writing to ISTA for or against the adoption of the proposal. The ISTA Secretariat will forward copies of all comments to the Technical Committee for resolution.
- 3.4.9.3 Methods accepted by a majority vote at the meeting will be published in the International Rules for Seed Testing.

3.4.10 Copyright

- 3.4.10.1 Copyright for any validated method published in the ISTA Rules or in an ISTA Method Validation Report is held by the International Seed Testing Association.

3.5 Related Documents

- 3.5.1 Performance Validated Methods
- 3.5.2 General Procedures for Method Validation (Appendix 2)
- 3.5.3 Instructions to Authors: Preparation of Test Method Descriptions (Appendix 3)
- 3.5.4 Test Method Validation Application Form (Appendix 4)
- 3.5.5 Instructions for Reviewers: Draft Test Plan (Appendix 5)
- 3.5.6 Instructions to Authors: Preparation of Validation Reports For New Test Methods or Revisions to Existing Test Methods (Appendix 6)
- 3.5.7 Instructions for Reviewers: the Validation Report (Appendix 7)

4. PERFORMANCE VALIDATED METHODS

4.1 Scope

4.1.1 Advances in laboratory technology for some components of seed quality have resulted in a shift from traditional direct testing to rapid, indirect bio-molecular techniques such as PCR. These test methods may be published and therefore freely available internationally, require the use of commercially available test kits, or have been developed “in-house” and used only by one laboratory.

4.1.2 The scope of test methods suitable for performance validation is primarily for test kits, but may also include non-test kit methods.

4.1.3 The performance validated methods procedure allows a review of performance claims.

4.2 Purpose

The purpose of this section is to give clear guidelines on the procedures to be followed for performance validated seed testing methods.

4.3 Parties Involved in Performance Validation and Their Responsibilities

4.3.1 Test Method Owner

The individual or company owning the test method and who makes an application to ISTA for a review of the performance claims.

4.3.2 Test Method User

A laboratory wishing to use the test method, and which makes an application to ISTA for a review of the performance claims.

4.3.3 Other Parties

See 3.3.1 to 3.3.6.

4.4 Procedure – Test Kit Methods

The Technical Committee in consultation with regulatory officials and the seed industry determines optimal realistic parameters of sensitivity and specificity for the test. Test kits which can supply the required parameters and which are found to be in conformance with their claims are granted Performance Validated Method status by ISTA. The Performance Validated Method status assures the test kit user that an independent assessment has been conducted, and that the kit performs as claimed.

4.4.1 Application and Submission

4.4.1.1 The applicant obtains an Application Package from the ISTA Secretariat. This package contains information about the programme, what types of data are required, application fees and the agreements. If the applicant is a test kit user, this information must first have been obtained from the test kit manufacturer.

4.4.2 Application Review

Two Expert Reviewers are recruited and selected by the ISTA Technical Committee to evaluate the product descriptive inserts to determine the completeness of the inserts (including: instructions for use; applicability; interpretation criteria; shelf life; detection limit and limit of quantification; sampling protocol) and to determine the exact performance claims made for the test kit.

4.4.3 Performance Tests

The Technical Committee appoints a test organiser to develop a test plan in accordance with Section 3 for collaborating independent testing laboratories, which then conduct performance testing of the test kit.

4.4.4 Technical Review of Test Data

The data generated by the collaborating independent testing laboratories are sent to the test organiser who will then distribute them back to the Expert Reviewers. The Expert Reviewers then review the data generated by the collaborating independent testing laboratories to determine whether these data corroborate the performance claims made for the test kit. Reviewers must return their reports to the ISTA Technical Committee within three weeks.

4.4.5 Acceptance

If the Expert Reviewers determine that the collaborating independent testing laboratories' data verify the manufacturer's performance claims for the product, and that the product descriptive inserts adequately describe the product performance, then the test method is accepted as a Performance Validated Method. Subsequent acceptance of the test method as an ISTA rule follows the preparation of a Rules proposal and a positive ISTA membership vote.

4.4.6 Publication and Distribution

Validation information is given on the ISTA Online website. Accepted performance validated methods will be published in the ISTA Rules.

4.4.7 Copyright

Copyright for any performance validated method published in the ISTA Rules is held by the International Seed Testing Association.

4.4.8 Changes in the Test Kit

It is the test kit manufacturer's contractual obligation to immediately notify the ISTA Secretariat if and when any changes are made to either the instructions for using the kit or the kit's performance. Failure to promptly notify the ISTA Secretariat of kit changes may result in removal of the test kit as a Validated Method. When the ISTA Secretariat receives notice of a kit change, the appropriate Technical Committee, in consultation with appropriate experts, will determine if the changes are of sufficient magnitude to warrant a complete re- evaluation of the kit. If so, the applicant of record must submit a complete application.

4.4.9 Complaints

User complaints may result in the Technical Committee initiating an inquiry and could lead to cancellation of the kit's Performance Validated Method status.

4.5 Procedure – Non-Test Kit Methods

Test methods which can supply the required performance parameters and are found to be in conformance with the claims of the sponsor are granted Performance Validated Method status by ISTA. Following the granting of this status, the relevant Technical Committee may forward the method as a Rules Proposal, which if accepted by the ISTA membership, means that the method will then be included in the ISTA Rules.

4.5.1 Application and Submission

Applicants are required to collect and submit data to the relevant ISTA Technical Committee that support the sponsor's claims for the test method. Applicants should follow the submission requirements provided in Appendix 8.

4.5.2 Application Review

See 4.4.2.

4.5.3 Confirmatory Tests

See 4.4.3.

4.5.4 Technical Review of Test Data

See 4.4.4.

4.5.5 Acceptance

See 4.4.5.

4.5.6 Distribution and Publication

See 4.4.6.

4.5.7 Copyright

See 4.4.7.

4.6 Related Documents

Collaboratively Validated Methods
Performance Validated Test Method Submission Requirements for Other than
Test Kits (Appendix 8)

5. MODIFICATIONS TO ISTA RULES

5.1 Method Modification Categories

Once a test method is accepted as an ISTA Validated Test Method and has been included in the ISTA Rules, modifications to the method will be addressed in these categories:

5.1.1. Editorial

Correction of errors and clarification of language or expression. (See 5.3) If a test method is modified editorially, it retains its ISTA Validated and ISTA Rules status.

5.1.2. Method extension

Change in the applicability statement. A test method with an extended applicability statement (e.g. the addition of a new species) retains its ISTA Validated status, but an ISTA Rules amendment will be required.

5.1.3. Procedure modification

Change in procedure, or a parameter that does not change the principle of the test method but may affect the performance parameters. A modified Validated Test Method must be proposed for acceptance into the ISTA Rules; a modified test method already in the ISTA Rules reverts to Validated status.

5.1.4. Substantive revision

Change that affects the procedure of the test method and thus the performance parameters. The old test method is repealed or declared surplus when a substantially revised (new) test method is accepted into the ISTA Rules. Statistical performance parameters for the modifications described in 5.1.2 – 5.1.4 must be reassessed.

5.2 Validation Requirements for Test Method Modifications

5.2.1. Editorial

No method validation is required, but the content of the presumed editorial changes must be approved by the Technical Committee.

Changes must be forwarded by the Technical Committee to the ISTA Rules Committee Chairperson who will determine that they are editorial and arrange for them to be made.

5.2.2. Minor change

By scientific judgment of the Technical Committee, literature or historical data are sufficient validation

Recommendation for a minor method revision must be approved by the Technical Committee.

Method revision will be forwarded to the ISTA Rules Committee Chairperson as an ISTA Rules change proposal.

5.2.3. Test method extension

Test method extension may be applied to a new species or within a species.

To add a new species or extend the application within a species, the extended test method requires a collaborative study and Method Validation Report. Following the receipt of a Method Validation Report, and acceptance by the Technical Committee a test method extension will be forwarded to the ISTA Rules Committee Chairperson as an ISTA Rules change proposal.

5.2.4. Procedure modification

A collaborative comparison of the old and new test procedure must be done and should demonstrate the same or an improved test method performance. The comparative study must be of sufficient scope to establish equivalency of performance. It is understood that the comparison may take place against original collaborative study data.

Following receipt of a Method Validation Report, the procedure modification must be approved by the Technical Committee, before being forwarded to the ISTA Rules Committee chairperson as an ISTA Rules change proposal.

5.2.5. Substantive revision

Any substantially revised test method must be submitted to a full multi-laboratory validation or performance validation study as appropriate.

All procedures for test method approval must be followed, including submission of the study plan. The Technical Committee must concurrently recommend repeal or surplus status of the test method to be replaced (old method).

The approved revised (new) validated test method will be forwarded to the ISTA Rules Committee chairperson as an ISTA Rules change proposal.

5.3 Examples of Editorial Changes That Correct or Clarify Language

- Addition of non-technical information for consistency.
- Updating of non-technical footnote or reference information.
- Changes to titles of sections, tables and figures.
- Rearrangement or renumbering of sections, tables and figures.
- Minor changes in sentence structure that do not change the original intent.
- Corrections to spelling or punctuation.
- Correction of typographical errors that would not have otherwise affected the technical interpretation of the method.

6. VALIDATION OF METHODS FOR TESTING FOR THE PRESENCE OF SPECIFIED TRAITS

6.1 Introduction

In April 2005, ISTA introduced the Performance Based Approach to the ISTA Rules (see ISTA Rules Chapter 8, clauses 8.1.2 and 8.2.3) whereby Performance Approved Methods can be used for the testing for the presence of specified traits. Under this approach, a laboratory may develop and/or use a method not published in the ISTA Rules (see “Principles and Conditions for Laboratory Accreditation under the Performance Based Approach”).

As these methods are not validated by ISTA, they can not be published in the ISTA Rules. However, the laboratories must be able to provide:

- A statement about how the method was validated and installed.
- Supporting documentation regarding the method’s fitness for the purpose (see “Performance Data Evaluation for Trait Purity”).

6.2 Previously Validated Method

The testing laboratory may choose a method already validated by an inter- laboratory study according to ISO5725 or other internationally or nationally recognised standards.

6.3 Inter-Laboratory Validation for a New or Modified Method

A testing laboratory developing a new method, or modifying an existing validated method, may choose not to retain the intellectual property (i.e. not keep the method “in-house”), which will allow inter-laboratory participation in the validation process. The test sponsor may choose to use the ISO5725 procedure, or the relevant sections of the ISTA procedure (see Section 4 of this book), omitting any official participation by an ISTA Technical Committee or the ISTA Secretariat.

6.4 Single Laboratory Validation for a New or Modified Method

A testing laboratory developing a new method, or modifying an existing validated method, may choose to retain the intellectual property (i.e. keep the method “in-house”). In this case a single laboratory method validation according to the IUPAC (see Pure Applied Chemistry 2002, 74(5), 835-855) should be performed.

Note: Method validation for tests for the presence of specified traits is the responsibility of the testing laboratory, and not ISTA. An ISTA auditor, when auditing a testing laboratory accredited by ISTA for testing for the presence of specified traits will require the laboratory to be able to demonstrate that the method(s) used is/are validated but ISTA has no role in the actual method validation process.

ANNEX

- Annex 1: Statistical Aspects of Method Validation
- Annex 2: General Procedures for Method Validation
- Annex 3: Instructions to Authors: Preparation of Test Method Descriptions
- Annex 4: Test Method Validation Application Form
- Annex 5: Instructions for Reviewers: Draft Test Plan
- Annex 6: Instructions to Authors: Preparation of Validation Reports for New Test Methods or Revisions to Existing Test Methods
- Annex 7: Instructions for Reviewers: the Validation Report
- Annex 8: Performance Validated Test Method Submission Requirements for Other Than Test Kits
- Annex 9: Glossary Terms ISTA Method Validation Studies

DISTRIBUTION LIST

ISTA website

REVISION HISTORY

Version #	Changes
2.0	Transfer to the new layout

Annex 1: Statistical Aspects of Method Validation

1.1. Introduction

This appendix contains explanations and example of things that can be addressed in method validation from a statistical perspective. The methods described are simple compared to more sophisticated possibilities, but provide comparable and objective support for validation studies. If a test organizer wishes to use more sophisticated methods, this should be in addition to the simple methods.

Method validation for a test method involves a planned experiment, from which data are obtained. The aim of the experiment is to obtain the objective supporting evidence required to allow validation of the test method.

The method must be developed to the stage that it is ready for routine use with a protocol and identified parameters to control (for instance substrate, temperature and duration for germination).

- a) The selection of appropriate parameters must have been done prior to submitting for method validation. If the test leader wishes to show evidence of ruggedness of the method, the simple experiment as described in **1.2 “Ruggedness testing procedure”** can be used.
- b) When data are collected and a statistical analysis is performed, support from a statistician is required. See **1.3 “Statistical support to establish and review the test”**.
- c) When planning the experiment it is important to consider **(1.4) “The number of samples, the number of repeats, the number of laboratories, and the true value of the samples prepared”**.
- d) Among the things that should be considered before the experiment are the differences in results that are acceptable among laboratories, and if a statistical test will be able to show them as significant; this is introduced in **1.5 “Benefits of simulation for a test plan design”**.
- e) The choice of statistical analysis will be driven in particular by **“The type of results and distributions”** (see **1.6**).
- f) A first step of the statistical analysis is to explore and check the data set; simple graphical representations of data are usually very helpful, and **1.7 “Detection of outliers”** is one of the aspects for which statistical tests can be used.
- g) There are a number of possibilities for statistical analysis. **1.8 “Statistical analysis, model and assumptions”** refers to applying a mathematical model to the data set, and **1.8.1 “ANOVA”** and **1.8.2 “GLM”** are given as classical statistical tools.
- h) Because tests will be performed on an international basis, **1.9 “Repeatability-reproducibility”** are two important features; they quantify the expected variability of results respectively within a laboratory, and within a group of laboratories.
- i) Accuracy, bias and uncertainty are also important features, but are not covered in this appendix.

1.2. Ruggedness Testing Procedure

(Derived from Youden and Steiner (1975) *Statistical Manual of the AOAC*, (AOAC International, Arlington, VA, USA)

When the test sponsor develops and standardizes the procedures of a method, data may be collected for a set of operations and equipment that is never varied. This process does not reveal what will happen during a method trial in a number of laboratories, each of which has its own set of reagents, equipment and routines. Preparation of standards, time and temperature variations, instrument calibration and performance, and analyst technique all contribute to minor variations even when a procedure is followed “exactly”. The only way to forecast the performance of a method under different laboratory conditions is to deliberately introduce reasonable variations and observe what happens. If the procedure is “rugged” the results should not be affected.

The suggested scheme does not study one alteration at a time, but introduces several changes at once in such a manner that the effects of individual changes can be ascertained.

Example:

Let A, B, C, D, E, F, and G denote the nominal values for seven different factors that might influence the results if their values were varied slightly. Let their alternative values be denoted by a, b, c, d, e, f, and g.

The conditions for running a determination will be completely specified by the seven letters, each letter being either capital or lower case. There are 2^7 or 128 different combinations, which is impractical to test. However, it is possible to choose a subset of 8 combinations that adequately balances upper and lower case conditions.

Eight combinations of seven factors used to test the ruggedness of an analytical method								
	<i>Combination or Determination No.</i>							
Factor Value	1	2	3	4	5	6	7	8
A or a	A	A	A	A	a	a	a	a
B or b	B	B	b	b	B	B	b	b
C or c	C	c	C	c	C	c	C	c
D or d	D	D	d	d	d	d	D	D
E or e	E	e	E	e	e	E	e	E
F or f	F	f	f	F	F	f	f	F
G or g	G	g	g	G	g	G	G	g
Observed result	s	t	u	v	w	x	y	z

The table specifies the value for seven factors to be used while running eight determinations. The results are designated s through z.

To find whether changing factor A to a had an effect, compare the average $(s+t+u+v)/4$ with the average $(w+x+y+z)/4$. Determinations 1,2,3, and 4 were run with factor level A, and determinations 5,6,7, and 8 with factor level a.

The seven differences for A-a, B-b, etc., can be computed. If one or two factors are having a bigger effect, their differences will be substantially larger than the group of differences associated with the other factors. The ranking of the differences is an indication of the method's sensitivity to the factors.

A rugged method is a method which is not much affected by changes that will most certainly be encountered among laboratories.

If there is no outstanding difference, the method is considered as tolerant to the variations introduced in the experiment.

The standard deviation of the 8 results can be computed to quantify the variability of the results when some conditions vary. It is a rough estimate of the analytical error.

It is suggested to always use 8 combinations. If only 3 factors vary, all 2³ possible combinations are in the experiment.

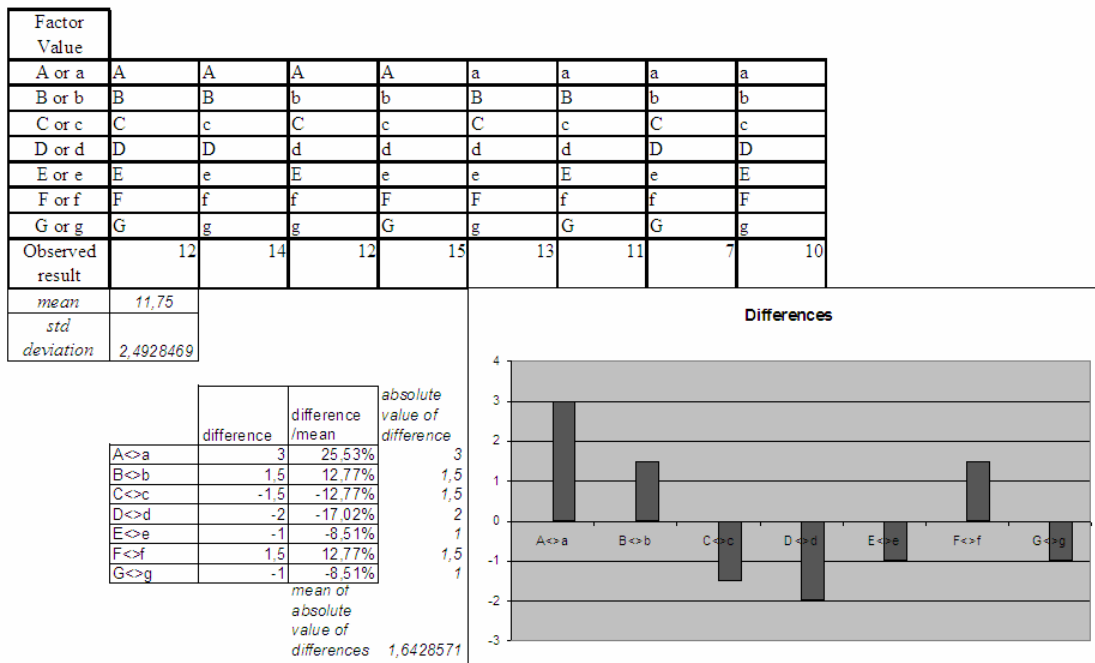
Eight combinations of seven factors used to test the ruggedness of an analytical method									
								Combination or Determination No.	
Factor Value		1	2	3	4	5	6	7	8
A or a		A	A	A	A	a	a	a	a
B or b		B	B	b	b	B	B	b	b
C or c		C	c	C	c	C	c	C	c

The same computations apply from 3 to 7 factors.

If some factors are not used, the differences computed for these factors are meaningless, but the mean and the standard deviation are still valid.

A spreadsheet providing the above computations is available on the ISTA website.

Eight combinations for 2 to 7 factors used to test the ruggedness of an analytical method



1.3 Statistical support to establish and review the test

Whenever possible it is advisable to identify two statisticians for each test: one will act as an **advisor** and provide advice and support throughout the test from the very early stages of planning and defining objectives through to preparation of the final report; the other will act as a **reviewer** and provide an independent review of the final report.

It is the responsibility of the test organizer to identify the two statisticians, although in case of difficulty assistance can be obtained from the chairperson of the Statistics Committee who will keep an updated list of persons who could perform these reviews.

The following check list provides an indication of the issues which should be considered/addressed by the statistical advisor/reviewers. Advisors/reviewers and test organizer are encouraged to keep a record of their correspondence/discussions and as a minimum should record the date on which each item on the check has been addressed. Items not discussed by the reviewers and the test organizer should be identified as “non addressed”. At the end of the review the check list shall be sent to the TCOM chair.

1.3.1 Check points for the statistical advisor

a) *Planning of the test*

- i Type of programme (multi-lab validated method, peer-validated method or performance validated method) and the type of test described in the Method Validation Programme.
- ii Potential number of laboratories
- iii Objective(s), aim, questions, hypothesis that test organizer wishes to address
- iv Look at results of previous tests or data the test organizer may have of a pre-test
- v Questionnaire addressed to the laboratories with the draft plan before final protocol is agreed
- vi Size of difference which would like to be distinguished by organizer. Illustrate with simulations if possible
- vii Availability of seed lots
- viii Technical difficulties (method, pathogen, measurements, sampling, stability of materials before analysis, ...)
- ix Cost and time schedule aspects x
Check if not too many factors
- xi Look how results could/will be analyzed
- xii Look at how the data will be collected (format)
- xiii Check on coding, blind testing, reference material identification. xiv
Level of confidentiality and communication during the work
- xv Check of the agreed protocol

b) *Analysis of the data received*

- i Check the results obtained by each laboratory, controls/reference materials, terms of protocol respected
- ii General information received from the organizer and the laboratories
- iii Graphical representations
- iv Look for outliers or exotic points
- v Transformation of data, if appropriate. vi
Analysis

- vii Questions and discussions with the test organizer (and laboratories if necessary)
 - viii Conclusions and questions
- c) *Reporting*
- i What to include in the report
 - ii Check of words, figures, values
 - iii Discuss coherence between the different parts of the report and agreement with objectives of the study
- N : The committee require the data are *in extenso* provided in the final report, and data corrected and/or discarded are identified.
- d) *Closing*
- i Archiving of data – identify what should be kept (further use) or on the contrary destroyed (confidentiality), how long, in which form (paper, computer file,...) [Both hard copies and computer text files of the raw data will be retained by the ISTA Secretariat and test organiser]
 - ii Availability of data and reports for further studies and on ISTA web site

1.3.2 Checkpoints for the statistical reviewer

- a) *Contact with the test organiser*
- i Agreement on reviewing, define possible date for availability of report, define pre-requisites
- b) *Review the report*
- i Is the analysis appropriate for the type of data ?
 - ii Are sufficient data presented to allow independent assessment ?
 - iii Are the conclusions justified ?
 - iv Are there questions that need a reply from the test organiser ?
- c) *Contact test organiser*
- i Send list of remarks and questions to the test organiser, receive answers from the test organiser
 - ii Send official review to the test organizer

1.4 Number of samples, number of repeats, number of labs, true value of the samples prepared.

Table 1 in Chapter 1 indicates some criteria, including the number of laboratories and the number of samples.

Programme Guidelines	Collaboratively Validated Methods		Performance Validated Test Methods
	Multi-laboratory Validated Test Methods	Peer Validated Test Methods	
Objective	Multi-laboratory characterisation of test method performance.	Verification of test method performance by peer laboratories	Verification of performance claims of test method
Laboratory Requirements	6-8 collaborating laboratories chosen by the test organiser	Minimum of two collaborating laboratories chosen by the test organiser	Minimum of one laboratory approved by the Technical Committee
Sample Requirements	Preferably six samples representing three levels of the quality component being assessed	Sample number determined by performance claims but should include samples of known characteristics chosen by test organiser. Could be similar to <i>Multi-laboratory Validated Methods</i>	Sample number determined by performance claims. Could be similar to <i>Collaboratively Validated Methods</i>

Example:

A test is made from 3 seed lots of different levels (of what shall be measured), and each laboratory receives 2 samples from each lot. This will result in $3 \times 2 = 6$ samples to test.

If in addition 3 samples with a zero level are introduced, each laboratory will receive $6 + 3 = 9$ samples to test.

The 9 samples are coded blind to the laboratory staff that perform the test.

The technical protocol will be known by the laboratories, but the design of the test can also remain blind to the laboratory staff.

1.4.1 Number of samples

In common practice for method validation studies at least 3 samples representing different levels of presence/quantity within the range of the method, and within the range of values that will be encountered in practice are tested:

- a low level
- an intermediate level a
- high level.

A number of parameters, among them accuracy, repeatability, reproducibility can vary over the range. Having a number of samples across the range helps to quantify this variation.

In tests where absence/zero level is of interest in practice, or to validate the method, a zero level needs to be added.

Whenever desired the organiser can have more levels in the range; for example 3 samples at low levels, an intermediate level and a high level (if the low levels are more likely to occur in routine tests, or need a more thorough check for the validation).

low level 1
low level 2
low level 3
intermediate level
high level

The test organiser can also “double a level”, which means for instance sending 4 samples, the same lot for low level being sent twice.

a low level (sent twice)
an intermediate level a
high level

This allows cross-checking of a number of parameters (among them accuracy, repeatability, reproducibility) for the level which is doubled. The staff who perform the tests should not know that a level has been doubled.

1.4.2 Number of repeats

A repeat in the context of method validation corresponds to the fact that the laboratory has to receive and test more than one sample for a given level. The absolute minimum is 2 samples per laboratory for a given level.

The purpose of repeats is to quantify the variability of the results. Repeats are a prerequisite to compute repeatability and reproducibility.

When possible and if desired, laboratories can also repeat measures on a given sample. For instance if the number of seeds in the sample is large, they can prepare two working samples and obtain two results.

For instance in a PCR test, a laboratory can conduct 2 separate DNA extractions from the sample received, and obtain 3 data points per DNA extraction, providing 6 data points for a given sample.

Unless all laboratories do the same repeats, the repeats in italics above are usually not part of the validation.

1.4.3 Number of laboratories

Whenever possible it is better to have a number of laboratories. This is a way to ensure uniformity in seed testing.

Six to 8 is the minimum for multi-laboratory validation. If only 2 to 6 laboratories can participate it becomes peer validation.

Laboratories must have the ability to perform the test (equipment, competent staff, etc). Data analysis will give values per laboratory. This enables the verification of uniformity among laboratories, or to detect any discrepancies among laboratories. In cases where a laboratory has different results from the others, the report of the parameters (repeatability, reproducibility) can be separated, one set of parameters with all laboratories, and another set of parameters where the laboratory in question is not included.

1.4.4 True value of the samples prepared

Depending on the type of test, the degree of precision of true value of the samples prepared varies.

For instance in germination, or seed health with naturally infected seeds, the true level is usually estimated through a test made on the lot. When a better estimate is needed more samples are tested, which gives a better estimate of the level, but also give information about the variability of the results on the same lot in a given laboratory.

In such tests, it is known that the true value of the samples sent for the test to the laboratories will not be exactly the same.

The true value can be better controlled when individual seeds having the trait to be detected are introduced in seed samples not having this trait. For instance, adding seeds from other species in a purity check, adding a number of infected seeds in samples of healthy seeds, etc.

In all circumstances, accuracy is computed as the deviation from the true value. The true value can be the value as known by the test organiser, or the mean (or median) of the results obtained by the laboratories. It is recommended to compute both.

1.5 Benefits of simulations for a test plan design

A simulation can be worthwhile to estimate if the goal of the test can be met statistically with the material present at that moment. A simulation gives you an idea about the outcome but is not completely related to all technical aspects that can occur during a test. On the website of DSS Research, <http://www.dssresearch.com/toolkit/default.asp>, a toolkit is freely available; sample size calculator gives an idea about the number of samples to be tested and sample error calculator results on the power of the test.

Some examples:

- a) Check the number of samples needed for 2 laboratories (peer validation) to give consistent results:
 - i. For instance, one laboratory finds 1% infected seeds and the other 3%. ii. How many seeds shall be tested to be able to declare 1% and 3% as different?
 - the computation indicates 642 seeds would be necessary.

- b) Check the power of the test according to the set up:
- i. Define a protocol validated by 2 laboratories
 - ii. In routine testing 400 individual seeds are checked
 - iii. According to the same sample size (400), is it possible to see a difference of 1% between the 2 laboratories?
 - the computation indicates a 21.4% statistical power. About 20 times out of 100, the statistical test will conclude there are no differences because the sample size is not big enough.

- i. How many seeds need to be checked, to see a difference between 1% and 2% in two laboratories?
 - the result is 4000 seeds in both laboratories are needed to get a statistical power of about 95% (95.7%)

researcher's toolkit print friendly

Averages, One Sample | Averages, Two Samples | Percentages, One Sample | **Percentages, Two Samples**

Two Sample Tests Using Percentage Values

One-Tail Test (Hypothesis that Sample 1's Percentage is greater than Sample 2's value or Sample 1's Percentage is less than Sample 2's value, but not both)

Two-Tail Test (Hypothesis that Sample 1's Percentage is not equal to Sample 2's Percentage value)

To calculate the minimum possible statistical power, set one percentage value to 50% which produces the largest possible variance (also known as the "Most Pessimistic Variance Assumption").

Sample 1 Percentage (%): (Value measured from Sample 1 or expected from this sample)

Sample 2 Percentage (%): (Value measured from Sample 2 or expected from this sample)

Size of Sample 1: (Size of Sample 1 or desired number of respondents)

Size of Sample 2: (Size of Sample 2 or desired number of respondents)

Alpha Error Level or Confidence Level: (Probability of incorrectly rejecting the null hypothesis that there is no difference in the percentage values). An Alpha of 5% corresponds to a 95% Confidence Interval.

Statistical Power: 95,7%

1.6 Different types of results and distributions

1.6.1 Direct tests

For direct tests the seeds are individually examined and the result is assessed without any transformation.

Tests like germination, purity, other seed determination, tetrazolium, moisture, vigour can deliver this type of data.

Example: germination results, 4 x 100 seeds for pepper

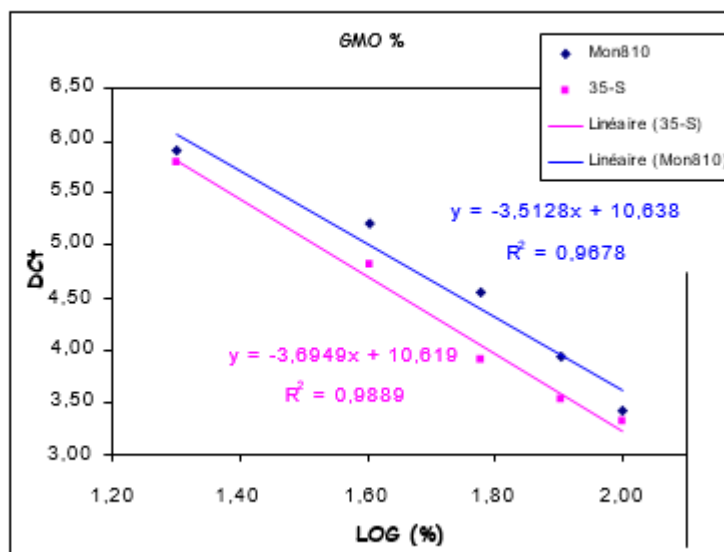
Rep	% normal	% abnormal	% dead seeds	% fresh ungerminated
A	85	7	6	2
B	86	6	6	2
C	83	7	7	3
D	87	5	5	3
Average	85	6	6	3

1.6.2 Indirect tests

In other tests data are obtained through computations and/or equipment calibration.

For instance % presence can be derived from Optical Density for ELISA in seed health tests, % presence of seeds with specified trait-(s) can be derived from a linear line drawn from delta CTs and Log of % presence of reference material.

Example: Real time PCR calibration:



Example: seed health results of a quantitative (ELISA values) and qualitative (presence / absence) test:

sample	lab 1			lab 2		
	average	result	average	result		
1	0,0470,054	0,051	neg	0,1420,144	0,143	neg
2	0,0340,043	0,039	neg	0,1210,125	0,123	neg
3	0,0830,083	0,083	neg	0,1940,159	0,177	neg
4	0,0480,053	0,051	neg	0,1040,145	0,125	neg
5	0,0430,046	0,045	neg	0,1290,112	0,121	neg
6	0,0570,061	0,059	neg	0,1200,125	0,123	neg
7	0,0610,058	0,060	neg	0,0950,109	0,102	neg
8	0,0640,060	0,062	neg	0,0910,094	0,093	neg
9	0,0420,046	0,044	neg	0,0860,094	0,090	neg
10	0,0560,063	0,060	neg	0,7480,742	0,745	pos
11	0,0440,041	0,043	neg	0,0760,086	0,081	neg
12	0,0400,046	0,043	neg	0,0900,086	0,088	neg
13	0,0290,025	0,027	neg	0,1250,107	0,116	neg
14	0,0370,040	0,039	neg	0,0840,090	0,087	neg
15	0,0440,044	0,044	neg	0,0940,101	0,098	neg
16	0,0440,044	0,044	neg	0,0790,083	0,081	neg
17	0,0530,050	0,052	neg	0,0860,091	0,089	neg
18	0,0420,044	0,043	neg	0,0880,103	0,096	neg
19	0,0450,043	0,044	neg	0,0640,074	0,069	neg
20	0,0390,032	0,036	neg	0,1310,141	0,136	neg

The results can either be **quantitative** (e.g. number of colonies found or percentage of GMO) or **qualitative** (presence and absence).

1.6.3 Distributions

The **statistical law of data distribution** also varies:
The main distributions encountered in seed testing are:

- the Binomial distribution where a seed or a pool of seeds can be classified in two categories (positive/negative, germinated/non germinated,...)
- the Poisson distribution for rare events (check for presence of other seeds to be avoided),
- the Normal distribution when a quantity is measured (1000 seed weight)

The choice of an appropriate statistical analysis will be driven in particular by the type of results and distribution.

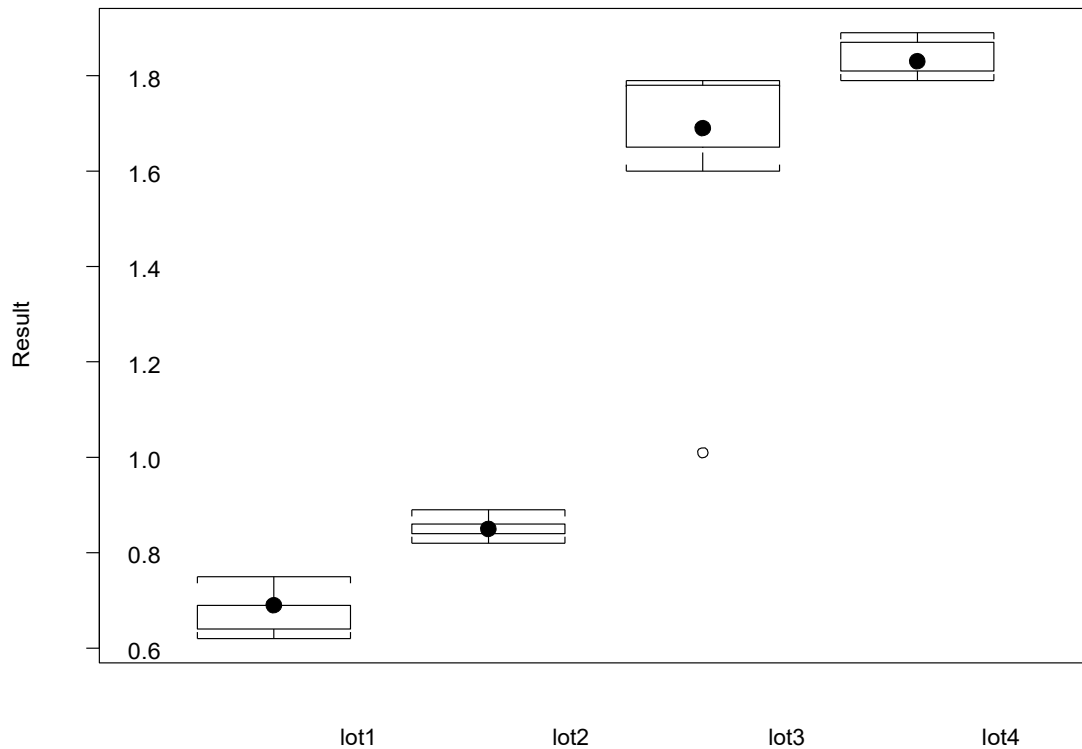
1.7 Detection of outliers

- a) Prior to any analysis, it is important to ensure that the data are of good quality. A good start in data quality assessment is to visualize the data using an appropriate graphic. A very useful tool to quickly get an overview of the data is the boxplot: it shows the location and spread of the data, as well as skewness and outside values. At the center of the boxplot is a dot that represents the median of the data (50% of the data are below the dot). Surrounding the dot is a box. The ends of the box represent the lower and upper quartiles (25% and 75% of the data are below these). Whiskers are drawn outward from the ends of the box a distance of $1.5 \times (\text{Upper quartile} - \text{Lower quartile})$ and then shortened to the nearest data point. Any values outside the whiskers are plotted as individual data points.

Example: Three laboratories ran a method validation programme and provided the following results:

Lot	lot1			lot2			lot3			lot4		
	1	2	3	1	2	3	1	2	3	1	2	3
Lab 1	0.71	0.69	0.75	0.85	0.83	0.82	1.6	1.65	1.66	1.87	1.88	1.89
Lab 2	0.64	0.62	0.69	0.84	0.85	0.86	1.69	1.71	1.01	1.83	1.82	1.81
Lab 3	0.64	0.65	0.69	0.89	0.87	0.86	1.79	1.78	1.78	1.8	1.83	1.79

In the following graphic, the boxplots are plotted for each lot using the same vertical scale.



We clearly see a lot effect but also we identify quickly a big outside value for lot 3; coming back to the data, there appear to be a re-transcription mistake in lab 2, rep 3.

b) There are also some numerical tools that can be useful for automatically flagging exotic values in a dataset.

i **Hampel's Method**

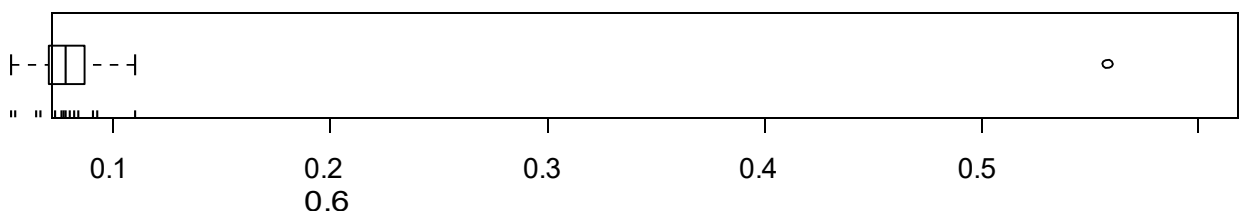
Hampel's method (Davies & Gather 1993) can be used for univariate samples.

Example: a laboratory runs a test on the same lot and obtains the following results for 15 replications:

Rep	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_i	0.1193	0.1038	0.0923	0.1173	0.1494	0.1229	0.1125	0.1061	0.0940	0.1213	0.1314	0.1151	0.1159	0.1298	0.5977

The question is:

Can it be assumed that all the data points x_i come from the same distribution? Inspecting the box plot below, there is an exotic value around 0.6:



To use Hampel's method on this dataset:

- compute the median: $\bar{x} = 0.1173$ and the Median Absolute Deviation:

$$\text{MAD} = \text{median} \{ x_i - \bar{x} \} = 0.0112$$

- then, identify x_i as an outlier if $|x_i - \bar{x}| > 5.2 \text{ MAD}$

In the above example, only one value satisfies the above criteria: $x_{15} = 0.5977$

ii Cochran's test

Sometimes identifying variance (or standard-deviation) outliers is required. The appropriate method for this purpose is the Cochran's test.

Example: An inter-laboratory test is performed to obtain an estimate of global intra-laboratory variance. There are 6 laboratories participating in the test and they report the following variances based on 5 results each:

	Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6
s^2_i	7.6	4.2	3.1	20.1	3.5	3.2

The question is:

Can all these variances be used to compute a global estimate of the within-laboratory variance?

The Cochran's test is performed as follows:

- Let $g = \frac{\max(s_i^2)}{\sum_{i=1}^s s_i^2} = 0.4808$

- Then if $g \leq g_\alpha$ where g_α is given in specific tables for levels of significance α , equal to 0.05 and 0.01 usually and which are entered with the number of variances (here 6) and the number of observations used to compute each variance (here 5), the hypothesis that the 5 variances are equal is not rejected. In our example, $g_{0.05}$ is equal to 0.4803 and thus the hypothesis that the 5 variances are equal is rejected: the variance from Lab 4 will not be used for the computation of global intra-laboratory variance.

1.8 Statistical analysis, model and assumptions

There are a number of available methods to analyze a dataset. Two of them are quoted here.

- ANOVA (Analysis of Variance) is still a widespread useful and robust technique.
- GLM (Generalized Linear Model) has similarities with ANOVA and can offer more flexibility.

Analysis of Variance

Assumptions:

- (Independent observations; no influence of other observations on the result of an observation)
 - (Same variability of comparable sub-data sets)
 - Normal distribution
- The 2 assumptions in brackets are classically stated, but recent software make them less important.

Example:

```
***** Analysis of variance *****
Variate: score
Source of variation      d.f.      s.s.      m.s.      v.r.      F pr.
lab                      5         48.951    9.790     6.35     <.001
lot                      3         21.619    7.206     4.67     0.004
incubation               1         31.674    31.674    20.53    <.001
lab.lot                  15         39.764    2.651     1.72     0.053
lab.incubation           5         38.295    7.659     4.96     <.001
lot.incubation           3         12.545    4.182     2.71     0.047
lab.lot.incubation       15         26.695    1.780     1.15     0.315
Residual                 144        222.145    1.543
Total                    191        441.688
```

Generalized Linear Model

Assumptions

- (Independent observations; no influence of other observations on the result of an observation)
- Results are linearly modeled; different transformations can be used to reach the linearity. By linking for normal, binomial, Poisson and gamma distribution a transformation is made to get linear results.

In recent software the assumption in brackets is not of concern (auto correlations are included in computations)

Example

```
Regression analysis
Response variate:  score28
Binomial totals:   100
Distribution:       Binomial
Link function:     Logit
Fitted terms:      Constant + lab + lot + lab.lot
Accumulated analysis of deviance
```

Change	d.f.	deviance	mean deviance	deviance ratio	approx chi	pr
+ lab	5	36.431	7.286	7.29	<.001	
+ lot	3	2442.347	814.116	814.12	<.001	
+ lab.lot	15	24.341	1.623	1.62	0.060	
Residual	168	170.855	1.017			
Total	191	2673.973	14.000			

Any statistical analysis consists of applying a model to the dataset, and for each method a number of assumptions are made. The check of the

assumptions is part of the analysis. When some assumptions are not met, robust analysis can still be performed, but the departure from the assumption must be noted in the report.

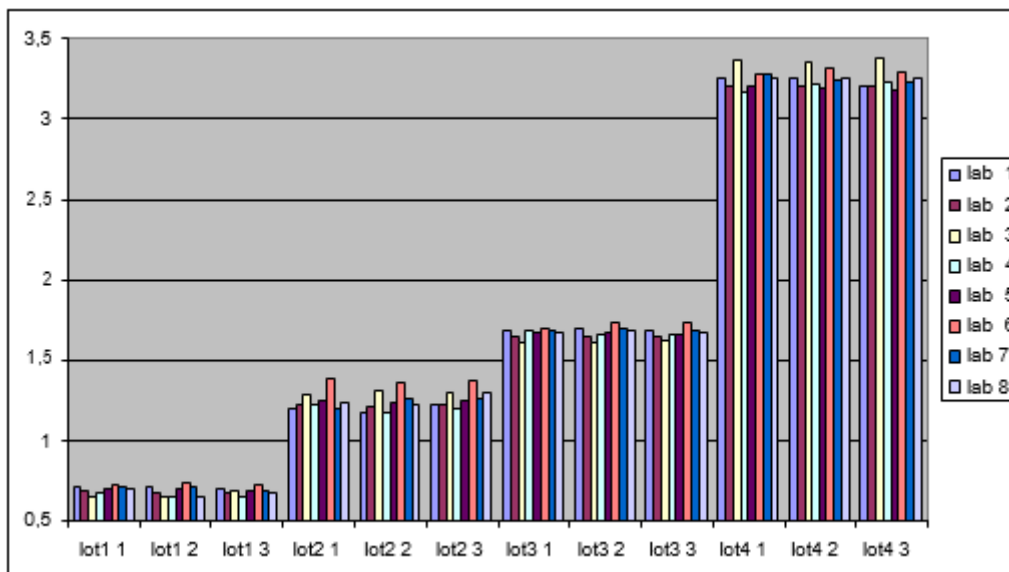
1.8.1 ANOVA (ANalysis Of VAriance)

- a) ANOVA is recommended as an exploratory and reporting tool when a data set is fitted for this kind of analysis (quantitative values).

Example: Eight laboratories ran a method validation program.

They received, blindly coded, 12 samples (3 repeats from 4 different lots covering the range of the method)

sample	rep	mean	lab 1	lab 2	lab 3	lab 4	lab 5	lab 6	lab 7	lab 8
lot1	1	0,70	0,71	0,69	0,66	0,67	0,70	0,73	0,71	0,70
lot1	2	0,68	0,71	0,67	0,65	0,65	0,70	0,74	0,71	0,65
lot1	3	0,69	0,70	0,68	0,69	0,66	0,69	0,73	0,69	0,68
lot2	1	1,25	1,20	1,22	1,28	1,23	1,25	1,39	1,20	1,24
lot2	2	1,24	1,18	1,21	1,31	1,18	1,24	1,36	1,26	1,22
lot2	3	1,27	1,23	1,22	1,30	1,20	1,25	1,37	1,26	1,30
lot3	1	1,67	1,68	1,64	1,61	1,68	1,67	1,70	1,69	1,67
lot3	2	1,67	1,70	1,64	1,61	1,66	1,67	1,73	1,70	1,68
lot3	3	1,67	1,68	1,65	1,62	1,66	1,66	1,73	1,68	1,67
lot4	1	3,25	3,26	3,20	3,37	3,16	3,20	3,27	3,27	3,25
lot4	2	3,25	3,26	3,20	3,36	3,22	3,19	3,31	3,24	3,26



- b) Test design:
 The simplest is recommended for method validation, only 2 factors:
 a laboratory factor (here with 8 modalities)
 a lot factor (here with 4 modalities)

As labs received more than 1 sample per lot (here 3 samples per lot), interaction lab*sample can also be computed.

ANOVA structures the variability in different components;

- laboratory,
- lot
- remaining variability; not controlled in the design, usually named residual.

c) Significant difference computation:

ANOVA tell whether the differences between the means from the different laboratories are significant for a given probability level (usually $\alpha=5\%$).

Usually the lots will be found significantly different, which is logical as the test organiser selected different levels.

If lots are supposed to be different (i.e. low, intermediate and high level) in the testing plan and ANOVA concludes there are no significant differences among lots, the results should be treated with caution.

Examples of Analysis of Variance table

Examples of Analysis of Variance table

- Source Value	Sum of Squares	Df	Mean Square	F-Ratio	P-

- MAIN FACTORS					
A:lab	0,0811791	7	0,011597	34,57	0,0000
B:sample	86,8585	3	28,9528	86319,08	0,0000
INTERACTIO					
N					
A*B	0,0943291	21	0,00449186	13,39	0,0000
RESIDUAL	0,0214667	64	0,000335416		

- TOTAL (CORRECTED)	87,0555	95			

A significant difference between laboratories is found when the ratio (variability explained by lab factor)/(non controlled variability) is great enough.

This ratio is called an F value.

In the example above the F ratio is 34.57 for the lab factor and the probability value is less than 5%, signifying a significant difference among laboratories.

d) Different situations:

There is not a unique rule to decide about the consequence of significance, or non significance among laboratories, but a statistician can provide assistance.

In method validation the aim is that the different laboratories will obtain comparable results (no significant difference among laboratories).

But, there are in particular two situations which need special attention.

- (i) When there are no significant differences among laboratories, but the power of the test is poor. This can occur when the non controlled variability is big. In that situation big differences among laboratories are not declared distinct. A retest is usually necessary, so that the variability of results on repeats can be reduced, and/or more repeats can be checked.
- (ii) When the laboratory variability of repeats is small, ANOVA can declare distinct differences which in practice are considered as too little to be of practical significance.

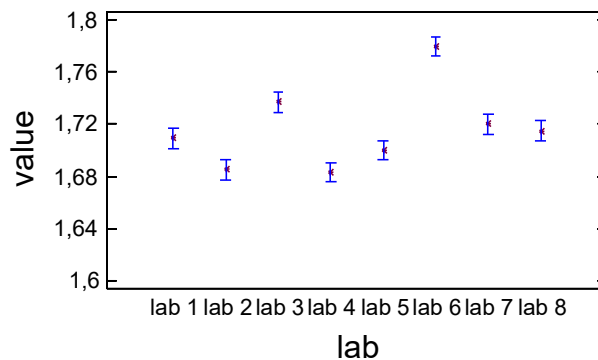
Note that any difference computed to be significant is a difference between means. In the case of differences among laboratories, the difference is computed from $X_{\text{samples}} * Y_{\text{repeats}}$ ($4*3=12$ results in our example). This difference can not be used directly to estimate the significant difference between 2 raw results from 2 laboratories.

- e) Statistical checks to perform in case ANOVA is used (check equivalence of variances, check residuals, transformation if necessary)

ANOVA assumes that all labs have about the same variability in their results.

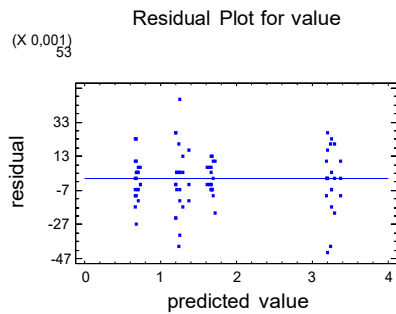
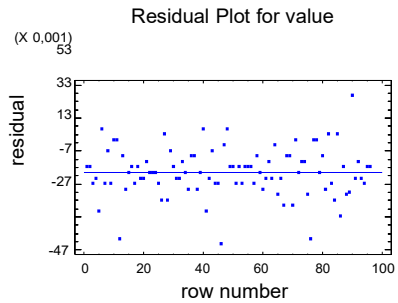
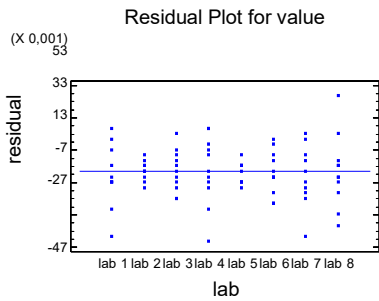
On the graph below this is illustrated by the fact that the width of the confidence interval is the same for all laboratories; ANOVA uses a pooled estimation of this variability to check differences among laboratories.

Means and 95,0 Percent LSD Intervals



Other types of computations can be performed if the assumption (similar variability) is not true. A Cochran's test can be performed to statistically check if variances are of the same magnitude.

A graphical representation of residuals computed by ANOVA is recommended, to be interpreted with the assistance of a statistician.



Sometimes transformation of data can be suggested by the statistician.

There is no free tool to compute ANOVA on the ISTA web site. Almost any statistical package has this capability.

1.8.2 GLM

...to be inserted in the future

1.9 Repeatability-reproducibility according to ISO5725-2 (quantitative data)

a) Introduction

When the results of the tests are appropriate ISO 5725-2 is recommended to compute repeatability and reproducibility.

Repeatability quantifies the average variability of results within each laboratory, when repeats are made on samples from a given lot.

Reproducibility is repeatability, increased by the variability of results from laboratory to laboratory.

Outliers are removed, to compute the repeatability and reproducibility values to report.

The model is $y = m + B + e$ where

m is the general mean

B is the laboratory component of bias under repeatability conditions

e is a random error occurring in every measurement under repeatability conditions

b) Repeatability conditions

In a design with L laboratories, l lots, r repeats; the test organiser needs to prepare $L*r$ samples from each given lot.

Each of the $L*r$ samples must be representative of the lot.

Usually the laboratory will obtain only one result for a given received sample.

To repeat r times the test on a given lot, the lab will receive r separate samples to test.

All samples are supposed to be tested following the technical protocol under the same conditions.

c) Reproducibility conditions

All labs are supposed to receive identical or similar samples, and to follow the same technical protocol.

At least 2 repeats are needed to compute repeatability and reproducibility. Usually 2 or 3 repeats are tested, sometimes 4 or 5.

d) Values

Repeatability and Reproducibility values are obtained for each lot tested.

For each lot, they take into account the results from the L participating laboratories.

Unless impossible, the test design must be the same for all laboratories.

Repeatability and Reproducibility can be expressed as variances, or as standard deviations. It is important to specify which in the report. Repeatability and Reproducibility are expressed using the units of measurement. Care must be taken when comparing values obtained in different studies.

Example of data set

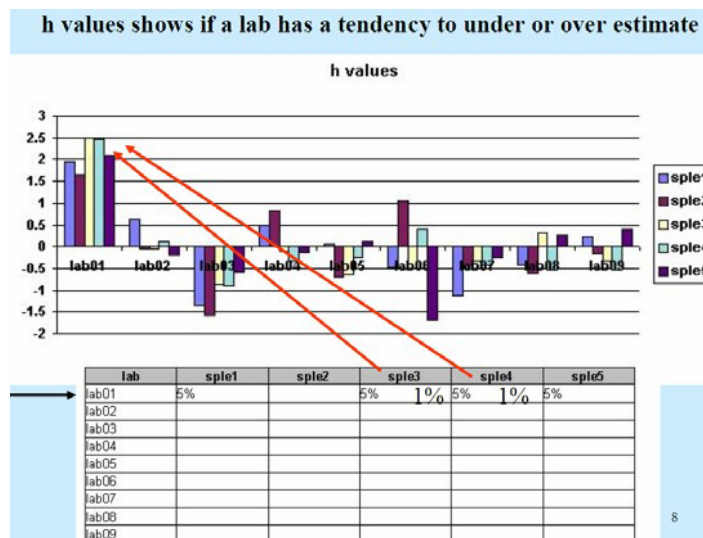
Each laboratory received 10 samples, 2 repeats from 5 different lots

laboratory	lot	repeat	value
lab01	lot1	rep1	4.44
lab01	lot1	rep2	4.39
lab01	lot2	rep1	9.34
lab01	lot2	rep2	9.34
lab01	lot3	rep1	17.4
lab01	lot3	rep2	16.9
lab01	lot4	rep1	19.23
lab01	lot4	rep2	19.23
lab01	lot5	rep1	24.28
lab01	lot5	rep2	24
lab02	lot1	rep1	4.03
lab02	lot1	rep2	4.23
lab02	lot2	rep1	8.42
lab02	lot2	rep2	8.33
lab02	lot3	rep1	14.42
lab02	lot3	rep2	14.5
lab02	lot4	rep1	16.06
lab02	lot4	rep2	16.22
lab02	lot5	rep1	20.4
lab02	lot5	rep2	19.91
etc...	etc...	etc...	etc...

Free software is available on the ISTA web site to compute repeatability and reproducibility values according to ISO 5725-2.

sample	nb of labs	observations mean	std dev repeatability sr2	std dev reproducibility sR2
high	9	3,246	0,020	0,060
low	9	0,803	0,597	0,580
medium1	9	1,252	0,023	0,059
medium2	9	1,670	0,009	0,031
medium 3	9	2,025	0,025	0,042

It also provides graphs with h and k values.



These values help to see which laboratories are over estimating or under estimating in comparison to the mean for all laboratories (h values), and to see the variability of results in each laboratory for each lot (k values).

Binomial and poisson data

Consider the following random effects model:

$$y_{ij} = \mu + L_i + e_{ij}$$

Where:

- . μ is the intercept
- . L_j is the effects of the i^{th} lab, these effects are iid $N(0, \sigma^2_{Lab})$
- . e_{jj} are the residuals, they are iid $N(0, \sigma^2)$

If we call s^2_{Lab} and s^2 the estimates of σ^2_{Lab} and σ^2 , then:

Repeatability variance = s^2

and reproducibility variance = $s^2 + s^2_{Lab}$

Now, let's consider the following binomial and poisson random effects model:

... to be inserted in the future

then we can define repeatability and reproducibility in the following way for binomial or poisson data *to be developed in the future*

Annex 2: General Procedures for Method Validation

2.1. Organization and Design of the Test Plan

It is the responsibility of the person organizing the comparative test to develop a scheme appropriate to the particular seed quality test. The test plan will include details of how the method is to be validated.

Details of the method validation

- a) The name and address of the test organiser and other personnel involved in the design and operation of the comparative test.
- b) Outline of the test method as follows:
 - i. The species.
 - ii. Scope of the method.
 - iii. Full details of the test protocol to be followed.
- c) Identification of the following parameters of the proposed protocol:
 - i. Accuracy.
 - ii. Precision (repeatability and reproducibility).
 - iii. Characterisation/specifications of the proposed method.
- d) Interferences (specificity):

Where relevant, the effects of any factor (e.g. seed weight, seed moisture content) on performance of the method.
- e) Critical steps or parameters:
 - i. instructions for handling and storage of samples and other materials used in the test, during all phases of the test.
 - ii. Control and recording of environmental conditions (temperature, humidity etc) during testing, if applicable.
 - iii. Comparison with other methods, if applicable.
- f) Quality control of method performance:
 - i. Information/specifications/requirements of the test material to be used.
 - ii. Minimum number of samples to be used in the test. iii. Reference samples to be included, if applicable.
 - iv. Reference reagents to be included, if applicable.
 - v. Preparation of samples (e.g. moisture content adjustment), if applicable.
- g) Identify and describe safety considerations while performing the test, if applicable:
 - i. Materials and chemicals that provide a significant health hazard.
 - ii. Potential hazards in handling or storage of reagents, samples, or standards.
 - iii. Any other hazards which might be important.
- h) Participating laboratories:

The test plan should include a statement of the following:

- i. The minimum number of participants to be included in the test (depending on the objectives and statistical analysis being used).
 - ii. Criteria which need to be met before participation is allowed.
 - iii. A list of potential participants.
 - iv. A clear indication of the expected amount of work for each participating laboratory.
- i) Data analysis and reporting:

An outline should be given of how the data are to be validated and the statistical methods to be used in the analysis of the data.
- j) Other details:
 - i. Problem solving and contingency plans.
 - ii. A time schedule for the test.
 - iii. A registration form for the participants which should include:
 - the species under study
 - objectives of the test
 - number of samples and replicates in the test
 - date samples sent to participants
 - date samples and other materials received
 - condition of samples and other materials on arrival
 - date test begun
 - date test completed
 - comments and remarks during testing
 - date results sent to test organiser
- k) Data record sheets and instruction as to how to complete these sheets.

2.3. Identification of Participating Laboratories

- a) Criteria for identification of participating laboratories are:
 - i. Only laboratories experienced in seed quality testing or in applying the specific techniques to be evaluated in the comparative test should be invited to participate in the programme.
 - ii. Participating laboratories must understand and accept their obligations in participating in a comparative test programme.
 - iii. In cases where there is some doubt as to the ability of a laboratory to perform the test a *pre-test* may be used. On the basis of the results of this test the test organiser will decide which laboratories can be involved in the comparative test.
 - iv. Participating laboratories must have the skilled personnel, appropriate facilities and equipment recommended for performing the test.

2.4. Preparation and Storage of Seed Samples

- a) The test organiser makes the final choice of the samples/lots to be included in the test.
- b) Whenever possible seeds having naturally determined levels of quality should be used in method validation tests. Samples should come from seed lots originating from normal seed production and storage systems. They may originate from more than one year of production. If difficulties are encountered in obtaining seeds having naturally different levels of quality, seeds in which quality differences have been artificially induced e.g. by artificially ageing seeds or by introducing a pathogen, may be used.

- c) Whenever possible the test organiser should collect and maintain relevant information pertaining to the sample(s) used in the test. This could include:
 - i. month and year of harvest.
 - ii. conditions and time of any storage period.
 - iii. nature of any seed treatment (chemical or other).
 - iv. seed moisture content (fresh weight basis).
 - v. laboratory germination.
- d) Three levels of seed quality should be used:
 - i. Low
 - ii. Medium
 - iii. HighSeed lots showing these quality differences should be identified by pre-tests carried out by the test organiser.
- e) The number of replicates will be determined by the test organiser and will depend on the protocol for the test method(s) under evaluation, the aim of the test and the statistical analysis. All replicates should be allocated a code number/letter randomly and known only to the test organiser so that all replicates are tested "blind".
- f) When not in use, samples should be stored under conditions to avoid any deterioration of the seed that might result in a reduction in seed quality (normally 5-10°C, in sealed moisture proof containers to minimise any change in seed moisture content from that on receipt of the seed).

2.5. Documentation

Detailed instruction covering all aspects of the protocol, which should be adhered to by the participating laboratories, should be provided. Instructions may include -

- i. Details concerning factors which could influence the testing of supplied samples.
- ii. Specific instructions on the recording of results (e.g. units, reporting basis, result headlines etc.).
- iii. The requirement to label samples legibly and without ambiguity

2.6. Distribution

- a) Shipping cartons should be well-packed and labelled properly to avoid chance misplacement, losses and transportation delays. Samples should be packed in a waterproof and shockproof cardboard box, and in a manner to preserve the integrity of the samples; they should be packed to minimise the possibility of breakage of containers. Collaborators should be notified of shipping arrangements, including waybill numbers, arrival time and required storage conditions. If necessary, use special transportation services. For international delivery, mark as "*Laboratory research samples-no commercial value*" or other designation required by customs regulations of the country to which the package is being sent. Phytosanitary certificates should be provided where necessary. In some instances special permits may be required.

- b) Include a return slip, to confirm safe receipt with each package. A copy of the methodology, and report forms should also be included.
- c) Provide instructions for proper storage of laboratory samples between unpacking and analysis and thereafter for remnant samples.

2.7. Obligations of Participants

- a) Receipt of materials
 - i. On receipt of samples and documentation, the participants must verify that the samples and documentation are in good condition. Forms confirming receipt of samples must be completed and returned to the test organiser.
 - ii. If there are any remarks, problems, or uncertainties the participant must contact the test organiser immediately for a solution.
 - iii. Between receipt of samples and actual start of the test, the samples and other test materials must be stored in optimal conditions in agreement with the instructions mentioned in the protocol or test plan.
- b) Analysis of material according to submitted protocol
 - i. In order to compare the results of the different laboratories it is important that the tests be performed in the same period and exactly according to the test protocol. Therefore participants should not start testing earlier or later than indicated in the time schedule of the protocol.
 - ii. The participant must follow the instructions and methods given in the protocol precisely. When there are deviations, problems or when a sample must be re-tested, this information must be recorded with the data on the test report.
 - iii. The participant must conduct exactly the number of determinations stated in the instructions. Any other number complicates the statistical analysis. Too few determinations may require discarding the results from that laboratory for that material or inserting missing values; too many values may require discarding the contribution of that laboratory or at least some of the values.
- c) Supply raw data, graphs, photographs or other documents as requested in the instructions.
- d) If analytical results appear unreasonable, investigate causes immediately.
 - i. Check for transcription and calculation mistakes, then conduct a re-analysis if necessary and permitted by the protocol. Contact test organiser to discuss suspicious values.
 - ii. Since participants have no basis for judging whether a value is an outlier, the results should be communicated to the test coordinator as soon as the protocol is complete so that repeat tests, if necessary, may be performed at once.
 - iii. If test organiser indicates a value may be an outlier, review the determination promptly to the extent possible, by recalculation and/or reanalysis. If time and materials are available, obtain new samples for repeat analysis.
 - iv. The most frequent causes of correctable outliers are:
 - Incorrect calculations and arithmetic errors
 - Errors in reporting, such as transposition of numbers
 - Misplacement of the decimal point or use of the wrong unit
 - Contamination of reagents, equipment or test materials

v. Following completion of repeat tests the participant should complete data record sheets and return forms to the test organiser.

2.8. Analysis of Results

- a) The test organiser must check the data for completeness and ensure that it is valid for further analysis. If not the test organiser contacts the participant to try and resolve the problem. When the results of a test are not clear, then the results from that participant must be rejected.
- b) When all the results are received and verified, statistical analysis of the results is arranged by the test organiser.

2.9. Statistical Approach

The statistical analysis must be carried out according to the statistical design in the test plan.

2.10. Final Report

- a) When the statistical analysis is completed the test organiser prepares the draft report with results, statistical analysis, conclusions, action points and recommendations.
- b) The results of each participant are included in the report under a coded identifier. Each participant is informed of their own code.
- c) The draft report should be sent to each of the participants.
- d) Participants have three weeks to send comments on the draft report back to the test organiser.
- e) The test organiser considers any comments from participants, makes amendments if necessary, and produces the final report.

2.11. Confidentiality and Ethical Considerations

- a) In reporting of results the identity of individual participants should remain confidential. The identity of participants should only be known to the minimum number of people involved in coordinating the test programme.
- b) For the purposes of discussion and mutual assistance in improvement, a group of participants may elect to waive confidentiality.

Annex 3: Instructions to Authors: Preparation of Test Method Descriptions

- 3.1 The test method must be written in English.
- 3.2 The test method must follow the same style and format as test methods in the current ISTA Rules with sections appropriate to the particular test method.
- 3.3 The test method must be supported by a technical report which clearly states the basis for the test method and gives estimates of the detection limits, reproducibility, repeatability, and uncertainty levels of the method.
- 3.4 The test method description must be clear and unambiguous, remembering that the test method may be followed by persons whose first language is not English.
- 3.5 The test method must be technically sound.
- 3.6 All critical steps must be identified.
- 3.7 All reagents and instruments must be described in performance terms with system suitability tests where relevant.
- 3.8 Sources of non-commodity materials and reagents should be given.
- 3.9 Quality control points must be identified and be adequate to ensure reliable test results.
- 3.10 All tables, figures and terms must be clearly explained.

Annex 4: Test Method Validation Application Form

4.1. Method class

Multi-laboratory Validation [] Peer Validation []
Performance Validation []

4.2. Method applicability

For most tests this refers to the plant species for which the method is intended to be used, but in some cases other information will be required (e.g. for health testing, the name of pathogen(s) and host(s)).

4.3. This method is considered a

New Method [] Additional Method []
Method Modification [] Replacement Method []

4.4. Brief description of the method

4.5. Submitter's information

Name: _____
Organisation: _____
Address: _____
City: _____ Postal/Zip Code: _____ Country: _____
Telephone: _____ Fax: _____
Email: _____

4.6. Proposed technical reviewer's information

a) Name: _____ Organisation: _____
Address: _____
City: _____ Postal/Zip Code: _____ Country: _____
Telephone: _____ Fax: _____
Email: _____

b) Name: _____ Organisation: _____
Address: _____
City: _____ Postal/Zip Code: _____ Country: _____
Telephone: _____ Fax: _____
Email: _____

4.7. List of accompanying documentation

(Give filenames of electronically submitted documents)

- Draft test method sheet []
- Technical report []
- Test plan []
- Other supporting documents []

Note:

It is the responsibility of the submitter to ensure that documents are presented in the correct format. Failure to do so may result in return of the documents to the submitter without review.

Submitters should check test method drafts against the evaluation criteria, which can be found in Appendix 5.

This application form can be downloaded from the ISTA Website.

Return completed application, with accompanying documentation, to the ISTA Secretariat.

Annex 5: Instructions for Reviewers: Draft Test Plan

Please review the enclosed draft test plan with reference to the evaluation criteria below, making comments on additional sheets as appropriate.

Test plan title:

Author:

Submission date:

Reviewer name:

Review request date:

Review returned date:

The method described in this draft test plan should be considered as a: New

Method Additional Method

Replacement Method Method Modification

Evaluation Criteria (not all aspects will necessarily apply):

	Yes	No	See Comments
Is the test plan presented in the correct format?			
Is the nomenclature/taxonomy correct?			
Is the purpose of the method and need for validation adequately explained?			
Is the method description clear and unambiguous?			
Are parameters for accuracy, repeatability, reproducibility and uncertainty of the test method identified?			
Are relevant safety precautions adequate?			
Are any reagents and apparatus described or defined in performance terms?			
Is the method described suitable for meeting the objective(s) of the test?			
Are relevant critical steps/parameters identified?			
Are parameters for quality control of method performance defined?			
Are potential participating laboratories identified?			
Are data analysis methods given appropriate?			
Is a participant registration form included?			
Are data record sheets and instructions for their completion included?			
Are all tables, figures and terms sufficiently explained?			
Are all references cited, and cited correctly?			

Please make any additional comments on a separate sheet.

Recommendation (delete as applicable).

- a) Approve the Draft Test Plan without revision.
- b) Approve the Draft Test Plan following minor revisions. c)
Defer a decision pending major revisions.
- d) Reject the Draft Test Plan.

Are you happy for your name to be revealed to the authors?

Yes

No

Annex 6: Instructions to Authors: Preparation of Validation Reports For New Test Methods or Revisions to Existing Test Methods.

6.1. The Validation Report:

- a) Should be a self-contained document suitable for publication in ISTA Method Validation Reports.
- b) Should follow the general style requirements of Seed Science and Technology (see instructions to authors in a recent edition of SST or on the ISTA Online web site). However it is understood that data sets may, when required, be attached as Appendices.
- c) Should clearly state the proposed new method or revisions that have been validated.
- d) Should present justifications for the new method or revisions based on the results of scientific studies contained in the report itself.

6.2. The report should normally contain the following sections:

- a) In all cases:
 - i) Title – which should begin: _____ Committee Technical report:
Validation of a revised/new method for
 - ii) Authors – names and addresses of authors.
 - iii) Summary – a short summary of the study and the validated method.
 - iv) Introduction – stating the problem, reasons for the study, the purpose of the method, pertinent background information and history of the method with reference to previously published information and if appropriate the objective(s) of the collaborative study(ies).
- b) If reporting the results of scientific studies directly:
 - i) Materials and methods – full details of the materials and methods used and design of the study, including the method(s) of statistical analysis.
 - ii) Results – of the study, statistical analysis and summaries of the data in the form of tables and/or figures, presented in sufficient detail and with appropriate measures of variation to allow the reader to draw independent conclusions. If appropriate, justifications for exclusion of raw data.
 - iii) Discussion – discussion of the method performance including comments from collaborators and how they were addressed.
- c) If referring to scientific studies published elsewhere:
 - i) Summaries and discussion of the external studies – results of external studies which have been published elsewhere should be summarised/reviewed and discussed in terms of method performance.
- d) In all cases:
 - i) Proposed changes and justification – if appropriate clearly identify proposed changes to an existing method with justification.
 - ii) Test results of repeatability and reproducibility – give estimates of the repeatability and reproducibility of the test method and how these were calculated.

- iii) Levels of uncertainty for the method – provide uncertainty estimates (tolerance table data for many methods) and how these were calculated.
- iv) Conclusions and recommendations – a clear statement of the conclusions of the validation and recommendations for actions.
- v) Acknowledgements – of collaborators (if not co-authors), funding bodies, etc. as appropriate.
- vi) References – details of all cited references.

6.3. Validation Reports which are not presented in the correct format and/or which do not fulfil these requirements will be returned for revision.

6.4. Raw data.

- a) A hard copy and electronic copy (spreadsheet; database) of the raw data should be deposited with the ISTA Secretariat. To maintain confidentiality, the identity of individual participating laboratories should be indicated by a coded identifier.

6.5. Copyright.

- a) Submission of a report implies that the work described has not been published elsewhere, except in the form of a poster, an abstract or a thesis, that it is not under consideration for publication elsewhere, and that all co-authors have approved the report. The International Seed Testing Association will retain the copyright of the method and the report.

Annex 7: Instructions For Reviewers: The Validation Report

Please review the enclosed validation report with reference to the evaluation criteria below, making comments on additional sheets as appropriate. Please indicate any aspects on which you do not feel qualified to comment.

Method:
 Author:
 Submission
 date:

Reviewer Name:
 Review request
 date: Review
 returned date:

The method should be considered as:

New Method Additional Method
 Replacement for Method Method Modification

Evaluation Criteria (not all aspects will necessarily apply):

	Yes	No	See comments
Is the title appropriate?			
Is the summary clear/adequate?			
Is the reason for the study clearly stated? (i.e. objective(s), aim, questions, hypothesis that test organiser wishes to address)			
Has/have previous literature/data been reviewed adequately?			
Is the cited literature appropriate, are there any omissions?			
In the case of inter-laboratory comparative test – is there evidence that the guidelines have been followed as far as possible?			
Have technical difficulties/problems identified during the validation process been highlighted?			
Have the comments of participants been reported/addressed?			
Was the design of the validation appropriate?			
Were the controls adequate to ensure repeatability and reproducibility of the data reported?			
Were reference materials included and are their results reported?			
Were steps taken to ensure the integrity of the data, i.e. blind testing/coding of samples?			
Were checks included to ensure that each participant followed the protocol?			
Has a statistical analysis been performed?			
Is the statistical analysis appropriate to the data, and has the approach been justified?			
Have sufficient data been presented to allow independent assessment?			
Is the exclusion of particular data/laboratories from the analysis justified?			

Has the accuracy, reproducibility and repeatability of the method(s) been estimated and clearly stated?			
Are the conclusions justified by the data and statistical analysis?			
Are all tables, figures, equations, and terms sufficiently explained?			
Are the summaries (graphs/tables) of the data appropriate?			
Could any figures or tables be explained by a simple statement?			
Have the conclusions and recommendations been clearly stated?			
Are the references correct?			
Are all the cited reports/data available?			
Is the method fully justified by the Method Validation Report?			
Have steps been taken to archive the raw data to ensure availability for re-analysis/future studies?			

Please make comments on an additional sheet.

Recommendations (Delete as applicable)

- a) Approve the validation report without revision.
- b) Approve the validation Report following revisions.
- c) Defer a decision pending major revisions.
- d) Reject the Validation Report

Are you happy for your name to be revealed to the author(s)? Yes No

Annex 8: Performance Validated Test Method Submission Requirements for other than Test Kits

To support the performance claim, the applicant should provide the following information:

- 8.0 Name of method
- 8.1. Scope of method
- i) Intended use of the method (e.g. quality control, enforcement).
 - ii) Type of method (e.g. screening, reference).
 - iii) Applicability of method (e.g. species).
- 8.2. Within-laboratory performance of method
- Provide data for:
- trueness or bias (systematic error)
 - recovery
 - limit of detection; limit of quantitation
 - repeatability
- 8.3. Characterisation/Specifications of method
- Provide information on:
- interferences (specificity): impurities, contaminants, additives, etc
 - performance specifications and acceptability criteria for media, reagents, instruments
 - suitability tests for systems
 - critical steps or parameters
 - comparison with other methods if available
- 8.4. Quality control of method performance.
- Provide information/specifications/requirements for:
- reference materials
 - standards
 - fortification samples (if applicable)
- 8.5. Safety information.
- i) Provide appropriate cautionary statements for any hazards associated with the method for health (organisms and/or their products, or substances which are carcinogenic, mutagenic, teratogenic, allergenic, pathogenic, radioactive, etc).
 - ii) Provide information for any special procedures required for the disposal of reagents or reaction products.
 - iii) Provide information on potential hazards associated with handling or storage of reagents, samples or standards.

Annex 9: GLOSSARY OF TERMS USED IN ISTA METHOD VALIDATION STUDIES

- Accuracy
- Bias
- Certification Mark
- Comparative Test
- Critical Control Point
- Cross Reactivity
- Error
 - Error of Measurement
 - Random Error
 - Systematic Error
- Fitness for Purpose
- Limit of Detection
- Limit of Quantitation
- Linearity
- Measurand
- Measurement
- Method (of a measurement)
- Multi-laboratory Study
- Peer Review
- Peer Validation
- Performance Approved Method
- Performance Based Method
- Performance Validated Method
- Precision
 - Intermediate Precision
 - Measurement Precision
- Proficiency Testing
- Quality
 - Quality Assurance
 - Quality Control
 - Internal Quality Control
- Recovery
- Reference Material
 - Certified Reference Material
- Repeatability
- Reproducibility
- Result of a Measurement
- Ruggedness Test
- Specified Traits
- Sponsor
- Standard Deviation
- Test Kit
- Test organizer
- Test Plan
- Test Report
- Traceability
- Trueness
- Uncertainty (of Measurement)
- Validation
 - Method Validation
- Value
 - Accepted Reference Value
 - True Value
- Verification

Accuracy:

'The closeness of agreement between a test result and the accepted reference value.'

Note: *The term accuracy, when applied to a set of test results, involves a combination of random components and a common systematic error or bias component.* [ISO 3534-1].

Accuracy (of Measurement):

'Closeness of the agreement between the result of a measurement and a true value of the measurand.'

Accuracy (of a Measuring Instrument):

'Ability of a measuring instrument to give responses close to a true value.'

Note: *In these contexts accuracy is a qualitative concept.* [IUPAC 'Orange' Book]. The term 'precision' should not be used for 'accuracy'. [VIM, 1993].

Bias:

'The difference between the expectation of the test results and an accepted reference value.'

Note: *Bias is the total systematic error as contrasted to random error. There may be one or more systematic error components contributing to the bias. A larger systematic difference from the accepted reference value is reflected by a larger bias value.* [ISO 3534-1]

Bias (of a Measuring Instrument):

'Systematic error of the indication of a measuring instrument.'

Note: *The bias of a measuring instrument is normally estimated by averaging the error of indication over an appropriate number of repeated measurements.* [VIM, 1993].

Certification Mark

'Identifier given by ISTA to a test kit manufacturer which signifies that the test kit has been granted Performance Validated Method status.'

Comparative Test

1. 'A comparison of different test methods to determine which one of these tests should be adopted as a validated test method.'
2. 'A multi-laboratory study of a test method used in method validation studies.'

Critical Control Point (CCP):

In the HACCP (Hazard Analysis and Critical Control Point) approach, CCP is 'a point, step or procedure in a process at which control can be applied, and an adverse event can, as a result, be prevented, eliminated or reduced to acceptable levels.'

Cross Reactivity:

'Response (of method) to analogues, metabolites, or other non-target components that may be present in the matrix(es).' [AOAC - PVMC].

Error (of Measurement):

'The value of a result minus the true value.' [IUPAC Compendium of Chemical Technology, 1985].

'Result of a measurement minus a true value of the measurand.' [VIM, 1993].

Note: *Since a true value cannot be determined, in practice a conventional true value is used.*

Random Error:

'Result of a measurement minus the mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions.'

Note: *Random error is equal to error minus systematic error. Because only a finite number of measurements can be made, it is possible to determine only an estimate of random error.* [VIM 1993].

Systematic Error:

'Mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions minus a true value of the measurand.'

Note: *Systematic error is equal to error minus random error. Like true value, systematic error and its causes cannot be known.* [VIM 1993].

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.' [IUPAC 'Orange' Book].

Limit of Detection:

'The lowest content that can be measured with reasonable statistical certainty.' [AOAC - PVMC].

Limit of Quantitation:

'The lowest concentration of an analyte that can be determined with acceptable precision (repeatability) and accuracy under the stated conditions of the test.' [NATA Tech Note #13].

Linearity:

'Defines the ability of the method to obtain test results proportional to the concentration of analyte.'

Note: *The Linear Range is by inference the range of analyte concentrations over which the method gives test results proportional to the concentration of the analyte.* [AOAC - PVMC].

Measurand:

'Particular quantity subject to measurement.'

Note: *Specification of a measurand may require statements about quantities such as time, temperature and pressure.* [VIM 1993].

Measurement:

'Set of operations having the object of determining a value of a quantity.' [VIM 1993].

Method (of a Measurement):

'Generic description of a logical sequence of operations used in a measurement.'

Multi-Laboratory Study

'A comparative test conducted by six or more laboratories during the method validation process.'

Peer Review

'A critical review of a document (e.g. test plan/test report) by an experienced researcher/statistician other than the author(s).'

Peer Validation

'A test method validation study conducted by a small (one to three) number of laboratories (c.f. multi-laboratory study).'

Performance Approved Method

'Method evaluated and approved according to the principles of the performance based approach for quality testing; usually restricted to bio-molecular tests and bioassays for testing for the presence of specified traits.'

Performance Based Method

'Synonym for Performance Approved Method.'

Performance Validated Method

'A performance approved/based method which has had its performance claims verified.'

Precision:

'The closeness of agreement between independent test results obtained under stipulated conditions.'

Note: Precision depends only on the distribution of random errors and does not relate to the true value or specified value. The measure of precision is usually expressed in terms of imprecision and computed as a standard deviation of the test results. Independent test results means results obtained in a manner not influenced by any previous result on the same or similar test object. Quantitative measures of precision depend critically on the stipulated conditions. Repeatability and Reproducibility are particular sets of extreme conditions. [ISO 3534-1].

Intermediate Precision:

'Intermediate precision expresses within laboratories variation: different days, different analysts, different equipment, etc.' [ICH Q2A, CPMP/ICH/381/95].

Measurement Precision:

'Closeness of agreement between quantity values obtained by replicate measurements of a quantity, under specified conditions.'

Note: Measurement precision is usually expressed numerically by measures of impression such as standard deviation, variance, or co-efficient of variation under the specified conditions of measurement. [VIM, 2004].

Proficiency Testing:

'A periodic assessment of the performance of individual laboratories and groups of laboratories that is achieved by the distribution by an independent testing body of typical materials for unsupervised analysis by the participants.' [IUPAC 'Orange' Book]

Quality:

'The totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs.' [ISO 9000:2000].

'Degree to which a set of inherent characteristics fulfils requirements.' [ISO/DIS 19011:2002].

Quality Assurance

'All those planned and systematic activities implemented within the quality system, and demonstrated as needed, to provide adequate confidence that an entity will fulfil requirements for quality.' [ISO 8402:1994].

'Part of quality management focused on providing confidence that quality requirements will be fulfilled.' [ISO 9000:2000].

Quality Control:

'The operational techniques and activities that are used to fulfil requirements of quality.' [ISO 8402:1994].

'Part of quality management focussed on fulfilling quality requirements [ISO 9000:2000].

Internal Quality Control:

'Set of procedures undertaken by laboratory staff for the continuous monitoring of operations and the results of measurements in order to decide whether results are reliable enough to be released.' [IUPAC 'Orange' Book].

Recovery:

'The fraction of analyte added to a test sample (fortified or spiked sample) prior to analysis of the unfortified and fortified samples; percentage recovery (%R) is calculated as follows:
 $\%R = [(CF-CU)/CA] \times 100$

Where CF is the concentration of analyte measured in the fortified sample; CU is the concentration of analyte measured in the unfortified sample; CA is the concentration of analyte added (measured value, not determined by method) in fortified sample.' [AOAC- PVMC].

Reference Material (RM):

'Material or substance one or more of whose property values are sufficiently homogeneous and well established to be used for the calibration of an apparatus, the assessment of a measurement method, or for assigning values to materials.'

Note: The term reference material describes materials which are often also called measurement standards, e.g. chemical substances used for calibration or identification purposes. Care is necessary when using the term 'standard' as it is commonly used in two different contexts. The term may refer to 'measurement standards' in the reference material sense, or it may refer to written standards, such as standard methods. It is important to ensure the distinction is always clear. [ISO/IEC Guide 30 - 1992, 2.1].

Certified Reference Material (CRM):

'Reference material, accompanied by a certificate, one or more of whose property values are certified by a procedure, which establishes its traceability to an accurate realisation of the unit in which the property values are expressed, and for which each certified value is accompanied by an uncertainty at a stated level of confidence.' [ISO/IEC Guide 30 - 1992, 2.2].

Repeatability:

Closeness of the agreement between the results of successive measurement of the same measurand carried out in the same conditions of measurement.' [IUPAC 'Orange' Book].

Reproducibility:

'Precision under reproducibility conditions, i.e. conditions where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment.

Note: A valid statement of reproducibility requires specification of the conditions changed. Reproducibility may be expressed quantitatively in terms of the dispersion of the results.' [ISO 3534-1].

Result of a Measurement:

'Value attributed to a measurand, obtained by measurement.'

Note: When the term result of a measurement is used, it should be made clear whether it refers to: the indication; the uncorrected result; the corrected result, and whether several values are averaged.

A complete statement of the result of a measurement includes information about the uncertainty of measurement. ' [VIM 1993].

Ruggedness Test:

'Intra-laboratory study to study the behaviour of an analytical process when small changes in the environmental and/or operating conditions are made, akin to those likely to arise in different test environments. Ruggedness testing allows information to be obtained on effects of minor changes in a quick and systematic manner.' [AOAC - PVMC].

Specified Traits

'A named characteristic which may be present in a seed lot.'

Sponsor

'Any method developer or test kit manufacturer who submits a test method for validation.'

Standard Deviation:

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

Test Kit

'A commercially packaged system of the principal or key components of a testing method. Test kits include directions for use and are often self-contained, complete analytical systems; however they may also require supporting equipment and supplies. The key components frequently represent proprietary elements or reagents that may be readily prepared only by the producer of the kit.'

Test Organiser

'A person designated by the test sponsor or appointed by the ISTA Technical Committee to develop and organise the validation test.'

Test Plan

'Detailed plan for execution of a comparative test by several laboratories on one or more test methods.'

Test Report

'Standard technical report of the results of a comparative test, with discussion of results, conclusions and recommendations prepared by the test organiser.'
'ISTA Method Validation Report.'

Traceability:

'Property of the result of a measurement or the value of a standard whereby it can be related with a stated uncertainty, to stated references, usually national or international standards (i.e. through an unbroken chain of comparisons).'

Note: *The standards referred to here are measurement standards rather than written standards.* [ISO/IEC Guide 30 - 1992, 3.8].

'Ability to trace the history, application or location of that which is under consideration.' [ISO 9000:2000].

'Property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties.'

Note: *The concept is often expressed by the adjective traceable. The unbroken chain of comparisons is called a traceability chain.* [VIM,].

Trueness:

'Closeness of agreement between the average that would ensue from an infinite number of quantity values obtained under specified measurement conditions and the true value of the measurand.'

Note: *Trueness cannot be expressed as a numerical value.
Trueness is inversely related to systematic error only.
The term 'trueness of measurement' should not be used for accuracy of measurement.*
[Draft International Vocabulary of Basic and General Terms in Metrology, VIM, April 2004].

Uncertainty (of Measurement) i.e. Measurement Uncertainty:

'Parameter associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand.'

Note: *The parameter may be, for example, a standard deviation (or a given multiple of it), or the width of a confidence interval. Uncertainty of measurement comprises, in general, many components. Some of these components may be evaluated from the statistical distribution of the results of a series of measurements and can be characterised by experimental standard deviations. The other components, which can also be characterised by standard deviations, are evaluated from assumed probability distributions based on experience or other information. It is understood that the result of the measurement is the best estimate of the value of the measurand and that all components of uncertainty, including those arising from systematic effects, such as components associated with corrections and reference standards, contribute to the dispersion.* [VIM 1993].

Validation:

'Confirmation by examination and provision of objective evidence that the particular requirements for a specified intended use are fulfilled.' [ISO 8402:1994].

Method Validation:

1. 'The process of establishing the performance characteristics and limitations of a method and the identification of influences which may change these characteristics and to what extent. Which analytes can it determine in which matrices in the presence of which interferences?
Within these conditions what levels of precision and accuracy can be achieved?'
2. 'The process of verifying that a method is fit for purpose, i.e. for use for solving a particular analytical problem.'
[Eurachem guide: The Fitness for Purpose of Analytical Methods]

Value:

Accepted Reference Value:

'A value that serves as an agreed-upon reference for comparison and which is derived as:

- a) a theoretical or established value, based on scientific principles;
- b) an assigned or certified value, based on experimental work of some national or international organisation;
- c) a consensus or certified value, based on collaborative experimental work under the auspices of a scientific or engineering group;
- d) when a), b), and c) are not available, the experimentation of the (measurable) quantity, i.e. the mean of a specified population of measurements.' [ISO 3534-1].

True Value:

'Value consistent with the definition of a given particular quantity.'

Note: *This is a value that would be obtained by a perfect measurement. True values are by nature indeterminate. The indefinite article a rather than the definite article the is used in conjunction with true value because there may be many values consistent with the definition of a particular quantity.* [VIM 1993].

Verification:

'Confirmation by examination and provision of objective evidence that specified requirements have been fulfilled.' [ISO 8402:1994].

'Confirmation, through provision of objective evidence, that specified requirements have been fulfilled.' [ISO 9000:2000].