एफएसएसएआई
fssai
भारतीय खाद्य सुरक्षा और मानक प्राधिकरण
Food Safety and Standards Authority of India

सत्यमेव जयते

स्वास्थ्य एवं
परिवार कल्याण मंत्रालय
MINISTRY OF
HEALTH AND
FAMILY WELFARE

RAFT Handbook

Vol. 2

FSSAI Guidelines for Validation of Analytical Methods and Rapid Analytical Food Testing Kits/Equipment/Methods

fssai RAFT
RAPID ANALYTICAL FOOD TESTING

March 2025

जी. कमलावर्धन राव
G. Kamala Vardhana Rao
मुख्य कार्यकारी अधिकारी
Chief Executive Officer

सत्यमेव जयते

एफएसएसएआई
fssai
भारतीय खाद्य सुरक्षा और मानक प्राधिकरण
Food Safety and Standards Authority of India
स्वास्थ्य और परिवार कल्याण मंत्रालय
Ministry of Health and Family Welfare

FOREWORD

Food safety is a public health priority as serious outbreaks of foodborne diseases have been documented globally in the past decade. As India increases its demand for food ingredients and raw materials, the need to safeguard public health throughout the entire value chain must be met. The current focus of Food Safety and Standards Authority of India (FSSAI) is to cover all scopes of the global food supply chain from farm to fork. FSSAI is working to address the challenges of an ever-increasing Indian population and ever-increasing food needs.

Food Safety and Standards Authority of India (FSSAI), in 2019, had commenced the policy for approval of Rapid Analytical Food Testing (RAFT) kits/equipment/methods in pursuance with the Food Safety and Standards (Laboratory and Sample Analysis) Regulations, 2011. FSSAI is facilitating rapid food testing methods to reduce the screening time of food products at the field level and accelerate surveillance activities. These rapid kits/equipment/methods are expected to become an integral part of quality assurance/quality control programs in the food industry as well as for use in regulatory and surveillance purposes, so that one can collectively achieve the goal of creating a robust food testing ecosystem in the country.

I am pleased to introduce this **Handbook on Rapid Analytical Food Testing Kits (RAFT) Volume 2.0 (FSSAI Guidelines for Validation of Analytical Methods and Rapid Analytical Food Testing Kits/Equipment/Methods)**. This handbook would provide guidance to desirous manufacturers/method developers, Food Laboratories, Research organizations/institutions, Start-ups, etc. for conducting the validation studies in India, thereby facilitating approval of RAFT kits/equipment/methods.

I am hopeful that this Handbook will enhance the understanding of validation procedures and encourage Research Organizations/Institutes/Start-ups to develop indigenous Rapid Food Testing Kits/Equipment/Methods.

March 2025

Shri G. Kamala Vardhana Rao,
Chief Executive Officer

एफडीए भवन, कोटला भवन, नई दिल्ली - 110002, दूरभाष - 011-23220995 / 996
FDA Bhawan, Kotla Road, New Delhi - 110002, Tel- 011-23220995/ 996
E-mail: ceo@fssai.gov.in, www.fssai.gov.in

Eat Right India
सही भोजन. बेहतर जीवन.

डॉ. सत्येन कुमार पंडा, एआरएस
Dr. Satyen Kumar Panda, ARS
सलाहकार
Advisor

सत्यमेव जयते

आज़ादी का अमृत महोत्सव

एफएसएसएआई
FSSAI
भारतीय खाद्य सुरक्षा और मानक प्राधिकरण
Food Safety and Standards Authority of India
स्वास्थ्य और परिवार कल्याण मंत्रालय
Ministry of Health and Family Welfare

PREFACE

Food testing is paramount in light of the adverse effects of poor safety and quality on human health. The use of traditional and sophisticated techniques for detecting food contaminants is a major challenge in modern times. Such techniques are relatively expensive, time-consuming, and labor-intensive. Rapid Analytical Food Testing kits for food testing are the need of the hour to ensure "*faster, better, cheaper*" real-time testing of food products throughout the supply chain. The rapid kits are advanced, easy to use, portable hand-held devices and they don't require any sophisticated equipment to conduct the tests. The tests can be performed anywhere by anyone without any specific/ advanced training requirements.

FSSAI has made a rapid progress towards the use of state-of-the-art food testing technologies since 2019. There is a need to popularize these Rapid Tool kits to ensure they are widely used. More importantly, there is a need to promote our startups who are working in the area of food testing so that our local innovations can be mainstreamed.

Validation is a crucial step for ensuring the workability of RAFT kits/equipment/methods. As per the RAFT guidelines, validation of the rapid food testing kit/equipment/method in accordance with national/international guidelines by a third party, such as an authorized FSSAI laboratory or international organizations is mandatory for approval.This **Handbook on Rapid Analytical Food Testing Kits (RAFT) Volume 2.0 (FSSAI Guidelines for Validation of Analytical Methods and Rapid Analytical Food Testing Kits/Equipment/Methods)** has been developed for the ease of understanding the procedures for validation of analytical methods and rapid kits/equipment/methods. These guidelines will serve as ready reference for authorized laboratories to conduct validation studies in India, thus promoting the indigenous Rapid Food Testing Kits/Equipment/Methods.

March 2025

Dr. Satyen Kumar Panda
Advisor (QualityAssurance)

एफडीए भवन, कोटला भवन, नई दिल्ली - 110002, दूरभाष-011-23217833
FDA Bhawan, Kotla Road, New Delhi - 110002, Tel - 011-23217833
E-mail: advisor.qa@fssai.gov.in, www.fssai.gov.in

Eat Right India
सही भोजन. बेहतर जीवन.

LIST OF CONTRIBUTORS

Dr. Lalitha Ramakrishna Gowda
Former Chief Scientist, CSIR- Central Food Technological Research Institute, Mysuru

Dr. Kaushik Banerjee
Director, ICAR- National Research Centre for Grapes, Pune

Dr. Kiran N Bhilegaonkar
Head, Division of Veterinary Public Health, ICAR- Indian Veterinary Research Institute, Izatnagar

Dr. Bhaskar Datta
Associate Professor, Indian Institute of Technology Gandhinagar

Dr. Ajit Dua
Chief Executive Officer, Punjab Biotechnology Incubator, Mohali

Dr. Anoop A. Krishnan
Assistant Director (Tech.), Export Inspection Council, New Delhi

Mr. Angshuman Saha
Assistant Director (Tech.), Export Inspection Agency, Kolkata

Dr. Deepak Kumar Sharma
Senior Principal Scientist, CSIR – Institute of Microbial Technology, Chandigarh

Dr. T.G. Shrivastav
Professor and Head (Retd.), Department of Reproductive Biomedicine, National Institute of Health and Family Welfare (NIHFW), New Delhi

Dr. Satyen Kumar Panda
Advisor (Quality Assurance), Food Safety and Standards Authority of India

Ms. Sweety Behera
Director, Food Safety and Standards Authority of India

Mr. Balasubramanian K
Joint Director, Food Safety and Standards Authority of India

Dr. Dinesh Kumar
Former Assistant Director (Tech.), Food Safety and Standards Authority of India

Ms. Gurpreet Kaur
Technical Officer, Food Safety and Standards Authority of India

Mr. Priyanka Meena
Technical Officer, Food Safety and Standards Authority of India

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

β-ETI	β-Expectation Tolerance Interval
σ	Standard Deviation
$\bar{x}$	Mean
µg	microgram
µL	microlitre
AL	Acceptability Limit
ANOVA	One-way analysis of variance
AOAC	Association of Official Analytical multi
CRM	Certified Reference Material
C_t	Cycle Threshold
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dPCR	Digital Polymerase Chain Reaction
ELISA	Enzyme Linked Immuno-Sorbent Assay
FN	False Negative
FP	False Positive
gDNA	Genomic Deoxyribonucleic Acid
ILV	Independent Laboratory Validation
ILS	Inter Laboratory Study
ISO	International Organization for Standardization
IUPAC	International Union of Pure and Applied Chemistry
LFIA	Lateral Flow Immunoassay
LOD	Limit of Detection
LOQ	Limit of Quantitation
MDC	Minimum Detectable Concentration
MLV	Multiple Laboratory Validation
ng	nanogram
NIST	National Institute of Standards and Technology

PA	Positive agreement
PD	Positive Deviation
POD	Probability of Detection
qPCR	Quantitative Polymerase Chain Reaction
RAFT	Rapid Analytical Food Testing
RLOD	Relative Level of Detection
RM	Reference material
RSD_r	Relative Standard Deviation Repeatability
RSD_R	Relative Standard Deviation Reproducibility
RT-PCR	Real Time- Polymerase Chain Reaction
T_m	Melting Temperature
TNA	Total Negative Agreement
TND	Total Negative Deviation

Chapter-1

INTRODUCTION

It is internationally recognized that a laboratory must take appropriate quality assurance measures to ensure that it is capable of and does provide data of the required quality. The Food Safety and Standards Authority of India (FSSAI), as a regulatory body tasked with ensuring the safety of the nation's food supply ensures that analytical measurement in food needed to support regulatory compliance, and enforcement actions meet the highest standards of analytical performance that are appropriate for their intended purposes. The following practices in analytical measurement are considered to constitute best practice: 1) the analytical measurements must satisfy an agreed requirement (i.e. to a defined objective), 2) must use methods and equipment which have been tested to ensure they are fit for purpose, 3) the personnel must be both technically qualified, competent and demonstrate that they can perform task properly, 4) regular independent third party assessment of the technical performance of a laboratory, 5) reproducibility of test results across food testing laboratories in the country and 6) must have well-defined quality control and quality assurance procedures.

Validation is the process used by the analytical community to acquire the necessary information to assess the ability of an analytical measurement to reliably obtain a desired result, determine the conditions under which such results can be obtained and determine the limitations of the procedure. Method validation is therefore an essential component of the measures that a laboratory should implement to allow it to produce reliable analytical data and is an important requirement in the practice of food analysis.

The validation process identifies the critical aspects of a procedure which must be carefully controlled and monitored to demonstrate and confirm that any new method/kit is suitable for its intended purpose. Intended purposes may include, but are not limited to:

i. qualitative or quantitative analyses
ii. screening or confirmatory analyses
iii. matrix or platform extensions
iv. emergency operations analyses

Method validation is achieved by conducting experiments to determine performance characteristics and quantify method performance and is performed after the initial method development and optimization. These performance characteristics of the method are documented, thereby demonstrating whether the method is fit for the analytical purpose. Typical performance characteristics of analytical methods are: applicability, selectivity (specificity), trueness (bias and recovery), precision (repeatability, intermediate precision and reproducibility), operating(working) range, sensitivity (limit of quantitation, limit of detection), and robustness (ruggedness). Validation applies to a defined protocol, for the determination of a specified analyte and range of concentrations in a particular type of test material, used for a specified purpose. In general, validation should check that the method performs adequately for the purpose throughout the range of analyte concentrations and test materials to which it is applied.

A full validation is required for:

i. A new or original qualitative /quantitative method/kit.
ii. Platform, matrix and analyte extensions to a method/kit.
iii. Modifications to a standard method/kit that may alter its performance specifications.

SCOPE

It is not always clear to manufacturers and developers of new method/ kits for food analysis in India the understanding of validation: what is involved, why it is important, when, and how it needs to be done and the quantum of data to be submitted for review. Therefore, the purpose of this handbook is to discuss the issues related to method validation and aims to direct the reader towards established protocols where these exist and where they do not.

The extent of validation will depend on the application and method going to be used. Therefore, different chapters on chemical, microbiological and biological methods are compiled. Each chapter provides details to enable one to design a validation strategy. These guidelines are expected to be used not only to a manufacturer/developer of a new kit but also to food analysts to ensure that the methods they use in the laboratory are adequately validated and fit for purpose.

Many a time it is difficult to decide which term to use in a validation study when several similar definitions are in use. For clarity it has been considered that the terms commonly used for each type of analysis will be defined at the end of each chapter.

METHOD VALIDATION PROCESS

The method validation exercises described in this document confirm by examination that the particular requirements for a method/kit have been fulfilled. All methods/kits used by FSSAI in support of its regulatory and compliance roles must be validated according to these guidelines. Three levels of scrutiny are defined and serve to demonstrate that the method can detect, identify and, where applicable, quantify an analyte(s) to a defined standard of performance. The hierarchy of criteria within the validation process also provides general characteristics on the method's utility and insights for its intended use.

The first step in the validation process for methods/kits designed for routine regulatory application is a comprehensive initial study carried out by the originating laboratory/manufacturer termed 'Single Laboratory Validation" also known as "in house validation" with defined performance characteristics and acceptance criteria. If available, a comparison is made to an existing reference method (Figure 1.1)

The second step is a single laboratory validation carried out by an independent laboratory termed as Independent Laboratory Validation (ILV). The purpose of an ILV is to determine if the method/rapid kit can be successfully used by a laboratory other than the originating laboratory/manufacturer. If available, in this validation a comparison is made to an existing reference/confirmatory method.

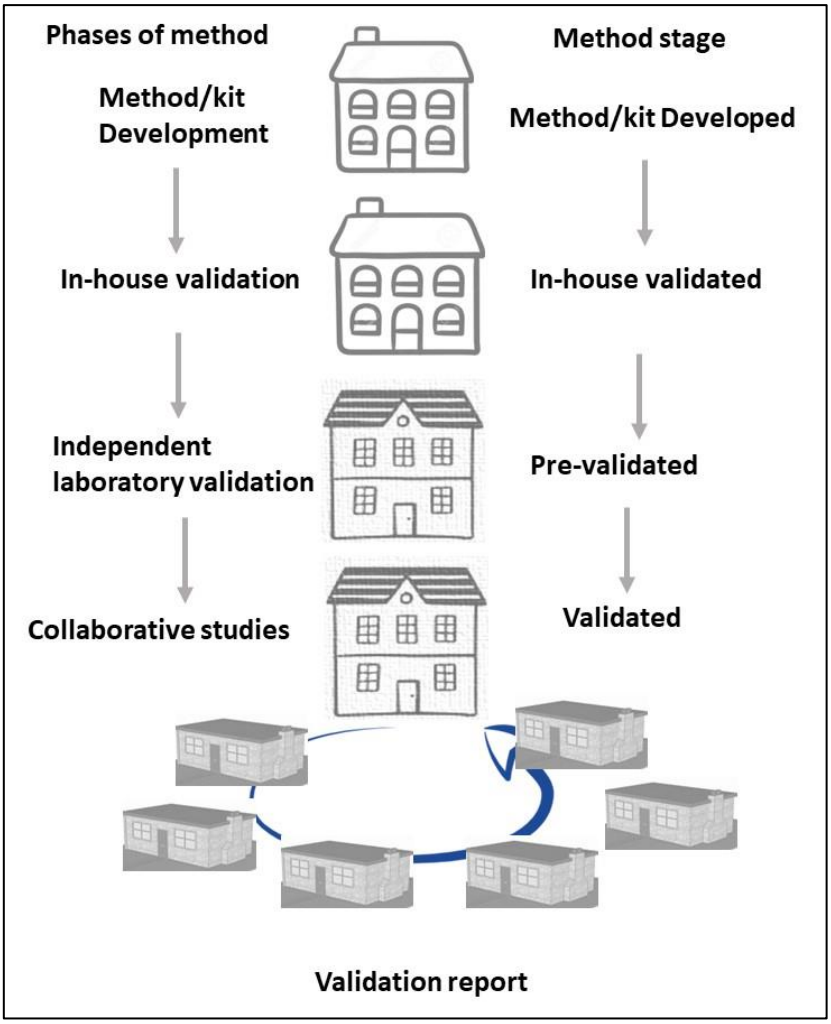


Figure 1.1: Phases in the kit/method validation procedure (Adapted from Zel, J.; Milavec, M.; Morisset, D.; Plan, D.; Van den Eede, G.; Gruden, K. How To Reliably Test for GMOs, 1st Edition; Springer Briefs in Food, Health and Nutrition; Springer: New York, 2012; 100 pp (ISBN 978-1-4614-1389-9)

This is followed by a Multiple Laboratory Validation (MLV) also known as 'Collaborative Studies'. This is an inter-laboratory study in which collaborators from multiple laboratories use the method /kit to analyse identical portions of homogeneous materials to assess the performance characteristics obtained for that method of analysis (W. Horwitz, IUPAC, 1987). It is designed to measure reproducibility, to determine if the method can be successfully performed by laboratories other than the originating laboratory/manufacturer.

The criteria defined for the levels of scrutiny by FSSAI (ILV and MLV) are closely aligned with other recognized and established validation criteria for collaborative studies e.g., Codex Alimentarius, International Organization for Standardization (ISO), Association of Official Analytical Collaboration(AOAC). All methods/kits validated to this level of scrutiny may be accepted for use in regulatory circumstances and surveillance and compliance support.

The applicant for FSSAI approval must submit:

i. evidence that the submitted method/kit fulfils the general principal conditions provided in the relevant section (for examples see Table 1.1).

ii. evidence that the submitted method meets the validation acceptance criteria for an ILV and MLV indicated in the respective sections. Such evidence will include supporting experimental data together with an indication of the reference values and experimental design chosen by the applicant during method testing and optimization.

Table 1.1 Validation methodology to be followed for different types of kits/methods

Type of kit/method	**Validation methodology to be adopted along with relevant Section**
Paper based Yes/No methods & kits (e.g. test for urea, formalin in milk)	Chapter 2: Guidelines for validation of chemical methods/kits (qualitative and quantitative) Section 4.5: Qualitative methods validation tools
Yes/No tests based on chemical reagents. (e.g. detection of sucrose, glucose in milk, rancidity, argemone oil)	Chapter 2: Guidelines for validation of chemical methods/kits (qualitative and quantitative) Section: 4.5: Qualitative methods validation tools
Dipstick chemical assays (color reaction)	Chapter 2: Guidelines for validation of chemical methods/kits (qualitative and quantitative) Section 4.5: Qualitative methods validation tools
Enzyme based lateral flow/paper-based assay (e.g. pesticide residues)	Chapter 2: Guidelines for validation of chemical methods/kits (qualitative and quantitative) Section: 4.5 Qualitative methods validation tools
Kits (chemical methods) for quantitative analysis (e.g. alcohol content)	Chapter 2: Guidelines for validation of chemical methods/kits (qualitative and quantitative) Section 4.4: Quantitative methods validation tools
DNA based methods/kits (e.g. GMO, meat speciation)	Chapter 4: Guidelines for Validation of DNA-based Methods/Kits *Relevant sections as per the method:* Section 5.3: Qualitative PCR methods Section 5.4: Quantitative PCR methods Section 6.0: Qualitative and quantitative multiplex assays
Lateral Flow immunoassays (Dipstick immunoassay) e.g. allergens in food, mycotoxins such as aflatoxins, ochratoxin, patulin, antibiotic residues etc.	Chapter 5: Guidelines for validation of immunoassay methods: enzyme linked immunosorbent assay (ELISA) and lateral flow immunoassay (LFIA) Section 4.3: Validation parameters and acceptance criteria for qualitative Lateral Flow Immunoassay (LFIA) devices.
Enzyme Immunoassays (ELISA kits) for quantitative analysis	Chapter 5: Guidelines for validation of immunoassay methods: enzyme linked immunosorbent assay (ELISA) and lateral flow immunoassay (LFIA) Section 6.0: Validation parameters and acceptance criteria for quantitative immunoassay methods (ELISA kits)
Methods/Kits (DNA based/Immunoassays/enzyme based) for qualitative and quantitative detection of	Chapter 3: Guidelines for Validation of Microbiological Methods/Kits *For Qualitative test*: Section 6.0: Validation for qualitative microbiological rapid test kits/equipment/ methods.

microorganisms	*For Quantitative test*: 7.0 Validation for quantitative microbiological rapid test kits/equipment/ methods *Other relevant chapters:* *For DNA based Methods/Kits* - Chapter 4: Guidelines for Validation of DNA-based Methods/Kits *For Immunoassay Methods/Kits* - Chapter 5: Guidelines for validation of immunoassay methods: enzyme linked immunosorbent assay (ELISA) and lateral flow immunoassay (LFIA) *For chemical/enzyme based methods/Kits* - Chapter 2: Guidelines for validation of chemical methods/kits (qualitative and quantitative)

It is the responsibility of the originating (developing) laboratory/kit manufacturer to ensure proper adherence to all criteria described in the different sections of this document.

In general, FSSAI's guidelines do not establish legally enforceable responsibilities. Instead, they describe current thinking of FSSAI and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word 'should/must' in these guidelines means that something is recommended.

All method/kit submission packages to be assessed by FSSAI are treated as confidential and ensure that transparency is upheld, while respecting the duty to confidentiality.

Chapter-2

GUIDELINES FOR VALIDATION OF CHEMICAL METHODS/KITS (QUALITATIVE AND QUANTITATIVE)

1. Introduction

This guideline presents a detailed discussion of elements for consideration during the validation of chemical analytical procedures. The objective of method validation of an analytical procedure/rapid kit is to demonstrate that it is suitable for the intended purpose (fit for purpose). Suitably characterized reference materials, with documented identity and purity or any other characteristics as necessary, should be used throughout the validation study. In practice, the experimental work can be designed so that the appropriate validation tests can be performed to provide sound, overall knowledge of the performance of the analytical procedure, for instance: specificity/selectivity, accuracy, precision, and ruggedness over the reportable range. Approaches other than those set forth in this guideline may be applicable and acceptable by FSSAI with appropriate science-based justification.

2. Purpose

FSSAI approved methods/kits employed in the notified food testing laboratories must meet the highest analytical performance standards. The purpose of this document is also to provide guidance on the data requirement for the validation of qualitative and quantitative chemical testing of foods. FSSAI establishes these criteria for validation, by which all submitted analytical methods/kits for chemical analyses in food, shall be evaluated.

These criteria are consistent with several related guidelines produced by international food standards setting organizations including, but not limited to, Codex Alimentarius, the International Organization for Standardization (ISO), AOAC International, European Union (EU) guidelines etc.

3. Scope

This guideline lays down procedures for validation of methods/kits to be used in chemical analysis of food for safety and quality. The guideline has been drawn up on the basis of current international standards and guidelines in the sector to help method developers and manufacturers of rapid food testing kits in validating them for seeking approval under the RAFT scheme of FSSAI.

The guidelines cover exclusively quantitative and qualitative analytical methods. However, it does not cover sampling in connection with the performance of these methods. This document establishes

validation criteria to be widely used in quantitative and qualitative analysis of food and food products covered by Food Safety and Standards Rules and Regulations (2011) including, but not limited to the following: colour, additives, decomposition products nutraceuticals, health supplements, ingredients/adulterants, elemental and metals, food and feed additives and preservatives, food allergens, naturally occurring toxins, intentional adulterants/poisons mycotoxins, nutrients persistent organic pollutants, pesticide residues, seafood and plant toxins, veterinary drug residues.

4. Method Validation

The process of method validation is intended to demonstrate that a method is fit-for-purpose. This means that when a test is performed by an adequately trained analyst using the specified equipment, materials, environmental conditions and exactly following the method protocol, accurate, reliable, and consistent results can be obtained within specified statistical limits for sample analysis. The validation should demonstrate the identity and concentration of the analyte, taking into account for matrix effects, provide a statistical characterization of recovery results, and indicate if the frequency of false positives and negatives are acceptable. When the method is followed using suitable analytical standards, results within the established performance criteria should be obtained on the same or equivalent sample material by a trained analyst in any experienced testing laboratory.

4.1 Performance characteristics

The following performance characteristics should be assessed in order to validate a method/kit which may vary depending on the intended use, type of method (See Table 2.1, quantitative vs qualitative).

1. Specificity/selectivity
2. Calibration range (Linearity)
3. Measuring range
4. Matrix effect
5. Limit of detection and Limit of Quantification
6. Accuracy (Trueness/Bias)
7. Precision (Repeatability and Reproducibility)
8. Ruggedness

The definitions of these characteristics are included in the Glossary.

Table 2.1 Performance Characteristics required for Validation of New Methods/Kits

Characteristic	**Quantitative**	**Qualitative**
Selectivity/specificity	Yes	Yes (True negative rate)
Sensitivity	Yes (LOD/LOQ)	Yes (True positive rate)
Limit of detection	Yes	Yes (minimum detectable concentration/probability of detection
Limit of quantitation	Yes	No

Linearity (or other calibration model)	Yes	No
Working Range	Yes	No
Accuracy	Yes	No
Precision	Yes	No
Measurement uncertainty,	Yes	No
Ruggedness	Yes	Yes
Confirmation of identity	Yes	Yes
Recovery/Trueness	Yes	No

Performance Characteristics for Validation of Method Extensions: Validating the extension of methods/applicability of the kits that have previously been validated requires a careful evaluation of the intended purpose of the extension. In cases where the sample preparation and/or the extraction procedure/analytical method is modified from the existing test procedure, it should be demonstrated that the modifications do not adversely affect the precision and accuracy of the data obtained. In order to implement the modified method, generally the standard or existing method is first performed. The modified method performance then is verified by comparison with that of the original method. Method applicability also can be carried out considering different matrices and varying concentration.

Confirmation of Identity Confirmation of identity for each analyte must be performed as part of the method validation for regulatory enforcement for both qualitative and quantitative methods. Unambiguous confirmation of identity usually requires analytically identifying key features of each analyte in the scope of the new method being validated such as with mass spectral fragmentation patterns or by demonstration of results in agreement with those obtained using an independent analysis.

4.1.1 General validation tools

The following provides some general guidelines/tools that should be used to assess method performance:

General Protocol: Prepare and analyze method blanks, matrix blanks, reference materials (if available) and matrix spikes (using matrix blanks if available) of known concentration as generally described under the Methods Validation Levels section and Table 2.2&2. 3 below. Accuracy or bias and precision are calculated from these results. Data will also be used to evaluate matrix effects and ruggedness/robustness of the method resulting from changes in the sample matrix. The following general validation tools should be used to generate method performance characteristics

Blanks: Use of various types of blanks enables assessment of how much of the result is attributable to the analyte in relation to other sources. Blanks are useful in the determination of limit of detection.

Reference materials (RM) and certified reference materials (CRMs): The use of known RMs/CRMs (when available and applicable) should be incorporated to assess the accuracy or bias of the method, as well as for obtaining information on interferences (*see Chapter 7 for more information on RMs and CRMs*)

Matrix Blank: This type of blank is a substance that closely matches the samples being analyzed with regard to matrix components. Matrix blanks are used to establish background level (presence or absence) of analyte(s) and to verify that sample matrix and equipment used does not interfere with or affect the analytical signal ensuring specificity of method.

Matrix Spikes (Laboratory spiked matrix): Recovery determinations can be estimated from fortification or spiking with a known amount of analyte and calculation of spike recoveries. (Note: spike recovery may not be accurately representative of recovery from naturally incurred analytes) Matrix effects can also be assessed with these samples. Accuracy or bias and precision are calculated from these results. The data can also be used to evaluate robustness of the method resulting from changes in the sample matrix.

Spiking is generally carried out at three levels) the level of concern (regulatory limit) or action level (X) as stated in the method and at levels corresponding to 1/2X and 2X.

Incurred Samples: This type of sample contains (not laboratory spiked) the analyte(s) of interest (if available e.g., peanuts naturally contaminated with aflatoxin, trace metals in food commodities) and can be used to evaluate precision and bias (if analyte concentration(s) are reliably known). Analyte recovery can also be evaluated through successive extractions of the sample and/or comparison to another analytical procedure with known bias.

Reagent Blank: This type of blank incorporates all reagents used in the method and is subjected to all sample processing operations. It serves to verify that reagents are analyte free and the equipment used does not interfere with or affect the analytical signal.

Replicate Analyses: The precision of the analytical process can be evaluated using replicate analyses. The originating laboratory should assure that adequate sample replicates are performed and that results from replicate measurements of each analyte are compared. Minimally, the method repeatability should be evaluated.

Interferences: Spectral, physical, and chemical interferences can be evaluated by analyzing samples containing various suspected interferences. Carryover should be evaluated using the incorporation of blanks immediately following standards and samples.

Statistics: Statistical techniques are employed to evaluate accuracy, trueness (or bias) precision, linear range, limits of detection and quantitation, and measurement uncertainty

4.2 Reference Method

A reference method is a method by which the performance of an alternate or new method is to be evaluated. For chemical analytes, an appropriate reference method is not always identifiable or available. However, there are some instances in which the use of a reference method is appropriate such as when replacing a method specified for use in a compliance program. Consultation between the

originating laboratory and validating laboratory/organization is suggested when deciding if the use of a reference method will be necessary.

Table 2.2. Key Validation Requirements for Chemical Methods/Kits (Qualitative and quantitative)

Criteria	Single Laboratory Validation Study Level 1	Independent Laboratory Validation Study Level 2	Multiple Laboratory Validation Study Level 3
Participating Laboratory	Originating Laboratory	Collaborator	Collaborators
Number participating labs	1	1	8 (quantitative) 10 (qualitative)
Number of matrices*	≥1	≥3 recommended where available	≥3 recommended where available
Number of analyte(s) spike levels for at least one matrix source**	≥2 spike levels + 1 matrix blank	≥3 spike levels + 1 matrix blank	≥3 spike levels + 1 matrix blank
Replicates required per matrix source at each level tested per laboratory	≥2 (quantitative) ≥3 (qualitative)	≥2 (quantitative) ≥3 (qualitative)	≥2 (quantitative) ≥3 (qualitative)
Replicates required at each level tested per laboratory if only one matrix source used	≥4 (quantitative) ≥6 (qualitative)	≥6 (quantitative) ≥9 (qualitative)	≥2 (quantitative) ≥6 (qualitative)

*If a variety of food matrices with differing physical and chemical properties are selected, the number of sources for each food category matrix may be one or more, but if only one food matrix is studied then ≥3 sources are recommended, where available. The number of matrix sources may be reduced, particularly if it is difficult to obtain blank matrix sources, as long as the total number of spike levels and matrix combinations are adequate (e.g., 6 replicates or greater at each spike level for quantitative methods and 9 replicates or greater for qualitative methods). Certified reference materials/ incurred tissues should be used, when available, and can replace one of your spiking levels.

** Number of spike levels is recommended for at least one source of matrix. Other similar sources of matrix (e.g., within the same food category) may be studied at one or two spike levels (e.g., at an action/guidance or tolerance level or close to the lower limit of quantitation/detection). CRMs/ incurred material should be used, when available, and can replace one of your spiking levels. For some analytes, spiking with pure standard alone does not sufficiently demonstrate method performance (i.e. BPA in can coatings contain oligomeric interferences, gluten in fermented/hydrolyzed products, protein-bound veterinary drug metabolites, sulfites binding

irreversibly after spiking). In these cases, reference materials and/or real samples must be used to demonstrate method performance

4.3 Method validation levels

Three standard levels of performance are defined in these guidelines for method/kit validation chemical analytes in foods (Appendix K, AOAC guidelines for single laboratory validation] and collaborative studies [AOAC Appendix D]. Key validation requirements for each level are summarized in Table 2.3. It is the responsibility of the originating (developing) /manufacturer to **validate at these three levels: Level 1 (Single Laboratory Validation), Level 2 (Independent Laboratory Validation and Level 3 (Full Collaborative Study) and submit data of Level 2 & 3 for approval by FSSAI under the RAFT scheme.**

4.3.1 Level one

This is an 'in-house' validation study, which is a 'single laboratory validation level (SLV). The manufacturer/originating laboratory generally conducts a comprehensive validation study, with performance criteria similar to an AOAC 'SLV study'. This study can be conducted in house or through an analytical laboratory that is accredited to ISO/IEC17025: 2017 having a valid scope for the said parameters and products. If appropriate, a comparison with an existing reference method has to be performed. If the method/kit is expected to be submitted for RAFT approval, its validation should be extended to at least Level Two.

4.3.2 Level two

This is also at the single laboratory validation level termed as 'Independent Laboratory Validation (ILV)'. The independent laboratory may be a specialized laboratory on a specific product/parameter i.e., Referral laboratories/Reference laboratories notified by FSSAI. The laboratory identified for the study should also be accredited to ISO/IEC17025: 2017 having a valid scope for the said parameters and products. The purpose of an ILV is to determine if the method/kit developed can be successfully performed by a laboratory other than the originating laboratory.

4.3.3 Level three

This validation level has criteria equivalent to a full AOAC or ISO Collaborative Study. This level of validation should be submitted for approval of a kit or method. This level is a study of the method/kit performance, not the laboratory. The method/instructions for use of the kit must be followed by all participating laboratories as closely as practicable, and any deviations by participants from the method described, no matter how trivial they may seem, must be noted on the report form.

4.4 QUANTITATIVE METHODS: PERFORMANCE CHARACTERISTICS TO BE EVALUATED (see Table 2.1)

4.4.1 Analytical selectivity

Analytical selectivity relates to "the extent to which the method can be used to determine particular analytes in mixtures or matrices without interferences from other components of similar behavior". The selectivity must be established for in-house developed methods/kits. It is important to establish during method validation that the test method kit is measuring only what it is intended to measure.

The selectivity of a method is the accuracy of its measurement in the presence of interferences such as competing non-target compounds, impurities, degradants and matrix components.

Selectivity can be shown by demonstrating that the identification and/or quantitation of an analyte is not impacted by the presence of other substances (e.g., impurities, degradation products, related substances, matrix, or other components present in the operating environment). Selectivity can be verified by demonstrating that the measured result of an analyte is comparable to the measured result of a second, well characterized analytical procedure (e.g., reference procedure).

Procedure: Analyse test samples, and RMs by candidate method/kit and compare with other independent methods. Analyse test samples containing various suspected interferences in the presence of the analytes of interest. Representative data (e.g., chromatograms, or spectra) should be used to demonstrate selectivity and individual components should be appropriately labelled.

4.4.2 Limit of Detection (LOD)

The determination of the limit of detection (LOD) or limit of quantitation (LOQ) is normally only required for methods intended to measure analytes at concentrations close to zero. LOD may be divided into two components, the method LOD and instrument LOD The method LOD is a term that should be applied to extraction and analysis methods developed for the analysis of specific analytes within a matrix. The LOD can be defined as the smallest amount or concentration of an analyte that can be reliably detected or differentiated from the background for a particular matrix (by a specific method). In other words, the LOD is the lowest value measured by a method that is greater than the uncertainty associated with it. All matrix interferences must be taken into account when determining the LOD.

The use of a signal to noise ratio for an analytical standard introduced to an instrument is a useful indicator of instrument performance but an inappropriate means of estimating the LOD of a method. Each instrument has a limitation on the amount of an analyte that they can detect, expressed as the instrument detection limit (IDL). IDL is defined as the smallest amount of an analyte that can be reliably detected or differentiated from the background (i.e. instrumental noise).

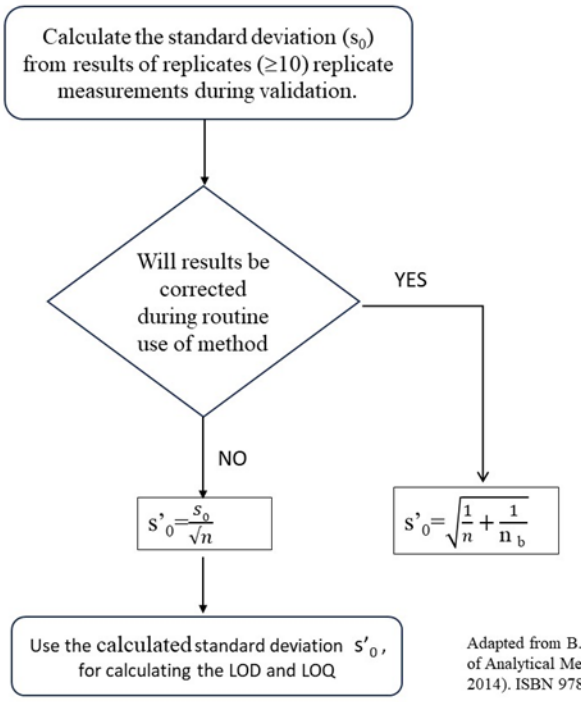
4.4.2.1 LOD based on the standard deviation of the blank

LOD may be determined by the analysis of a large number of blanks or independent sample blanks fortified at lowest acceptable concentration. This method is applicable to

Procedure: Carry out replicate measurements of blank samples, i.e. matrices containing no detectable analyte or replicate measurements of test samples with low concentrations of analyte ($n \geq 10$ is recommended). Calculate the standard deviation (s_0) of the results. Calculate s'_0 from s_0 following the flow chart in Box 1. **Calculate LOD = 3 × $s_{0'}$**.

The samples used should preferably be either a) blank samples, i.e. matrices containing no detectable analyte, or b) test samples with concentrations of analyte close to or below the expected LOD. Blank samples can be used for analysis involving titration, spectrophotometry and atomic spectroscopy. However, for techniques such as chromatography, which rely on detecting a peak above the noise, samples with concentration levels close to or above the LOD are required. These can be prepared by, for example, spiking a blank sample.

Box 1: Flow chart for calculation of the standard deviation, s'_0 to be used for estimation of LOD and LOQ.



S_0 is the estimated standard deviation of replicates at or near zero concentration.
s'_0 is the standard deviation used for calculating LOD and LOQ.
n is the number of replicate observations averaged when reporting results where each replicate is obtained following the entire measurement procedure.
n_b is the number of blank observations averaged when calculating the blank correction according to the measurement procedure.

Adapted from B. Magnusson and U. Örnemark (eds.) Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics, (2nd ed. 2014). ISBN 978-91-87461-59-0. Available from www.eurachem.org."

If blank samples or test samples at low concentrations are not available, reagent blanks can often be used. When these reagent blanks do not go through the whole sample preparation and measurement procedure, and are presented directly to the instrument, the calculation based on these measurements will give the instrument LOQ/LOD.

For methods with scope covering different food matrices, it may be necessary to determine the standard deviation for each matrix separately.

The number of replicates used for calculation should be sufficient to obtain an adequate estimate of the standard deviation. Typically, between 6 and 15 replicates are considered necessary; 10 replicates are recommended.

4.4.2.2 LOD based on the range in which the applies calibration equation

If data on samples near or at the LOD are not available, parameters of the calibration equation can be used to estimate the instrumental LOD. Using the estimate of LOD as the blank plus three standard deviations of the blank, the instrument response to a blank is taken as the intercept of the calibration (c), and the standard deviation of the instrument response (slope m) is taken as the standard error of the calibration. The calibration equation y=mx+c is widely used in analytical chemistry to determine instrument LOD. LOD is given by $3 \times \sigma$ (the standard deviation of the mean of the slopes).

Procedure: Prepare three calibration curves covering concentrations across the range Evaluate the instrument response (e.g., peak area, absorbance etc) at each level minimum of three times. Plot the instrument response (peak area/absorbance/titre value) was against the corresponding standard concentration. Use linear regression analysis to generate the standard curves y= mx+c. Calculate the mean of slopes 'm' for each curve

Calculate the LOD and limit of quantitation (LOQ) based on the standard deviation of the peak area response and slope of the linear calibration curves.

$$LOD = \frac{3.3\ d}{S}$$

$$LOQ = \frac{10\ d}{S}$$

where *d* is the standard deviation of the *y*-intercept of the regression line
S is the mean of slope of the regression lines.
However, because this is an extrapolation, the results cannot be as reliable as those from experiments made near the expected LOD.

4.4.2.3 LOD based on signal-to-noise

In the case of instrumental analytical procedures that exhibit background noise, a common approach is to compare measured signals from samples with known low concentrations of analyte with those of blank samples and establishing the minimum concentration at which the analyte can reliably be detected. Typically, acceptable signal-to-noise ratios are 2:1 or 3:1.

LOD = 3×Noise/Signal × Lowest concentration of the linearity samples

LOQ =10×Noise/Signal × Lowest concentration of the linearity samples

4.4.3 The limit of quantitation (LOQ)

The LOQ is the lowest level of analyte that can be reliably determined and quantified with acceptable performance. ('Acceptable performance' is variously considered by different guidelines to include precision, precision and trueness, or measurement uncertainty. In practice, however, LOQ is calculated by most conventions to be the analyte concentration corresponding to the obtained standard deviation (s'_0) at low levels multiplied by a factor, k_Q. The IUPAC default value for k_Q is 10 and if the standard deviation is approximately constant at low concentrations this multiplier corresponds to a relative standard deviation (RSD) of 10 %.

Procedure: Carry out replicate measurements of blank samples, i.e. matrices containing no detectable analyte or replicate measurements of test samples with low concentrations of analyte (n ≥ 10 is recommended).

Calculate the standard deviation, *s*0 of the results.

Calculate s'_0 from s_0 following the flow chart in Box 1

Calculate LOQ as

LOQ = *k*Q × s'_0

If blank samples or test samples at low concentrations are not available, reagent blanks can often be used. When these reagent blanks do not go through the whole sample preparation and measurement procedure, and are presented directly to the instrument, the calculation based on these measurements will give the instrument LOQ.

For methods with a scope covering different food matrices, it may be necessary to determine the standard deviation for each matrix separately.

The number of replicates used for calculation should be sufficient to obtain an adequate estimate of the standard deviation. Typically, between 6 and 15 replicates are considered necessary; 10 replicates are generally recommended.

4.4.4 Working and linear range

The 'working range' is the interval between the upper and lower levels which the method/kit demonstrates results with an acceptable uncertainty. The lower end of the working range is bounded by the limit of quantification LOQ. The upper end of the working range is defined by concentrations at which no significant anomalies in the analytical sensitivity are observed (e.g., the plateauing effect at high absorbance values in UV-Vis spectroscopy or chromatography using UV-Visible detector). This range will be the concentration range in which the linearity test is done. Carry out procedure as follows:

1. Measure blank plus calibration standards, at 6-10 concentrations evenly spaced across the ***range of interest.***
2. Plot response (y axis) against concentration (x axis) and carry out a regression analysis.
3. Visually examine to identify approximate linear range and upper and lower boundaries of the working range for the instrument.
4. Measure blank plus calibration standards, three times at 6-10 concentrations evenly spaced across the ***linear range***.
5. Plot response (y axis) against concentration (x axis). Visually examine for outliers, which may not be reflected in the regression.
6. Carry out a regression analysis. R^2 should be$\geq$0.999

Working range and linearity are assessed by visual inspection of the plot, supported by statistics from a linear regression. The method working range needs to be established for each matrix covered in the method scope. This is because interferences can cause non-linear responses, and the ability of the method to extract/recover the analyte may vary with the sample matrix.

4.4.5 Accuracy

The accuracy of the method/kit is the closeness of agreement between the value that is accepted either as a conventional true value or an accepted reference value and the value found.

Method validation seeks to investigate the accuracy of results by assessing both systematic and random effects on single results. Accuracy is, studied as two components: 'trueness' and 'precision'.

4.4.6 Trueness (Bias)

Trueness is defined as 'the closeness of agreement between the average value obtained from a large set of test results and an accepted value reference value. The measure of trueness is usually expressed as 'Bias'.

The bias of a method is the difference between the mean of a number of measured values and the true value. Three general methods can be used a) analysis of reference materials, b) recovery experiments using spiked samples, and c) comparison with results obtained with another method

Bias studies should cover the method scope and therefore require the analysis of different matrices and/or different analyte levels.

4.4.6.1 Determining bias using RM

To determine the bias using a RM, the mean and standard deviation of a series of replicate measurements are determined and the results compared with the assigned property value of the RM. The ideal RM is a certified matrix reference material with property values close to those of the test samples of interest. CRMs are generally accepted as providing traceable values.

It is also important to remember that a particular RM should only be used for one purpose during a validation study. For example, an RM used for calibration shall not also be used to evaluate bias. Availability of RMs is limited, but the chosen material must be appropriate to the use. For regulatory work, a relevant certified material, ideally matrix-matched if available, should be used. For methods used for long-term in-house work, a stable in-house material can be used to monitor bias but a CRM should be used in the initial assessment (*see Chapter 7 for more details on RMs*).

Measure a minimum of 10 replicates of the RM by the method/kit. Compare mean value, $\bar{x}$ with reference value x_{ref} for the RM. Calculate bias, *b*, per cent relative bias, *b* (%) or the relative per cent recovery (apparent recovery).

$b=\bar{x}-x_{ref}$

$$b\ (\%) = \frac{\bar{x} - x_{ref}}{x_{ref}} \times 100$$

$$R(\%) = \frac{x}{x_{ref}} \times 100$$

4.4.6.2 Determining bias using recovery experiments

In the absence of suitable RMs, recovery studies (spiking experiments) may be used to give an indication of the likely level of bias. Determine the bias using a minimum of 10 determinations over a minimum of the three concentration levels, covering the specified range.

1. Measure matrix blanks or test samples unspiked and spiked with the analyte of interest over a range of concentrations
2. Prepare spiked samples (n=10) at three concentrations over the range of 50 to 150% of the target analyte concentration.
3. When it is impossible or difficult to prepare known placebos, use a low concentration of a known standard.

4. Compare the difference between mean spiked value $\overline{x'}$ and mean unspiked value $\bar{x}$ with the added concentration x_{spiked}

5. Calculate the relative spike recovery R' (%) at the various concentrations:

$$R'\,(\%) = \frac{\overline{x'} - \bar{x}}{x_{spike}} \times 100$$

6. For each spiked sample, report the theoretical value, assay value, and percent recovery.

7. The mean recovery will be within 90 to 110% of the theoretical value.

8. Acceptable criteria for recoveries of quantitative methods are listed in Table 2.3.

Spiked samples should be compared with the same sample unspiked to assess the net recovery of the added spike. Recoveries from spiked samples or matrix blanks will usually be better than for routine samples in which the analyte is more tightly bound.

4.4.7 Precision

The precision of an analytical method is the degree of agreement among individual test results when the method is repeated multiple times. The precision of an analytical procedure is usually expressed as the standard deviation (σ) or Relative Standard Deviation (RSD) (coefficient of variation) of a series of measurements. It is indicated by RSD, which should generally not be more than 2%.

It is not necessary to know exact concentration of the analyte in the sample used in the precision study. However, it should be stable, homogenous and should be close to the test sample in relation to analyte concentration and matrix. The types of precision estimate are: 1) Repeatability, 2) Reproducibility and 3) Intermediate precision. Both repeatability and the reproducibility are expressed in terms of standard deviation and are generally dependent on analyte concentration.

4.4.7.1 Repeatability

Repeatability refers to the use of the analytical procedure within a laboratory over a short period of time using the same analyst with the same equipment. Repeatability represents the tightest extreme of replicate measurements. It describes the precision for a set of replicate measurements made one after the other in a single laboratory by a single analyst using a single instrument.

It is thus recommended that the repeatability is determined at different concentrations (minimum three concentrations) across the working range, by carrying out 10 repeated determinations at each concentration level.

Record all the information (e.g. titer value, absorbance, the retention time, peak area, and peak the datasheet. Calculate the mean (($\bar{X}$), standard deviation (σ), and RSD_r. The acceptance criterion is based on the method level which is defined for the analyte(s)/sample matrix(ces) combination as a maximum level, minimum level, normative level or concentration range depending on the intended use of the method.

The RSD_r. varies with concentration, C expressed as a mass fraction. Acceptable values are given in Table 2.3 or calculated by the formula:

RSDr = $C^{-0.15}$

unless there are reasons for using tighter requirements.

4.4.7.2 Reproducibility

Reproducibility represents the widest extreme of precision. Reproducibility expresses the precision between laboratories (collaborative studies, usually applied to standardization of methodology) and demonstrated by means of an inter-laboratory trial. It describes the variation within a set of measurements made on a sample over an extended time period, in several laboratories, by a number of different analysts on different instruments. Reproducibility reflects variation in the method from all possible sources. Reproducibility represents the expected variation in results when the method is used by several laboratories (e.g., regulatory laboratories).

Reproducibility is determined at different concentrations (minimum three concentrations) across the working range, by carrying out 10 repeated determinations at each concentration level by different analysts/laboratories.

For more details see Chapter 6: Guidelines for Multi-Laboratory Validation (Collaborative Study)

4.4.7.3 Intermediate Precision

Intermediate precision (within laboratory reproducibility) represents the variation in results obtained within a single laboratory over an extended period of time. Intermediate precision is the result from within lab variations due to random events such as different days, different analysts, different equipment, etc. It is carried out within the same laboratory using different analysts, different sets of equipment and on different days. Intermediate precision gives a better estimate of the precision of the method in routine use and is the most appropriate precision value for setting quality control limits. Intermediate precision needs to be evaluated over different analyte concentration (minimum three concentrations covering the working range) and representative range of sample matrices.

Calculate the mean ($\bar{X}$), standard deviation (σ), and RSD_R among the operators and instruments. The acceptance criterion is based on the method level which is defined for the analyte(s)/sample matrix(ces) combination as a maximum level, minimum level, normative level or concentration range depending on the intended use of the method. For acceptance criteria see Table 2.3.

4.4.7.4 Acceptability criteria for quantitative methods/kit

The acceptability ranges based on various national and international organizations and their sources are provided in Table 2.3. Acceptable spike recoveries vary with analyte concentration as indicated in (e.g., recoveries may fall in approximately the 80- 120% range for quantitative methods at the 1 µg/g (ppm) concentration). Repeatability and reproducibility also vary with analyte concentration.

Table 2.3 Method Criteria for Quantitative Method Levels at Increasing Orders of Magnitude (reproduced from USFDA (2019) and Codex Procedural Manual (2013))

ML* unit	0.001 mg/kg	0.01 mg/kg	0.1 mg/kg	1 mg/kg	10 mg/kg	100 mg/kg	1 g/kg	10 g/kg

Alternative ML* unit	1 ppb	10 ppb	100 ppb	1 ppm	10 ppm	100 ppm	0.1%	1 %
Concentration ratio of ML (CML)	10^{-9}	10^{-8}	10^{-7}	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-2}
LOD (≤ mg/kg)	0.0002	0.002	0.01	0.1	1	10	100	1000
LOQ (≤ mg/kg)	0.0004	0.004	0.02	0.2	2	20	200	2000
RSDr **	22%	22%	11%	8%	6%	4%	3%	2%
$PRSD_R$ #	22%	22%	22%	16%	11%	8%	6%	4%
RSD_R ##	≤ 44%	≤ 44%	≤ 44%	≤ 32%	≤ 22%	≤ 16%	≤ 12%	≤ 8%
Recovery	40%-120%	60%-115%	80%-110%	80%-110%	80% - 110%	90% - 107%	95% – 105%	97%-103%

* ML is a method level and can be defined for the analyte(s)/sample matrix(ces) combination as a maximum level, minimum level, normative level or concentration range depending on the intended use of the method.

**The RSD_r or repeatability precision refers to the degree of agreement of results when conditions are maintained as constant as possible within a short period of time (e.g., relative standard deviation of replicates or best precision exhibited by a single laboratory). Typically, acceptable values for RSD_r are between ½ and 2 times the value shown (Horwitz Ratio ($HorRat_r$) = RSD_r (found, %)/ RSD_r (calculated %)). For concentration ratios ≥ 10^{-7} Horwitz theory is applied for concentration ratios < 10^{-7}, Thompson theory is applied.

The $PRSD_R$ or Predicted Relative Reproducibility Standard Deviation is based on the Horwitz/Thompson equation. For concentration ratios < 10^{-7}, Thompson theory is applied.

The RSD_R or Reproducibility Precision refers to the degree of agreement of results when operating conditions are as different as possible (e.g., same test samples in different laboratories) and should be calculated from the Horwitz/Thompson equation. When the Horwitz/Thompson equation is not applicable (for an analytical purpose or according to a regulation) or when "converting" methods into criteria then it should be based on the RSD_R from an appropriate method performance study. The ratio between the found and predicted value should be ≤ 2. ($HorRat_R$ = RSD_R / $PRSD_R$ ≤ 2)

The acceptability ranges in Table 2.3 provide approximate target ranges for quantitative methods and are not rigid binding guidelines. It is recognized that for some situations such as with difficult matrices, extremely low analyte concentrations (e.g., chlorinated dioxins, persistent organic pollutants), multi-residue methods/kits these general acceptability ranges may not be achievable.

4.4.8 Robustness

The robustness of an analytical procedure is a measure of its capacity to remain unaffected by small but deliberate variations in procedural parameters listed in the procedure documentation and is an indicator of its suitability during normal usage. Robustness may be determined during development of the analytical procedure.

If measurements are susceptible to variations in analytical conditions, the analytical conditions should be suitably controlled or a precautionary statement should be included in the procedure. Robustness is the degree to which a method/kit is affected by small changes in the operating conditions. The method should ideally be unaffected by changes in instruments, reagent supplier and environmental conditions. Robustness testing helps to identify the parameters which have a significant effect on the performance of a method and how closely these parameters need to be controlled for the method/kit. A robust method is one whose performance is not affected by minor changes in the experimental limits.

A robustness study involves designing experiments by deliberately introducing changes. The results are analysed under differing conditions and statistically evaluated to determine whether these changes have an effect on the method. These parameters will vary depending on the principle of the method/kit. Mean and %RSDs are compared against the acceptance criteria to evaluate impact of changing experimental parameters.

Examples of typical variations are:

- stability of analytical solutions;
- extraction time.
- incubation temperature
- incubation time

In the case of liquid chromatography, examples of typical variations are:

- influence of variations of pH in a mobile phase;
- influence of variations in mobile phase composition;
- different columns (different lots and/or suppliers);
- oven temperature;
- flow rate.

In the case of gas-chromatography, examples of typical variations are:

- different columns (different lots and/or suppliers);
- temperature;
- flow rate.

4.4.8 Recovery

Recovery provides some information about the accuracy of the analytical method and also the efficiency of extraction of the analyte from a sample with matrix. In principle, recoveries are estimated

by the analysis of matrix reference materials. The recovery is the ratio of the concentration of analyte found to that stated to be present. Where (certified) reference materials are unavailable, the recovery of analyte can be estimated by studying the recovery of the analyte added as a spike.

If a matrix blank is available the analyte can be spiked into that and its recovery determined after application of the normal analytical procedure. If no matrix blank is available, the spike can be added to an ordinary test portion that is analysed alongside an unspiked test portion. The difference between these two results is the recovered part of the added analyte, which can be compared with the known amount added. Matrix mismatch occurs when a recovery value is estimated for one matrix and applied to another.

Conduct recovery experiment by measuring matrix blanks or test samples unspiked and spiked with the analyte of interest over a range of concentrations as follows:

1. Prepare spiked samples (n=10) at three concentrations over the range of 50 to 150% of the target analyte concentration.
2. When it is impossible or difficult to prepare known placebos, use a low concentration of a known standard.
3. Compare the difference between mean spiked value $\bar{x'}$ and mean unspiked value $\bar{x}$ with the added concentration x_{spiked}
4. Calculate the relative spike recovery R' (%) at the various concentrations:

$$R'\,(\%) = \frac{\bar{x'} - \bar{x}}{x_{spike}} \times 100$$

5. For each spiked sample, report the theoretical value, assay value, and percent recovery.
6. Acceptable criteria for recoveries of quantitative methods are listed in Table 2.3

4.5 QUALITATIVE METHODS: PERFORMANCE CHARACTERISTICS TO BE EVALUATED

Qualitative methods are intended to detect biological and chemical compounds. Qualitative analysis, instead of quantitative analysis, is therefore recommended primarily for screening purposes using low-cost methods or at analyte concentrations near to the limit of detection (LOD). A 'qualitative method' gives effectively a 'Yes'/No' answer at a given cut-off concentration of an analyte.

Validation involves identification of the cut-off concentration in order to classify/diagnose a condition, e.g. the presence or absence of an adulterant in milk defining which cut-off concentration applies. Properties of the qualitative method should be determined at a number of concentrations, below and at the cut-off concentration using number of samples as given in Table 2.4. For qualitative methods, precision cannot be expressed as a standard deviation or relative standard deviation, but can be expressed as true and false positive rates.

The following parameters need to be evaluated:

4.5.1 Probability of Detection (POD) limit

The POD or detection limit for qualitative analysis is defined as "the lowest concentration of the analyte at which the test can reliably detect as positive in the given matrix". This implies that we should consider only the probability of a β -type error or false negative rate, usually at ≤ 5%. Using different concentrations of the analyte the false negative rate is determined. The lowest concentration where the false negative rate is ≤ 5% is considered as detection limit (*See Box 2*)

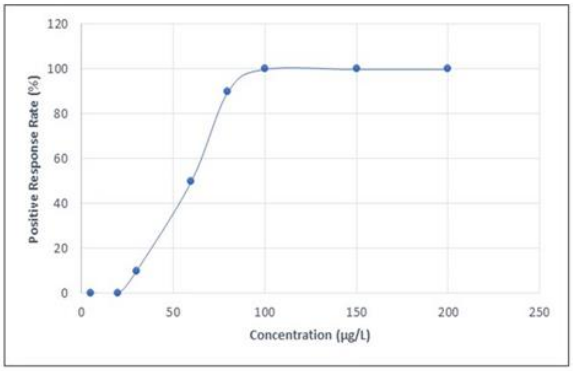
4.5.2 Cut-off value

The cut-off value is a special performance parameter, this value means the concentration level where the qualitative method differentiates the samples with a certain probability of error, usually 5%. The cut-off value is related to the sensitivity, as it corresponds to the concentration level at which the sensitivity rate is 95%, when the β-type error probability has been set at 5%.

The cut-off concentration can be determined by establishing the false positive and negative rates at a number of levels below and above the expected cut-off concentration. The cut-off limit is where false negative rates for concentrations above the limit are low – with a stated probability, e.g. 5 %. During validation the proposed cut-off limit given in the documented procedure is assessed.

Box 2 Determination of POD and cut-off concentration

Concentration (µg/L)	No. of positive/ negative results	% Positive
200	10/0	100
150	10/0	100
100	10/0	100
80	9/1	90
60	5/5	50
30	2/8	20
20	0/10	0
5	0/10	0



1. Measure, in random order, sample blanks spiked with the analyte at a range of concentration levels.
2. Positive or negative observations by visual inspection for ten replicates are recorded at each concentration level.
3. A response curve with fraction (in %) of positive results versus concentration is constructed.
4. With a criterion of < 5 % false negative results (95% positive), the POD is determined between 80 and 100 µg /L.
5. To obtain a narrower range experiment can be repeated between 80-100 with using 5 µg intervals

Adapted from "B. Magnusson and U. Ornemark (eds.) Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics, (2nd ed. 2014). ISBN 978-91-87461-59-0. Available from www.eurachem.org."

Measure, in random order, sample blanks spiked with the analyte at a range of concentration levels. Use 10 replicates for each concentration. Construct a response curve of % positive or negative results versus analyte concentration. Determine, by inspection, the threshold concentration at which the test becomes unreliable (Box 2).

4.5.3 False Positive and Negative Rates

False positive rate is ''the probability that a test sample is a known negative, given that the test sample has been classified as positive by the method'' also called Type 1 error (scoring a false positive).

$$\text{false positive rate } (\alpha) = \frac{\text{fp}}{\text{tn} + \text{fp}}$$

where fp are false positive test samples and tn are known true negative test samples.

Acceptance criterion: α = ≤5%

False negative rate is ''the probability that a test sample is a known positive, given that the test sample has been classified as negative by the method''

$$\text{false negative rate } (\beta) = \frac{\text{fn}}{\text{tp} + \text{fn}}$$

where fn are false negatives samples and tp known true positive test samples.

Acceptance criterion: $\beta = \leq 5\%$

4.5.4 Sensitivity and specificity

When dealing with qualitative methods, sensitivity is ''the ability of a method to detect truly positive samples as positive', so the sensitivity rate ''is the probability, for a given concentration, that the method will classify the test sample as positive, given that the test sample is a 'known' positive' (see Box 3). It can be calculated as:

$$\text{Senstivity} = \frac{\text{truly positive samples (tp)}}{\text{truly positive samples(tp)} + \text{false negative smples (fn)}}$$

The same occurs with specificity, which is 'the ability of a method to detect truly negative samples as negative'. In the same way, the specificity rate 'is the probability, for a given concentration, that the method will classify the test sample as negative, given that the test sample is a 'known' negative' (see Box 3), so it can be expressed as:

Specificity = True negative rate

$$\text{Specificity} = \frac{\text{truly negative samples (tn)}}{\text{truly negative samples(tn)} + \text{false positive smples (fp)}}$$

4.5.5 Number of qualitative tests required to confirm false negative (FN) and false positive (FP) rates as <5%

Zero acceptance number sampling is a statistical approach commonly used to test a hypothesis (or criteria) for the frequency of defective items in a population (e.g., such as FN or FP rates with repeated testing). For this approach, all tested samples must have the correct response in order to accept the hypothesis (i.e., accept only when zero "no" responses observed). The minimum number of samples that must be tested depends on the criteria for the negative (no) rate and the level of statistical confidence:

$$n = \frac{\text{Log}(\propto)}{\text{Log}\,(1 - \rho)}$$

where 1-α is the confidence level and ρ is the maximum acceptable FN or FP rate. Sample sizes to assess selected criteria for FN or FP rates with varying levels of confidence are listed in Table 2.4. Both FN and FP rates should be determined as <5% using a valid statistical approach.

Box 3: Example of Determining Sensitivity and Specificity of a Binary paper test for formalin in milk				
		Milk adulterated with urea		
Urea paper test		Positive	Negative	
	Positive	TP=4	FP=18	Positive predicted value =TP/(TP+FP) =4 /(4+18) =4/22 = 18%
	Negative	FN=2	TN=102	Negative predicted value =TN/(FN+ TN) =102/(2+ 102) =102/104 =98.1%
False positive rate (%) =FP/(FP+TN) =18/(18+102) =18/120 =15%= 1-specificity	False negative rate (%) =FN/(TP+FN) =2/(4+2) =2/6 =33.3%= 1-sensitivity	Sensitivity = TP/(TP+FN) =4/(4+2) =4/6 =66.67%	Specificity =TN/(FP+TN) =102/(18+102) =102/120 =85%	
With a large number of false positives and very few false negative, this paper test is poor in confirming the adulteration of milk with formalin. It will detect 66.7% of the adulterations . The paper test correctly identifies 85% of the milk not adulterated				

Table 4 Test Sample numbers recommended for assessing FN or FP rates

False Positive/ Negative rate	Confidence level			
	80%	90%	95%	99%
< 1%	161	230	299	459
<2%	80	114	149	228
<5%	**32**	**45**	**59**	**59**
<10%	16	22	29	44

The acceptance criterion is to have 95% confidence that the FN rate is <5% then test 59 samples with the target analyte present at the concentration of interest, typically the detection limit or a relevant level of concern, in a range of matrices. The criteria are satisfied if all 59 test results are positive for the target.

This sample size formula is related to the Clopper-Pearson confidence interval for Binomial proportions and frequently used for zero defect acceptance sampling plans for commodity lots. The rationale for the sample size is that when the probability of a false positive/negative response is ρ for each sample then $(1 - \rho)$ n is the probability that n samples will have the correct response.

4.5.6 Unreliability region

For qualitative methods, having binary responses of the YES/NO type, there is no meaning for a number associated with the result so uncertainty is expressed not as a numerical value but as a region of probabilities of committing error. The 'unreliability region' corresponds to the region in which false responses are obtained (either false positive or false negative). This region is defined by an upper and a lower concentration limit between which the qualitative method can provide false responses. As these false responses can be either positive or negative, the upper and lower limits that define this unreliability region depend on the probability of obtaining these false responses, which is fixed by the analyst.

4.5.7 Ruggedness

Ruggedness is an important parameter related to how the method performs under variations in the operational, environmental, etc., conditions. Carry out the analysis by introducing appropriate limits to method parameters that are likely to impact the results if not controlled such as incubation time, temperature, volume of sample. Ruggedness can be caried out by using: Single variable tests (test and re-test with small changes to a method parameter) or Multivariable tests (Plackett-Burman designed experiment (see Ruggedness *in Chapter- 5 Guidelines for validation of immunoassay methods: Enzyme Linked Immunosorbent Assay (ELISA) And Lateral Flow Immunoassay (LFIA)*

4.5.8 Cross reactivity

Another parameter to be considered is cross-reactivity or the presence of interferences. For test kits, in particular, it is recommended to check whether the presence of analytes of the same family as the one under study might modify the result of the analysis. These checks are mandatory for manufacturers of the test kits. e.g., if the kit is for histamine detection it should be evaluated for spermine, spermidine, putrescine and any other diamine and false positive rate should be $\leq 5\%$.

5.0 FOOD MATRIX SELECTION

Food matrix and sample source selection should be based on the types of foods most likely to be used in the analysis or based on risk of contamination. The number of food categories to be used will depends on the intended use of the method.

The number of different food categories to be validated depends on the applicability and intended use of the method. Depending on how many categories the RAFT application is for, a minimum of 1 – 3 representative matrices per category listed below should be selected for SLV.

Example: If the application is intended for one category 'Fish and fish products', choose the matrix from the category seafood. Include a raw fish, partially processed and highly processed fish

A list of foods that can be used based on the applicability are:

a. **Meat and meat products:** Fresh meat, Frozen meat, Raw marinated/minced/comminute meat, Semi-cooked /Smoked Meat, partially heat treated and/ or smoked meat and meat product, Canned/Retorted meat product, Chilled meat, Cooked Meat/meat product, Cured/pickled meat products, Dried/Dehydrated meat/meat products, Fermented meat products sausage, lunch meat, meat substitutes

b. **Seafood:** Chilled/Frozen Finfish, crustaceans, cephalopods, mollusks, bivalves, dried or Salted and dried fish products, thermally processed, fermented, smoked, canned fish products. Fish sticks, surimi, raw fish filet, raw oysters, raw mussels, raw clams, cooked crawfish, crabmeat (fresh or pasteurized), battered and breaded fish products,

c. **Fruits, Vegetables, and Nuts**: Fresh / frozen /dehydrated or dried fruits and vegetables, fresh fruit juice, apple cider, tomato juice, fruit cubes, berries, peanut butter, coconut, fruit powders almonds, minimally processed lettuce, spinach, kale, collard greens, cabbage, bean sprouts, seed sprouts, peas, mushroom, green beans and other minimally processed fruit and vegetable products, thermally processed fruits and vegetables (sauce and chutney)

d. **Dairy and Dairy products**: Dahi, Yogurt, Paneer, Khoa, Channa, hard and soft cheeses, raw or pasteurized liquid milk, co-mingled milk, full cream, butter, ghee infant formula, coffee creamer, ice cream, milk powders, casein, whey, whey powder, non-fat dry milk/dry whole milk,

e. **Confectionary:** Chocolate / bakery ware Frosting and topping mixes, candy and candy coating, milk, chocolate, cake mixes,

f. **Egg and egg products**: Shell eggs, liquid whole eggs, dried whole egg or dried egg yolk, dried egg whites,

g. **Herbs and spices**: Oregano, pepper, paprika, black pepper, white pepper, celery seed or flakes, chili powder, cumin, parsley flakes, rosemary, sesame seed, thyme, vegetable flakes, onion flakes, onion powder, garlic flakes, allspice, masala powders

h. **Whole/Processed grains and legumes**: Food grains/Flours, grits, rice corn meal, soy flour, dried yeast, cereal based complementary food, uncooked noodles, macaroni, spaghetti, soygurt, tofu, soy beverage.

6.0 INFORMATION TO BE PROVIDED WITH VALIDATION DOCUMENTS

Appended below is the information that should be provided from the developer/manufacturer when the results of a ILV are submitted for review by FSSAI. This list of information is in addition to the results submitted for the validation criteria described in Table 2.3

1. The scope of the method/kit should be defined and documented
2. Purpose of measurement (what is to be identified)?
3. Detailed procedure.
4. What are the likely sample matrices?
5. What are the expected concentration levels or ranges?
6. Does it meet the regulatory requirements?

7. Are there any specific equipment accommodation and environmental conditions that need to be considered?
8. What type of equipment is to be used? Is the method for one specific instrument, or should it be used by all instruments of the same type?
9. Method used for the preparation, sub-sampling, procedure

7.0 GLOSSARY

Accuracy - Closeness of agreement between a measured quantity value and a true quantity value of a measurand (JCGM200:2008)

Action level: Level of concern or target level for an analyte that must be reliably identified or quantified in a sample.

Analyte: The component of a sample or test item which embodies a quantity or quality that is ultimately determined directly or indirectly. The term 'analyte' in this document is applied to any substance or material to be analysed (e.g. chemical constituent, residue, contaminant etc.).

Bias: Estimate of a systematic measurement error (JCGM200:2008).

Blank: A blank value is obtained as a result of analysis of a matrix which does not, as far as possible, contain the analyte(s) in question. Use of various types of blanks (to which no analytes have been added) enables assessment of how much of the measured instrument response is attributable to the analyte and how much to other causes. Various types of blank are available to the user:

Certified Reference Material (CRM): Reference material accompanied by documentation (certificate) issued by an authoritative body and providing one or more specified property values with associated uncertainties and traceability, using valid procedures. Note: Standard Reference Material (SRM) is the trademark name of CRMs produced and distributed by the National Institute of Standards and Technology (NIST).

Collaborating Laboratory: The collaborating laboratory refers to the laboratory (or laboratories) other than the originating laboratory involved in Independent /Multi-Laboratory method validation studies. A collaborating laboratory analyzing collaborative study samples, work independently of one another.

Confirmation of Identity: Unambiguous identification of an analyte(s) by a highly specific technique such as mass spectrometry or by demonstration of results from two or more independent analyses in agreement.

Confirmatory Analysis/Method: Independent analysis/method used to confirm the result from an initial or screening analysis. A different method is often used in confirmation of screening results.

Cut-off Concentration: In qualitative analysis, the concentration of the analyte that is either statistically lower than the level of concern (for limit tests) or at which positive identification ceases (for confirmation of identity methods).

Detection limit: The detection limit for qualitative analysis is defined as "the lowest concentration of the analyte at which the test can reliably detect as positive in the given matrix".

False Negative Rate: In qualitative analysis, a measure of how often a test result indicates that an analyte is not present, when, in fact, it is present or, is present in an amount greater than a threshold or designated cut-off concentration.

False Positive Rate: In qualitative analysis, a measure of how often a test result indicates that an analyte is present, when, in fact, it is not present or, is present in an amount less than a threshold or designated cut-off concentration.

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, 2000).

Incurred Samples: Samples that contain the analyte(s) of interest, which were not derived from laboratory spiking but from sources such as exogenous exposure or endogenous origin.

Intermediate Precision Condition of Measurement, Intermediate Precision Condition: Condition of measurement, out of a set of conditions that includes the same measurement procedure, same location, and replicate measurements on the same or similar objects over an extended period of time, but may include other conditions involving changes (JCGM200:2008).

Level of Concern: Level of concern is the concentration of an analyte in a sample that has to be exceeded before the sample can be considered non-compliant. This concentration can be a regulatory limit, safe level, action level, guidance level.

Limit of Detection: Measured quantity value, obtained by a given measurement procedure, for which the probability of falsely claiming the absence of a component in a material is β, given a probability α of falsely claiming its presence (JCGM200:2008). Note: IUPAC recommends default values for α and β equal to 0.05

Limit of Quantitation: Refers to the smallest analyte concentration or mass, which can be quantitatively analysed with a reasonable reliability by a given procedure

Matrix: The predominant material, component or substrate which contains the analyte of interest.

Matrix Blank: A substance that closely matches the samples being analyzed with regard to matrix components. Ideally, the matrix blank does not contain the analyte(s) of interest but is subjected to all sample processing operations including all reagents used to analyze the test samples. The matrix blank is used to determine the absence of significant interference due to matrix, reagents and equipment used in the analysis

Measuring interval, working interval: Set of values of quantities of the same kind that can be measured by a given measuring instrument or measuring system with specified instrumental uncertainty, under defined conditions (JCGM200:2008). Note: The lower limit of a measuring interval should not be confused with detection limit.

Measuring Range: The concentration range within which the analytical parameter in question can be determined with specifically determined trueness and precision.

Method blank*:* A substance that does not contain the analyte(s) of interest but is subjected to all sample processing operations including all reagents used to analyze the test samples. An aliquot of reagent water is often used as a method blank in the absence of a suitable analyte-free matrix blank.

Method Validation: The process of establishing the performance characteristics and limitations of a method and the identification of the influences which may change these characteristics and to what extent. 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? The process for verifying that a method is fit for purpose; i.e. for use of solving a particular analytical problem (EURACHEM, 1998)

Minimum Detectable Concentration (MDC)*:* In qualitative analysis, an estimate of the minimum concentration of analyte that must be present in a sample to ensure at a specified high probability (typically 95% or greater) that the measured response will exceed the detection threshold, leading one to correctly conclude that an analyte is present in the sample

Originating Laboratory*:* The originating laboratory refers to the laboratory/manufacturer that has developed the method and has completed the Single laboratory validation requirements

Precision*:* Closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions. Measurement precision is used to define measurement repeatability, intermediate measurement precision, and measurement reproducibility (JCGM200:2008).

Qualitative test results*:* Results of tests not numerically derived (e.g. visual examinations or binary classification tests such as absence/presence, positive/negative, reactive/non-reactive, etc). Qualitative test results based on a numerical outcome, e.g. based on thresholds, are often described as semiquantitative or semi-qualitative and it is expected that method validation or verification is in line with quantitative procedures.

Quantitative test results: Numerically derived test results.

Reagent blanks*:* Reagents used during the analytical process (including solvents used for extraction or dissolution) are analysed in isolation in order to see whether they contribute to the measurement signal. The measurement result arising from the analyte can then be corrected accordingly.

Recovery: The extraction efficiency of an analytical process, reported as (a percentage of) the known amount of analyte carried through the sample extraction and processing steps of the method.

Reference material*:* A material, sufficiently homogenous and stable with respect to one or more specified properties, which has been established to be fit for its intended use in a measurement process or in examination of nominal properties.

Repeatability (RSD_r)*:* Precision obtained under observation conditions at a specific concentration/spike level where independent test results are obtained with the same method on identical test items in the same test facility by the same operator using the same equipment within short intervals of time. Should be included in all quantitative collaborative laboratory (Level 2) reports

Reproducibility (RSD_R)*:* Precision obtained at a specific concentration/spike level under observation conditions where independent test results are obtained with the same method on identical test items in different test facilities with different operators using different equipment. Should be included in all quantitative Collaborative (Level 2) reports.

Measurement Reproducibility: Measurement precision under reproducibility conditions of measurement response from standard in matrix extract and standard in solvent

Ruggedness/Robustness: The degree of independence of the method of analysis from minor deviations in the experimental conditions of the method of analysis.

Selectivity: The extent to which a method can determine particular analyte(s) in a mixture(s) or matrix(ces) without interferences from other components of similar behavior. Selectivity is generally preferred in analytical chemistry over the term Specificity.

Sensitivity: The change in instrument response which corresponds to a change in the measured quantity (e.g., analyte concentration). Sensitivity is commonly defined as the gradient of the response curve or slope of the calibration curve at a level near the LOQ.

Specificity: In quantitative analysis, specificity is the ability of a method to measure analyte in the presence of components which may be expected to be present. The term Selectivity is generally preferred over Specificity.

Spike Recovery*:* The fraction of analyte remaining at the point of final determination after it is added to a specified amount of matrix and subjected to the entire analytical procedure. Spike Recovery is typically expressed as a percentage. Spike recovery should be calculated for the method as written. For example, if the method prescribes using deuterated internal standards or matrix-matched calibration standards, then the reported analyte recoveries should be calculated according to those procedures.

Trueness: The degree of agreement of the mean value from a series of measurements with the true value or accepted reference value. This is related to systematic error (Bias).

Uncertainty: Non-negative parameter characterizing the dispersion of the values being attributed to the measured value.: Non-negative parameter characterizing the dispersion of the values being attributed to the measured value.

Verification: Confirmation by examination and provision of objective evidence that specified requirements have been fulfilled (ISO 9000:2005).

8.0 REFERENCES

Analytical Quality Control and Method Validation Procedures for Pesticide Residues Analysis in Food and Feed SANTE 11312/2021

AOAC International (2007), *How to meet ISO/IEC 17025 requirements for method verification*, ALACC Guide, www.aoac.org/alacc_guide_2008.pdf.

Validation and verification of quantitative and qualitative test methods. NATA, January 2018

AOAC International, “Appendix K: Guidelines for Dietary Supplements and Botanicals, Part 1 AOAC Guidelines for Single-Laboratory Validation of Chemical Methods for Dietary Supplements and Botanicals”, 2013. (http://www.eoma.aoac.org/app_k.pdf

AOAC International, “Appendix D: Guidelines for Collaborative Study Procedures to Validate Characteristics of a Method of Analysis”, 2005. http://www.aoac.org/imis15_prod/AOAC_Docs/StandardsDevelopment/eoma_appendix_d.pdf

Codex Alimentarius (2014), Codex Alimentarius Commission, Procedural Manual, Twenty-first ed., “Principles for the Establishment of Codex Methods of Analysis”, 2014. Accessed 10/11/2018. http://www.fao.org/3/a-i3243e.pdf

EURL Guidance Document on Screening Method Validation Version 1.1, 21 September 2023. https://sitesv2.anses.fr/en/system/files/Guidance_screening_20230921_v1_0.pdf

Magnusson, B. and U. Örnemark (eds.) EURACHEM Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics, (2nd ed. 2014). ISBN 978-91-87461-59-0. Available from http://www.eurachem.org

Trullols, E., Ruisa´nchez, I. and Xavier Rius F. (2004) Trends in Analytical Chemistry, 23, 137-145

USFDA (2019) Guidelines for the Validation of Chemical Methods in Food, Feed, Cosmetics, and Veterinary Products 3rd Edition U.S. Food and Drug Administration Foods Program October 2019

APPENDIX 2.1

a) Examples for carrying out Single Laboratory Validation (SLV)- only one lab

Work flow for Quantitative Method validation

- Identify the matrix for method validation (eg. Peanut)
- Identify the analyte for method validation (eg. Aflatoxin B1)
- Number of analyte spike level for at least one matrix/product - ≥ 2 spike level (x, 1/2x) + 1 matrix blank where "x" can be tolerance limit/regulatory limit eg. 3, 10 µg/kg of Aflatoxin B1
- Replicates required per matrix source - ≥ 2 (quantitative) and ≥ 3 (Qualitative) eg. Peanut sources considering which for quantitative it would be 4 samples and qualitative 6 samples
- If only one matrix source only is used the replicates to be used is ≥ 4(quantitative) and ≥ 6 (Qualitative) e.g Peanut
- Performance characteristic to be derived as per Table 2.1 considering whether method is qualitative or quantitative.

Day	Matrix	Analyte and fortification levels	Replicates for qualitative analysis	Replicates for quantitative analysis
1	Peanut 1, Peanut 2, Peanut 3	Aflatoxin B1 3, 10 µg/kg	Blank+ three samples at each fortification level	Blank+ two samples at each fortification level
2	Peanut 1, Peanut 2, Peanut 3	Aflatoxin B1 3, 10 µg/kg	Blank+ three samples at each fortification level	Blank+ two samples at each fortification level
3	Peanut 1, Peanut 2, Peanut 3	Aflatoxin B1 3, 10µg/kg	Blank+ three samples at each fortification level	Blank+ two samples at each fortification level

b) Example for carrying out an Independent Laboratory Validation (ILV)- only one lab

Work flow for method validation

- Identify the matrix for method validation (eg. Peanut)
- Identify the analyte for method validation (eg. Aflatoxin B1)
- Number of analyte spike level for at least one matrix/product (three source/three commercial brand)- ≥ 3 spike level (x, 1/2x and 2X) + 1 matrix blank where "x" can be tolerance limit/regulatory limit eg. 3, 10, 20 µg/kg of Aflatoxin B1

- Replicates required per matrix source - ≥ 2 (quantitative) and ≥ 3 (Qualitative) eg. Peanut source 1, peanut source 2 and peanut source 3, considering which for quantitative it would be 6 samples and qualitative 9 samples
- If only one matrix source only is used the replicates to be used is ≥ 6(quantitative) and ≥ 9 (Qualitative) e.g Peanut source 1
- Performance characteristic to be derived as per Table 2.1 considering whether method is qualitative or quantitative.

Day	**Matrix**	**Analyte and fortification levels**	**Replicates for qualitative analysis**	**Replicates for quantitative analysis**
1	Peanut 1, Peanut 2, Peanut 3	Aflatoxin B1 3, 10, 20 µg/kg	Blank+ three samples at each fortification level	Blank+ two samples at each fortification level
2	Peanut 1, Peanut 2, Peanut 3	Aflatoxin B1 3, 10, 20 µg/kg	Blank+ three samples at each fortification level	Blank+ two samples at each fortification level
3	Peanut 1, Peanut 2, Peanut 3	Aflatoxin B1 3, 10, 20 µg/kg	Blank+ three samples at each fortification level	Blank+ two samples at each fortification level

In case of MLV study (Level 3) the samples have to be analyzed in 8 labs for quantitative and 10 labs for qualitative. For details see Chapter 6: **Guidelines for Multi-laboratory Validation (Collaborative Study)**

APPENDIX 2.2 – Example for qualitative method validation

Work flow for Method validation

Performance characteristic to be derived as per Table 2.1 for a qualitative test/kit

- Identify the matrix for method validation (eg. Milk)
- Identify the analyte for method/kit validation (eg. Formalin)
- **Determine POD and cut-off value for the binary test** (yes/No) for Cow's milk

  i. Prepare in random order sample blanks, spiked samples with analyte (e.g.urea 0.10, 0.9. 0.8, 0.75, 0.7, 0.65 0.6, 0.5, 0.4, 0.2 % and u %) with a range of concentrations (8-10 concentrations).

  ii. Test minimum 10 replicates at each concentration

  iii. Draw a response curve of concentration (µg or mg/L) versus positive response (%) as shown in Box 2

  iv. With a criterion of 95% positive rate estimate POD.

  v. If need be use a narrower concentration range.

  vi. Repeat with different source (e.g. buffalo milk and mixed milk)

  vii. Tabulate results and present POD curve for cow's, buffalo, and mixed milk

  viii. Estimate POD.

- **Determine Cut-off value**

The cut-off concentration can be determined by establishing the false positive and negative rates at a number of levels below and above the expected cut-off concentration. The cut-off limit is where false negative rates for concentrations above the limit are low – with a stated probability, e.g. 5 %.

  i. Measure, in random order, sample blanks spiked with the analyte at a range of concentration levels.

  ii. Test 10 replicates for each concentration (total number of analyses must be > 60 for each method) using the kit and reference method

  iii. Construct a response curve of % positive results versus analyte concentration.

  iv. Determine, by visual inspection, the threshold concentration at which the test becomes unreliable.

  v. *You can use the POD curve to determine cut-off value The concentration of analyte at 95% positive rate*

- **Determine False positive rate and False negative rate**

  i. Prepare five different concentrations of the analyte in matrix near the POD (e.g formalin 0.09, 0.07, 0.065,0.06,0.02 % & Blank).

ii. Analyze 20 replicates at each concentration using kit and reference method

iii. Determine the false positive, false negatives, true positive and true negative. An explanation for TP, TN, FP and FN is given in table below,

iv. Repeat the above procedure with different milk (buffalo, toned, pasteurised and mixed milk

v. Tabulate the results as shown

Real situation	**Analyst decision = Yes**	**Analyst decision = No**
Present (yes)	true positive (TP)	false positive (FP)
Absent (No)	false negative (FN)	true negative (TN)

- Determine sensitivity from tabulated data (*see Box 3)*
- Determine specificity (*see Box 3)*
- Replicates required per matrix source - ≥ 3 (Qualitative) eg. source 1 Pasteurised milk; source 2: buffalo milk and source 3: Skimmed milk. Analyse 20 samples each
- If only one matrix source is used the replicates to be used is 60 (Qualitative) e.g Cow's milk source

S.No	Matrix	Analyte concentration level (%)	Number of replicates	Positive/ Negative rate		Sensitivity		Specificity	
				Kit	Ref method	Kit	Ref method	Kit	Ref method
	Source 1	Blank	20						
		0.05	20						
		0.065	20						
		0.07	20						
		0.09	20						
		0.15	20						
	Source 2	Blank	20						
		0.05	20						
		0.065	20						
		0.07	20						
		0.10	20						

		0.15	20						
	Source 3	Blank	20						
		0.05	20						
		0.065	20						
		0.07	20						
		0.10	20						
		0.15	20						

- Carry out ruggedness tests

Carry out the analysis by introducing appropriate limits to method parameters that are likely to impact the results if not controlled such as incubation time for color change, environment temperature, changes in ratio of volume of milk ((volume of milk e.g. 0.8 mL instead of 1.0 mL or 1.2 mL) to reagent. Calculate the positive rate/negative rate for 20 replicates each for every change introduced and tabulate. The *acceptance criterion*: = ≤5% of the original.

APPENDIX 2.3

Checklist for Single Laboratory Validation (Chemical)

These checklists provide guidance for the validation of the analytical procedures included as part of applications submitted to FSSAI

Please take note the following checklists are not exhaustive FSSAI reserves the right to request additional data whichever it deems necessary. Furthermore, these checklists do not intend to provide direction on how to accomplish validation of analytical procedures. Please take note also all documents submitted to FSSAI must be arranged and labelled accordingly.

Table 2.5 shows the parameters checklist required for the validation of quantitative analytical methods/kits.

Table 2.6 illustrates the parameters required for the validation of procedures kits using qualitative analytical methods/kits.

TABLE 2.5

TEST	QUANTITATIVE TEST METHOD (CHEMICAL)		
PARAMETER	No.	DOCUMENTS REQUIRED	AVAILABILITY (Yes/No)
Selectivity/Specificity	1	Testing Method	
	2	Acceptance criteria	
	3	Chromatogram/spectrum for the following:	
		a) Standard	
		b) Sample	
		c) Blank/Placebo	
		d) Spike solution	
		e) System suitability tests	
	4	Peak purity (LC Method)	
	5	Relative Retention Time (RRT) (chromatographic method only)	
Sensitivity	1	Testing Method	
	2	Acceptance criteria	
	3	Calculations	

Linearity	1	Testing Method	
	2	Acceptance criteria	
	3	Minimum five (5) levels of standard solutions	
	4	Data for peak area/absorbance, linear regression equation, Y-intercept, slope, r^2 and linearity graph	
Range	1	For the assay: 80% - 120% of working concentration (WC)	
Accuracy	1	Testing Method	
	2	Acceptance criteria	
	3	Minimum three (3) levels of concentration in triplicates covering the specified range	
	4	Result: reported as percent recovery by the assay of known added amount of analyte in the sample, OR as the difference between the mean and the accepted true value together with the confidence intervals	
Precision (Repeatability	1	Testing Method (using sample/product as the test solution)	
	2	Acceptance criteria	
	3	Minimum three (3) levels of concentration in triplicates covering the specified range, OR minimum six (6) replicates at 100% of the WC	
	4	Result: SD, RSD and confidence Interval	
Precision (Intermediate) Reproducibility	1	Testing Method	
	2	Acceptance criteria	
	3	Minimum three (3) levels of concentration in triplicates covering the specified range	
	4	Cover at least 2 parameters among variation of analyst, date and equipment	

	5	Result: σ, RSD and confidence Interval	
Limit of detection	1	Testing Method: visual observation / signal-to-noise / standard deviation of the response and the slope	
	2	If based on standard deviation of the response and the slope method	
		a) Minimum five (5) levels of standard solution	
		b) Peak area values for all concentrations	
		c) Data for linear regression equation, Y-intercept, slope, r^2 and linearity graph	
	3	Calculation/formulation where applicable	
	4	Related Chromatogram(s) at LOD	
	5	Value of detection limit	
Limit of Quantitation	1	Testing Method: visual observation / signal-to-noise / standard deviation of the response and the slope	
	2	If based on standard deviation of the response and the slope method	
		a) Minimum five (5) levels of standard solution	
		b) Peak area values for all concentrations	
		c) Data for linear regression equation, Y-intercept, slope, R^2 and linearity graph	
	3	Calculation/formulation where applicable	
		Value of quantitation limit	

TABLE 2.6			
TEST	QUALITATIVE TEST METHOD (CHEMICAL)		
Parameter	**No.**	**Documents Required**	**Availability (Yes/No)**
False positive/False negative rate	1	Testing Method	
	2	Acceptance criteria	
	3	Calculations	
Specificity	1	Testing Method	
	2	Acceptance criteria	
	3	Calculations	
Sensitivity	1	Testing Method	
	2	Acceptance criteria	
	3	Calculations	
Probability of detection/Detection limit	1	Acceptance criteria	
	2	Calculations	
	3	Related test results at POD	
	4	Value of detection limit	
	5	Value of cut-off concentration	
Confidence interval levels, LCL (lower control limit) and UCL (upper control limit)	1	For the assay: 80% - 120% of working concentration (WC) Minimum three (3) levels of concentration in triplicates covering the specified range	
Ruggedness	1	Testing Method	
	2	Acceptance criteria	
	3	Parameters studied	
Applicable Food matrix	1	Matrix 1	
	2	Matrix 2	
	3	Matrix 3	

Chapter-3

GUIDELINES FOR VALIDATION OF MICROBIOLOGICAL METHODS/KITS

1.0 INTRODUCTION

Several new proprietary methods/ kits have been developed to assess the microbiological safety and quality of raw materials and finished food products. These methods/kits are often faster and less cumbersome to perform than the corresponding standardized/reference method under the Food Safety and Standards Rules and Regulations, 2011. The developers, end users, and regulatory authorities need a reliable validation of such alternative methods. A reliable validation study will enable FSSAI to make decisions for adoption of such method/kits for regulatory compliance.

2.0 PURPOSE

This chapter is intended to provide specific protocols and guidelines for the validation of rapid kits/proprietary methods intended to be used by FSSAI for compliance testing. The general principle and technical protocols for the validation of rapid kits/methods for microbiology in the food chain against a reference method are described. These guidelines can also be the basis for the certification of the method/kit by independent organizations/laboratories.

3.0 SCOPE

These guidelines are applicable for the validation of new methods/rapid kits for the analysis (detection and/or enumeration/quantification) of microorganisms (bacteria and fungi) in the food products intended for human consumption. Some clauses could be applicable to other microorganisms or their metabolites on a case-by-case-basis.

The validating laboratory/organisation must be competent to perform microbiology analysis using both the reference method and the new rapid kit/ method.

NOTE Competence can be demonstrated in different ways, e.g. for the reference method an ISO/IEC 17025 accreditation and for the Rapid kit/method a documented training.

4.0 APPLICABLE STANDARD

IS 17113: Part 1: 2019/ISO 16140-1:2016 - Microbiology of the food chain - Method validation - Part 1: Vocabulary

IS 17113: Part 2: 2019/ISO 16140-2:2016 - Microbiology of the food chain - Method validation - Part 2: Protocol for the validation of rapid (proprietary) methods against a reference method

5.0 GENERAL PRINCIPLES FOR THE VALIDATION OF RAPID KIT/METHODS

The validation protocol comprises two phases:

1) *Method Comparison Study* of the rapid kit/method against the reference method carried out in the organizing laboratory

2) *Interlaboratory Study* of the rapid kit/method against the reference method carried out in different laboratories

Paired and Unpaired Study

The validation study can be of two types depending upon the use of same or test portions/samples:

- Paired Study- The study where the same test portion is used for both the reference and the rapid kit/method due to both methods having exactly the same first step in the (enrichment) procedure
- Unpaired Study- The study where different test portions need to be used for the reference and the rapid kit/method due to different enrichment procedures

The choice of having a **paired** study or an **unpaired** study depends on the protocols of the reference and rapid kit/method. If there is a common initial step in the (enrichment) procedures, a **paired** study design is mandatory.

6.0 VALIDATION PROTOCOL FOR QUALITATIVE MICROBIOLOGICAL RAPID TEST KITS/EQUIPMENT/METHODS

6.1. Method comparison study

The method comparison study is the part of the validation process that is performed in the organizing laboratory. It consists of three parts namely the following:

- Sensitivity study
- Relative Level of Detection (RLOD) study
- Inclusivity/Exclusivity study

The test portions size to be used will be as written for the reference method.

6.1.1. Sensitivity study

The sensitivity study aims to determine the difference in sensitivity between the reference and the rapid/new method. This study is conducted using naturally and/or artificially contaminated samples. Different food categories and products must be tested. Acceptability Limits have been defined for the maximum acceptable difference depending on the type of study (**paired**/**unpaired**) and the number of categories tested.

6.1.1.1. Selection of categories

If the rapid kit/method is to be validated for a restricted number of food categories, then only those categories need to be studied.

If the Rapid kit/method is to be applied for a broad range of foods, then at least five major food categories (see ***Annexure 3.1***) must be studied.

For each of the selected five major food categories, a minimum of three different food products per category must be included in the study (5×3=15 food products).

Following shall be included in the study while selecting different food products (see ***Annexure 3.1***):

- both high and low (natural) background microflora,
- different types of stresses due to processing, and
- raw (unprocessed) food products.

6.1.1.2. Number of samples

For each category being examined, a minimum of 60 individual samples shall be tested made up of at least three food products with at least 20 samples representative for each product (three food products × 20 samples for each product = 60 samples).

Fractional positive results by either the reference or rapid method (i.e. samples should not be all positive or all negative) shall be obtained for each product tested. Ideally, fractional recovery should range between 25 % and 75 %. For each category, at least 30 samples shall have a positive result by the reference and/or the rapid kit/method.

6.1.1.3. Rapid-method result and confirmation

Many rapid-method protocols contain two steps, the first being the detection step and the second being the confirmation of the detection result. The end result of such rapid kit/method is the result after step two.

All results obtained with the rapid kit/method in an **unpaired** study shall be confirmed. In a **paired** study, only the positive results obtained with the rapid kit/method, for which the corresponding result with the reference method was negative, shall be confirmed. This confirmation is needed to determine whether the result is a true-positive or false-positive result. These test(s) can be based on the confirmation procedure of the reference or any other method that is able to isolate and confirm the identity of the target analyte.

The rapid/new method shall be evaluated for a defined test portion size (e.g. 25 g, 200 g, 375 g) during the validation study. The method is considered to be validated for any test portion size up to the validated test portion size if the testing protocol (dilution ratio, incubation time and incubation temperature) is the same as that used during the validation study.

6.1.1.4. Calculation and interpretation for sensitivity

The results obtained for the reference and rapid kit/methods shall be described for a **paired** study according to Table 3.1 and for an **unpaired** study according to Table 3.2.

The interpretation of the results (positive agreement, negative agreement, etc.) is based on a comparison of the reference method result (column 1 in Tables 3.1 and 3.2) and the rapid method result, including any confirmations as described in the rapid method protocol, (column 2 in Tables 3.1 and 3. 2). When positive or negative deviations are obtained, a footnote should be included at the end

of each table to provide additional explanations for the interpretation of the deviations. The footnotes indicate if the result is due to a false-positive or false-negative result of the rapid method. The footnote is a comparison of the results of the rapid method (including any confirmations as described in the rapid-method protocol) (column 2 in Tables 3.1 and 3.2) and the confirmed rapid method (by any means) (column 3 in Tables 3.1 and 3.2).

Table 3.3 is prepared for the summarized sample results for all Food products per category (≥60 samples) and per food product (≥20 samples) for both **paired** and **unpaired** studies.

Table 3.1 Comparison and interpretation of sample results between the reference and rapid kit/methods for a paired study

Result of the (reference or rapid) method per sample			
Reference method	**Rapid method** (including any confirmations as described in the rapid method protocol)	**Confirmed rapid method** (by any means)[a]	**Interpretation** (based on the confirmed rapid-method result)
+	+	Not needed[b]	Positive Agreement (PA)
-	-	Not needed[b]	Negative Agreement (NA)
+	-	Not needed[b]	Negative Deviation due to false negative rapid kit/method result ($ND_{FN(rapid)}$)
-	+	+	Positive Deviation (PD)
-	+	-	Positive Deviation due to false positive rapid kit/method result ($PD_{FP(rapid)}$)[c]

[a] Confirmation of the rapid kit/method result is done according to 6.1.1.3
[b] No need for additional confirmation test(s). Confirmed rapid kit/method result is the same as the rapid kit/method result
[c] This false positive result (FP) shall also be used to calculate the false positive ratio
Table adapted from ISO16140-2:2016/Amd 1:2024

Table 3.2 — Comparison and interpretation of sample results between the reference and rapid kit/methods for an unpaired study

Result of the (reference or rapid) method per sample			
Reference method	**Rapid method** (including any confirmation as described in rapid method protocol)	**Confirmed rapid method** (by any means)[a,b]	**Interpretation** (based on the confirmed rapid kit/method result)
+	+	+	Positive Agreement (PA)
+	+	-	Positive Agreement due to false positive rapid kit/method result ($PA_{FP(rapid)}$)[c]
-	-	-	Negative Agreement (NA)
-	-	+	Negative Agreement due to false negative rapid kit/method result ($NA_{FN(rapid)}$)
+	-	-	Negative Deviation (ND)
+	-	+	Negative Deviation due to false negative rapid kit/method result ($ND_{FN(rapid)}$)

-	+	+	Positive Deviation (PD)
-	+	-	Positive Deviation due to false positive rapid kit/method result ($PD_{FP(rapid)}$)[c]

[a] Confirmation of the rapid kit/method result is done according to 6.1.1.3
[b] Confirmation by any means is only required when the result of the rapid kit/methods does not produce viable organisms. This will be used as the confirmed rapid method result in comparison to the reference method result
[c] This false positive result (FP) shall also be used to calculate the false positive ratio
Table adapted from ISO16140-2:2016/Amd 1:2024

Determine the Total Negative Deviation (TND) and Total Negative Agreement (TNA) for the validation study (*See Box 1*).

Box 1 Determination of Total Negative Deviation (TND) and Total Negative Agreement (TNA)

Paired evaluation	
Total Negative Deviation	$TND = ND_{FN(rapid)}$
Total Negative Agreement	$TNA = NA + PD_{FP(rapid)}$
Unpaired evaluation	
Total Negative Deviation:	$TND = ND + ND_{FN(rapid)} + PA_{FP(rapid)}$
Total Negative Agreement	$TNA = NA + NA_{FN(rapid)} + PD_{FP(rapid)}$

Where
ND is Negative deviation
NA is Negative agreement
$ND_{FN(rapid)}$ is Negative Deviation due to false negative by rapid kit/method result
$PA_{FP(rapid)}$ is Positive Agreement due to false positive by rapid kit/method result
$NA_{FN(rapid)}$ is Negative Agreement due to false negative by rapid kit/method result
$PD_{FP(rapid)}$ is Positive Deviation due to false positive by rapid kit/method result

Table 3.3 — Summary of results obtained with the reference and rapid kit/methods of all samples for each category

	Rapid kit/method positive (A+)	Rapid kit/method negative (A-)
Reference-method positive (R+)	+/+ Positive Agreement (PA)	+/- Total Negative Deviation (TND)
Reference-method negative (R-)	-/+ Positive Deviation (PD)	-/- Total Negative Agreement (TNA)

Based on data summarized in Table 3.3 for the combined categories per category and per type, calculate the values for sensitivity of 1) the rapid method (SE_{rapid}) 2) the reference method (SE_{ref}), 3) relative trueness (*RT*) and 4) false positive ratio (FPR) and false negative ratio (FNR) for the rapid method after the additional confirmation of the results (see Box 2 for formulae).

Box 2: Formulae for determining Sensitivity, Relative Trueness and False Positive/Negative ratio.	
Sensitivity for the rapid method	$SE_{rapid} = \frac{(PA + PD)}{(PA + TND + PD)} \times 100\%$
Sensitivity for reference method	$SE_{ref} = \frac{(PA + TND)}{(PA + TND + PD)} \times 100\%$
Relative trueness	$RT = \frac{(PA + TNA)}{N} \times 100\%$
False positive ratio (FPR) and false negative ratio (FNR) for the rapid method	
Paired evaluation	$\mathrm{FPR} = \frac{PD_{FP(rapid)}}{TNA} \times 100\%$
Unpaired evaluation	$\mathrm{FPR} = \frac{PA_{FP(rapid)} + PD_{FP(rapid)}}{TNA} \times 100\%$
Paired evaluation	$\mathrm{FNR} = \frac{\mathrm{NA_{FN(rapid)}} + \mathrm{ND_{FN(rapid)}}}{\mathrm{PA + TND + PD}}$
where *N* is the total number of samples (PA + PD + TND + TNA), FP is the false positive results FN is the false negative results TND is the Total Negative Deviation TNA is the Total Negative Agreement PA is Positive agreement PD is Positive Deviation $ND_{FN(rapid)}$ is Negative Deviation due to false negative by rapid kit/method result $PA_{FP(rapid)}$ is Positive Agreement due to false positive by rapid kit/method result $NA_{FN(rapid)}$ is Negative Agreement due to false negative by rapid kit/method result $PD_{FP(rapid)}$ is Positive Deviation due to false positive by rapid kit/method result	

The confirmed rapid-method results must be used to determine whether the rapid method produces comparable results to the reference method.

Calculate the difference between (TND – PD) for both **paired** and **unpaired** studies and the sum of (TND + PD) for **paired** studies. Check whether the difference and/or sum of PD and TND conform to the Acceptability Limit (AL) stated in Table 3.4 with respect to the type of study (**paired** or **unpaired**) and the number of categories used in the evaluation.

NOTE 1 Acceptability Limits (AL) are based on data and consensus expert opinion. The ALs are not based on statistical analysis of the data.

The interpretation of results must be done per category and for all categories used in the validation study. An interpretation of results per enrichment protocol is necessary in case different protocols are

used for different types of samples. A sensitivity study can also exist of a partly paired and unpaired study. In such a case the results for (TND + PD) must be evaluated based on the number of positive samples obtained for the categories tested using the paired study design. The results for (TND – PD) must be evaluated based on the number of positive samples obtained for the full study (so all categories belonging to both the paired and unpaired study design).

The AL is not met when the observed value is higher than the AL. When the AL is not met, if the number of positive samples is higher than expected according to the number of categories (e.g. having 60 or more positive samples for one single category), it is possible to use the second column of Table 3.4 and switch to higher AL. When the AL is not met, investigations should be made (e.g. root cause analysis) in order to provide an explanation of the observed results. Based on the AL and the additional information, it is decided whether the rapid method is regarded as not fit for purpose for the category or categories involved. The reasons for acceptance of the rapid method in case the AL is not met shall be stated in the study report.

Table 3.4: Acceptability limit parameters and values for a paired and unpaired study design in relation to the number of positive samples obtained

Number of categories	Number of positive samples (N+)	Paired Study		Unpaired Study	Mixed Study[c]	
		(TND[a] – PD[b])	(TND + PD)	(TND – PD)	(TND – PD)	(TND + PD)
1	30 to 59	3	6	3	3	6
2	60 to 89	4	8	4	4	8
3	90 to 119	5	10	5	5	10
4	120 to 149	5	12	5	5	12
5	150 to 179	5	14	5	5	14
6	180 to 209	6	16	6	6	16
7	210 to 239	6	18	7	7	18
8	240 to 269	6	20	7	7	20
9	270 to 299	7	22	8	8	22
10	300 to 329	7	24	8	8	24
11	330 to 359	7	26	9	9	26
12	360 to 389	8	28	9	9	28
13	390 to 419	8	30	10	10	30
14	420 to 449	8	32	10	10	32

15	450 to 479	9	34	11	11	34
16	480 to 509	9	36	11	11	36
17	510 to 539	9	38	12	12	38
18	540 to 569	10	40	12	12	40
19	570 to 599	10	42	13	13	42
20	600 to 629	10	44	13	13	44
21	630 to 659	11	46	14	14	46
22	660 to 689	11	48	14	14	48
23	690 to 719	11	50	15	15	50
24	720 to 749	12	52	15	15	52
25	750 to 779	12	54	16	16	54

[a] TND = total number of samples with Negative Deviation results

[b] PD = number of samples with Positive Deviation results

[c] Mixed study includes paired and unpaired study design

NOTE 2: A negative value for (TND – PD) is acceptable as thus indicate a better performance of the rapid method compared to the reference method.

NOTE 3 Information on differences observed between results of the rapid method before and after confirmation of the results according to the rapid method protocol is commonly presented in the validation report as additional information but is not used in the overall assessment of the rapid method performance.

6.1.2. Relative level of detection study (RLOD)

The level of contamination shall be determined. This allows calculation of the LOD50 of the rapid method, which is required in order to verify the performance of the rapid method upon implementation of the validated method in a laboratory in accordance with **IS 17113 : Part 3:2022** /ISO 16140-3:2021. The level of contamination is determined by performing a most probable number (MPN) analysis on the (stabilized) inoculated samples (preferably) and/or through the enumeration of the inoculum at the time of inoculation.

6.1.2.1. Selection of categories, number of samples, and replicates tested

The same categories will be used as selected for the sensitivity study. For each category, one relevant food product is selected. The samples shall be artificially inoculated as per procedures specified in IS 17113: Part 2: 2019/ISO 16140-2:2016. Each food product will be inoculated with a different strain.

A minimum of three levels per type shall be prepared consisting of at least a negative control level, a low level, and a higher level. Ideally, the low level shall be the theoretical detection level (i.e., 0.7 cfu per test portion) and the higher level just above the theoretical detection level (e.g. 1 cfu to 1.5 cfu per test portion). For fastidious bacteria, the low level might significantly exceed the theoretical detection level, thus, low and high level should be adjusted appropriately. A fixed ratio (e.g., 1:2) between the low- and high-level contamination should be used to aid in determining the final contamination levels. At least the low level should provide fractional recovery by either the reference method or the rapid method (fractional recovery at the low level should be between 25 % and 75 % of the number of samples tested). An evaluation shall be performed to ensure the relationship and consistency of the number of positives of the intermediate and high level. In case the rapid method produces fractional recovery at the low level and the reference method produces all positive results, the results of the RLOD study are not valid and a root cause analysis shall be performed.

The level of contamination of the sample used (except for the negative control) shall be determined. At the negative control level, at least five replicate samples should be tested by both methods. For the second (low) level (theoretical detection level), at least 20, and for the third (higher) level, at least five replicates samples should be tested by both methods. The negative control level shall not produce positive (by isolation of the target organism) results. When positive results are obtained, the experiments have to be repeated for all levels.

6.1.2.2. Calculation and interpretation of the RLOD

The RLOD is defined as the ratio of the LODs of the rapid kit/method and the reference method:

$$RLOD = \frac{\mathrm{LOD_{alt}}}{\mathrm{LOD_{ref}}}$$

Calculate for each item *i* the RLODi as described in ISO 16140-2:2016. Tabulate the results as indicated in Table 3.5.

The AL for **paired** study data is set at 1.5, meaning that the LOD for the rapid kit/method shall not be higher than 1.5 times the LOD of the reference method.

The AL for **unpaired** study data is set at 2.5, meaning that the LOD for the rapid kit/method shall not be higher than 2.5 times the LOD of the reference method

Based on the AL and the additional information, it is decided whether the rapid kit/method is regarded as not fit for purpose for the item or category involved.

An Excel®-based program for calculating RLOD value is freely available for download at

https://www.wiwiss.fu-berlin.de/fachbereich/vwl/iso/ehemalige/wilrich/index.html (download the file "RLOD-ver4.xlsm")

Table 3.5: Presentation of RLOD before and after confirmation of the rapid kit/method results

	RLOD using the rapid-method	RLOD using the confirmed rapid-

	results	method results
Item (category) (*i*)	RLOD*i*	RLOD*i*
1		
2		
k		
Combined		
Table adapted from ISO16140-2:2016/Amd 1:2024		

6.1.2.3. Calculation of the LOD_{50} [Optional]

LOD50 shall be calculated for each category for the rapid method and optionally for the reference method. The LOD50 is used in method verification (see IS 17113: Part 3: 2022/ISO 16140-3).

For each category evaluated, determine the contamination of the low level by performing a 3-level MPN for the particular (food) item tested using the reference method (preferably) at the time of the RLOD experiment and/or by the enumeration of the inoculum at the time of inoculation using a non-selective medium. When enumeration is performed, the inoculum for the low and high contamination levels should be determined.

The MPN and 95 % confidence interval shall be determined for the fractional level only. For the low (or fractional) contamination level, analyse 20 test portions plus 5 test portions at approximately 2 times the test portion size and 5 test portions at approximately ½ the test portion size evaluated in the validation study (e.g. if the reference method test portions were 25 g, evaluate 5 test portions at 50 g and 5 test portions at 10 g). To each test portion, add a proportionate amount of enrichment broth as described in the reference method to maintain the enrichment volume to mass ratio (e.g. a reference method with a 1:10 enrichment ratio, add 450 ml to the 50 g test portions and 90 ml to the 10 g test portions). Analyse the test portions following the reference method from enrichment to confirmation. Use the number of positive results per test portion size to calculate the MPN value.

An Excel®-based program1 for calculating MPN values is freely available for download at https://standards.iso.org/iso/7218 (download the file “MPN calculation Excel program”).

Enumerate the inoculum at the time of inoculation by plating onto non-selective agar (see ISO 7218 for guidance on plating). Agar plates should be incubated under conditions to allow for optimal growth of the target microorganism.

Use the number of positive results per test portion size and the MPN value or the results of the enumeration of the inoculum to calculate the LOD_{50} and 95 % confidence interval.

An Excel®-based program for calculating LOD_{50} values is freely available for download at https://www.wiwiss.fu-berlin.de/fachbereich/vwl/iso/ehemalige/wilrich/index.html (download the file “PODLOD_ver12”).

The LOD_{50} value is calculated per category tested in the RLOD study and shall be expressed as cfu/test portion.

NOTE: *The 20 test portions from the low level are the same as the 20 test portions used in the RLOD study. Therefore, only 5 test portions at 2 times and 5 at approximately ½ the test portion size are analysed in addition to the RLOD study*.

6.1.3. Inclusivity/Exclusivity Study

6.1.3.1. Selection and number of strains

The strains used should take into account the measurement principle of the rapid kit/method (e.g. culture-based, immunoassay-based, and molecular). Different measurement principles may require the use of different test panels of strains.

The original source of all strains should be known and they should be held in a local (e.g. expert laboratory), national, or international culture collection to enable them to be used in future testing, if required.

For inclusivity testing, at least 20 pure cultures of (target) microorganisms shall be tested. For testing the inclusivity for *Salmonella* methods, at least 40 pure cultures of different serotypes of *Salmonella* shall be tested.

For exclusivity testing, at least 15 pure cultures of (non-target) microorganisms shall be tested.

NOTE F*or some microorganisms, it will be difficult to obtain the required number of strains for inclusivity and exclusivity. In these cases, an agreed set of test strains should be selected by the parties involved in the validation study.*

6.1.3.2. Inoculation of target strains (inclusivity)

Each test is performed once and only with the rapid kit/method (including a confirmation step if prescribed in the rapid-method protocol). Inoculation of a suitable non-selective growth medium is carried out with a dilution of a pure culture of each test strain. No sample is added.

The inoculum level shall be 10 times to 100 times greater than the minimum detection level of the rapid kit/method being validated to provide high cell populations in a stationary phase. If the rapid kit/method includes more than one (enrichment) protocol (e.g. for different sample types), then use the most challenging one with the complete panel of test strains. When negative or doubtful results are obtained, the test should be repeated and with the reference method included.

6.1.3.3. Inoculation of non-target strains (exclusivity)

Each test is performed once and only with the rapid kit/method (including a confirmation step if prescribed in the protocol). Inoculation of a suitable non-selective growth medium is carried out with a dilution of a pure culture of each test strain. No sample is added.

If the rapid kit/method involves a growth in a selective medium before a detection step, then for the purposes of exclusivity testing, the selective medium should be replaced by a non-selective medium. If the rapid kit/method gives a positive or doubtful result, then the test should be repeated using the complete (enrichment) protocol recommended in the instructions of the rapid method, if mentioned

so. Additionally, the reference method should be used to check that the strain could not be detected with the reference method.

6.1.3.4 Expression and interpretation of the results

Tabulate the results as shown in Table 3.6.

Table 3.6: Presentation of the results for inclusivity and exclusivity		
A. Inclusivity		
Microorganisms	Rapid Kit/Method	
Inclusivity (target strains)	Test result	Confirmed result
1		
2		
Etc.		
B. **Exclusivity**		
Microorganisms	Rapid Kit/Method	
Exclusivity (non-target strains)	Test result	Confirmed result
1		
2		
Etc.		

6.2. INTER LABORATORY STUDY

6.2.1. General considerations

The aim of the inter laboratory study is to determine the difference in sensitivity between the reference and the rapid kit/method when tested by different collaborators using identical samples (reproducibility conditions). The conditions for conducting the inter laboratory study should reflect, as much as possible, the normal conditions used by the individual collaborators in order to fulfill reproducibility conditions as far as possible. The inter laboratory study is managed by the organizing laboratory.

The inter laboratory study shall be conducted with collaborators belonging to more than one region. The collaborators shall be competent to perform both the reference method as well as the rapid method.

NOTE Competence can be demonstrated in different ways, e.g. for the reference method an ISO/IEC 17025 accreditation and for the rapid method a documented training.

6.2.2. Measurement protocol

The inter laboratory study must produce at least 10 valid data sets from at least 05 collaborators. The collaborators shall come from a minimum of five different organizations, excluding the organizing laboratory. A maximum of TWO data sets can be produced by one organization.

NOTE Laboratories in different locations, but belonging to one company or institute, are accepted as different organizations.

The protocol is as follows:

- In cases where different (enrichment) protocols for the rapid kit/method exist, a challenging (enrichment) protocol shall be selected, e.g. the protocol having the shortest incubation time or the most selective conditions. The relevant food product used to prepare the test samples should contain a natural background microflora;
- The food product shall be inoculated with the target organism. Samples shall be prepared by the organizing laboratory to ensure homogeneity;

At least three different levels of contamination shall be used: a negative control (*L0*) and two levels (*L1* and *L2*). At least one of these shall produce fractional positive results. The level of contamination needed to obtain fractional recovery shall be based on the RLOD study data of the reference method in the method comparison study. Theoretically, an average level of contamination of 1 cfu/sample is adequate to obtain fractional recovery. The level of contamination shall be determined. This allows calculation of the LOD50 of the rapid method, which is required in order to verify the performance of the rapid/Rapid method upon implementation of the validated method in a laboratory in accordance with IS 17113: Part 3: 2022/ISO 16140-3. The level of contamination is determined by performing an MPN analysis at the time of the start of the inter laboratory study,

- At least eight blind replicates at each level of contamination are analysed by each collaborator by both methods, so in total, a minimum of 48 results (eight replicates × three levels × two methods) per collaborator;
- Similar protocol needs to be followed for confirmation in case of paired and unpaired study as described in Single Laboratory Study.
- The combination “number of levels of contamination/number of replicates/numbers of non-outlier collaborators” shall be selected so that at least 240 results (120 by each method) are generated for use in the calculations for each method.

When the inter laboratory study is completed, all the information on data sheets and the results shall be submitted to the organizing laboratory, which shall determine which results are suitable for use in analysing the data as follows:

- disregard data from collaborators if transit conditions and times fall outside the specified acceptable tolerances (the limits for transport time and temperature have to be set before the samples are shipped);
- disregard data from collaborators that received samples/test kits, etc. that were damaged during transportation;

- disregard data from collaborators using media formulation that are not in accordance with the (reference) method;
- disregard data from collaborators if the questionnaire suggests that the laboratory has deviated from either the standard protocol or the critical operating conditions.

6.2.3. Calculations and Summary of Data

The results obtained by the individual collaborators in the interlaboratory study are summarized in Table 3.7 and in Table 3.8.

Table 3.7: Positive results by the reference method

Collaborators	**Contamination Level**		
	L_0	L_1	L_2
Collaborator 1	/8[a]	/8[b]	/8[c]
Collaborator 2	/8	/8	/8
Collaborator 3	/8	/8	/8
Etc.	/8	/8	/8
Total	P_0	P_1	P_2

[a] Number of positive reference-method results at level 0.
[b] Number of positive reference-method results at level 1.
[c] Number of positive reference-method results at level 2.

Table 3.8: Positive results (before and after confirmation) by the rapid kit/method

Collaborators	**Contamination Level**					
	L_0		L_1		L_2	
Collaborator 1	/8[a]	/8[b]	/8[c]	/8[d]	/8[e]	/8[f]
Collaborator 2	/8	/8	/8	/8	/8	/8
Collaborator 3	/8	/8	/8	/8	/8	/8
Etc.	/8	/8	/8	/8	/8	/8
Total	P_0	CP_0	P_1	CP_1	P_2	CP_2

[a] Number of positive rapid-method results at level 0.
[b] Number of confirmed rapid -method results at level 0.
[c] Number of positive rapid -method results at level 1.
[d] Number of confirmed rapid -method results at level 1.
[e] Number of positive rapid -method results at level 2.

[f] Number of confirmed rapid -method results at level 2.

6.2.3.1 Calculation of specificity

Calculate the percentage specificity (SP) of the reference method and the rapid kit/method, using the data after confirmation, based on the results of level L_0 (see Box 3)

Box 3: Percentage specificity (SP) of the reference method	
Specificity for the reference method	$SP_{ref} = \left[1 - \left(\frac{P_0}{N_-}\right)\right] \times 100\%$
Specificity for rapid kit/method	$SP_{rapid} = \left[1 - \left(\frac{CP_0}{N_-}\right)\right] \times 100\%$
where N− is the number of all L_0 tests; P_0 is the total number of false-positive results obtained with the blank samples before confirmation; CP_0 is the total number of false-positive results obtained with blank samples	

6.2.3.2 Interpretation and calculations for sensitivity

The results for all the collaborators for each of the levels L_1 and L_2 are combined and summarized as per the interpretations given in Table 3.9 for a paired study and in Table 3.10 for an unpaired study design, Summarize the results of all the collaborators as per the Table 3.11.

Table 3.9 Interpretation of results for all collaborators for a paired study

Reference method	**Rapid method**[a]	**Confirmed rapid method**[b]	**Interpretation** (based on the confirmed rapid-method result)
+	+	Not needed[c]	Positive Agreement (PA)
-	-	Not needed[c]	Negative Agreement (NA)
+	-	Not needed[c]	Negative Deviation due to false negative rapid kit/method result ($ND_{FN(rapid)}$)
-	+	+	Positive Deviation (PD)
-	+	-	Positive Deviation due to false positive rapid kit/method result ($PD_{FP(rapid)}$)[d]

[a] The rapid kit/method results includes any confirmations as described in the rapid method protocol.
[b] The confirmed rapid method result is the result after additional confirmation as described in the protocol for the validation study.
[c] No need for additional confirmation test(s). Confirmed rapid kit/method result is the same as the rapid

kit/method result
[d] This false positive result (FP) shall also be used to calculate the false positive ratio
Table adapted from ISO16140-2:2016/Amd 1:2024

Table 3.10 Interpretation of results for all collaborators for an unpaired study

Reference method	**Rapid method**[a]	**Confirmed rapid method**[b]	**Interpretation** (based on the confirmed rapid-method result)
+	+	+	Positive Agreement (PA)
+	+	-	Positive Agreement due to false positive rapid kit/method result ($PA_{FP(rapid)}$)
-	-	-	Negative Agreement (NA)
-	-	+	Negative Agreement due to false negative rapid kit/method result ($NA_{FN(rapid)}$)
+	-	-	Negative Deviation (ND)
+	-	+	Negative Deviation due to false negative rapid kit/method result ($ND_{FN(rapid)}$)
-	+	+	Positive Deviation (PD)
-	+	-	Positive Deviation due to false positive rapid kit/method result ($PD_{FP(rapid)}$)

[a] The rapid kit/method results includes any confirmations as described in the rapid method protocol.
[b] The confirmed rapid method result is the result after additional confirmation as described in the protocol for the validation study.
Table adapted from ISO16140-2:2016/Amd 1:2024

Determine the Total Negative Deviation (TND) and Total Negative Agreement (TNA) for the validation study (see Box 4).

Box 4: Determination of Total Negative Deviation (TND) and Total Negative Agreement (TNA)	
Paired evaluation	
Total Negative Deviation	$TND = ND_{FN(rapid)}$
Total Negative Agreement	$TNA = NA + PD_{FP(rapid)}$
Unpaired evaluation	
Total Negative Deviation:	$TND = ND + ND_{FN(rapid)} + PA_{FP(rapid)}$
Total Negative Agreement	$TNA = NA + NA_{FN(rapid)} + PD_{FP(rapid)}$
Where ND is Negative deviation NA is Negative agreement $ND_{FN(rapid)}$ is Negative Deviation due to false negative by rapid kit/method result $PA_{FP(rapid)}$ is Positive Agreement due to false positive by rapid kit/method result $NA_{FN(rapid)}$ is Negative Agreement due to false negative by rapid kit/method result $PD_{FP(rapid)}$ is Positive Deviation due to false positive by rapid kit/method result	

Table 3.11 Summary of results for all collaborators obtained with the reference and rapid kit/method (after confirmation) for Level L_1 or L_2

	Rapid kit/method positive (A+)	Rapid kit/method negative (A-)
Reference method positive (R+)	+/+ Positive Agreement (PA)	+/- Total Negative Deviation (TND)
Reference method negative (R-)	- /+ Positive Deviation (PD)	-/- Total Negative Agreement (TNA)

Based on data summarized in Table 3.11, calculate the values for sensitivity of the rapid kit/method (SE_{rapid}) and reference method (SE_{ref}), as well as relative trueness (RT), false positive ratio and false negative ratio as shown in Box 5. The confirmed rapid kit/method results shall be used to determine whether the rapid kit/method produces comparable results to the reference method.

6.2.3.3 Interpretation of trueness data

Paired study

For a paired study, calculate the difference between (TND – PD) and the sum of (TND + PD) for the level(s) where fractional recovery was obtained (so *L1* and possibly *L2*). The values found for (TND – PD) and (TND + PD) shall not be higher than the Acceptability Limits (ALs) given in Table 3.12 with respect to the number of participating laboratories (N_{lab}).

Box 5 : Formulae for determining Sensitivity, Relative Trueness and False Positive/Negative ratio for interlaboratory study	
Sensitivity for the rapid method	$SE_{rapid} = \frac{(PA + PD)}{(PA + TND + PD)} \times 100\%$
Sensitivity for reference method	$SE_{ref} = \frac{(PA + TND)}{(PA + TND + PD)} \times 100\%$
Relative trueness	$RT = \frac{(PA + TNA)}{N} \times 100\%$
False positive ratio (FPR) and false negative ratio (FNR) for the rapid method	
Paired evaluation	$\mathrm{FPR} = \frac{PD_{FP(rapid)}}{TNA} \times 100\%$
Unpaired evaluation	$\mathrm{FPR} = \frac{PA_{FP(rapid)} + PD_{FP(rapid)}}{TNA} \times 100\%$
Paired evaluation	$\mathrm{FNR} = \frac{\mathrm{NA_{FN(rapid)}} + \mathrm{ND_{FN(rapid)}}}{\mathrm{PA + TND + PD}}$
where *N* is the total number of samples (PA + PD + TND + TNA), FP is the false positive results FN is the false negative results TND is the Total Negative Deviation TNA is the Total Negative Agreement PA is Positive agreement PD is Positive Deviation $ND_{FN(rapid)}$ is Negative Deviation due to false negative by rapid kit/method result $PA_{FP(rapid)}$ is Positive Agreement due to false positive by rapid kit/method result $NA_{FN(rapid)}$ is Negative Agreement due to false negative by rapid kit/method result $PD_{FP(rapid)}$ is Positive Deviation due to false positive by rapid kit/method result	

Table 3.12: Acceptability limits for a paired study design in relation to the number of collaborating laboratories

N_{lab}	(TND-PD)	(TND+PD)
10	3	4
11	4	4
12-13	4	5
14-16	4	6
17	4	7
18	5	7
19-20	5	8

The AL is not met when the observed value is higher than the AL. When the AL is not met, investigations should be made (e.g. root cause analysis) in order to provide an explanation of the observed results. Based on the AL and the additional information, it is decided whether the rapid method is regarded as not fit for purpose. The reasons for acceptance of the rapid method in case the AL is not met shall be stated in the study report.

Unpaired study

For the unpaired study, calculate the difference between (TND=PD) for the level(s) where fractional recovery was obtained (so L_1 and possibly L_2). The observed value found for (TND-PD) shall not be

higher that the AL. The AL is defined as [$(TND\text{-}PD)_{max}$] and calculated per level where fractional recovery was obtained as shown in Box 6 using the following three parameters : $(p+)_{ref}$, $(p+)_{rapid}$ and $(TND\text{-}PD)_{max}$.

Box 6: Evaluation of Acceptability Limits (AL) for unpaired study	
$(p+)_{ref} = \frac{P_{x(ref)}}{N_{x(ref)}}$	Where $P_{x(ref)}$ is number of samples with a positive result obtained with the reference method at level x (L_1 or L_2) for all laboratories; $N_{x(ref)}$ is number of samples tested at level x (L_1 or L_2) with the reference method by all laboratories.
$(p+)_{rapid} = \frac{CP_{x(rapid)}}{N_{x(rapid)}}$	where $CP_{x(rapid)}$ is number of samples with a confirmed positive result obtained with the rapid kit/method at level x (L_1 or L_2) for all laboratories; $N_{x(rapid)}$ is number of samples tested at level x (L_1 or L_2) with the rapid kit/method by all laboratories.
$((TND-PD)_{max} = \sqrt{3N_{x(ref)} \times \left((p+)_{ref} + (p+)_{rapid} - 2\left((p+)_{ref} \times (p+)_{rapid}\right)\right)}$	

The AL is not met when the observed value is higher than the AL. When the AL is not met, investigations should be made (e.g. root cause analysis) in order to provide an explanation of the observed results. Based on the AL and the additional information, it is decided whether the rapid method is regarded as not fit for purpose. The reasons for acceptance of the rapid method when the AL is not met shall be stated in the study report.

6.2.3.4. Calculation of the Relative Level of Detection (RLOD)

Additionally, for both a paired and unpaired study, an evaluation should be made for the difference between the relative levels of detection (RLOD) between laboratories. This evaluation can be conducted according to protocol specified in Annex F of IS 17113 : Part 2 : 2019/ISO 16140-2:2016. The ALs for the RLOD of paired and unpaired studies are found in 5.1.2.2.

An Excel®-based program for calculating RLOD values from collaborators is freely available for download at https://www.wiwiss.fu-berlin.de/fachbereich/vwl/iso/ehemalige/wilrich/index.html (download the file "PODLOD-Interlab_ver2").

Note: Worked out examples are furnished in Annexure 3.2

7.0 VALIDATION PROTOCOL FOR QUANTITATIVE MICROBIOLOGICAL RAPID TEST KITS/EQUIPMENT/METHODS

7.1. Method comparison study

7.1.1 General considerations

The method comparison study is the part of the validation process that is performed in the organizing laboratory. It consists of four parts.

1) Relative trueness study
2) Accuracy profile (AP) study
3) Limit of quantification (LOQ)
4) Inclusivity/Exclusivity study of the rapid kit/method

The results (tables and calculations) of the different parts and the interpretation of the results, including discrepant results, shall be given in a study report.

The organizing laboratory shall be competent to perform both the reference method as well as the rapid method.

NOTE Competence can be demonstrated in different ways, e.g. for the reference method an ISO/IEC 17025 accreditation and for the rapid method a documented training.

7.1.2. Relative trueness study

The relative trueness study is a comparative study between the results obtained by the reference method and the results of the rapid kit/method. This study is conducted using naturally and/or artificially contaminated samples. Different categories, types, and items will be tested for this.

7.1.2.1 Selection of categories to be used

If the method is to be validated for a restricted number of food categories, then only those categories can be studied.

If the method is to be applied for a broad range of foods, then at least five major categories (Please see ***Annexure 3.1***) of food shall be studied.

For all selected categories, at least three different food products per category shall be included in the study.

When selecting samples for the study, it is of the highest priority to find those that are naturally contaminated. If it is not possible to acquire a sufficient number of naturally contaminated samples, artificial contamination of samples is permissible.

Following shall be included in the study while selecting different food products:

- both high and low (natural) background microflora,
- different types of stresses due to processing, and
- raw (unprocessed) food products.

7.1.2.2 Number of samples

For each category being examined, a minimum of 15 samples shall be tested. At least three food products within a category should be used. For each food product, at least five samples representative for this product shall be tested (three food products × 5 samples for each product = 15 samples).

The samples should be contaminated at a level that is representative for the natural variation in level of contamination. The reference and rapid methods shall be performed with, as far as possible, exactly the same sample.

7.1.2.3 Calculation and interpretation of relative trueness study

The results obtained are analysed using the Bland-Altman method. Plot the data for each sample per category and for each sample in all categories and draw the line of identity on which all points would lie if the two methods gave identical results for each sample analysed. The plot for all categories should show the results of each category tested with a distinct symbol. This provides a rapid visual assessment of the extent to which the two methods (do not) agree.

Determine the average of each pair of data values and the difference between the values as in Table 3.13 and plot these differences against the corresponding averages per category and for all categories to illustrate the degree of agreement between the reference method and the rapid method. Figure 3.1 shows the line of identity (zero difference), the line of bias (average difference) as well as the upper and lower 95 % prediction limits of the individual sample-specific bias values. These limits of the prediction range are called limits of agreement.

Table 3.13 — Summarized results for all categories

Food Category	Food Product	Sample	Log_{10} cfu		Mean	Difference
			Reference method result	Rapid kit/method result		
1	1	1	R1	A1	(R1+A1)/2	D1=A1-R1
		2	R2	A2	(R2+A2)/2	D2=A2-R2
		3	R3	A3	(R3+A3)/2	D3=A3-R3
		4	R4	A4	(R4+A4)/2	D4=A4-R4
		5	R5	A5	(R5+A5)/2	D5=A5-R5
1	2	6	R6	A6	(R6+A6)/2	D6=A6-R6
		7	R7	A7	(R7+A7)/2	D7=A7-R7
		8	R8	A8	(R8+A8)/2	D8=A8-R8
		9	R9	A9	(R9+A9)/2	D9=A9-R9
		10	R10	A10	(R10+A10)/ 2	D10=A10-R10
1	3	…	…	…	…	…
Average category 1					$\overline{D_1}$	

Standard deviation category 1				S_{D1}		
…	…	…	…	…	…	…
x4			Rx	Ax	(Rx+Ax)/2	Dx=Ax-Rx
Average category x				$\overline{D_x}$		
Standard deviation category x				S_{Dx}		
Average all categories				$\overline{D_{all}}$		
Standard deviation all categories				S_{Dall}		

Compute per category and for all categories the average difference $\overline{D}$, the standard deviation of differences *sD* and the limits of agreement using the following formula:

$$\left[\overline{D} \pm T\sqrt{\frac{s_D^2}{n}}\right]$$

Where n is number of data pairs, *T* is the percentile of a student's t distribution for the prediction probability β (β=95% is used) and for (n-1) degrees of freedom that is: $T_{\left(\frac{1-\beta}{2}\right);(n-1)}$

Plot as in Figure 3.1 the individual sample differences against the mean values on a graph that shows the line of identity (zero difference), the line of bias (average difference) as well as and the upper and lower 95 % prediction limits (limits of agreement) of the individual sample-specific bias values (both separately per category and across all categories). This illustrates the degree of bias and the (lack of) agreement of the data.

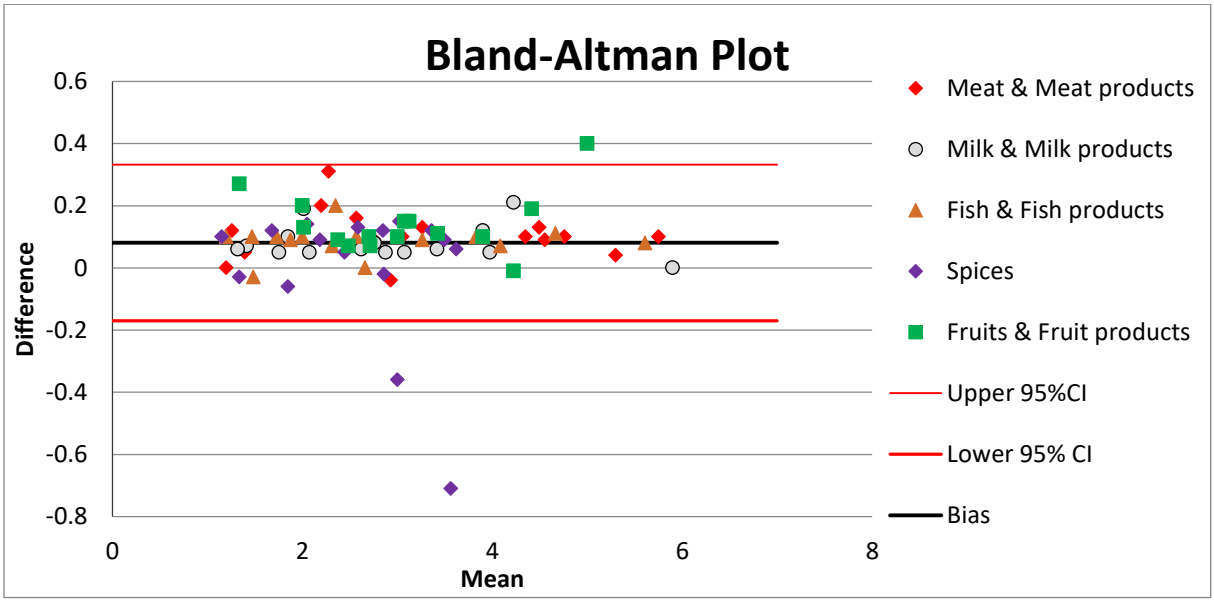


Figure 3.1: Bland-Altman difference plot for all categories

The results of the difference and scatter plot shall be interpreted based on a visual observation of the overall bias, the spread of the individual sample-specific bias values, and any bias values lying outside the limits of agreement. If the individual bias values are normally distributed, it can be expected that 1 in 20 values will lie outside the limits of agreement. Discrepancies can be seen as an indication that the individual bias values do not follow a normal distribution, e.g. due to the presence of outliers. Any such departures from expectations should be documented.

7.1.3. Accuracy Profile (AP) study

The accuracy profile study is a comparative study between the results obtained by the reference method and the results of the rapid kit/method. This study is conducted using artificially contaminated samples. One product per food category is to be tested.

The detail description on principle and application of Accuracy Profile is prescribed in ISO 16140-2:2016 Annex G & Annex H.

7.1.3.1 Selection of categories to be used

See 7.1.2.1.

7.1.3.2 Number of samples

For each food category being examined, at least one product shall be tested using six samples per food product. Of the six samples, there should be two at a low level, two at an intermediate level, and two at a high level of contamination. These levels should cover the whole range of contamination of the selected type. For each sample, five replicates representing five different test portions from the same sample shall be used.

7.1.3.3 Calculation and interpretation of accuracy profile study

The calculations of accuracy profile shall be as per tabulated results as in Table 3.14 based on log-transformed counts.

Table 3.14 Results of the accuracy study (in log_{10} cfu/g)

Food Category	Food Product	Item (level)	Reference method					Rapid Kit/Method				
			t.p.[a]1 (x_1)[b]	t.p.2 (x_2)	t.p.3 (x_3)	t.p.4 (x_4)	t.p.5 (x_5)	t.p.1 (y_1)	t.p.2 (y_2)	t.p.3 (y_3)	t.p.4 (y_4)	t.p.5 (y_5)
Category 1	Product 1	Sample 1 (low)										
		Sample 2 (low)										
		Sample 3 (intermediate)										
		Sample 4 (intermediate)										
		Sample 5										

		(high)										
		Sample 6 (high)										
…	…	…										
Category x	Product x	Sample 1-6										

[a] t.p. = test portion.
[b] (x_a) = $\log_{10}$ test result for the reference method (x) for test portions 1 to 5.
[c] (y_a) = $\log_{10}$ test result for the rapid kit/method (y) for test portions 1 to 5

The accuracy profile is used to check the requirement that the rapid kit/method produces a result for a sample that differs from the value produced by the reference method by less than a certain acceptability criterion. For notation used and calculations see Box 7.

Box 7: Notations used for calculating accuracy profile.	
i	Refers to sample
q	number of samples ($1 \leq i \leq q$)
j	Refers to test portion
n	Number of test portions ($1 \leq j \leq n$)
xij	the log*10* transformed test result of sample *i* for replicate *j* with $1 \leq i \leq q$ and $1 \leq j \leq n$ using the reference method
yij	the log*10* transformed test result of sample *i* for replicate *j* with $1 \leq i \leq q$ and $1 \leq j \leq n$ using the rapid method.

For each sample, measurements are made under repeatability conditions for both methods. *yi* values are assumed to be normally distributed. The β-expectation tolerance interval (β-ETI) of the *yi* values is computed and it is assumed that a combined standard deviation can be calculated that holds for all item samples.

The Acceptability Limit is set at: AL = ±0.5 $\log_{10}$ units. It is expressed as a difference between the reference and the rapid kit/method.

Calculations are performed per category/type as the following sequence of operations:

Step 1	For each sample *i,* calculate the central value X_i as the median of the log*10* transformed counts obtained with the reference method, *xij*. These values are the reference values of the validation samples $X_i = median(x_{ij})$
Step 2	For each sample *i*, calculate the central value *Yi* as the median of the log*10* transformed

	counts obtained with the rapid method *yij*. These values are the rapid values of the validation samples: $Y_i = median(y_{ij})$
Step 3	For each sample *i*, calculate the standard deviation *salt,i* as follows: $s_{alt,i} = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(y_{ij} - \bar{y}_i)^2}$
Step 4	Calculate the combined standard deviation s_{alt} as follows $s_{alt} = \sqrt{\frac{1}{q}\sum_{i=1}^{q} s_{alt,i}^2}$
Step 5	Calculate the combined standard deviation of the reference method s_{ref} (analogous to step 3 and step 4) as follows: $s_{ref,i} = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(x_{ij} - \bar{x}_i)^2}$ $s_{ref} = \sqrt{\frac{1}{q}\sum_{i=1}^{q} s_{ref,i}^2}$
Step 6	For each sample *i*, compute the absolute bias as the difference of the medians calculated for both methods $B_i = Y_i - X_i$ This is an estimate of the lack of trueness of the rapid method in comparison to the reference method
Step 7	For each sample *i*, compute the limits of the β-ETI. This is the interval where the expected proportion of future results will fall is β. For each sample, β-ETI is expressed as $\left[B_i \pm T.s_{alt}\sqrt{1+\frac{1}{n}}\right]$ where *T* is the percentile of a Student-*t* distribution for β the chosen probability and *q(n-1)*1degrees of freedom (24 in de-requested setup), that is: $T_{\left(\frac{1-\beta}{2}\right);q(n-1)}$ For the purpose of this part of the guideline, β is set at 80%. *T* is the coverage factor of the

β-ETI of the validation sample. It defines the upper limit *Ui* and the lower limit L_i.

$$U_i = B_i + T.s_{alt}\sqrt{1+\frac{1}{n}}$$

$$L_i = B_i - T.s_{alt}\sqrt{1+\frac{1}{n}}$$

Step 8 For each category, tabulate the different values calculated for the samples as in Table 3.15.

Table 3.15 — Presentation of the statistical results of the comparison study

Category	Sample	Central value (Ref)	Central value (Rapid)	Bias	Upper β-ETI	Lower β-ETI	Upper AL	Lower AL
Category 1	Sample 1	X_i	Y_i	B_i	U_i	L_i	+AL	-AL
	Sample 2							
	Sample 3							
	Sample 4							
	Sample 5							
	Sample 6							

Make a graphical representation of computed results as follows:

— the horizontal axis is for reference values X_i in log10 units;

— the vertical axis is for the bias, the Acceptability Limits, and the tolerance interval limits U_i-X_i and L_i-X_i all expressed in log10 units as differences to the corresponding reference value of the sample.

Make a graphic representation like the example given in Figure 3.2. The upper and lower tolerance-interval limits are connected by straight lines to interpolate the behaviour of the limits between the different levels of the validation samples. The horizontal line represents the reference values obtained with the reference method. The differences between reference values and average levels of contamination are represented by black dots. Whenever no biases exist, these recovered values are located on the horizontal reference line. In addition, Acceptability Limits are represented by two dashed horizontal lines and β-ETI limits as broken full lines.

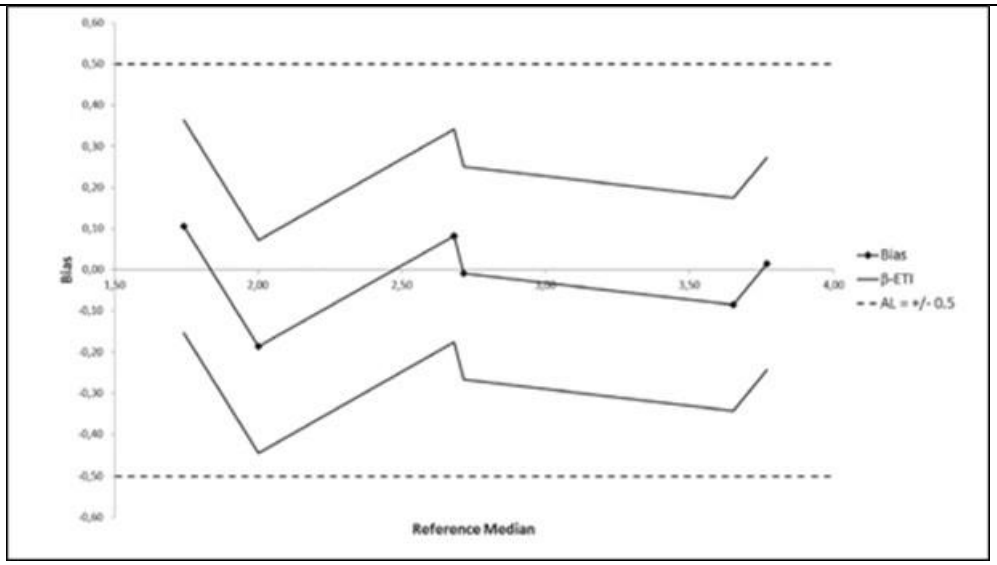


Figure 3.2 — Example of accuracy profile for a category in the method comparison study

If for all *i* in the accuracy profile *Ui* ≤ AL and *Li*_−AL, the rapid method is accepted as being equivalent to the reference method for the individual categories and the combined categories.

If any of the upper or lower limits exceeds the Acceptability Limits and the standard deviation, *S*ref > 0,125 but < 0,25, the following additional evaluation procedure is followed:

In case the *Sref* > 0,25 no recalculation of the AL is allowed. In this case investigations should be made (e.g. root cause analysis) in order to provide an explanation of the high *Sref*, correct possible issues and,as an option, repeat the experiment.

Step 9

Calculate new Acceptability Limits as a function of the standard deviation: Als = 4. S_{ref} If for all *i* in the accuracy profile *Ui* ≤ AL*s* and *Lis*_−AL, the rapid method is accepted as being equivalent to the reference method for the given combination category and type.

NOTE If S_{ref} ≤ 0,125, the new Acceptability Limits would be smaller than or equal to 0,5. The second evaluation would yield the same results in that case.

The rapid kit/method is accepted as being equivalent to the reference method if it is equivalent for all individual and combined categories. Based on investigations, it is decided whether the methods are regarded as equivalent or not for the category or categories involved.

For the Accuracy Profile calculations, Excel® spreadsheets is freely available at http://standards.iso.org/iso/16140 and then download the file -

AP_calculation tool MCS 16140-2_clause 6-1-3-3 ver_31-07-2018.xlsx

7.1.4. Limit of Quantification (LOQ) Study

The LOQ is only relevant when the measurement principle of the rapid kit/method is not based on counting visible colonies of the target microorganism and shall therefore be determined in these cases. This study is only done for instrumentally-based methods. Examples of methods for which the LOQ needs to be determined are the instrumental measurement of fluorescence which is related to the growth of the microorganism.

7.1.4.1 Selection of categories to be used

Select the same food categories and food products as used for the accuracy profile study.

7.1.4.2 Number of samples

Blank samples are tested per food product/category used. A minimum of 10 test portions from the same sample shall be used. Examine the test portions with the rapid kit/method.

7.1.4.3 Calculation and interpretation of limit of quantification study

Calculate the standard deviation s_0 of the n results as follows:

$$s_0 = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}\left(y_j - \bar{y}\right)^2}$$

where

n is the total number of test portions used;

y_j is the log10 transformed result of test portion j;

$\overline{y}$ is the average log10 transformed result of all test portions.

The limit of quantification is calculated as LOQ = 10 s_0.

7.1.5 Inclusivity and exclusivity study

Inclusivity and exclusivity testing is not required for general enumeration methods such as total plate count (TPC) and yeast and mould (Y&M) methods. It is required for enumeration methods designed for specific microorganisms (e.g. *Listeria, Enterobacteriaceae*).

7.1.5.1 Selection and number of test strains

Select the strains as per criteria specified for inclusivity and exclusivity study for Qualitative Microbiological Rapid Test Kits/Equipment/Methods.

For inclusivity testing, at least 20 pure cultures of (target) microorganisms shall be tested.

For exclusivity testing, at least 15 pure cultures of (non-target) microorganisms shall be tested.

7.1.5.2 Target microorganisms (inclusivity)

Each test is performed once and with the rapid kit/method, the reference method, and a non-selective agar. The inoculum level should be at least 100 times greater than the minimum level for quantification

of the rapid kit/method being validated. When using a plate method as the rapid kit/method, the inoculum level shall obtain a countable number on the plate.

7.1.5.3 Non-target microorganisms (exclusivity)

Each test is performed once and with the rapid and the reference method. The inoculum level should be similar to the greatest level of contamination expected to occur in any of the categories being used. No sample is added. The pure culture should be grown in a suitable non-selective broth under optimal conditions of growth for at least 24 h and diluted to an appropriate level before testing begins. If the organism cannot be cultured, a stock suspension should be diluted to an appropriate level before use.

7.1.5.4 Expression and interpretation of results

Tabulate the results as in Table 3.16 for the inclusivity tests and Table 3.17 for the exclusivity tests. The interpretation shall be done by the laboratory in charge of the method comparison study. The report should state any anomalies from the expected results.

Table 3.16 Presentation of results for inclusivity

Microorganisms	Reference Method	Rapid Kit/Method	Non-selective Agar
1			
2			
Etc.			

The interpretation of the inclusivity data for a rapid kit/method using a plating medium is made on a qualitative basis. However, quantitative data should facilitate the interpretation of the data.

Table 3.17 Presentation of results for exclusivity

Microorganisms	Reference Method	Rapid Kit/Method
1		
2		
Etc.		

7.2. Inter laboratory study

7.2.1 General considerations

See 6.2.1

7.2.2 Measurement protocol

The inter laboratory study shall produce at least eight valid data sets from at least FOUR collaborators. The collaborators shall come from a minimum of four different organizations, excluding the organizing laboratory. A maximum of TWO data sets can be produced by one organization.

NOTE Laboratories in different locations, but belonging to one company or institute, are accepted as different organizations.

The inter laboratory study shall be conducted with collaborators belonging to more than one REGION/LOCATION. The collaborators shall be competent to perform both the reference method as well as the rapid method.

NOTE Competence can be demonstrated in different ways, e.g. for the reference method an ISO/IEC 17025 accreditation and for the rapid method a documented training.

The accuracy and precision estimates should be calculated from a large number of duplicate test results. This figure should be a minimum of 48 results for the one item chosen consisting of FOUR collaborators, three levels of contamination, two methods of enumeration (reference and rapid), and duplicate measurements, i.e. 4 × 3 × 2 × 2 = 48 [One sample of each level to be tested in duplicate in 4 labs].

The organizer is responsible for the preparation of the test protocol and a data sheet for the recording of all experimental data and critical experimental conditions used by each laboratory.

The protocol is as follows

- A relevant food product is used to prepare the test samples. The product should contain a natural background microflora.
- The selected food product can be inoculated with the target organism. Samples shall be prepared to ensure homogeneity between samples. Homogeneity tests and criteria for acceptance are described in ***IS 17385: 2020/ ISO 22117: 2019 Microbiology of the Food Chain-Specific Requirements and Guidance for Proficiency Testing by Inter laboratory Comparison.***
- At least three different levels of contamination shall be used cover at least the lower, middle, and upper levels of the entire range of the rapid kit/method. A negative control level should be included in addition.
- Duplicate samples (blind coded) are tested by each collaborating laboratory at the three levels of contamination.
- When the inter laboratory study is completed, all the information on data sheets and the results shall be submitted to the organizing laboratory, which shall determine which results are suitable for use in analysing the data as per criteria mentioned in clause 1.2.2.

7.2.3 Calculations, summary, and interpretation of data

The $\log_{10}$ transformed test results of the different collaborators for both the reference and rapid kit/method are presented in Table 3.18. Note the data as follows:

- x_{ijk}, the $\log_{10}$ transformed test result on level i for replicate j of collaborator k with $1 \le i \le q$, $1 \le j \le n$ and $1 \le k \le p$ using the reference method;
- y_{ijk}, the $\log_{10}$ transformed test result on level i for replicate j of collaborator k with $1 \le i \le q$, $1 \le j \le n$ and $1 \le k \le p$ using the rapid kit/method.

Where

i refers to the level and *q* is the number of levels ($1 \leq i \leq q$); *j* refers to the replicate and *n* is the number of replicates ($1 \leq j \leq n$); k refers to the collaborator and *p* is the number of collaborators ($1 \leq k \leq p$).

Table 3.18 Summary of the results of the inter laboratory study per each analyte level (k)					
		Reference Method x_{ijk}		Rapid Kit/Method y_{ijk}	
Collaborators (k)	Level (k_i)	Result		Result	
1	Blank				
2	Blank				
Etc.	Blank				
(1)	Blank				
		Duplicate 1	Duplicate 2	Duplicate 1	Duplicate 2
1	Low				
2	Low				
Etc.	Low				
(1)	Low				
1	Medium				
2	Medium				
Etc.	Medium				
(1)	Medium				
1	High				
2	High				
Etc.	High				
(1)	High				

The detailed calculations of Acceptance Limits are available in ISO 16140-2:2016 clause 6.2. The Acceptability Limit is set at ±0,5 $\log_{10}$ units. The AL is not met when the observed value is higher than the AL. Based on the AL and the additional information, it is decided whether the rapid kit/method is regarded as not fit for purpose for the category or categories involved.

For the Accuracy Profile calculations, Excel® spreadsheets is freely available at http://standards.iso.org/iso/16140 and then download the file - AP_calculation_tool_ILS_(clause_6-2-_Calculations_summary_and_interpretation_of_data).

NOTE: Worked out examples are furnished in Annexure 3.2.

Annexure 3. 1

Food Categories for Rapid Microbiological Analysis

S. No.	Food categories	Type	Example of most challenging food products
1	Fish and Fish Products	Major	Raw shrimp Fermented fish products Dry fish
2	Milk and Milk Products	Major	Raw milk Milk powder Cottage cheese
3	Spices and Herbs	Major	Onion Powder Black pepper Seasoning powder
4	Fruits and Vegetables and their products	Major	Pickle, Fruit juice with natural color Jam
5	Meat and Meat Products	Major	Fresh meat Sausage Fermented meat products
6	Yeast- Baker's (compressed and dried), Brewer's yeast, Wild yeast	Other	Baker yeast powder
7	Non-Carbonated Water Based Beverages (Non -alcoholic)	Other	Low pH beverages
8	Eggs and Egg Products	Other	Egg powder
9	Cereals and Cereal Products	Other	Flour Dry cereal like oats Low moisture foods
10	Health Supplements, Nutraceuticals, Food for Special Dietary Use, Food for Special Medical Purpose, Functional Food and Novel Food	Other	Low moisture foods in the form of tablets, capsules, powders etc. High moisture foods in the form of liquids, syrups, jellies etc.
11	Alcoholic beverages	Other	Beverages with high alcohol content e.g. brandy, whisky, vodka etc. Beverages with low alcohol content e.g. wine, beer etc.
12	Ayurveda Aahara	Other	Low moisture foods in the form of tablets, capsules, powders etc. High moisture foods in the form of liquids, syrups, jellies etc.

Annexure 3. 2

WORKED OUT EXAMPLES

Example 1: Qualitative Methods

A rapid Salmonella Testing Kit is submitted for validation in meat products, milk products, fish and vegetables.

The reference method used was IS 5887 : Part 3 : Sec 1 : 2020/ISO 6579-1:2017 Horizontal method for the detection of *Salmonella* spp.

The rapid method has 2 enrichment protocols:

Protocol 1: unprocessed raw/frozen food, with high background microflora: in Buffered Peptone Water (BPW) supplemented with selective reagents for 22 h ± 2 h at 41.5 °C ± 1°C

Protocol 2: processed food with low background microflora: in BPW for 22 h ± 2 h at 37 °C ± 1°C

Study Design:

Protocol 1- Unpaired study design because reference method and rapid kit have different primary enrichment procedures

Protocol 2- Paired study design because reference method and rapid kit have common primary enrichment procedures

Method comparison study

Sensitivity study

Number and nature of samples:

412 samples- 254 samples with Protocol 1 and 158 samples with Protocol 2

Category		**Type**	**Positive Samples**	**Negative samples**	**Total**
1	Cereal & cereal products	Flour	14	10	24
		Dry cereal	21	20	41
		low moisture	9	11	20
2	Meat & meat products	a. Raw meat	12	11	23
		b. Raw poultry meat	10	10	20
		c. Delicateessen (raw and cooked)	12	16	28
3	Milk & Milk products	a1. Pasteurised products	3	9	12
		a2. Milk powder	7	4	11
		b. Fermented/ acidified products	11	11	22
		c. Raw milk-based products	12	12	24
4	Fish & fish	a. Cooked shrimp	8	12	20

	products	b. Fish pickle	11	9	20
		c. Dry fish	11	9	20
5	Egg and egg products	a. Egg powders	10	10	20
		b. Liquid egg products	11	11	22
		c. Egg based products (mayonnaise, custard etc.)	9	11	20
6	Fruits & Vegetables and their products	a. Raw products	11	14	25
		b. Low moisture products	8	12	20
		c. Heat processed products	13	7	20
Total			203	209	412
Protocol 1			122	132	254
Protocol 2			81	77	158

182 samples were artificially contaminated, using 63 different strains. 172 give a positive result. Most of the inoculation levels, after injury protocols on the inoculum, were lower or equal to 5 CFU/sample. 15.3 % of the samples are naturally contaminated.

Confirm the positive results of Rapid Salmonella kit by a subculture in RVS broth for 24h at 41.5 °C, followed by streaking onto selective agar plates XLD and BGA/chromogenic agar.. Confirm the isolated colonies by:

- Latex tests directly on isolated colonies
- Tests described in the reference method

Calculation of Relative Trueness (RT), Sensitivity (SE) and False positive ratio (FPR):

Category		Type	Protocol	PA	NA	PD	ND	$ND_{FN(rapid)}$	$PD_{FP(rapid)}$	$PA_{FP(rapid)}$	SE_{rapid} %	SE_{ref} %	RT%	FPR%
1	Cereal & cereal products	a.	2	12	10	0	2	2	0	0	85.7	100	100	0
		b.	2	21	20	0	0	0	0	0	100	100	100	0
		c.	2	9	11	0	0	0	0	0	100	100	100	0
2	Meat & meat products	a.	1	4	11	6	2	0	0	0	83.3	50	65.2	0
		b.	1	8	10	2	0	0	0	0	100	80	90	0
		c.	1	9	16	2	1	0	0	0	91.7	83.3	89.3	0
3	Milk & milk products	a1.	1	2	9	0	1	0	0	0	66.7	100	91.7	0
		a2.	2	6	4	0	1	1	0	0	85.7	100	90.9	0
		b.	1	7	11	4	0	0	0	0	100	63.6	81.8	0
		c.	1	12	12	0	0	0	0	0	100	100	100	0
4	Fish & fish products	a.	1	8	12	0	0	0	0	0	100	100	100	0
		b.	1	6	8	4	1	0	1	0	90.9	63.6	75	12.5
		c.	1	8	9	2	1	0	0	0	90.9	81.8	85	0
5	Egg and egg products	a.	2	10	10	0	0	0	0	0	100	100	100	0
		b.	2	11	11	0	0	0	0	0	100	100	100	0
		c.	2	9	11	0	0	0	0	0	100	100	100	0
6	Fruits & Vegetables and their products	a.	1	3	14	3	5	0	0	0	54.5	72.7	68	0
		b.	1	6	11	1	1	0	1	0	87.5	87.5	90	9.1
		c.	1	13	7	0	0	0	0	0	100	100	100	0
All Categories				164	207	24	15	3	2	0	92.6	88.2	90.5	1
Protocol 1				86	130	24	12	0	2	0	90.2	80.3	85.8	1.5
Protocol 2				78	77	0	3	0	0	0	96.3	100	98.1	0

Analysis of discordant results

Total Negative Deviations (TND): 18 Total Negative Deviations obtained: 8 on artificially contaminated samples and 7 on naturally contaminated samples.

Protocol 1 (unpaired study)- 12 TND

Protocol 2 (paired study)- 3 TND

For 9 samples, presence of *Salmonella* spp. confirmed in enrichment broth.

Positive deviations: 13 positive deviations obtained: 10 on artificially contaminated samples and 3 on naturally contaminated samples.

The presence of *Salmonella* spp. confirmed in the enrichment broth of rapid method for 2 negative samples.

Analysis of discordant results:

Unpaired study design											
Category		**Type**	**Protocol**	**PD**	**ND**	**$ND_{FN(rapod)}$**	**$PA_{FP(rapid)}$**	**N+**	**TND-PD**	**AL**
2	Meat & meat products	a	1	1	2	0	0	12	1	
		b	1	2	0	0	2	10	0	
		c	1	2	2	0	1	12	1	
		Total		5	4	0	3	34	2	3
3	Milk & milk products	a1	1	0	1	0	0	3	1	
		b	1	1	0	1	0	11	0	
		c	1	0	0	1	1	12	2	
		Total		1	1	2	1	26	3	3
4	Fish & fish products	a	1	0	0	1	0	8	1	
		b	1	1	1	0	1	11	1	
		c	1	2	1	0	0	11	-1	
		Total		3	2	1	1	30	1	3
6	Fruits & Vegetables and their products	a	1	3	2	0	0	11	2	
		b	1	1	1	0	0	8	0	
		c	1	0	0	0	0	13	0	
		Total		4	3	0	0	32	2	3
Protocol 1 (unpaired)				13	10	3	5	122	5	5

Paired study design											
Category		Type	Protocol	PD	$ND_{FNrapid)}$	$(PD_{FP(rapid)})$	N+	TND-PD	AL	TND+PD	AL
1	Cereals & cereal products	a	2	0	2	0	14	2		2	
		b	2	0	0	0	21	0		0	
		c	2	0	0	0	9	0		0	
		Total		0	2	0	44	2	3	2	6
3	Milk & milk products	a2	2	0	1	0	7	1		1	
		Total		0	1	0	7	1	3	1	6
5	Egg and egg products	a	2	0	0	0	10	0		0	
		b	2	0	0	0	11	0		0	
		c	2	0	0	0	9	0		0	
		Total		0	0	0	30	0	3	0	6
Protocol 2				0	3	0	81	3	4	3	8

For the unpaired study design, the observed values for (TND-PD) are lower than the AL for each category and all the categories.

For the paired study design, the observed values for (TND-PD) and (TND+PD) are lower than the AL for each category and all the categories.

Confirmation

For 15 samples, typical colonies observed only on XLD plates; 13 samples artificially contaminated with Salmonella strain. In these cases, atypical colonies observed on BGA/Chromogenic plates.

Relative Level of Detection (RLOD)

Test on Four levels. Test Six replicates for each combination.

- Level 0
- Level providing between 0 and 50% positive results
- Level providing between 50 and 75% positive results
- Level providing 100% positive results

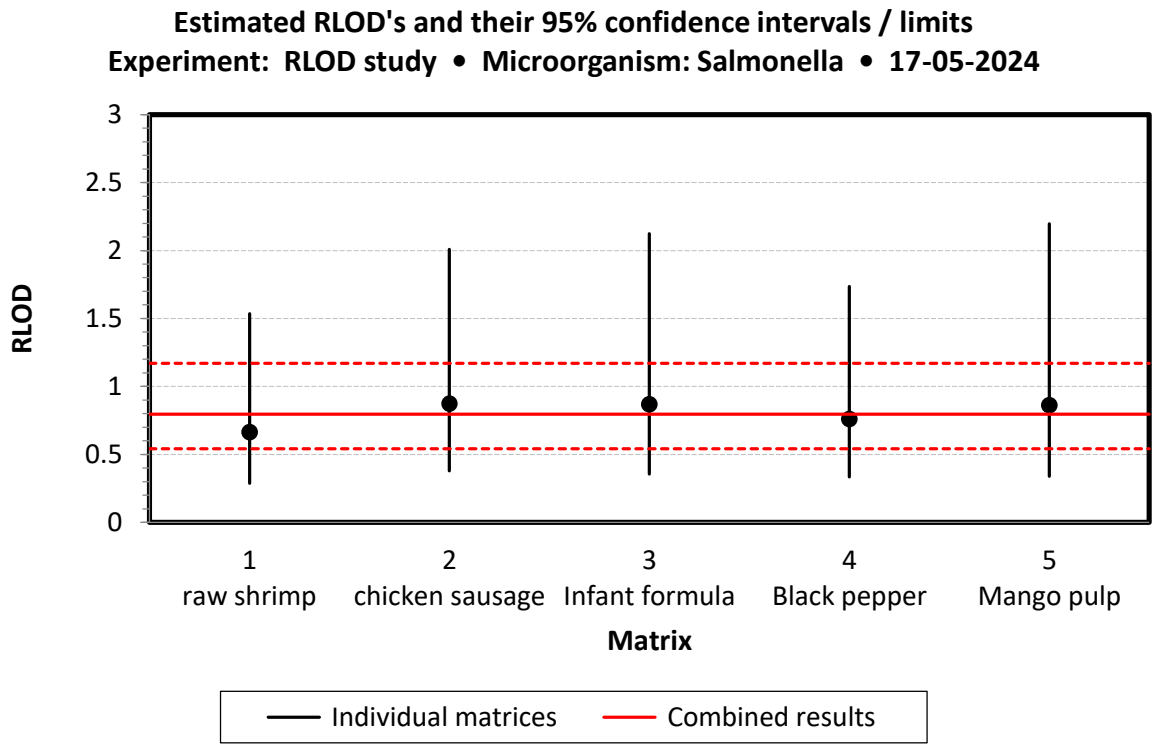
Analyze six (matrix/strain) pairs by the reference and rapid method.

Calculation and interpretation of RLOD

Perform RLOD calculations using Excel sheet available at:

https://www.wiwiss.fu-berlin.de/fachbereich/vwl/iso/ehemalige/wilrich/index.html

RLOD calculation results • RLOD study • Salmonella • 17-05-2024							
Matrix		Auxiliary values			95 % confidence limit		Test statistic [3]
No. i	Name	ln RLOD $\hat{g}$	sd(ln RLOD) $s_{\hat{g}}$	RLOD $\hat{G}$	lower G_L	upper G_U	|z|
1	raw shrimp	-0.411	0.420	0.663	0.286	1.536	0.978
2	chicken sausage	-0.136	0.417	0.873	0.379	2.008	0.327
3	Infant formula	-0.142	0.447	0.868	0.355	2.123	0.316
4	Black pepper	-0.273	0.412	0.761	0.334	1.735	0.663
5	Mango pulp	-0.148	0.467	0.862	0.339	2.197	0.316
		-0.227	0.193	0.797	0.542	1.171	1.180
3 For information on the test statistic |z| and the combined results, see the sheet 'Info'.							



Remarks: The RLOD are lower than the AL fixed at 2.5 for the unpaired study design and at 1.5 for the paired study design for all the tested matrix/strains.

Inclusivity/Exclusivity

Inclusivity

For inclusivity study, perform Salmonella strain cultures in BHI medium at 37°C. Make dilutions in order to inoculate between 10 to 100 cells/ 225ml in supplemented BPW (Protocol 1) without adding any sample. Then inoculate the broths for 20 h at 37°C ± 1°C and T perform the rapid method.

Test 40 different *Salmonella* strains using the rapid kit.

All the strains give a positive result.

Exclusivity

For exclusivity study, test the negative strains in BHI at 37°C. Dilute in order to inoculate between 10^5 cells/ ml in BPW. Inoculate the broths for 24 h at 37°C ± 1°C without adding any sample. Then perform the rapid method.

No cross reaction was observed with 15 non-target strains.

Conclusion of Method Comparison Study

- The protocol of the rapid kit shows 24 positive deviations (PD) and 15 negative deviations (ND). For unpaired study design, the observed values for ((ND+PPND)-PD) are lower that AL for each category and all the categories. For the paired study design, the observed values for ((ND+PPND+PD) are lower than the AL for each category and all the categories.
- The RLOD are lower than the AL fixed at 2.5 for the unpaired study design and at 1.5 for the paired study design
- The inclusivity and exclusivity study give the expected results for40 target strains and 15 non-target strains

INTER-LABORATORY STUDY

Study organization

Send the samples to 16 laboratories. Perform the study with cooked chicken samples contaminated with *Salmonella typhimurium* A00C060.

The targeted inoculation levels:

- 0 CFU/ 25 g
- 1-10 CFU/ 25 g
- 5-50 CFU/ 25 g

Provide 8 replicates by tested contamination level.

Strain stability

Before inoculation- Perform IS 5887: Part 3: Sec 1: 2020**/**ISO 6579-1:2017 on fresh chicken test portions (25 g). All results negative.

Sample stability- Inoculate the matrix at 500 CFU/g and 5 CFU/g. Enumerate for the high contamination level and perform detection analyses for the low contamination level. Analyze the triplicates.

Day	Reference method (research)			CFU/g (XLD)			Aerobic mesophilic flora (CFU/g)
	Sample 1	Sample 2	Sample 3	Sample 1	Sample 2	Sample 3	
Day 0	+	+	+	570	480	550	$1.4\ 10^3$
Day 1	+	+	+	600	510	460	$1.4\ 10^3$
Day 2	+	+	+	460	520	470	$1.2\ 10^3$

No evolution observed during storage at 4°C.

Contamination levels

Level	Samples	Theoretical target level (b/25 g)	True level (b/25 g sample)	Low limit / 25 g sample	High limit / 25 g sample
Level 0	3 – 9 – 11 – 12 – 15 – 19 – 22 – 24	0	/	/	/
Low level	4 – 7 – 10 – 13 – 14 – 20 – 21 – 23	5	4.8	4.0	5.5
High level	1 – 2 – 5 – 6 – 8 – 16 – 17 – 18	25	30	26.1	34.6

Logistic conditions

Laboratories	Temperature measured by the probe (°C)	Temperature measured at receipt (°C)	Receipt date and time	
A	1.0	7.0	09/04/2013	*11h30*
B	1.0	4.8	09/04/2013	*11h25*
C	0.5	10.0	09/04/2013	*08h00*
D	1.0	2.6	09/04/2013	*10h00*
E	1.0	5.7	09/04/2013	*09h30*
F	1.0	2.5	09/04/2013	*10h15*
G	1.5	4.7	09/04/2013	*09h00*
H	0.5	6.0	09/04/2013	*11h30*
I	*The Lab didn't realize the analyses*			
J	1.5	7.5	09/04/2013	*08h15*
K	2.5	7.2	09/04/2013	*12h15*
L	3.5	3.4	10/04/2013	*11h00*
M	1.5	7.0	09/04/2013	*13h30*
N	3.0	3.1	09/04/2013	*14h00*
O	1.0	2.2	09/04/2013	*09h00*
P	/	5.5	09/04/2013	*11h45*

No problem encountered during transport or receipt for 16 labs. However, lab I opts out of the ring trial.

Result analysis

Expert lab results

Level	Reference method	Alternative method
L0	0/8	0/8
L1	8/8	8/8
L2	8/8	8/8

Results of collaborative laboratories

Enumeration of aerobic mesophilic bacteria- varied from 3.8×10^2 to 2.1×10^3 CFU/g

Salmonella detection- 15 labs participated in the study.

Lab E encounters some problem with incubation temperature. Lab J obtains 2 positive results at Level 0. Exclude these labs for interpretation.

Table- Positive results by reference method

Laboratory	Contamination level		
	L0	L1	L2
A	0	8	8
B	0	8	8
C	0	8	8
D	1	8	8
F	0	7	8
G	0	8	8
H	0	8	8
K	0	8	8
L	1	8	8
M	0	6	8
N	0	8	8
O	1	8	8
P	0	8	8
Total	**$P_0 = 3$**	**$P_1 = 101$**	**$P_2 = 104$**

Table-Positive results (before and after confirmation) by rapid kit

Laboratory	Contamination level					
	L0		L1		L2	
	Before confirmation	*After confirmation*	*Before confirmation*	*After confirmation*	*Before confirmation*	*After confirmation*
A	1	0	8	8	8	8
B	0	0	8	8	8	8
C	0	0	8	8	8	8
D	0	0	8	8	8	8
F	0	0	8	8	8	8
G	0	0	8	8	8	8
H	0	0	8	8	8	8
K	0	0	8	8	8	8
L	0	0	8	8	8	8
M	1	0	8	8	8	8
N	0	0	8	8	8	8
O	1	0	8	8	8	8
P	0	0	8	8	8	8
Total	**$P_0 = 3$**	**$CP_0 = 0$**	**$P_1 = 104$**	**$CP_1 = 104$**	**$P_2 = 104$**	**$CP_2 = 104$**

Calculation and interpretation

Calculation of Specificity percentage (SP)- The percentage specificities using the data after confirmation, based on the results of level L0 are the following:

Specificity for the reference method	$SP_{ref} = \left(1 - \left(\frac{P_0}{N\_}\right)\right) x\ 100\ \% =$	97 %
Specificity for the alternative method	$SP_{alt} = \left(1 - \left(\frac{CP_0}{N\_}\right)\right) x\ 100\ \% =$	100 %

N: number of all L0 tests

P_0 = total number of false-positive results obtained with the blank samples before confirmation

CP_0 = total number of false-positive results obtained with the blank samples

Calculation of Sensitivity (SE_{rapid}), sensitivity for reference method (SE_{ref}), relative trueness (RT) and false positive ratio for the rapid kit (FPR)

Fractional positive results obtained for low inoculation level (L1), which are retained for calculation.

Response	Reference method positive (R+)	Reference method negative (R-)
Rapid kit/method positive (A+)	+/+	- /+

	Positive Agreement (PA) PA=101	Positive Deviation (PD) PD=3
Rapid kit/method negative (A-)	+/- Total Negative Deviation (TND) TND=0	-/- Total Negative Agreement (TNA) TNA=0

Sensitivity of the rapid kit/method	$\mathrm{SE_{rapid}} = \frac{(\mathrm{PA}+\mathrm{PD})}{(\mathrm{PA}+\mathrm{TND}+\mathrm{PD})} \times 100\%$	100.0%
Sensitivity of the reference method	$\mathrm{SE_{ref}} = \frac{(\mathrm{PA}+\mathrm{TND})}{(\mathrm{PA}+\mathrm{TND}+\mathrm{PD})} \times 100\%$	97.1%
Relative Trueness	$\mathrm{RT} = \frac{(\mathrm{PA}+\mathrm{TNA})}{\mathrm{N}} \times 100\%$	97.1%
False Positive Ratio for the rapid kit/method	$\mathrm{FPR} = \frac{PA_{FP(rapid)} + PD_{FP(rapid)}}{\mathrm{TNA}} \times 100\%$	0%
False Negative Ratio	$\mathrm{FNR} = \frac{NA_{FP(rapid)} + ND_{FN(rapid)}}{\mathrm{PA+TNA+PD}}$	0%

Three positive deviations observed for Level 1. Fractional recovery is observed at this level.

N_x	104
(p+)ref	0.97
(p+)alt	1.00
AL= $(TND-PD)_{max}$	3.00
TND-PD	-3
Conclusion	TND-PD<AL

The requirements of ISO 16140-2 are fulfilled as (ND-PD) is below AL.

Evaluation of RLOD between laboratories

Calculate the RLOD using Excel sheet available at

https://www.wiwiss.fu-berlin.de/fachbereich/vwl/iso/ehemalige/wilrich/index.html

RLOD	RLODL	RLODV	b=ln(RLOD)	sd(b)	z Test statistic	p.value
0.644	0.332	1.251	-0.440	0.332	1.326	1.815

.

Conclusion of Inter-laboratory study (ILS):

The data and interpretation comply with **Guidelines for Validation of Microbiological Methods/Kits** requirements. The Salmonella rapid kit is considered comparable performance to the reference method IS 5887: Part 3: Sec 1 : 2020/ISO 6579-1:2017 for detection of Salmonella in the broad range of foods

EXAMPLE 2: QUANTITATIVE METHOD

A rapid kit for enumeration of *Staphylococcus aureus* (coagulase-positive) in five major food categories is submitted for validation.

Reference method: IS 5887: Part 8: Sec 1:2023/ ISO 6888-1:2021 Microbiology of food and animal feeding stuffs- Horizontal method for of coagulase-positive staphylococci (*Staphylococcus aureus* and other species) - Part 1: Technique using Baird-Parker agar medium

Rapid Method- Method to be validated

Food Categories included:

- Fish & Fishery products
- Meat &Meat products
- Milk & Milk products
- Spices & Herbs
- Fruits and Vegetables and their products.

1. Method comparison study

A. Relative trueness

Analyze the samples by the reference and the rapid kit methods in order to have at least 15 interpretable results per food category, and at least 5 interpretable results per tested food product/type by the two methods.

Test samples: The tables below show a total of 5 categories were included in this validation study. Test a minimum of 15 items for each food category by both the reference method and the rapid method, with a minimum of 15 interpretable results per category. Each category made up of 3 products/types, with at least 5 items representative for each product/type.

Of the 77 samples taken (Table 3A) for study, 67 samples were artificially contaminated; 10 samples were naturally contaminated. (Artificial spiking- Annex)

Table 3A.- Samples of Food categories and products				
Category	**Product/Types**		**No. of samples analysed**	**No. of samples with interpretable results**
Fish & fishery products	a	Raw shrimp	5	5
	b	Fermented fish products	5	5
	c	Dry fish	5	5
Meat & Meat products	a	Fresh chicken	5	5
	b	fermented meat	6	6
	c	Sausage	5	5
Milk & milk products	a	Raw milk	5	5
	b	Milk powder	5	5
	c	Cottage cheese	5	5
Spices & Herbs	a	Chilli powder	5	5
	b	Cardamom	5	5
	c	Black pepper	6	6
Fruits and vegetables and their products	a	Pickle,	5	5
	b	Fruit juice with natural color	5	5
	c	Jam	5	5
TOTAL			77	77

Protocols applied during the validation study:

- **Sample quantity-** 10 g each. Perform the reference method and rapid methods with the same sample quantity. The study is therefore a paired study design.
- Make 1 in 10 initial dilutions of the product sample containing the target organism with an appropriate diluent and homogenize. Make appropriate serial 10-fold dilutions, and analyze all relevant dilutions using the reference method and rapid method.
- ***Reference method***: Plates were incubated at 37±1ºC for a total of 48±4h. In all cases, use the minimum incubation times.
- ***Rapid Method***: As per its protocol

Test results calculation and interpretation of relative trueness study:

Analyze the obtained data using the scatter plot. The graphs are provided with the line of identity (y = x). Figures 3.1 – 3.5 shows the scatter plots for the individual categories and Figure 3.6 for all categories.

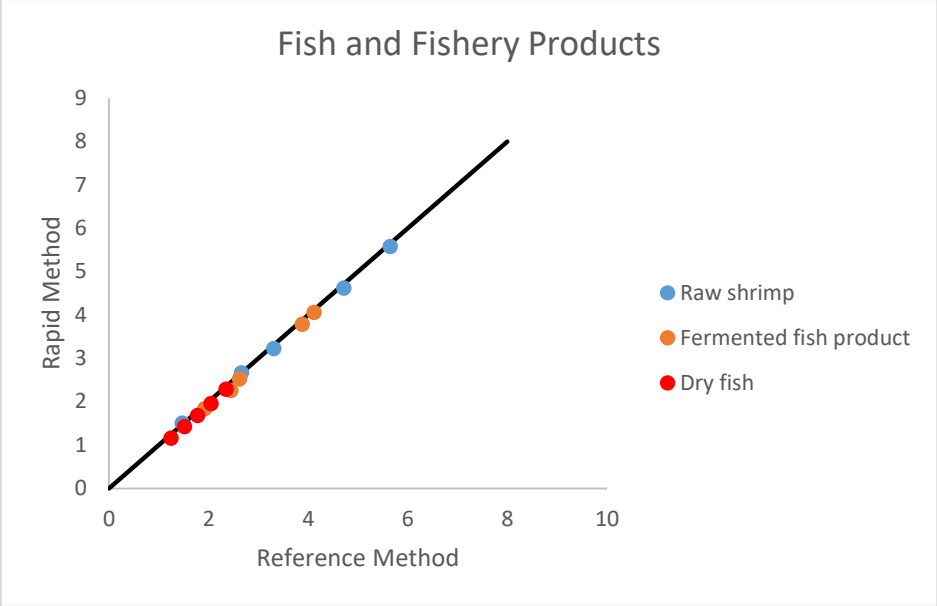


Figure 3.2 - Scatter plot of the reference method versus rapid kit results for Fish & Fishery Products

Figure 3.3- Scatter plot of the reference method versus rapid kit results for Meat & Meat products

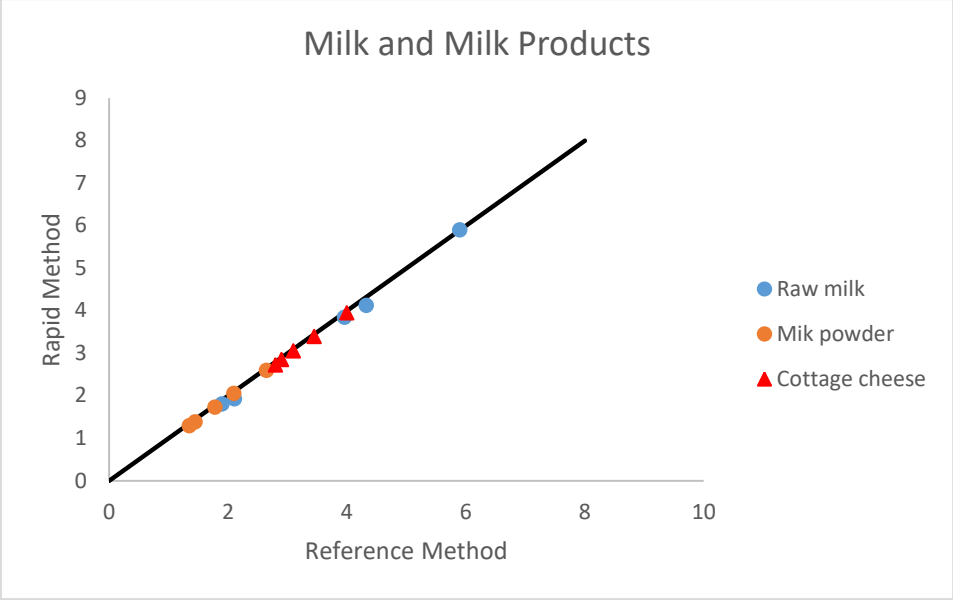


Figure 3.4- Scatter plot of the reference method versus rapid kit results for Milk & Milk products

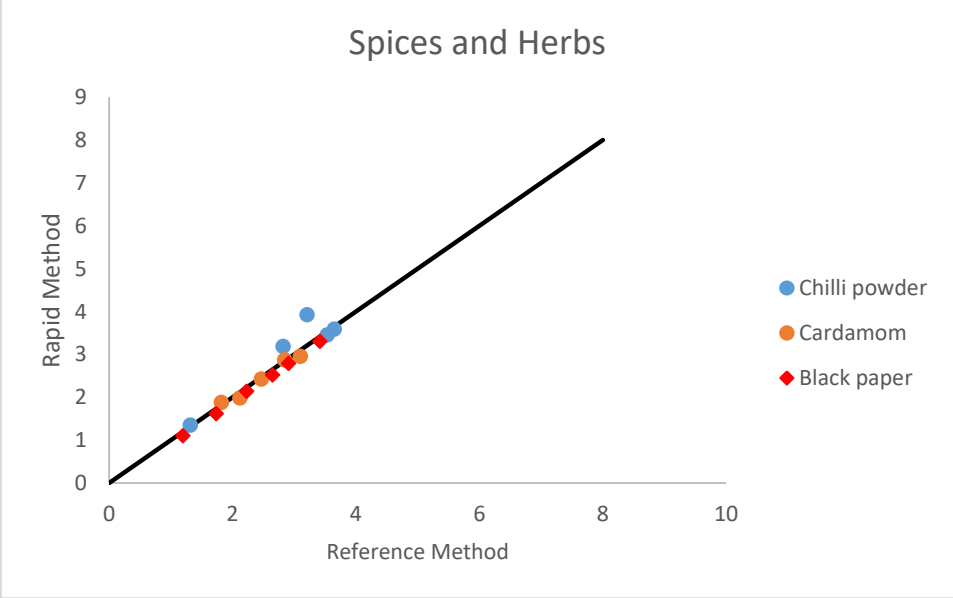


Figure 3.5- Scatter plot of the reference method versus rapid kit results for Spices & Herbs

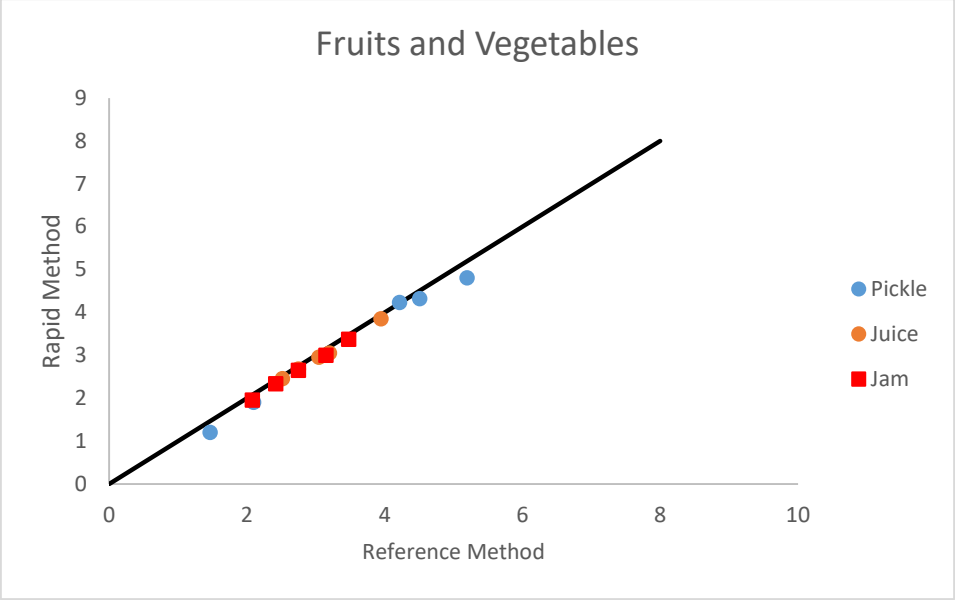


Figure 3.6- Scatter plot of the reference method versus rapid kit results for Fruits and vegetables and their products

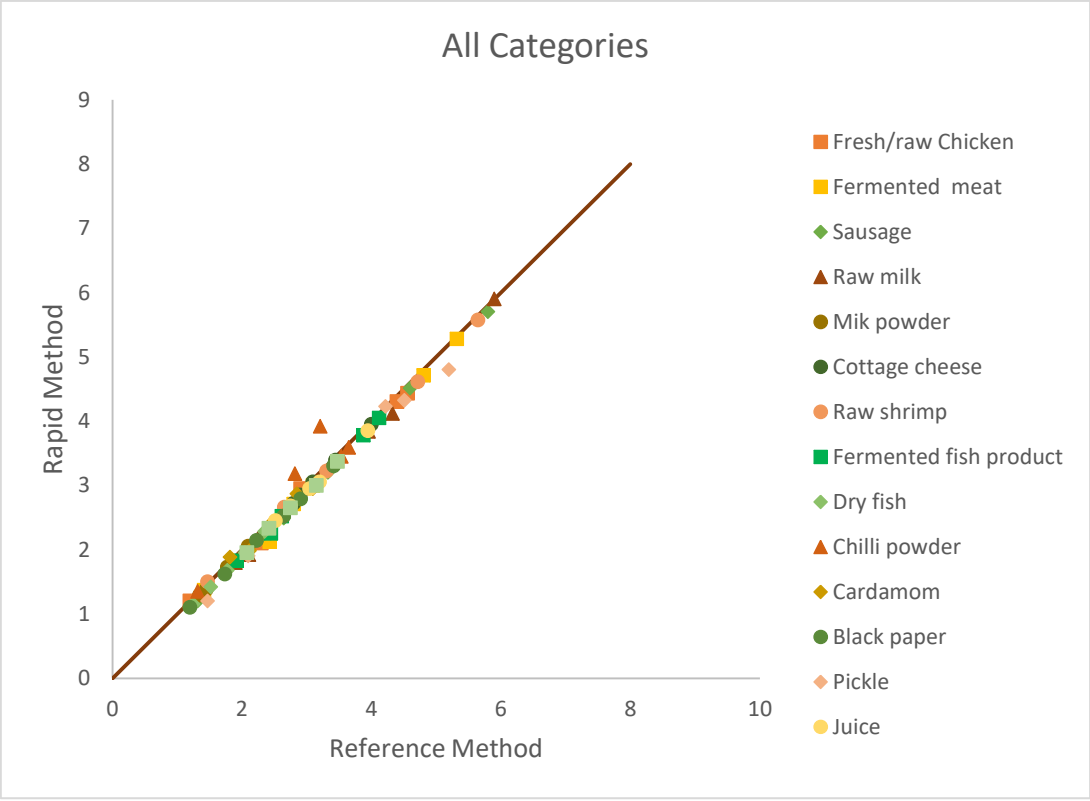


Figure 3.7 - Scatter plot of the reference method versus rapid kit results for all categories

Interpret the results of the scatter plot based on visual observation on the amount of bias and extreme results. The data appears acceptable on the whole but there is some evidence of a positive bias for rapid method in two chili powder samples (Figure 3.2 and positive bias for the reference method for fruit and vegetable product i.e. pickle, this can be seen from the individual product and (Figure 3.7) and from the all categories (Figure 3.8).

Table 3B - Summary of the calculated values per category					
Category	**n**	$\bar{D}$	S_D	**95% lower limit**	**95% upper limit**
Fish & fishery products	15	0.0853	0.079	-0.014	0.184
Meat & meat products	16	0.105	0.078	-0.051	0.261
Milk & Milk products	15	0.08	0.056	-0.028	0.188
Fruits and vegetables and their products	15	0.141	0.097	-0.136	0.331
Spices & Herbs	16	-0.0006	0.227	-0.446	0.444
All Categories	77	0.081	0.128	-0.17	0.332

$\bar{D}$ =Average Difference; S_D= Standard deviation of differences; n= number of samples

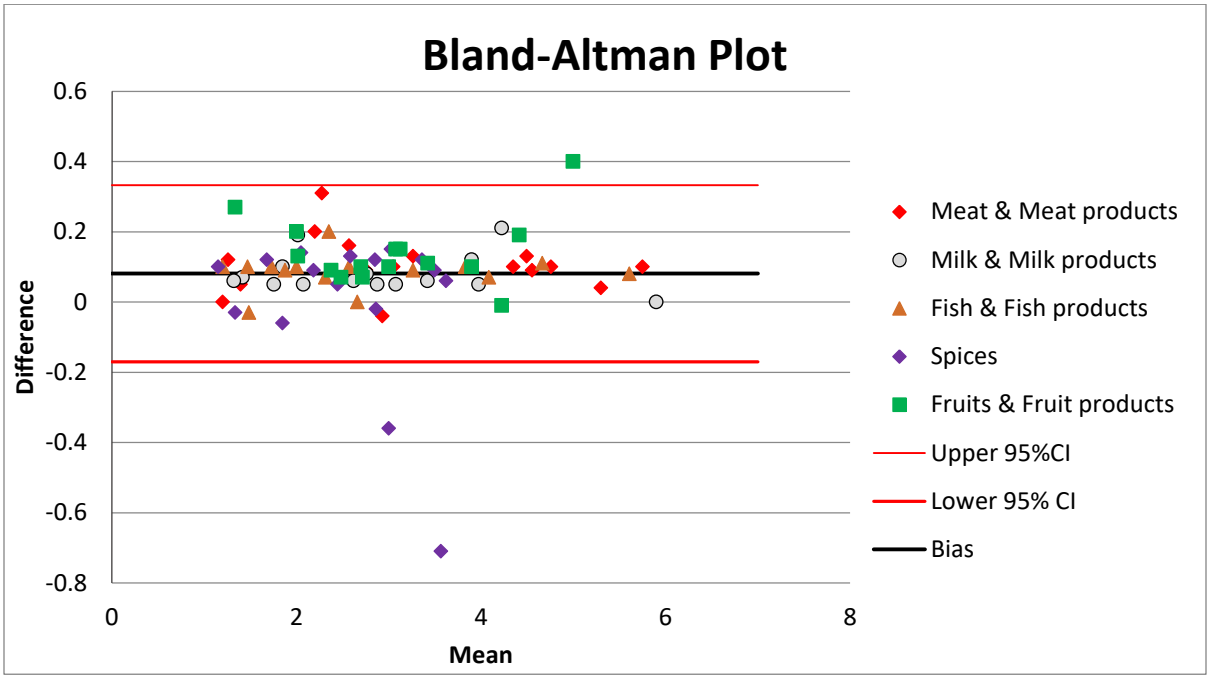


Figure 3.8 – Bland-Altman difference plot for all the samples

For this data set there are 3 in 77 data values which lie outside the CLs (All categories plot). This is within the acceptable limits of expectations i.e. less than one in 20. The three points which were outside of the CLs were shown below in Table 3. There were no identifiable trends in these data and they covered 2 different food categories and 2 different inoculated strains. The two of these points are

concerned with samples which have been inoculated with heat stressed strains immediately prior to analysis

Table 3C - Data which are outside of the accepted limits					
Sample code	**Major Food Category**	**Food Products**	**Strain no of *S. aureus***	**Spiking protocol**	**Difference log cfu/g (rapid – reference)**
32	Fruit and Vegetable product	Pickle	1223	55°C/5mim heating	0.4
14	Spices and herbs	Chilli Powder	1238	55°C/5mim heating	-0.36
15	Spices & Herbs	Chilli Powder	1223	Chill/2 days	- 0.71

Conclusion (RT study) - The relative trueness of the rapid kit for S. *aureus* (coagulase-positive) is satisfactory.

B. Accuracy profile

Conduct using artificially contaminated samples, using one food product per food category.

Categories, products and strains:

- Test a single product of each food category using 6 samples per product.
- Spike bulk food samples with the test stain as per following protocol: 2 samples at a low level (500 CFU/g), 2 at intermediate level (10000 CFU/g) and 2 at a high level (100000 CFU/g)).
- For each sample, test 5 replicates (5 different test portions) with both reference and rapid methods.
- A total of 30 samples per food product per category will be analysed by reference as well as rapid method (Table 3D).

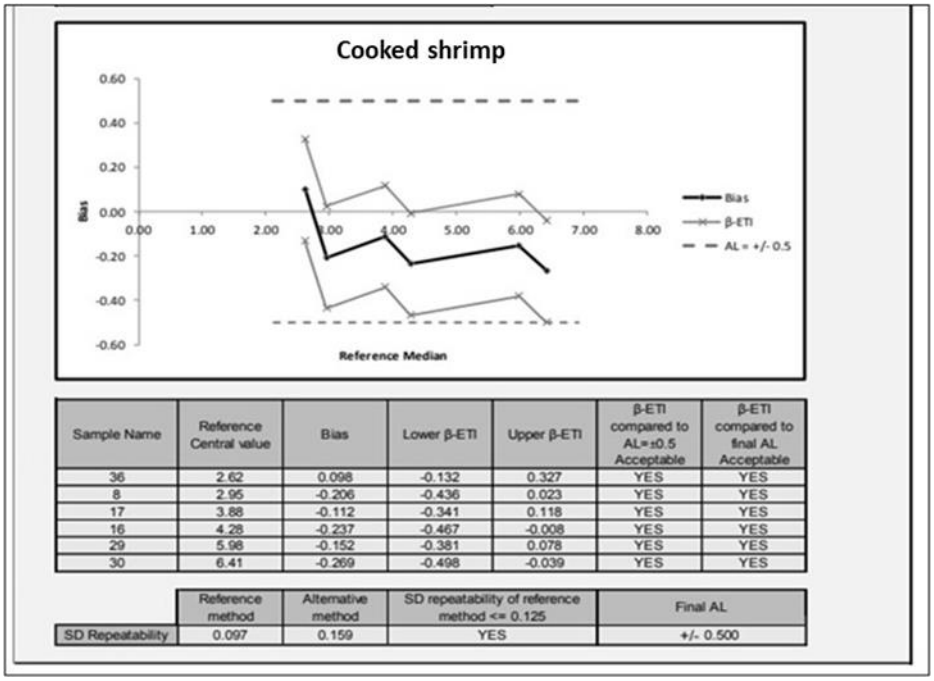
Table 3D- Details of food category and level of contamination						
Category	**Type**	**Strain**	**Product**	**Level (CFU/g)**	**Test portion for reference method**	**Test portion for rapid method**
Fish & fishery products	Major	*S. aureus* ATCC6538	Cooked shrimp sample 1	Low – 500	5	5
				Medium - 10000	5	5
				High – 100000	5	5
			Cooked	Low – 500	5	5

			shrimp Sample 2			
				Medium - 10000	5	5
				High – 100000	5	5
Meat & Meat products	Major	*S. aureus* ATCC6538	Fresh meat sample 1	Low – 500	5	5
				Medium - 10000	5	5
				High – 100000	5	5
			Fresh meat sample 2	Low – 500	5	5
				Medium - 10000	5	5
				High – 100000	5	5
Milk & Milk products	Major	*S. aureus* ATCC6538	Milk Powder sample 1	Low – 500	5	5
				Medium - 10000	5	5
				High – 100000	5	5
			Milk Powder sample 2	Low – 500	5	5
				Medium - 10000	5	5
				High – 100000	5	5
Spices & Herbs	Major	*S. aureus* ATCC6538	Black pepper sample 1	Low – 500	5	5
				Medium - 10000	5	5
				High – 100000	5	5
			Black Pepper Lot 2	Low – 500	5	5

				Medium - 10000	5	5
				High – 100000	5	5
Fruits and vegetables and their products	Major	*S. aureus* ATCC6538	Mixed Fruit Jam sample 1	Low – 500	5	5
				Medium - 10000	5	5
				High – 100000	5	5
			Mixed Fruit Jam sample 2	Low – 500	5	5
				Medium - 10000	5	5
				High – 100000	5	5
			Total		150	150

Calculations and interpretation of accuracy profile study

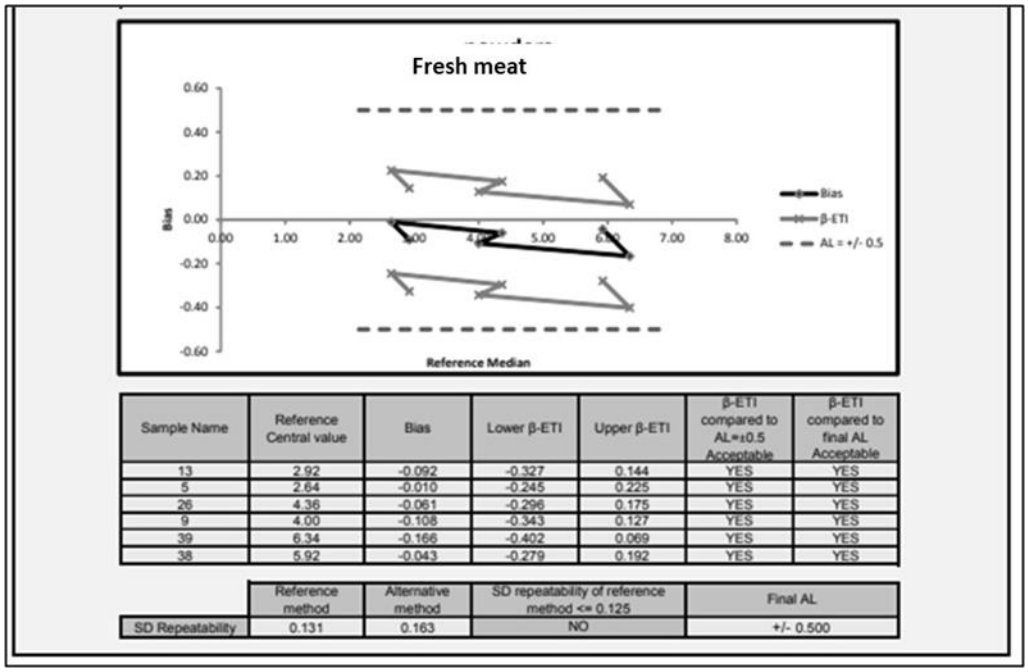
The statistical results and the accuracy profiles are provided in Figures 3.9 -3.13. Perform the calculations using the AP Calculation Tool MCS available on http://standards.iso.org/iso/16140.



Sample Name	Reference Central value	Bias	Lower β-ETI	Upper β-ETI	β-ETI compared to AL=±0.5 Acceptable	β-ETI compared to final AL Acceptable
36	2.62	0.098	-0.132	0.327	YES	YES
8	2.95	-0.206	-0.436	0.023	YES	YES
17	3.88	-0.112	-0.341	0.118	YES	YES
16	4.28	-0.237	-0.467	-0.008	YES	YES
29	5.98	-0.152	-0.381	0.078	YES	YES
30	6.41	-0.269	-0.498	-0.039	YES	YES

	Reference method	Alternative method	SD repeatability of reference method <= 0.125	Final AL
SD Repeatability	0.097	0.159	YES	+/- 0.500

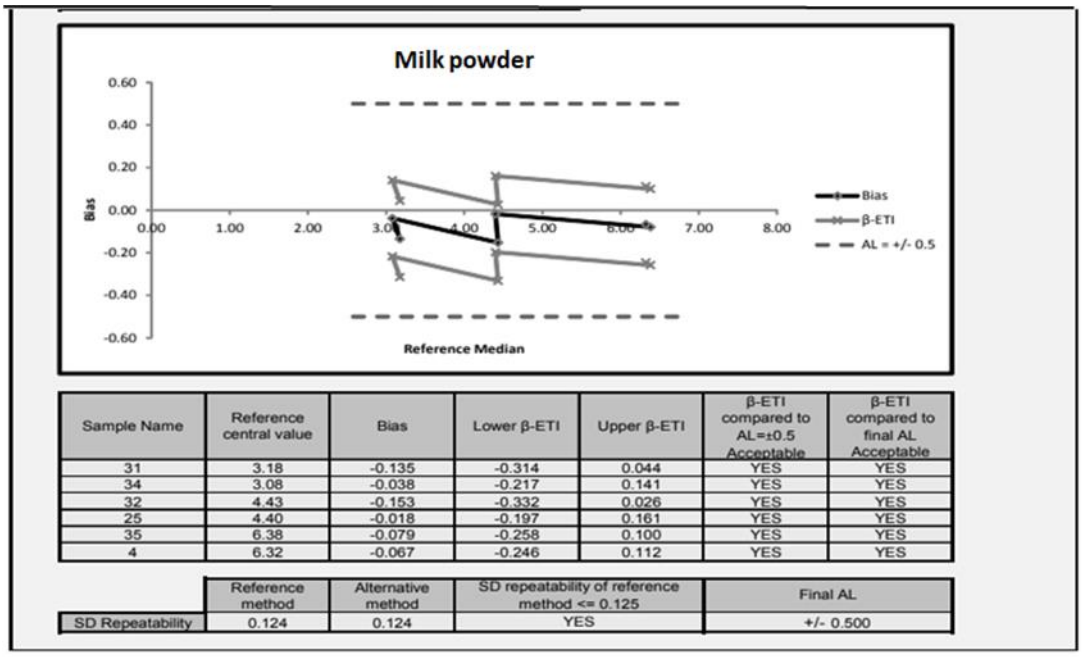
Figure 3.9 – Accuracy Profile for Category: Fish and Fishery products (Cooked shrimp)



Sample Name	Reference Central value	Bias	Lower β-ETI	Upper β-ETI	β-ETI compared to AL=±0.5 Acceptable	β-ETI compared to final AL Acceptable
13	2.92	-0.092	-0.327	0.144	YES	YES
5	2.64	-0.010	-0.245	0.225	YES	YES
26	4.36	-0.061	-0.296	0.175	YES	YES
9	4.00	-0.108	-0.343	0.127	YES	YES
39	6.34	-0.166	-0.402	0.069	YES	YES
38	5.92	-0.043	-0.279	0.192	YES	YES

	Reference method	Alternative method	SD repeatability of reference method <= 0.125	Final AL
SD Repeatability	0.131	0.163	NO	+/- 0.500

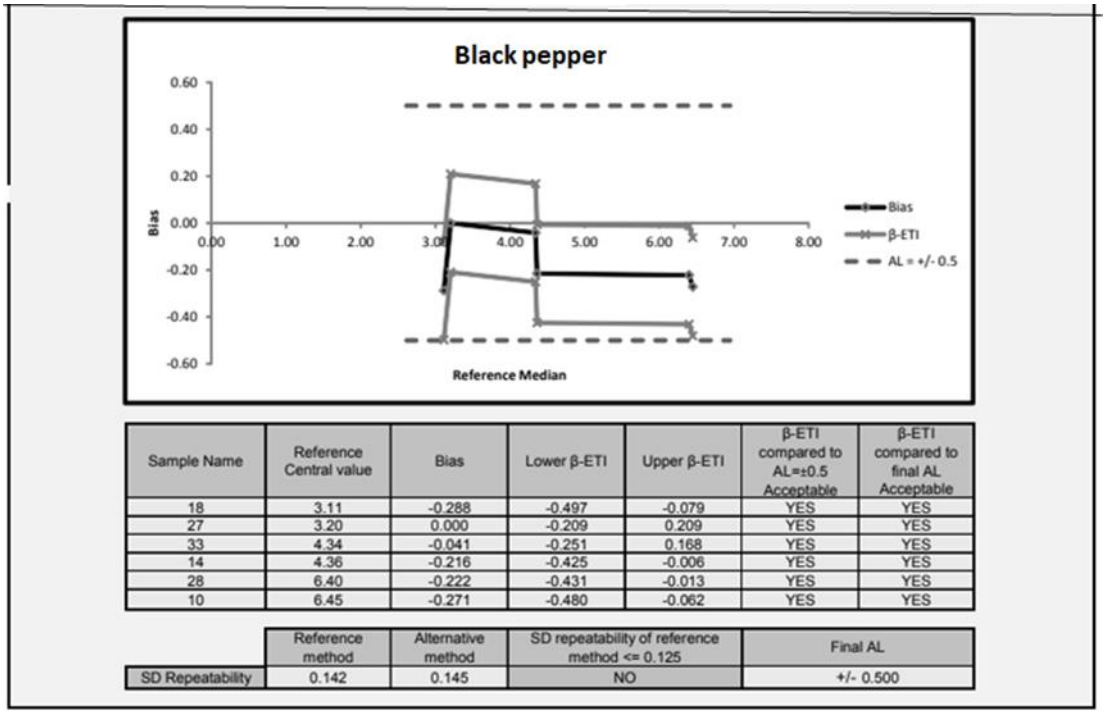
Figure 3.10 Accuracy profile for Category: Meat and Meat Products (Fresh meat samples)



Sample Name	Reference central value	Bias	Lower β-ETI	Upper β-ETI	β-ETI compared to AL=±0.5 Acceptable	β-ETI compared to final AL Acceptable
31	3.18	-0.135	-0.314	0.044	YES	YES
34	3.08	-0.038	-0.217	0.141	YES	YES
32	4.43	-0.153	-0.332	0.026	YES	YES
25	4.40	-0.018	-0.197	0.161	YES	YES
35	6.38	-0.079	-0.258	0.100	YES	YES
4	6.32	-0.067	-0.246	0.112	YES	YES

	Reference method	Alternative method	SD repeatability of reference method <= 0.125	Final AL
SD Repeatability	0.124	0.124	YES	+/- 0.500

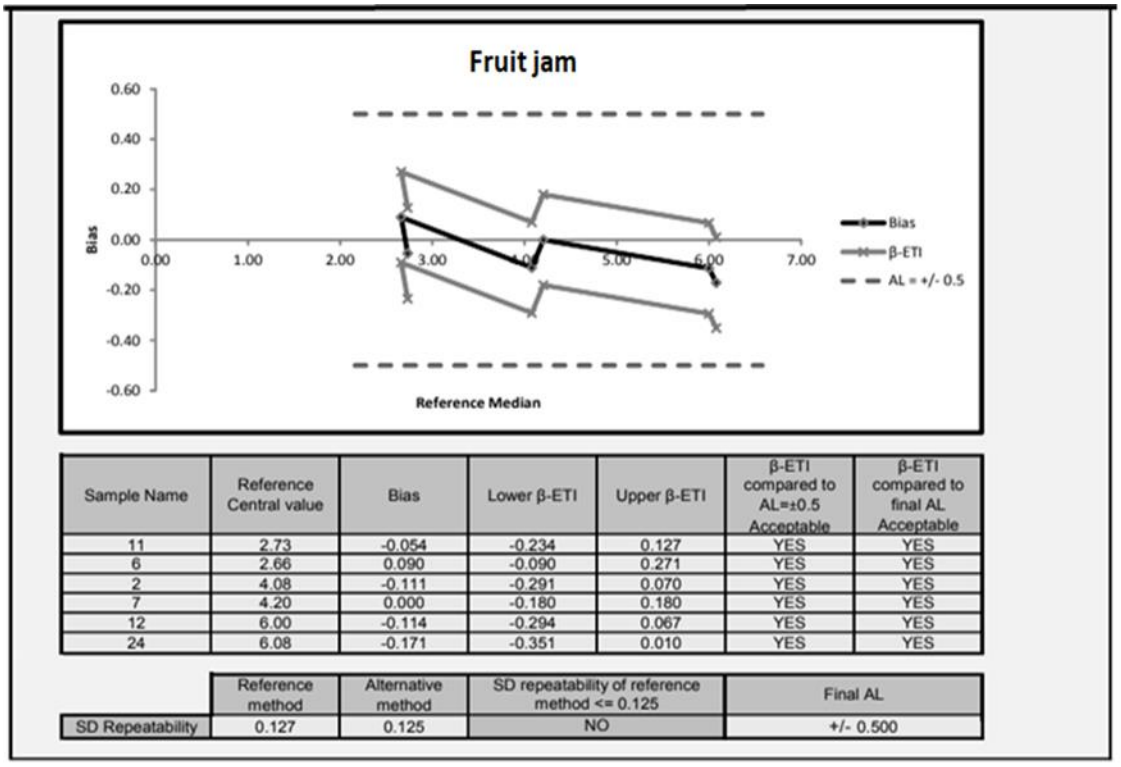
Figure 3.11 Accuracy profile for Category: Milk and Milk products (Milk powder)



Sample Name	Reference Central value	Bias	Lower β-ETI	Upper β-ETI	β-ETI compared to AL=±0.5 Acceptable	β-ETI compared to final AL Acceptable
18	3.11	-0.288	-0.497	-0.079	YES	YES
27	3.20	0.000	-0.209	0.209	YES	YES
33	4.34	-0.041	-0.251	0.168	YES	YES
14	4.36	-0.216	-0.425	-0.006	YES	YES
28	6.40	-0.222	-0.431	-0.013	YES	YES
10	6.45	-0.271	-0.480	-0.062	YES	YES

	Reference method	Alternative method	SD repeatability of reference method <= 0.125	Final AL
SD Repeatability	0.142	0.145	NO	+/- 0.500

Figure 3.12 Accuracy profile for Category: Spices and Herbs (Black pepper)



Sample Name	Reference Central value	Bias	Lower β-ETI	Upper β-ETI	β-ETI compared to AL=±0.5 Acceptable	β-ETI compared to final AL Acceptable
11	2.73	-0.054	-0.234	0.127	YES	YES
6	2.66	0.090	-0.090	0.271	YES	YES
2	4.08	-0.111	-0.291	0.070	YES	YES
7	4.20	0.000	-0.180	0.180	YES	YES
12	6.00	-0.114	-0.294	0.067	YES	YES
24	6.08	-0.171	-0.351	0.010	YES	YES

	Reference method	Alternative method	SD repeatability of reference method <= 0.125	Final AL
SD Repeatability	0.127	0.125	NO	+/- 0.500

Figure 3.13 Accuracy profile for Category: Fruit and Vegetable Products (Fruit jam)

If any of the upper or lower limits for the six samples exceeds the 0.5 log Acceptability Limits (ALs) and the standard deviation, S_{ref}> 0.125, then an additional evaluation procedure is followed:

New ALs are calculated as a function of the standard deviation: ALs = 4_ S_{ref}. If for all i in the accuracy profile Ui ≤ ALs and Li _ −ALs , the rapid kit is accepted as being equivalent to the reference method for the given combination category and type.

C. Inclusivity / exclusivity

Grow cultures (20 target strains and 15 non-target strains) as per appropriate conditions, then make serial decimal dilutions, and enumerate target and non-target strains by the rapid kit, the reference method and a non-selective agar (TSA).

Results - Of the 20 inclusivity strains tested, 20 strains are detected using both methods and 2 strains gave typical colonies on both media but did not confirm using the coagulase test. Of the 15 exclusivity strains tested, none are detected by the alternate method and 2 are detected by the reference method.

Conclusion

Overall, the conclusions for the Method Comparison are:

- The rapid kit for enumeration of *S. aureus* in foods method showed satisfactory trueness
- The rapid kit for enumeration of *S. aureus* in foods method showed satisfactory accuracy profile.
- The rapid kit for enumeration of *S. aureus* in foods in foods method was found to be specific and selective

2. INTER LABORATORY STUDY

Study organisation

- **Collaborators** –Samples are sent to 6 organizations with 2 collaborators for each organization involved in the study making a total of 12 collaborators.
- **Matrix** –Inoculate milk powder with *S. aureus* CRA 1292 isolated from milk.
- **Sample preparation** –Spike the samples (10 g) with the desired level of organism as shown below in Table 3E. Perform a homogeneity test first and then stability test of the inoculated samples during 72 hours chilled transportation. Please refer IS17385/ISO 22117 for homogeneity test and stability test.

Table 3E: Contamination levels

Contamination level	Sample code Set 1	Sample code Set 2
Uninoculated	04	8
Low (10^2 cfu/g)	01	13
Low (10^2 cfu/g)	05	14
Medium (10^3 cfu/g)	02	10

Medium (10^3 cfu/g)	06	12
High (10^4 cfu/g)	03	09
High (10^4 cfu/g)	07	11

Logistic conditions

Note logistic conditions

Table 3F: Sample logistics conditions			
Collaborators	**Date received**	**Temperature of sample upon receipt (^{0}C)**	**Average storage temperature over entire transport period (^{0}C)**
1	06/02/2024	2.8	2.5
2	06/02/2024	3.6	3.0
3	06/02/2024	4.4	3.5
4	06/02/2024	5.0	3.6
5	06/02/2024	2.5	2.4
6	06/02/2024	7.2	3.9
7	06/02/2024	4.9	3.7
8	06/02/2024	3.8	3.2
9	06/02/2024	3.6	3.0
10	06/02/2024	5.4	4.0
11	06/02/2024	6.0	3.8
12	06/02/2024	2.8	2.5
Organizing laboratory	06/02/2024	--	--

Calculation and summary of data -

Results obtained by the collaborative laboratories -

The calculations and interpretations were performed using the Excel spreadsheet (http://standards.iso.org/iso/16140).

Table 3G: Summary of the results of the inter-laboratory study per analyte level					
Collaborator		Reference method (cfu/g)		Rapid kit cfu/g)	
		Duplicate 1	Duplicate 2	Duplicate 1	Duplicate 2
1	low	100	90	110	90
2	low	110	100	100	130
3	low	90	100	120	130

Table 3G: Summary of the results of the inter-laboratory study per analyte level					
Collaborator		Reference method (cfu/g)		Rapid kit cfu/g)	
		Duplicate 1	Duplicate 2	Duplicate 1	Duplicate 2
4	low	110	80	140	100
5	low	140	130	90	80
6	low	180	160	100	90
7	low	120	130	110	150
8	low	100	120	140	160
9	low	120	90	120	100
10	low	80	70	70	90
11	low	130	100	100	120
12	low	110	80	100	90
1	medium	800	920	930	1700
2	medium	850	950	1100	1100
3	medium	510	680	730	530
4	medium	810	810	1000	530
5	medium	860	860	830	730
6	medium	680	800	630	730
7	medium	630	790	1300	580
8	medium	900	800	920	1000
9	medium	730	900	900	1100
10	medium	1100	890	900	830
11	medium	800	820	830	1100
12	medium	840	930	900	1000
1	high	8900	9000	12000	9000
2	high	10000	8300	12000	11000
3	high	7000	7100	9100	12000
4	high	7800	8900	12000	11000
5	high	8500	9700	11000	7800
6	high	8700	8500	11000	12000
7	high	7000	5900	9100	21000
8	high	7800	9000	21000	25000
9	high	8600	8800	12000	15000
10	high	7000	7400	9100	7800
11	high	7800	9000	12000	11000

Table 3G: Summary of the results of the inter-laboratory study per analyte level					
Collaborator		Reference method (cfu/g)		Rapid kit cfu/g)	
		Duplicate 1	Duplicate 2	Duplicate 1	Duplicate 2
12	high	9100	10000	10000	12000
1	blank	<10		<10	
2	blank	<10		<10	
3	blank	<10		<10	
4	blank	<10		<10	
5	blank	<10		<10	
6	blank	<10		<10	
7	blank	<10		<10	
8	blank	<10		<10	
9	blank	<10		<10	
10	blank	<10		<10	
11	blank	<10		<10	
12	blank	<10		<10	

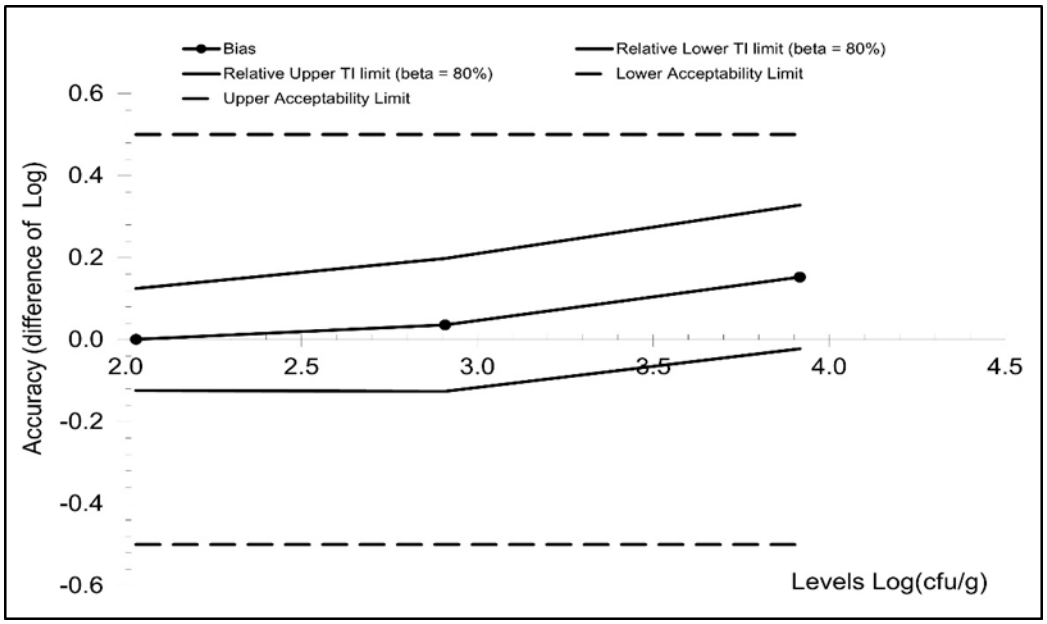
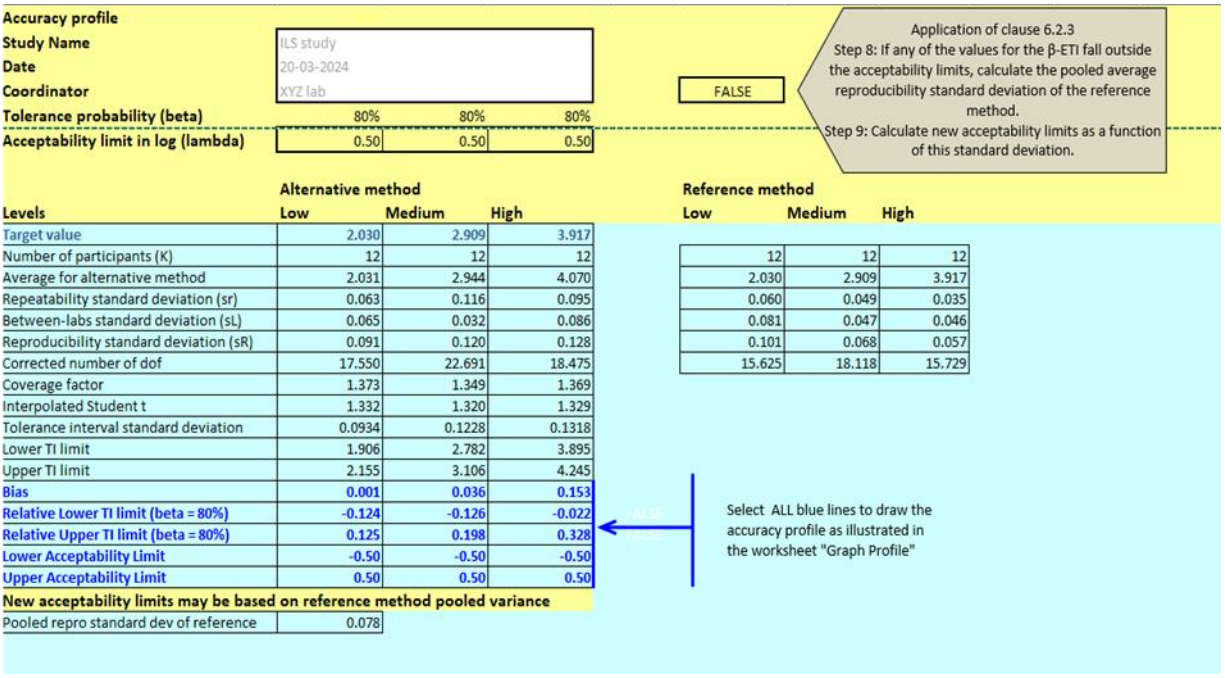


Figure 3.14. Accuracy profile of MC Media Pad SA from the ILS

Table 3H. Statistical analysis of the ILS data according to the ISO spreadsheet



Accuracy profile						
Study Name	ILS study					
Date	20-03-2024					
Coordinator	XYZ lab			FALSE		
Tolerance probability (beta)	80%	80%	80%			
Acceptability limit in log (lambda)	0.50	0.50	0.50			

Application of clause 6.2.3
Step 8: If any of the values for the β-ETI fall outside the acceptability limits, calculate the pooled average reproducibility standard deviation of the reference method.
Step 9: Calculate new acceptability limits as a function of this standard deviation.

	Alternative method			Reference method		
Levels	Low	Medium	High	Low	Medium	High
Target value	2.030	2.909	3.917			
Number of participants (K)	12	12	12	12	12	12
Average for alternative method	2.031	2.944	4.070	2.030	2.909	3.917
Repeatability standard deviation (sr)	0.063	0.116	0.095	0.060	0.049	0.035
Between-labs standard deviation (sL)	0.065	0.032	0.086	0.081	0.047	0.046
Reproducibility standard deviation (sR)	0.091	0.120	0.128	0.101	0.068	0.057
Corrected number of dof	17.550	22.691	18.475	15.625	18.118	15.729
Coverage factor	1.373	1.349	1.369			
Interpolated Student t	1.332	1.320	1.329			
Tolerance interval standard deviation	0.0934	0.1228	0.1318			
Lower TI limit	1.906	2.782	3.895			
Upper TI limit	2.155	3.106	4.245			
Bias	0.001	0.036	0.153			
Relative Lower TI limit (beta = 80%)	-0.124	-0.126	-0.022			
Relative Upper TI limit (beta = 80%)	0.125	0.198	0.328			
Lower Acceptability Limit	-0.50	-0.50	-0.50			
Upper Acceptability Limit	0.50	0.50	0.50			
New acceptability limits may be based on reference method pooled variance						
Pooled repro standard dev of reference	0.078					

Select ALL blue lines to draw the accuracy profile as illustrated in the worksheet "Graph Profile"

Conclusion of Inter laboratory study

The data and interpretation comply with **Guidelines for Validation of Microbiological Methods/Kits** requirements. The rapid kit for enumeration of *S. aureus* (coagulase-positive *Staphylococci*) showed comparable performance to the reference method IS 5887: Part 8: Sec 1:2023/ISO 6888-1 for enumeration of *S aureus* (coagulase-positive) in the broad range of foods.

Chapter-4

GUIDELINES FOR VALIDATION OF DNA-BASED METHODS/KITS

1.0 INTRODUCTION

The development of advanced analytical methods to ensure food safety and quality are more relevant now because of rapid changes in the quantity, diversity, and mobility of food. The underpinning technique of polymerase chain reaction (PCR) has transformed the field of nucleic acid analysis, due to its robustness and simplicity. Therefore, nucleic acid-based assay is among the most significant advances in food diagnostics as it provides rapid, reliable, and both qualitative and quantitative results. Technological advances in instrumentation have resulted in a wide range of PCR-based analytical approaches. Performing nucleic acid-based assays to the high standard of analytical quality can be challenging. Use of nucleic acid target sequences is a cross-cutting fundamental measurement that broadly impacts many aspects of food analysis.

2.0 PURPOSE

FSSAI approved methods/kits employed in food testing laboratories must meet the highest analytical performance standards. The aim of this document is to provide guidance on the data requirement for the validation of qualitative and quantitative nucleic acid-based testing of foods. The RAFT committee establishes these criteria for validation, by which all submitted analytical methods/kits for targeted nucleic acid sequence-based analyses in food and feed, shall be evaluated.

These criteria for DNA-based methods are consistent with several related guidelines produced by international food standards setting organizations including, but not limited to, Codex Alimentarius, the International Organization for Standardization (ISO), the European Union Reference Laboratory (EURL)for GM Food and Feed.

3.0 SCOPE

These guidelines describe the performance characteristics and the minimum performance criteria which should be considered when conducting a single-laboratory validation study for qualitative detection, identification and quantification of specific DNA sequences including those derived from modern biotechnology using either digital (dPCR) or real-time PCR (qPCR) amplification technologies.

Templates used for analysis may include, but are not limited to, genomic DNA (gDNA), chloroplast and/or mitochondrial DNA, and reverse-transcribed RNA (cDNA). Currently, this document does not

provide guidance related to the validation of methods that employ Next-Generation Sequencing (NGS) technology, or other nontargeting sequencing methods

Method validation is required for:

I. Submission of a new or original method or new kit.
II. Platform, matrix and analyte extensions.
III. Modifications to an existing method that may alter its performance specifications.

FSSAI, places an emphasis on the acceptance of methods of analysis, which have been validated through a collaborative trial conforming to an internationally accepted protocol according to ISO 5725:1994 or the AOAC/IUPAC Harmonized Protocol. In this area of nucleic acid-based testing there may be a need to adopt a formal single-laboratory validation as an interim measure in the absence of collaborative trial data. Validation data determined by a single laboratory will also help FSSAI to decide whether the method/kit in question should be validated in the framework of a collaborative study and be included into FSSAI's Manual of Methods

4.0 DEFINITIONS OF VALIDATION TOOLS

The analysis of method blanks, matrix blanks, reference materials (standards) and spikes will be used for the calculations of accuracy, bias and precision, as well as determining ruggedness, depending on the type of method being evaluated. The following general validation tools are used. Some of these items are not applicable to all the method types covered in this document.

Extraction Blank: This type of blank incorporates all the reagents and steps of the nucleic acid extraction and is processed simultaneously with the samples. Extraction controls are used to verify that the extraction reagents are free of contamination. Additionally, these controls are used to demonstrate that no cross-contamination between samples has occurred.

*Matrix Blank***:** This type of blank is a substance that closely matches the samples being analyzed with respect to matrix components. Matrix blanks are used to verify that sample matrix and equipment used does not interfere with or affect the analytical signal.

Positive Control: Any reliable source of well characterized positive sample material, containing intact target nucleic acid sequences for PCR. Reference DNA or DNA extracted from a certified reference material/reference material is generally used to demonstrate that PCR reagents are working as intended

Negative DNA Target Control Well-characterized DNA preparation material that does not contain target nucleic acid

Internal Amplification Control: Internal amplification controls should be included in the PCR assays design to ensure that PCR inhibitors are not present. Internal controls are amplified using different primer and probe sets from those used to amplify assay targets and may be based on exogenous DNA or endogenous DNA.

Matrix Spikes: Matrix effects can be assessed by spiking known amounts of analyte into a matrix of interest. Accuracy or bias and precision are calculated from these results. The data can also be used to evaluate robustness/ruggedness of the method resulting from changes in the sample matrix.

No Template Control (for PCR): This type of blank incorporates all reagents used in the PCR except the template DNA but including the internal control. It serves to verify that reagents are analyte-free, and the equipment used does not interfere with or affect the analytical signal. Instead of the template DNA, for example, a corresponding volume of nucleic acid free water is added to the reaction.

*Reference Materials and Certified Reference Materials***:** The use of known reference materials (when available and applicable) should be incorporated to assess the accuracy or bias of the method, as well as for obtaining information on interferences

Replicate Analyses: The precision of the analytical process can be evaluated using replicate analyses. The originating laboratory should assure that adequate sample replicates are used and that results from replicate measurements of each analyte are compared. PCR performed on the same DNA extraction replicate analysed in different reaction wells is shown (see Figure 4.1)

Statistics: Statistical techniques are employed to evaluate accuracy, trueness (or bias) precision, linear range, limits of detection and quantitation, and measurement uncertainty.

For additional definitions, please see Glossary.

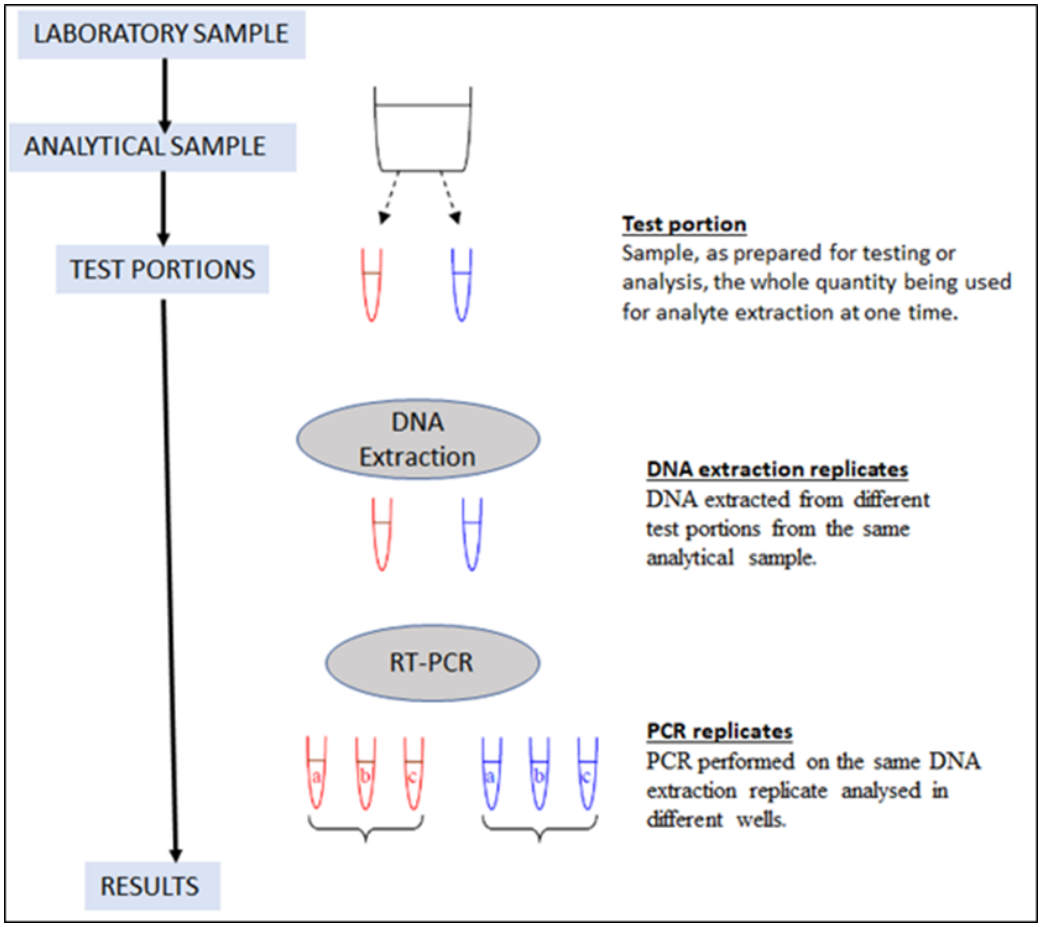


Figure 4.1: Illustration of replicates terminology. Adapted from 'Verification of analytical methods for GMO testing when implementing interlaboratory validated methods Guidance document from the European Network of GMO laboratories (ENGL). EUR 24790 EN – 2011'

4.1 Reference method

A reference method is a method by which the performance of an alternate or new method may be measured or evaluated. In some cases, an appropriate reference method is not available. However, the use of a reference method is appropriate when replacing or recommending a method specified for use in the regulatory compliance program. Consultation between the originating laboratory, and the validating laboratory/agency is essential when deciding if the use of a reference method is necessary.

5.0 CRITERIA TO BE EVALUATED BY TYPE OF METHOD

Performance characteristics of the validation protocols will vary depending on the intended use of the method, the type of method, and the degree to which it has been previously validated. A new DNA-based method evolves through several steps. After the initial development and optimization phases, the developing laboratory performs an in-house validation on the method to ensure that the method is fit for the intended purpose. Before the method can be accepted as an FSSAI/International standard, it needs to be validated in several laboratories.

New qualitative methods: These methods are developed to detect analytes (DNA/RNA) from a matrix of interest in a strictly qualitative way. The following performance characteristics should be assessed: sensitivity, specificity, false positive rate, false negative rate, limit of detection, and ruggedness/robustness.

New quantitative methods: These methods are developed to detect analytes (DNA/RNA) from a matrix of interest in a quantitative manner. The following performance characteristics should be assessed: accuracy, precision, specificity, limit of detection, limit of quantitation, linearity (efficiency), range, measurement uncertainty, robustness/ruggedness, confirmation of identity and extraction efficiency.

New identification methods: These methods are developed to unequivocally identify specific analytes (based on their nucleic acid sequences e.g., identification of a genetically modified organism (GMO). The following performance characteristics should be assessed: accuracy, precision, specificity, robustness, confirmation of identity.

Method, matrix, and platform extensions: Validating the extension of methods that have previously been validated requires a careful evaluation of the intended purpose of the extension. To implement the modified method, generally the standard or existing method is first performed. The modified method performance then is verified by comparison with that of the original method. The proposed new changes must be compared to the existing reference method.

5.1 The process of method validation

Method validation protocols confirm by examination that the requirements for a method have been fulfilled. All DNA based methods/rapid kits used by the FSSAI in support of its regulatory and

compliance roles must be validated according to these guidelines. The validation consists of different levels. This approach is based on AOAC guidelines for single-laboratory validation and collaborative studies. Key validation parameters for each level are summarized in Table 4.1. It is the responsibility of the developing laboratory to perform validations up to Multiple Laboratory Validation (MLV) and through the single independent laboratory validation levels (ILV). It is recommended that developing laboratories work with an expert in the field when determining the appropriate level of validation beyond the SLV level.

Table 4.1: Design for PCR-Based Methods for different levels of validation

S.no	Level 1 SLV	Level 2 ILV#	Level 3 MLV#
Number of labs	1	1	10-12*
**Matrix sources per matrix	≥1	≥3	≥3 if food = 1; ≥1 if food >1
Replicates if matrix source = 1			
Quantitative	≥4	≥6	≥2
Qualitative	≥6	≥9	≥6
Replicates if matrix source > 1			
Quantitative	≥2	≥2	≥2
Qualitative	≥2	≥3	≥3

*A minimum of **eight laboratories** contributing usable data for statistical analysis at the end of the study is mandatory

**See Appendix 1 for Matrix and matrix sources

Data from Level 2 (ILV) and Level 3 (MLV) will be considered for RAFT scheme approval

Level one: This is a SLV level carried out by the originating laboratory. Examples of where Level One validation would be acceptable include single-occurrence samples, and application of a method developed for a specific sequence of a matrix not previously validated in response to a real or perceived threat to food safety or public health.

Level two: This is also a complete SLV by an independent laboratory known as Independent Laboratory Validation (ILV). The developer has conducted a Level One validation and has compared the new method with an existing reference method if available and appropriate. The method/kit is validated by an independent laboratory at this level and is expected to be of widespread use, or long term, or of high public visibility.

Level three: This validation level has criteria equivalent to a full validation. It is designed to measure inter-laboratory reproducibility, so that it can be determined if the method can be successfully performed by multiple laboratories other than the developing laboratory. All methods validated to this level of scrutiny are acceptable for use in all regulatory circumstances including confirmatory analyses, routine regulatory surveillance, and compliance. For details see chapter on Guidelines for Collaborative Trials.

Table 4.1 provides guidance for the design of validation studies for the PCR amplification at all levels. Table 4.2 shows the performance parameters that need to be evaluated at all the four levels for both qualitative and quantitative methods. If the method to be validated is designed to detect a prokaryotic microorganism, also refer, and adhere to all General Guidelines for the Validation of Methods for Microbial Analysis (Chapter 2).

For approval under the RAFT scheme, it is mandatory to validate the method/kit at Level 2 and Level 3 and submit all data for review

5.3 Qualitative PCR methods

DNA-based methods that are used for the detection of a specific sequence which could be part of a mixture of related targets should allow for the unequivocal detection of that sequence, which is specific to the target organism, group or sub-set of organisms (family, genus, etc.), or transformation event in the case of genetically altered organisms. For instance, target-specific methods that are used for detection of a single transformation event

should allow for unequivocal detection, identification and/or confirmation of a nucleic acid sequence that is unique or specific to that transformation event (event or inserted DNA). For food authentication, the specific target sequence/s should uniquely define the target as required e.g., in meat, horse, pork, fish etc. For qualitative PCR methods, the basic performance characteristics are:

i. Applicability
ii. Practicability
iii. DNA Extraction Efficiency
iv. Specificity (Selectivity)
v. Sensitivity (Limit of Detection-LOD)
vi. False Negative and Positive Rates
vii. Robustness/Ruggedness

5.3.1 Applicability

Definition: The description of analytes, sample materials (matrices) and concentrations to which the method/kit can be applied. The applicability consists in demonstrating during the full validation that the method is applicable to different food matrices with the same detection capabilities (False negative rates and LOD).

Acceptance criteria: The applicability statement should provide information on the scope of the module and include all data demonstrating the fitness for purpose of the method/kit with respect to the scope.

The applicability statement should contain complete information on the scope of the method i.e., which target, which matrix and DNA amount have been tested by the manufacturer. It can be evaluated using different matrices (see Appendix A, raw/processed material, food/feed, genomic (gDNA)). Some methods that can be applied to a single raw matrix may not be necessarily applicable to complex matrices and/or processed food since the DNA may be altered. Additionally, warnings on the interference with other analytes and its inapplicability to certain matrices and conditions should be included when identified.

Applicability of the methods could be determined by confirming whether the methods may be used in the intended foods with the required performance, and it should be clearly stated. In principle the method/kit should be applicable to the matrix of concern. In the case of "general purpose" methods to identify and quantify DNA sequences in a range of food matrices, at least one extraction method applicable to each food matrix should be available. The applicant should provide test results of the whole method submitted (i.e. from sample preparation to quantification including, where available, reference materials). The analytes, matrices, and concentrations for which a method of analysis may be used should be stated clearly.

5.3.2 Practicability

Definition: The ease of operations, the feasibility and efficiency of implementation, the associated unitary costs (e.g. cost/sample) of the module.

Acceptance criteria: The module should generally be practicable in line with other applications for a similar purpose. More specifically the module is deemed unacceptable, unless suitable justification is supplied, if:

it requires a new type of apparatus (not generally available) or expensive equipment; or for PCR modules, the temperature-time programme used for the amplification is different between the assay targeting a specific sequence and the assay targeting the reference-specific sequence; or the resources required to perform the analysis (time, workload, reagents, costs) are considerably higher than the resources required to perform other analyses for similar purpose; or the method should not involve the use of hazardous if suitable alternative solutions are available

5.3.3 DNA Extraction Efficiency and Purity

The aim of a DNA extraction method is to provide DNA of suitable quality and quantity for subsequent analysis. DNA quality depends on the length, structural integrity, and physical-chemical purity of the extracted DNA. For each matrix being validated; it is necessary to demonstrate the DNA extracted is sufficiently pure, adequate in quantity and method is reproducible. Extraction efficiency for a given matrix can be determined by spiking known amounts of a pure DNA into that matrix prior to extraction. Extraction blanks should be included to ensure that cross-contamination does not occur during the extraction protocol.

If the DNA extraction method is provided with the kit, for a SLV the extraction is to be carried out on 6/9 test portions for SLV/ILV (replicates see Table 4.1) and repeated on different days (minimum three days), giving a total of 18/27 DNA extractions. if possible, with different operators. The DNA purity and yield must be reported.

DNA extraction methods applied to one food matrix may not be suitable for other matrices. This procedure may need to be carried out on a range of food matrices that approval is required for. The number of matrices for a SLV and MLV are shown in Table 4.1.

In agreement with international guidelines (ISO 21571, ISO 24276) acceptance criteria used to assess performance of a DNA extraction method listed below should be verified on the same working DNA concentration, i.e., the DNA concentration used in subsequent PCR analysis. Dilution will reduce the

concentration of the analyte. The structural integrity and purity of the DNA must therefore be satisfactory at the concentration to be applied in PCR analyses.

5.3.3.1 DNA Concentration

Definition: Amount of DNA per volume unit of DNA solution.
Acceptance criterion*:* The DNA concentration should be appropriate for the subsequent PCR analyses. The DNA concentration should be higher than the working concentration described in the detection method.
Example: if the PCR protocol indicates 40 ng/µL as the DNA concentration of the DNA solution to be added to the master-mix, the average concentration of the DNA extract should be > 40 ng/µL.
The determination and expression of the DNA concentration can be done with reference to mass of DNA (ng/mL)

5.3.3.2 DNA Yield

Definition: Total amount of DNA in the extract.
Acceptance criterion: The yield should be at least as much as is required for all the subsequent PCR analyses.
The DNA extraction module should provide similar yields for the same matrix.

5.3.3.3 DNA Structural Integrity

Definition: Breakage of genomic (high molecular weight) DNA into smaller DNA fragments
Acceptance criterion: The minimum size of most DNA fragments should be larger than the size of the amplicon produced by the PCR module used in subsequent analyses. This can be evaluated by agarose gel electrophoresis with DNA ladder. The DNA extraction module should not significantly reduce the structural integrity of the DNA or exhibit significant biased selectivity of DNA fragments.
Structural integrity is determined for two purposes: 1) validating a DNA extraction module, and 2) to verify that a DNA extract has provided DNA fit as template for e.g. quantitative PCR analysis.

5.3.3.4 Purity of Extracted DNA

Definition: The absence of coextracted compounds in a DNA sample impairing the efficiency of the PCR reaction and leading to a delay in the onset of the exponential phase of the amplification profile.
Acceptance criterion: the difference (ΔCt) average between the measured Ct value and the extrapolated Ct value of the first diluted sample of the inhibition test should be <0.5 [(measured Ct – extrapolated Ct)] and the slope of the inhibition curve should be in the range of - 3.1 ≤ slope ≤ - 3.6, corresponding to amplification efficiencies of 110% to 90%.
The total amount of DNA in the first sample of the dilution series should not be less than the total amount of DNA used in the submitted method
The preferred qPCR module for the inhibition test is a validated endogenous-specific reference system (e.g., lectin for soybean DNA or Act b for chicken DNA).

Note: In case of specific samples from which it may be difficult to extract genomic DNA of high quality (e.g. processed food/feed samples, meat patties, refined oils, lecithin), a slope of the inhibition curve within -4.1 and 3.1 is acceptable.

Example: DNA purity can be demonstrated by analysing two PCR replicates using four points of four-fold serial dilutions (1:4, 1:16, 1:64 and 1:256) of each DNA extraction replicate.

1. The DNA extract is first brought to a level corresponding to the highest DNA concentration intended to be used called 'undiluted' sample (working dilution e.g., 25-40 ng/µL).
2. Four-fold dilution series is prepared (from 1:4. 1:16, 1:64, and 1:256).
3. RT-PCR for endogenous reference gene carried out
4. The Ct values of the four serially diluted samples are plotted against the logarithm of the dilution factor.
5. The Ct value of the 'first diluted' sample extrapolated from the linear regression equation is compared with the Ct measured for it.

5.3.4 Specificity

Definition: Property of a method to respond exclusively to the characteristic or analyte of interest.

Acceptance criteria: The PCR method/kit should only produce amplification products with the target sequence for which it was developed.

A description of the verification of the amplification product should be included in the PCR method. Appropriate techniques are probe hybridisation, DNA sequence analysis or restriction enzyme digestion or other sequence verification techniques.

The method should be tested with DNA from closely related or potentially co-occurring non-target species/varieties and DNA from the reference species/variety material.

The specificity of a novel assay can be accomplished theoretically and experimentally.

Theoretical test for specificity- Carry out a computer-aided ("in-silico") test, examining the oligonucleotide sequences (primer, probe) as well as the amplicon. Search for e similarities to other sequences by using suitable databases (e.g. BLASTn).

Experimental test for specificity- The method must be tested with DNA from non-target species/varieties (e.g., DNA form sheep/cow meat when the target is horse meat) and DNA from the reference species/variety (inclusivity) material. Demonstrate the absence of amplification products when the target sequence specific assay is applied to pure genomic DNA of:

i. · A representative collection of the closest related species;
ii. · The most important food/feed crops.

The tests should be conducted with approximately 2500 copies of non-target DNA and with at least 100 copies of target DNA.

Table 4.2 Parameters to be evaluated for validation PCR methods. The acceptance criteria are given between brackets.

Parameter	**Quantitative**	**Qualitative**
Method acceptance parameters		

Applicability	Yes	No
Practicability	Yes	Yes
Extraction efficiency	Yes	Yes
Specificity	Yes	Yes (no false +ve/-ve)
Sensitivity (LOD)	No	Yes (≤0.05%)
Sensitivity (LOQ)	Yes (≤0.1%)	No
PCR Amplification efficiency	Yes (90%-110%)	only for multiplex (80-100%)
Linearity (R^2)	Yes (R^2 ≥0.98)	Only for multiplex (R^2 ≥0.98)
Trueness	Yes (±25%)	No
Precision (Repeatability)	Yes (RSD_R ≤25%)	No
Robustness	Yes (≤30%)	Yes (correct +ve/-ve classification)
Method performance parameters		
False Positive/Negative rate	No	✓ (≥ 5%)
Precision (Reproducibility)	Yes (RSD_R ≤25%)	No
Measurement of uncertainty	√Yes (≤50%)	No

In microbiological DNA based methods the number of bacterial species used for inclusivity/exclusivity testing will vary with the analysis being conducted and the target microorganism. Samples for inclusivity assessment should be chosen to reflect the genetic diversity of species on which the assay will be used; samples for exclusivity testing should be chosen to reflect related and potentially cross-reactive organisms and species, as well as those likely to co-occur in food products. Both inclusivity and exclusivity testing should be performed on purified samples and amounts of DNA should be equal between inclusivity samples and exclusivity samples. Samples used in specificity testing should be traceable to the source.

5.3.5 Sensitivity: Limit of Detection (LOD)

Definition: The limit of detection (CCβ) is the lowest amount or concentration of nucleic acid in a sample, which can be reliably detected, but not necessarily quantified.

Acceptance criterion: The LOD should be < 0.05% with a level of confidence of 95%, ensuring ≤ 5% false negative results.

By their very nature, qualitative test results refer to the identification above/below a detection limit. Data obtained from testing the method at different concentrations of the target sequence to determine the sensitivity of the method should be provided. Limits of detection (LOD) should be defined using samples comprised of single ingredients only. The LOD is usually understood as the

concentration of the target DNA at which an amplification product is detected with a probability of 95%.

The LOD should be determined by means of a dilution series of the target DNA. To reach the required level of confidence, a suitable number of replicates should be tested. As an example, the number of replicates tested per amount or concentration may be 60, with the LOD set at the lowest concentration yielding at least 59 positive results (Zar 1999). The amounts tested may include approximately 2, 1, 0.05, 0.02, 0.01 %. This data may be represented as DNA weight/reaction (ng or pg etc.) or the target DNA (%)/reaction or copy number.

After the LOD of the assay is determined using a dilution series of the target DNA, the originating lab should perform experiments to estimate the LOD of the assay in various food matrices. The limit of detection for a qualitative method can be defined as the concentration at which a positive sample yields a positive result at least 95% of the time. This results in a rate of false negative results of 5% or less.

A real-time PCR method may employ a Ct cut off value above which a result is considered negative. It is the responsibility of the originating laboratory to determine if a cut off value should be established and if so, what cut off value should be used. The decision shall be based upon validation data and if available, results of testing naturally incurred material.

If all replicates at a lowest dilution are positive, this infers that the LOD_{rel} is below or equal to that dilution of the positive control material level.

5.3.6 False Positive/Negative rate

5.3.6.1 False Positive Rate

Definition: The false-positive rate is a measure of the probability that a method will classify a known negative test sample as positive also called Type 1 error (scoring a false positive).

For convenience, this rate can be expressed as percentage:

$$\%\ \text{false positive rate}\ (\alpha) = \frac{100\ \text{x number of misclassified known negative samples}}{\text{total number of known negative samples}}$$

Acceptance criterion: $\leq 5\%$.

5.3.6.2 False Negative Rate

Definition: The false-negative rate is a measure of the probability that a method will classify a known positive test sample as negative also known as Type II error (scoring a false negative).

For convenience, this rate can be expressed as percentage

$$\%\ \text{false negative results}(\beta) = \frac{100\ \text{x number of misclassified known positive samples}}{\text{total number of known positive samples}}$$

Acceptance criterion: $\beta \leq 5\%$

To demonstrate the false negative rate for qualitative assay, a series of samples with a constant, known concentration of positive material in a pool of negative material must be analysed and the results evaluated. It is important to note that the concept of confidence intervals and statistical uncertainty

needs to be applied to the risk of false positive and/or false negative results as well. The desired level of confidence determines the size and number of pools that need to be tested.

5.3.6.3 Statistical approach to confirm false negative and false positive rates as <5%

The minimum number of samples that must be tested depends on the criteria for the defect rate and the level of statistical confidence is calculated using the formula

$$n = \frac{\text{Log}(\propto)}{\text{Log}\,(1-\rho)}$$

where 1-α is the confidence level and ρ is the maximum acceptable FN or FP rate. Sample sizes to assess selected criteria for FN or FP rates with varying levels of confidence are listed in Table 4.3. Both FN and FP rates should be determined as <5% using a valid statistical approach.

Table 4.3 Sample sizes recommended for assessing FN or FP rates

False Positive/ Negative rate	Confidence level			
	80%	90%	95%	99%
< 1%	161	230	299	459
<2%	80	114	149	228
<5%	**32**	**45**	**59**	**59**
<10%	16	22	29	44

For example, if the goal is to have 95% confidence that the FN rate is <5% then test 59 samples with the target DNA present at the concentration of interest, typically the LOD or a relevant level of concern, in a range of matrices. The criteria are satisfied if all 59 test results are positive for the target.

This sample size formula is related to the Clopper-Pearson confidence interval for Binomial proportions and frequently used for zero defect acceptance sampling plans for commodity lots. The rationale for the sample size is that when the probability of a false positive/negative response is ρ for each sample then (1 - ρ) n is the probability that n samples will have the correct response.

5.3.7 Robustness

Definition: The robustness (ruggedness) of a method is a measure of its capacity to remain unaffected by small, but deliberate deviations from the experimental conditions described in the procedure.

Acceptance criterion: the method should provide the expected results when small deviations are introduced from the experimental conditions described in the procedure. For quantitative methods/kits the acceptance criterion of 25% for the relative repeatability standard deviation (RSDr) and trueness, the RSDr and trueness calculated for a combination of changes should not exceed 30% (Table 4.3).

For qualitative modules all replicates should give positive results.

To determine robustness, measurement results corresponding to small, deliberate changes in the measurement conditions are collected. The method should provide the expected results when small deviations are introduced from the experimental conditions described in the procedure.
For quantitative methods, the target amount/concentration to be tested should be at the LOQ.
For qualitative methods the target amount/concentration to be tested should be 3 times the LOD.
The following factors are potential examples of robustness/ruggedness for a qualitative PCR method:

a. Use of different brands and models of a thermal cycler
b. Use different PCR reaction mix kits (Buffer, dNTPs, Mg^{2+})
c. Changes in Master mix reaction volume (e.g., 20 µL to 25 µL)
d. Changes in probe and primer concentrations
e. Changes in the thermal cycling parameters

The following Table 4.4 provides guidance for a multifactorial robustness test that may be used to validate for robustness of PCR assays. The target sequence used in the tests should be at the LOQ of the method. If the LOQ is not known, a simplified estimation can be done by multiplying the LOD by three. The target DNA should be diluted in 5 ng/µL non-target DNA.

Table 4.4: Robustness Testing Matrix* (n≥3)

Factor	Combination							
Thermocycler	A	A	A	A	B	B	B	B
PCR kit	X	X	X	X	Y	Y	Y	Y
Primer concentration	NC	-30%	NC	+30%	NC	-30%	NC	+30%
Probe concentration	NC	-30%	+30%	NC	-30%	NC	NC	+30%
Master mix volume	-5%	-5%	+5%	+5%	+5%	+5%	-5%	-5%
Annealing temperature (°C)	+1	-1	+1	-1	-1	+1	-1	+1

NC=No change
Note: If negative PCR results are observed for any combination(s) they should be repeated once. If the negative results are confirmed in the second test, the
outcome indicates insufficient robustness of the PCR method.
*Adapted from **From: Guidelines for the single-laboratory validation of qualitative real-time PCR methods-Bundesamt fur Verbraucherschutz und Lebensmittelsicherheit-March 2016 and* Guidelines for the Validation of Analytical Methods for Nucleic Acid Sequence-Based Analysis of Food, Feed, Cosmetics and Veterinary Products *Edition 1.1* U.S. Food and Drug Administration Foods Program October 2020

5.4 Quantitative PCR methods

The analysis of DNA, especially in processed foods, requires the detection of very small amounts of target-specific DNA. The result of a quantitative PCR analysis is often expressed in % as the amount of target nucleic acid relative to an endogenous control or a taxon-specific PCR product, therefore, this measurement involves two PCR-based determinations – that of the target-specific DNA sequence and that of the comparator. Each of these determinations has its own uncertainties, and the two are likely to have different measurement characteristics. It is thus important that both measurements are properly validated, and that the assay is fit for purpose.

For quantitative PCR methods, the basic performance characteristics are:

i. Extraction Efficiency-See Section 5.3.3
ii. Standard curve
    a. PCR Efficiency
    b. Linear Dynamic Range (Range of quantification)
iii. Specificity (Selectivity)- See Section 5.3.4
iv. Sensitivity (Limit of Detection-LOD)-See Section 5.3.5
v. Sensitivity (Limit of Quantification-LOQ)
vi. Precision-Repeatability and Reproducibility Standard Deviations
vii. Robustness/Ruggedness-See Section 5.3.7
viii. Accuracy (Trueness/Bias)

5.4.1 Standard curves

The generation of standard curves is required for the optimization of quantitative PCR analyses. At the initial stages of method development, standard curves should be derived from pure samples of target nucleic acid. Assays used for the analysis of food should also include standard curves generated using a relevant food matrix or matrices (see section 8.0 below for selection of relevant matrices). These standard curves will be used to determine the PCR efficiency, dynamic range, and limit of quantification, which are discussed below.

The average values of a minimum of two standard curves should be taken. The slope of the standard curve is calculated using the equation $y = mx + b$ (where y = C_T value and x = log target DNA concentration).

Example 1: Two calibration curves minimum requirements 5 calibration points with 3 PCR replicates each (triplicates)

Slope of the curve must be in the range of -3.6 to -3.1 and all R^2 values should be ≥0.98. (30 PCR reactions)

Example 2: Four calibration curves-5 calibration points with 2 PCR replicates each (duplicates) average of 4 slopes (-3.6 to -3.1) and R^2 (be ≥0.98) are used to verify the acceptance. (40 PCR reactions)

Example 3: Two calibration curves;8 calibration points in 5 PCR replicates (pentaplicates) also covering the low concentrations for LOD and LOQ. Average of the part above LOQ for slope and R^2 are used to verify the acceptance. (80 PCR reactions)

5.4.1.1 PCR efficiency

PCR efficiency is a measure of how close the observed reaction is to a true statistical doubling of amplified product over successive cycles and is based on C_T values. The efficiency is tested by preparing a standard curve of the template DNA and determining the C_T-value for each dilution. For an amplification efficiency of 100%, a two-fold reduction in template DNA should result in an increase in the C_T value of one cycle. Therefore, if DNA is diluted 10-fold, the theoretical difference in C_T values between the two concentrations of template nucleic acid should be approximately 3.32 cycles. Most current real-time PCR analysis software packages will calculate the slope and PCR efficiency of an assay based on the C_T values of the standard curve. In rare cases, quantification can still be performed using assays with efficiency outside the optimal range. The PCR efficiency is calculated as

Efficiency (%) = $-1+10^{(-1/\text{slope})}$

Acceptance criteria: The average slope of the standard curve should be between -3.1 and -3.6, which corresponds to an efficiency of 90%-110%.

Significant deviations in PCR efficiency may indicate the presence of PCR inhibitors or that the assay needs to be further optimized. Reaction efficiencies determined using known amounts of target spiked into a food matrix will also indicate extraction efficiency for that matrix.

5.4.1.2 Linear dynamic range

The dynamic range is the concentration range over which the target DNA sequence will be reliably detected. This desired concentration range defines the standard curves which will be used for quantification. Generally, a minimum five-log concentration range with three replicates each is necessary to determine the dynamic range of the assay; a well-designed and well-optimized real time PCR assay will have a linear range of 6-8 orders of magnitude. However, each method should be validated for a dynamic range that is relevant to the application. If a method is validated for a given range of values, the range may not be extended without further validation. The R^2 value of the standard curves used to determine the dynamic range should be ≥0.98.

5.4.2 Sensitivity: Limit of Quantification (LOQ)

Definition: The limit of quantification is the lowest amount or concentration of analyte in a sample that can be reliably quantified with an acceptable level of trueness and precision.

Acceptance criterion: The LOQ should be ≤ the lowest amount or concentration included in the dynamic range (i.e., 0.09%).

There are multiple experimental approaches to determine the LOQ, such as assaying spiked samples that have a known amount of target DNA, or by analyzing several samples that contain known amounts of DNA.

The limit of quantification is the minimum DNA concentration for which all 12 replicates give a positive result with coefficient of variability (CV) of no more than 0.5 C_t. The quantification should be determined by spiking the target (e.g., GMO/meat/fish) into a relevant food matrix prior to sample preparation and DNA extraction. Quantification should be expressed in units which are relevant to the

intended purpose of the method, for example as mg/kg, parts per million, or percentage in a food matrix.

5.4.4 Precision

5.4.4.1 Relative repeatability standard deviation (RSDr)

Definition: Repeatability conditions are conditions where test results are obtained with the same method, on identical test items, in the same laboratory, by the same operator, using the same equipment within short intervals of time.
Acceptance criterion: The relative repeatability standard deviation should be ≤25% over the whole dynamic range of the PCR modules individually for all genes and in combination.
Note: Estimates of repeatability submitted by the applicant should be obtained on a sufficient number of test results, at least 15, based on ISO 5725-3.
RSDr is not applicable to qualitative method
Repeatability should be available for all tested food matrices.

5.4.4.2 Relative reproducibility standard deviation (RSDR)

Definition: The relative standard deviation of test results obtained under reproducibility conditions. Reproducibility conditions are conditions where test results are obtained with the same method, on identical test items, in different laboratories, with different operators, using different equipment. Reproducibility standard deviation describes the inter-laboratory variation and is used in collaborative trials.
Acceptance criterion: The relative reproducibility standard deviation RSD_R should be <35% over the whole dynamic range.
Note: RSDr is usually calculated only for quantitative methods.

5.4.5 Trueness (Bias)

Definition: The closeness of agreement between the average value obtained from a large series of test results and an accepted reference value. The measure of trueness is usually expressed in terms of bias.
Acceptance criterion: The trueness should be within ±25% of the accepted reference value over the whole dynamic range.
Trueness compares the obtained value from a series of samples to the actual or reference value.
Trueness is not applicable to qualitative methods.

6.0 QUALITATIVE AND QUANTITATIVE MULTIPLEX ASSAYS

For multiplex assays, all method verification must be carried out in multiplex and performance requirement described above (Table 4.2) must be reported for each individual target as it performed under multiplex conditions. For probe-based assays the signals from the fluorophores on different targets must not interfere with each other. Multiplex intercalating dye-based assays will not be

considered quantitative because intercalating dyes do not distinguish between different targets in a multiplex assay.

For multiplex qualitative modules, specificity should be evaluated for each target sequence. The same acceptance criteria as for the corresponding single PCR method should be applied.

For multiplex qualitative PCR modules an asymmetric LOD (LOD_{asym}) should be determined.

The LOD_{asym} is determined by testing the analyte target at low amount or concentration (corresponding or close to the absolute LOD) in the presence of increasing amounts or concentrations of the other target(s) in the multiplex assay.

The LOD_{asym} is expressed as the minimum ratio between the copy number of the tested analyte target and the copy number level of the other target(s) for which the analyte target can still be detected with a level of confidence of 95%, ensuring ≤ 5% false negative results.

Example: in a duplex PCR module, if 20 copies of the target DNA are detected with a level of confidence of 95%, in presence of 20,000 copies of the other target, the LOD_{asym} is then a ratio below 1:1000.
In case of multiplex PCR modules, the LOD_{asym} can be determined by testing e.g. 20 copies of each target sequence in presence of a background of all other targets summed at the level of 20,000 copies.

As described for the LOD the required level of confidence can be achieved with a suitable number of replicates that should be tested. As an example, the number of replicates tested per amount or concentration may be 60, with the LOD_{asym} set at the lowest concentration of the target in presence of the other target at the higher concentration yielding at least 59 positive results (Cochran 1977, Zar 1999).

7.0 MATRIX EXTENSIONS

The validation of method performance with a new matrix is intended to assure that the method will continue to produce accurate and reliable results. Matrix extensions (Level 1 in Table4.1) are intended wherein a validated method is used with a new matrix not previously validated in response to requirement by FSSAI. Matrix extensions of validated methods that are intended to increase the regulatory scope and applicability, such as running the method on a recurring basis, would minimally fall under Level 2 validation in Table 4.1. It is generally assumed that the more closely related a new food matrix is to a previously validated matrix for a defined analyte, the greater the probability that the new matrix will behave similarly. Sec 8.0 provides guidance on commodity categories. The number of different food categories to be validated depends on the applicability and intended use of the method. Depending on how many categories will be validated, a minimum of 1 – 3 representative matrices from each category should be selected, depending on the level of validation required and the number of food categories being tested

7.1 Platform extensions

Expanding the use of a validated method to include another significantly different instrument or platform requires further validation. Such instances include the use of an instrument or platform similar in scope and function to that currently validated for approved use. However, it may have differences in throughput, configuration, chemistry, or detection methodology. Platform extension validation should be performed as described in Table 4.1, Level 2. In planning platform extensions, acceptance criteria for these comparisons must be established by the validating laboratory. Care must be taken to ensure that the new platform produces equivalent results to the originally validated method.

7.2 Reference materials (RMs)

All reference materials used must have traceable Sequences included in a reference library should be collected from a wide variety of organisms or groups likely to be encountered based on the intended purpose of the method. These reference libraries must be curated for quality and accuracy. Optimally, reference sequences should be generated under a defined set of conditions to assure quality. If reference sequences are derived from public databases, it is incumbent on the originating laboratory to verify the validity and quality of the sequence used. In addition, accession numbers for all sequences used for the development of molecular assays (and any supporting documentation) should be included in the validation package

8.0 FOOD MATRIX AND SAMPLE SELECTION

Food matrix and sample source selection should be based on the types of foods most likely to be used in the analysis or based on risk of contamination. A PCR method intended for use in processed foods should be tested on samples subjected to similar processing. Processing conditions such high-temperature and high-pressure treatments (e.g., canning) and low pH (e.g. tomato-based products) that have adverse effects on DNA such as degradation should not be chosen. The number of food categories to be used will depends on the intended use of the method.

The number of different food categories to be verified depends on the applicability and intended use of the method. Depending on how many categories will be verified, a minimum of 1 – 3 representative matrices from each category listed below should be selected,

A list of foods that can be used based on the applicability are:

i. **Meats:** Fresh meat, Frozen meat, Raw marinated/minced/comminute meat, Semi-cooked /Smoked Meat, partially heat treated and/ or smoked meat and meat product, Canned/Retorted meat product, Chilled meat, Cooked Meat/meat product, Cured/pickled meat products, Dried/Dehydrated meat/meat products, Fermented meat products sausage, lunch meat, meat substitutes etc ("meat" means all edible parts (including edible offal) of any food animal slaughtered in an abattoir that are intended for or have been judged as safe and suitable for, human consumption; "meat food products" means any product prepared from meat and other ingredients through various processing methods in which meat should be the major ingredient of all the essential ingredients

j. **Seafood:** Chilled/Frozen Finfish, crustaceans, cephalopods, molluscs, bivalves, dried or Salted and dried fish products, thermally processed, fermented, smoked, canned fish products. Fish

sticks, surimi, raw fish filet, raw oysters, raw mussels, raw clams, cooked crawfish, crabmeat (fresh or pasteurized), battered and breaded fish products,

k. **Fruits, Vegetables, and Nuts**: Fresh / frozen /dehydrated or dried fruits and vegetables, fresh fruit juice, apple cider, tomato juice, fruit cubes, berries, peanut butter, coconut, fruit powders almonds, minimally processed lettuce, spinach, kale, collard greens, cabbage, bean sprouts, seed sprouts, peas, mushroom, green beans and other minimally processed fruit and vegetable products

l. **Dairy products**: Dahi, Yogurt, Paneer, Khoa, Channa, hard and soft cheeses, raw or pasteurized liquid milk (skim, 2% fat, whole, buttermilk), infant formula, coffee creamer, ice cream, milk powders, casein, whey, non-fat dry milk/dry whole milk. dried buttermilk, dried cheese spray

m. **Confectionary:** Chocolate / bakery ware Frosting and topping mixes, candy and candy coating, milk, chocolate, cake mixes,

n. **Egg and egg products**: Shell eggs, liquid whole eggs, dried whole egg or dried egg yolk, dried egg whites,

o. **Herbs and spices**: Oregano, pepper, paprika, black pepper, white pepper, celery seed or flakes, chili powder, cumin, parsley flakes, rosemary, sesame seed, thyme, vegetable flakes, onion flakes, onion powder, garlic flakes, allspice,

p. **Processed grains and legumes**: Flours, grits, rice corn meal, soy flour, dried yeast, cereal based complementary food, Uncooked noodles, pasta, macaroni, spaghetti, soygurt, tofu, soy beverage.

9.0 INFORMATION TO BE PROVIDED WITH METHOD VALIDATION DOCUMENTS

Appended below is the information that should be provided from the developing laboratory/manufacturer when the results for a validation study are prepared for review by FSSAI. This listed information is in addition to the results submitted for the validation criteria described in Table 4.2.

9.1 Assay Design

i. Type of assay: oligonucleotide probe-based or double stranded DNA staining dye-based
ii. Name of target gene or region
iii. Internal control/amplification control type: exogenous or endogenous
iv. Exact oligonucleotide sequences for all primers and probe(s)
v. Length of PCR product (amplicon)
vi. Dye and probe characteristics. For probe-based assays, provide a brief description of the probe chemistry and the identities and locations of fluorophores and quenchers, including internal quenchers. For dye-based assays, state which dye is being used.
vii. Any other reporter molecules.

9.2 Sample preparation and nucleic acid extraction

i. Form and quantity of sample required. Include information on subsampling or sample compositing as well as relevant aspects of handling and storage.

ii. Method or kit used for DNA extraction. Include any relevant modifications as well as information on RNAse treatment. Independent assessments of DNA/RNA quality and quantity are not required as long as the method is shown to yield acceptable/reliable PCR results. For quantification of targets in complex food matrices, normalization of total DNA/RNA amounts prior to PCR may result in higher quality data.

9.3 PCR conditions

i. Reaction: reaction volume; identities and concentrations of all reaction components, including buffer or master mix, all primers, all probes and/or dyes, template DNA/RNA, $Mg^{2+,}$ and additives (e.g. Bovine serum albumin (BSA), DMSO, or glycerol).

ii. Instrument: State make and model of real-time PCR platform as well as name and version of accompanying software. Include brief descriptions of physical format (e.g., 96 well thermal block or other) and optical system.

iii. Thermal cycling conditions. Include PCR cycling conditions for both dye-and probe-based assays; also include melt conditions for dye-based assays. Optimal cycling conditions should be determined empirically and not through software-based calculations of primer or probe annealing temperature, as annealing temperatures can be significantly affected by specific reaction conditions.

9.4 Data analysis

i. Specify which software program and version was used for data analysis.

ii. Report and explain any adjustments made to baseline and threshold determination, or other software default analysis parameters.

iii. For dsDNA dye-based assays, analysis of melt curves must be performed to confirm the presence of a single, sharp melting peak optimally with a melting temperature (Tm) of approximately 80-90°C in all samples and standard.

The following checklist may be used for submission of date

Checklist for validation of RT-PCR showing different steps including the essential (E) and desired (D) information to be reported in validation dossier			
Target information	Name of target	Importance	Comment
qPCR target information	Name of target	E	
	Full name of targeted sequence (accession number	D	
	Length of amplicon	E	

	Sequence of amplicon	E	
	Location of amplicon on gene	D	
	Target containing materials	E	
	Non-target containing materials	D	
qPCR oligonucleotides	Primer sequences	E	
	Probe sequence	E	n/a for SYBR-Green methods
	Probe labelling	E	Give details e.g. FAM, Scorpion
	Purification method	D	
	Manufacturer	D	
Nucleic acid extraction	Method/Kit used	E	
	Sample size required	E	
	Sample treatment if any (e.g., defatting etc.)	E	
	Sampling procedure	D	
	Subsampling	D	
	Purity (A260/A280)	E	
	Yield	E	
	Storage conditions	E	
qPCR reaction	qPCR cycling conditions	E	
	Melting curve conditions	E	Only for SYBR-Green chemistry
	Type of chemistry	E	
	Singleplex/multiple	E	
	qPCR reagents	E	
	qPCR volume	E	

	qPCR mix setup	E	
	qPCR instrument (Make and Model)	E	
	DNA amount/reaction	E	
	Purity of DNA (A260/A280)	E	
	qPCR cycling programme	E	
Data analysis	qPCR analysis programme	E	
	Outlier test	E	
	Inhibition test	E	
	Results of controls (Premise control, No template control	E	
Overview of PCR-validation			
Applicability	Scope of the method	E	
	Type of matrix	E	
	Inhibition information	E	
	Analytical interference	E	
Practicability	Easy combination with other methods	E	
	Costs	E	
	Need for training	E	
	Analytical interference	E	
	Blind sample analyses	D	
	Method transfer	D	
	Equipment needs	D	Essential if e.g special chemistry or conditions are used
	In silico specificity (BLAST etc)	E	Give details of database used parameters used

Specificity	In situ specificity against target containing materials	E	Give details about the material used (CRM/RM/incurred material/documented)
	In situ specificity against non-target containing materials	E	Give details about the material used (CRM/RM/incurred material/documented)
	In situ specificity against other materials (bacteria. virus)	D	Give details about the material used (CRM/RM/incurred material/documented)
	Confirmation (gel electrophoresis/ sequencing etc)	E	
	Results expressed as false positive/false negative rate	E	
Sensitivity	LOD 95%	E	
	PCR efficiency	E	
	PCR linearity	E	
Robustness	Change in primer concentration	E	
	Change in probe concentration	E	Only for TaqMan and other FRET probes
	Change in reaction volume	E	
	Change in annealing temperature	E	
	Change of instrument make	D	
	Change of master mix/ qPCR reagents	D	
	Change in DNA concentration	E	
Collaborative	Number of participating laboratories	E	

trials	Choice of laboratories	E	
	False positive/False negative rate	E	
	Types of samples	E	
	Level and content of DNA in samples	E	
	Number of samples	E	
Adapted from Broeders et al., (2014) Guidelines for validation of qualitative real-time PCR methods. Trends in Food Technology, 37, 115-126			

10.0 GLOSSARY

Accuracy: Closeness of agreement between a measured quantity value and a true quantity value of a measurand.

Amplicon: DNA sequence produced by a DNA-amplification technology, such as PCR.

Amplification plot: graph representing the generation of a reporter (usually fluorescent) signal during a qPCR or dPCR reaction. The amplification plot shows the relationship between cycle number (x-axis) and fluorescence signal (y-axis).

Amplification efficiency: The rate of amplification calculated from the slope of the standard curve obtained after a decadic semi-logarithmic plot of Ct values over the quantity. The efficiency (in %) can be calculated by the following equation

Efficiency = $(10^{(-1/slope)}-1) \times 100$

Analyte: Component of a system to be analyzed

Analytical sample: Sample prepared from the laboratory sample by grinding/ homogenization

Calibration: Operation that establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and uses this information to establish a relation for obtaining a measurement result from an indication.

Certified Reference Material (CRM): Reference material, accompanied by documentation issued by an authoritative body and providing one or more specified property values with associated uncertainties and traceability, using valid procedures.

Cross-Reactivity: Degree to which binding occurs between an antibody and antigenic determinants, or primers and a target sequence, which are not the analyte of primary interest.

Cycle Threshold (Ct) also known as quantification cycle is defined as the fractional cycle number at which the fluorescence generated by the amplification of a target DNA in a real time PCR experiment reaches a fixed threshold and so allows the quantification of the amount of target DNA.

Denaturation: Process of partial or total alteration of the native structure of a macromolecule resulting from the loss of tertiary and/or secondary structure that is a consequence of the disruption of stabilizing weak bonds of DNA: DNA that has been converted from double-stranded to a single-stranded form by a denaturation process such as heating

Deoxyribonuclease/Ribonuclease (DNase/RNase): Enzyme that catalyses the hydrolytic cleavage of deoxyribonucleic acid/ribonucleic acid that may produce a single nucleotide residue by cleavage at the end of the chain or a polynucleotide by cleavage at a position within the chain.

Deoxyribonuclease/Ribonuclease Inhibitor: Substance that either fully or partially blocks deoxyribonuclease/ribonuclease activity.

DNA extraction replicates (see Figure 4.1): DNA extracted from different test portions from the same analytical sample.

Deoxyribonucleotide triphosphate (dNTP) Generic term referring to a deoxyribonucleotide that includes: deoxyadenosine nucleotide triphosphate (dATP), deoxycytidine nucleotide triphosphate (dCTP), deoxyguanosine nucleotide triphosphate (dGTP), deoxythymidine nucleotide triphosphate (dTTP) and deoxyuridine nucleotide triphosphate (dUTP).

DNA Extraction: Sample treatment for the liberation and separation of DNA from other cellular components.

DNA Polymerase: Enzyme that synthesizes DNA by catalysing the addition of deoxyribonucleotide residues to the free 3'-hydroxyl end of a DNA molecular chain, starting from a mixture of the appropriate triphosphorylated bases.

DNA Probe: Short sequence of DNA labelled isotopically or chemically that is used for the detection of a complementary nucleotide sequence.

Dynamic range: The range of concentrations over which the method provides a linear correlation between the measurement and the amount of the target, with an acceptable level of trueness and precision.

End-Point PCR Method: where the amplicons are detected at the end of the PCR reaction, typically by gel electrophoresis and the amplified product is visualized with a fluorescent dye.

False Negative Error of failing to reject a null hypothesis when it is in fact not true.

False Negative Rate: Probability that a known positive test sample has been classified as negative by the method. The false negative rate is the number of misclassified known positives divided by the total number of positive test samples.

$$\%\ \text{False negative} = \frac{\#\ \text{of misclassified positive samples}}{\#\ \text{of positive test resulst(inclusive of missclassified}} \times 100$$

False Positive Rate: Probability that a known negative test sample has been classified as positive by the method. The false positive rate is the number of misclassified known negatives divided by the total number of negative test samples.

$$\%\ \text{False positive} = \frac{\#\ \text{of misclassified negative samples}}{\#\ \text{of negative test resulst(inclusive of missclassified}} \times 100$$

Fitness for Purpose: Applicability of a prescribed method or the degree to which data produced by a measurement process enables a user to make technically and administratively correct decisions for a stated purpose

Fluorescence Resonance Energy Transfer/FRET: Distance dependent energy transfer from a donor molecule to an acceptor molecule resulting in enhanced fluorescence of the acceptor molecule after excitation with electromagnetic radiation of a defined wave length.

Fluorescent Probe: Oligonucleotide: oligonucleotide analogue of defined sequence coupled with one or more fluorescent molecules emitting a fluorescent signal after specific hybridization to the target nucleic acid sequence which can be detected by the specific equipment.

Laboratory sample: Sample as received by the laboratory and intended for inspection or testing.

Limit of detection (LOD): LOD is the lowest amount or concentration of analyte in a sample, which can be reliably detected but not necessarily quantified. Experimentally, methods should detect the presence of the analyte for at least 95 % of the cases (samples) at the LOD, ensuring ≤5 % false negative results.

Limit of quantification (LOQ): LOQ is the lowest amount or concentration of analyte in a sample, which can be reliably quantified with an acceptable level of precision and trueness consistently under the experimental conditions specified in the method.

Linearity: Ability of a method of analysis, within a certain range, to provide an instrumental response or results proportional to the quantity of analyte to be determined in the laboratory sample.
Matrix: All relevant components of a sample inclusive of analyte.

Multiplex PCR: PCR technique that employs multiple pairs of primers combined within a single reaction mixture to produce multiple amplicons.

Measurand: Quantity intended to be measured.

Measurement uncertainty: Non-negative parameter characterizing the dispersion of the quantity values attributed to a measurand based on the information used. Measurement uncertainty includes components arising from systematic effects, the assigned quantity values of measurement standards, as well as the definitional uncertainty. It is understood that the measurement uncertainty is associated

with a stated quality value attributed to the measurand. A modification of this value results in a modification of the associated uncertainty

Melting Curve Analysis: describing the dissociation characteristics of double-stranded DNA observed during heating. The information gathered can be used to infer the presence and identity of single-nucleotide polymorphisms.

Melting Temperature (Tm): Temperature at which 50% of double-stranded DNA helices are dissociated since a DNA helix melts in a temperature range rather than at one very specific temperature

Probability of detection (POD): The probability of a positive (i.e., presence detected) analytical outcome for a qualitative method for a given matrix at a given concentration. It is estimated by the expected ratio of positive to negative results for the given matrix at the given analyte concentration.

Precision (Relative repeatability standard deviation (RSD_r): The relative standard deviation of test results obtained under repeatability conditions. Repeatability conditions are conditions where test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time. RSDr is calculated by dividing the repeatability standard deviation by the mean of results.

Passive Reference Dye: Fluorescent molecules present in the reaction mix used to normalize the signal and may be coupled with nucleic acid sequences or other molecules not taking part in the reaction.

PCR Target Sequence Specific: region of DNA that becomes selectively amplified during PCR-based detection, identification and/or quantification. The PCR target sequence is characterized by being located between the primers, and in the case of real-time PCR, may include the probe hybridization site.

Percent Error: Relative error expressed as a percentage.

Polymerase Chain Reaction (PCR) In vitro enzymatic technique to increase the

number of copies of a specific DNA fragment by several orders of magnitude.

Primer: Strand of nucleic acid sequence that serves as a starting point for DNA synthesis.

Qualitative Method: Method of analysis that yields a binary result.

Quality Assurance: Planned and systematic actions necessary to provide adequate confidence that analytical results will satisfy given requirements for quality.

Quantitative Analysis: Analyses in which the amount or concentration of an analyte may be determined and expressed as a numerical value in appropriate units.

Reference Method: A reference method is a method by which the performance of an alternate or new method may be measured or evaluated. In some cases, an appropriate reference method may not be available. However, there are some instances in which the use of a reference method is appropriate

such as when replacing a method specified for use in a compliance program. Use of a reference method is required for microbiological methods when available.

Reference (Endogenous) gene: gene target present in each sample at approximately constant concentration that is resistant to response fluctuations due to changes in biological or experimental conditions, or stable within a particular species or taxon. Reference genes have, historically, been referred to as housekeeping genes.

Repeatability standard deviation (RSDr): Standard deviation of test results obtained under repeatability conditions.

Reproducibility: Measurement precision under reproducibility conditions of measurement.

R^2coefficient: R^2 is the coefficient of determination, which is calculated as the square of the correlation coefficient (between the measured Ct-value and the decadic logarithm of the concentration) of a standard curve obtained by linear regression analysis.

Robustness: The robustness of a method is a measure of its capacity to remain unaffected by small, but deliberate deviations from the experimental conditions described in the procedure.

Specificity: The property of the method to respond exclusively to the characteristic or the analyte of interest.

Template: strand of DNA or RNA that specifies the base sequence of a newly synthesized strand of DNA or RNA, the two strands being complementary.

Test portion: Sample, as prepared for testing or analysis, the whole quantity being used for analyte extraction at one time (Figure 4.1)).

Test result: A test result is a Ct value or copy number concentration originating from a PCR replicate.

Trueness: The closeness of agreement between the average value obtained from a large series of test results and an accepted reference value. The measure of trueness is usually expressed in terms of bias.

Validation of method: Validation is the confirmation by examination and provision of objective evidence that the particular requirements for a specific intended use are fulfilled. Method validation criteria may include: sensitivity, accuracy, trueness, reproducibility and robustness/ruggedness, precision

Verification of method: Provision of objective evidence that a laboratory can adequately operate a method, achieving the performance requirements for the sample matrices to which the method is being applied.

Working DNA concentration: The highest DNA concentration intended to be used in PCR analysis

11.0 REFERENCES

AOAC International, “Appendix D: Guidelines for Collaborative Study Procedures to Validate Characteristics of a Method of Analysis”, 2005. Accessed 12/4/2014. http://www.aoac.org/aoac_prod_imis/AOAC_Docs/StandardsDevelopment/Collaborative_Study_Validation_Guidelines.pdf

AOAC International, “Appendix K: Guidelines for Dietary Supplements and Botanicals, Part 1 AOAC Guidelines for Single-Laboratory Validation of Chemical Methods for Dietary Supplements and Botanicals”, 2013. Accessed 12/4/2014. http://www.eoma.aoac.org/app_k.pdf

Bundesamt fűr Verbraucherschutz und Lebensmittelsicherheit, 2016. Guidelines for the single-laboratory validation of qualitative real-time PCR methods.

Broeders,S., Huber, I., T., Grohmann, T., Berben, G., Taverniers, I., Mazzara, M., Roosens,N., Morisset, D. (2014) Guidelines for validation of qualitative real-time PCR methods. Trends in Food Technology, 37, 115-126

Cochran, W.G. *Sampling Techniques.* New York: John Wiley and Sons, 1977.

Feldsine, P. et al. JAOACI 2002. AOAC International Methods Committee Guidelines for Validation of Qualitative and Quantitative Food Microbiological Official Methods of Analysis.

Grohmann L, Kiesecker H, Uhlig S, Mankertz J, Broll H, Hübert P, Lieske K, Mäde D, Reiting R, Schulze M, Speck B, Waiblinger H-U, Woll K, Zur K (2012) Practical Approach for single-laboratory validation of qualitative real-time PCR methods. In preparation.

Guidelines for the Validation of Analytical Methods for Nucleic Acid Sequence-Based Analysis of Food, Feed, Cosmetics and Veterinary Products Edition 1.1 U.S. Food and Drug Administration Foods Program October 2020

Horwitz W., Protocol for the Design, Conduct and Interpretation of method Performance Studies, Pure & Appl. Chem. 67: 331-343, 1995.

International Standard (ISO) 16577: 2016, Molecular biomarker analysis – Terms and definitions.

International standard (ISO) 21569:2005/Amd 1:2013, Foodstuffs. Methods of analysis for the detection of genetically modified organisms and derived products – Qualitative nucleic acid-based methods.

International standard (ISO) 21570:2005/Amd 1:2013, Foodstuffs. Methods of analysis for the detection of genetically modified organisms and derived products – Quantitative nucleic acid based methods.

International standard (ISO) 21571:2005/Amd 1:2013, Foodstuffs. Methods of analysis for the detection of genetically modified organisms and derived products – Nucleic acid extraction.

International standard (ISO) 24276:2006/Amd 1:2013, Foodstuffs. Methods of analysis for the detection of genetically modified organisms and derived products – General requirements.

International Standard (ISO) 5725:1994. Accuracy (trueness and precision) of measurement methods and results.

International Standard (ISO/IEC) 17025:2005, General requirements for the competence of testing and calibration laboratories.

JRC Technical report: Definition of Minimum Performance Requirements for Analytical Methods of GMO Testing. European Network of GMO Laboratories (ENGL) 2015

Zar, J. Biostatistical Analysis. (4th ed.) Upper Saddle River, NJ, Prentice Hall. 1999.

Chapter-5

GUIDELINES FOR VALIDATION OF IMMUNOASSAY METHODS: ENZYME LINKED IMMUNOSORBENT ASSAY (ELISA) AND LATERAL FLOW IMMUNOASSAY (LFIA)

1. INTRODUCTION

Immunochemical bioanalytical methods represent one of the most versatile that give highly specific and sensitive results. Most screening and rapid methods are based on immunoassays. An immunoassay is based on a highly specific interaction of an analyte/antigen with a selective antibody. The efficacy of immunoassay is mainly based on the efficiency of antigen-antibody complex formation. ELISA, is an immunoassay procedure coupled with an enzymatic detection method and in theory offers unparalleled qualitative detection and quantitative capability. ELISAs vary in assay format, target analyte(s), antibodies, extraction protocols, quantitative ranges, selection/ availability of reference standards, and quantification units used for immunoreactivity.

Lateral flow immunoassay (LFIA) based on the principles of immunochromatography exist for a wide array of target analytes. LFIA, is one of the most successful analytical platforms for rapid or point-of-need testing requiring little to no supporting infrastructure. In an LFIA the sample is added on a standalone device and the result is obtained in a few minutes.

Availability of validated ELISA/LFIA kits is critical for both developers and end-users. For kit/method developers, validation of an analytical procedure is used to demonstrate that it is suitable for its intended purpose. For end-users, validated methods help to ensure reliability, repeatability, accuracy, and precision of the results generated using the kit. Food matrix and food processing are also known to affect ELISA/LFIA performance. These factors must be taken into consideration when validating the performance.

This document provides guidelines for carrying out validation of ELISA kits and LFIA. ELISA/LFIA are applicable in various domains of analyses related to food safety. ELISA is most commonly used for quantitative methods, whereas LFIA are used in screening of:

- Antibiotics (tetracycline, penicillin, gentamicin etc.) and other veterinary drug residues in animal derived foods.
- Food pathogens such as *Salmonella enterica, Bacillus cereus, E. coli* etc.
- Crop contaminants such as aflatoxins, DON, fumonisin etc in cereals

- Food allergens such as peanut, egg, milk, gluten, soybean, crustaceans, tree nuts etc

2. PURPOSE

FSSAI approved rapid kits etc. employed in food laboratories are required to meet the highest analytical performance standards. The validation of an ELISA/LFIA protocol is a crucial step in ensuring the accuracy, reliability, and reproducibility of the assay to support regulatory decision. The guidelines are aimed at providing guidance on how to establish ELISA/LFIA methods to detect and identify target analytes in food by defining appropriate validation criteria, and whether or not the kit complies with these criteria based on the performance characteristics of a method. The guidelines specify the relevant criteria and give explanations on how to consider these criteria, i.e.: -by providing the rationale for the most relevant criteria and -by showing how to find out whether or not a method fulfils the given criteria requirements. The Rapid Analytical Food Testing (RAFT) committee establishes these guidelines and criteria for submission of immunoassay-based kits (ELISA and LFD) validation data.

3. SCOPE

The scope of these guidelines is to provide details on the parameters and acceptance criteria for a complete validation of ELISA/LFIA kits for food analysis. These kits may be applicable to a wide range of analytes e.g., food allergens, mycotoxins, antibiotics, veterinary drug residues, pesticides/adulterants etc. in foods. The validation information is essential for FSSAI to decide whether the kit/method is suitable for use in food testing for regulatory compliance. Adherence to the principles presented in this guidance will ensure the quality and consistency of the data in support of the approval process. This guideline does not cover sampling in connection with the performance of a method. Full method validation is recommended for the primary food matrix intended to support submissions. Additional matrices should be validated as necessary. These guidelines apply for validation of all analytical platforms based on ELISA/immunochromatography.

4.0 METHOD VALIDATION

Method validation is a process to establish the performance characteristics and limitations of the Kit. The results of the validation process describe which analytes can be determined in what kind of food matrices in the presence of which interference. Method validation is important to ensure the acceptability of assay performance and the reliability of analytical results.

In-house validation: The manufacturer must conduct a comprehensive validation study, with performance criteria similar to Level 1 (Independent Laboratory Validation) study. If appropriate, a comparison with an existing reference method must be performed. A kit/method validated at this level is expected to have use that is widespread, long term, of high public visibility or potentially involved in regulatory, its validation should be extended to the two levels described below.

A complete validation of an ELISA kit (quantitative) /LFIA (qualitative) or any other platform for the FSSAI approval submission package should include the following two levels:

Level 1 Independent Single Laboratory Validation (ILV): After an in-house validation by the manufacturer/kit developer, a single laboratory validation of the kit/method should be carried out by a

third party e.g., FSSAI laboratory or international organizations like AOAC, AFNOR etc. This validation would ensure that a kit/method performs in a manner, which allows for a successful conclusion and provides evidence about the suitability of the kit/method for its intended purpose. This validation should preferably be carried out for the elements listed in Table 5.1 unless otherwise justified.

Level 2 Multi Laboratory Validation (MLV): Validation through a collaborative trial, which usually follows only after the method has shown acceptable performance at Level 1. In a collaborative trial, the method performance should comply with the relevant parts of the method acceptance criteria and fulfil the method performance requirements specifically set for the collaborative trial. In particular, the compliance with the criteria for 1) Sensitivity 2) Repeatability (RSD_r), and 3) Reproducibility (RSD_R) should be assessed.

Table 5.1 Parameters to be evaluated during validation of ELISA kits and Lateral Flow Immunoassay (LFIA)

Validation parameter	**Semi-quantitative and Quantitative ELISA Kit**	**Qualitative ELISA kit and LFIA**
Acceptance characteristics		
Selectivity (Specificity)	Yes	Yes
Matrix study	Yes	Yes
Sensitivity	Yes LOD & LOQ	Yes LOD &POD
Linearity (R^2) (Standard curve)	Yes	No
Linearity-of-dilution	Yes	No
Accuracy (Trueness)	Yes	No
False Positive/Negative rate	No	Yes
Precision (Repeatability)	Yes	No
Ruggedness (Robustness)	Yes	Yes
Matrix effect	Yes	Yes
Stability	Yes	Yes
Recovery	Yes (80%-120%)	No
Performance characteristics		
Repeatability RSD (RSDr)	Yes	Yes
Reproducibility RSD (RSD_R)	Yes	Yes

4.1 Tools to validate immunoassay methods

To develop the experimental processes involved in the validation of qualitative (test) methods, there are several common tools, the most significant of which can be used are

1. Blanks of reagents and samples which are essential in order to know the proportion of the measured signal attributable to the analyte for its identification. The availability of proper sample

blanks is crucial because many of the experiments involve the use of sample blanks spiked with different amounts of target analyte.

2. Samples: Test materials: Stable food samples which give rise to a well-established binary yes/no response obtained by repeated analysis in the laboratory or by analyzing it using a confirmatory method.
3. Spiked samples/materials/solutions: Spiking with analyte giving rise to an increase in its response (e.g. from no to yes). The spiking process should be adequate in order to obtain a sample with similar conditions to the original ones.
4. Incurred samples: which are spiked at their origin (e.g. in the detection of aflatoxin B1 in peanuts naturally contaminated samples from the field).
5. Measurement standards, which can range from reference materials (RMs) to certified reference materials (CRMs). They are essential tools for validation but their unavailability for most qualitative methods is the main problem in this respect. In addition, there is also the need to have RMs and CRMs at different concentration levels of the analyte close to the threshold or cut-off limits.
6. Other tools such as statistics.

4.2 Specific information of method/kit

In order to evaluate a kit/method, information concerning both the kit and the method testing is required. The package should include a complete and detailed description of all the components of the kit/method. The description should include information on the:

1. Scope of the method
2. Unit of measurement
3. Type of assay: direct, indirect, sandwich, competitive immunoassay
4. Name of analyte (e.g., protein allergen/aflatoxin/pesticide/adulterant etc.)
5. Nature of antibody used (polyclonal/monoclonal)
6. Reporter molecule tagged to secondary antibody for ELISA: enzyme/substrate characteristics (e.g. Horse radish peroxidase/alkaline phosphatase etc.)
7. Details of particular labels or the use of particular molecular recognition element for LFIA
8. Detection method: UV/Vis/Fluorescence, etc.
9. Assay applicability: The applicability statement should contain complete information on the scope of the method i.e., which target matrix and allergen/ analyte the kit /method is applicable to. Food processing will generally lead to degradation or denaturation of the target protein, which may result in a substantial change in immunoreactivity. ELISA /LFIA kits should be evaluated for applicability to the target analyte in processed products. Empirical results from testing the method for applicability for target in processed foods should be provided. Indication of the matrix (e.g., processed food, raw materials, etc.), the type of samples and the range to which the method can be applied should be given. It can be evaluated using different matrices (raw/processed material, food/feed). Additionally, warnings on the interference with other analytes and its inapplicability to certain matrices and conditions should be included

when identified. Applicability of the methods could be determined by confirming it can be used in the intended foods with the required performance and it should be clearly stated.

10. Limitations: Relevant limitations of the method should also be addressed (e.g., interference by other analytes or inapplicability to certain situations). Limitations may also include, as far as possible, possible restrictions due to the costs, equipment or specific and non-specific risks implied for either the operator and/or the environment.
11. Operator skills requirements: A description of the practical skills and expertise necessary to properly apply the proposed method should be provided
12. Confirmatory method used: For immunoassays, antibodies may cross-react with other proteins present in the matrix; thus, it is necessary to demonstrate the selectivity of assays using another method as a confirmatory method. Empirical results from testing both methods with aliquots of the same analytical samples of known concentration should be provided.
13. Hook Effect In an antibody-based lateral flow device and plate format assay, a hook (saturation) effect could lead to a false negative result. A thorough demonstration that the working concentration range comfortably covers the practical need of target analytical samples is necessary. Therefore, empirical results from testing for a hook effect (Dilution linearity) in target matrices should be provided.

5.0 VALIDATION PARAMETERS AND ACCEPTANCE CRITERIA FOR QUALITATIVE LATERAL FLOW IMMUNOASSAY (LFIA) DEVICES (LATERAL FLOW STRIPS/DIP STICK IMMUNOASSAY ETC.)

A binary qualitative method is one that produces one out of two possible responses when it is used (e.g., visual inspection of a LFIA/ dip stick: present or absent) The parameters to be validated are listed in Table 5.1.

5.1 Selectivity-Specificity (Confirmation of Identity)

Selectivity is the degree to which a method can quantify the analyte accurately in the presence of interferents. The selectivity of a qualitative method is an ordinal concept: the extent to which analytes other than the one included in the specification interferes with the analysis. This fundamental feature of the method can also be defined as its ability to produce results which are not influenced by matrix effects. The better the selectivity, the better the certainty of identity and sample classification. The selectivity study is a single-laboratory study designed to demonstrate that the LFIA does not detect nontarget compounds, and at the same time demonstrate the ability to detect the related compounds
Organize a “selectivity” test panel of related compounds that are expected to give a positive result and non-target compounds expected to give a negative result. Document the source and origin of each test panel analyte/food (see Table 5.2 for suggested analytes/foods
Procedure
Prepare at least 20 blank samples (5 samples per combination matrix/species) and the same 20 blank samples spiked with analyte at POD concentration shall be analysed. The specificity for the tested analytes would be determined if less than 5 % of the spiked sample are negative.

Prepare at least five replicate of each non-target compound from the selectivity panel at an appropriate POD concentration.

Blind code and randomly mix the target and non-target test panel. An analyst (or analysts) not involved in the preparation of the test panel should conduct the test and record the results.

Conduct the assay using the LFIA procedure. If an individual test panel compound yields an incorrect result (a negative in the case of a target analyte or a positive in the case of a non-target compound) then the analyte may be retested with a larger number of replicates (Minimum 20).

The selectivity study (for food allergens should be designed to demonstrate that the LFIA does not produce positive results with other common food ingredients (cross-reactivity), and also to demonstrate that the method is able to detect only the target allergen in the presence of other food ingredients (lack of interference). The non-target food materials must be tested, alone and also with a spike of 10 µg/Kg (ppm) targeted allergen.

The reference materials used in the spiking must be traceable (e.g Egg powder (NIST RM-8445), Whole egg: NIST 8445, Egg white powder, LGC7422, Milk: Non-fat milk powder (NIST RM-1549), Skimmed milk powder (LGC7421) , Peanut: SRM 2387 (peanut butter), Hazelnut flour (NIST-8405) Almond flour (NIST-8404), AFM1 contaminated reference materials: .ERMVR-BD282 (<0.02 mg/kg in the powder, IRMM-JRC, Geel, Belgium); ERMVR-BD283 (0.111 mg/kg in the powder, IRMMJRC, Geel, Belgium);

Table 5.2: Suggested Test Panel for selectivity (Cross reactivities)

Target analyte/allergen	Non target analytes/foods
Aflatoxin M1	Ochratoxin A, Zearalenone, , Fumonisins, Aflatoxin M2, Aflatoxin B1, B2, G1, G2, Ochratoxin A, Patulin, nivalenol (NIV), trichothecenes (T2, Deoxynivalenol (DON), DAS, HT-2, Zearalenone), Sterigmatocystin Ampicillin, Apramycin, Albendzole, Chlortetracycline, Oxytetracycline, Tetracycline, Ceftiofur, Cefphactril , Doramectin, Diminazene, Flunixin, Febantel/Fenbendazole/Oxyfendazole, Lincomycin, Monensin, Meloxicam, Oxyclozanide, Parbendazole Sulfaquinoxaline, Sulfadiazine, Sulfadimidine, Thiabendazole, Trimethoprim . Tylosin, Virginiamycin
Aflatoxin B1	Ochratoxin A, Zearalenone, , Fumonisins, Aflatoxin M1, M2, Aflatoxin B2, G1, G2, Ochratoxin A, Patulin, nivalenol (NIV), Trichothecenes (T2, Deoxynivalenol (DON), DAS, HT-2, Zearalenone), Sterigmatocystin.
Peanut	Almond, Cereals /pseudocereals (Buckwheat, Rice, Rye, Oats), Beans/lentils/Peas, Brazil nut, Cashew, Cashew (roasted), Chestnut, Corn Coconut, Macadamia, Pecan nut, Pistachio nut, Pumpkin seed, Sesame seed, Sunflower, Soyabean, Skim milk powder, Walnut, Hazelnut, Brazil nut. Pine nut, , Wheat, Adapted from Journal of AOAC INTERNATIONAL, 105(3), 2022, 784–801

Milk*	Almond, Barley, Brazil nut. Beef, Buckwheat, Cashew. Chick peas, Cocoa, Corn meal. Crustacean/prawn, Egg, Fish, Hazelnut, Lecithin Lima bean, Oat, Peas. Peanut. Pecan, Pine nut, Pistachio, Poppy seed, Pumpkin seed, Rice–white and brown, Rye, Sesame seed, Soy bean. Split peas. Sunflower seed, Walnut, Wheat Adapted from ABBOTT ET AL. (2010) JOURNAL OF AOAC INTERNATIONAL 93, 442–450
Egg	Adzuki beans, Almond, Barley, Beef, Brazil nut, Buckwheat Cashew, Chestnut, Chick peas, Chicken, Cocoa, Coconut, Corn, Crustacean/prawn/shrimp, Duck, Fish, Gelatin (bovine), Hazelnut, Kidney beans, Kiwi, Lecithin, Lentils, Lima beans, Linseed, Macadamia nut, Milk, Oats, Octopus, Peanut, Peas, Pecans, Pine nut, Pistachio, Poppy seeds, Pork, Pumpkin seed, Rice—white and brown, Rye, Sesame, Soybean, Split peas, Sunflower seed, Turkey, Walnut, Wheat Adapted from ABBOTT ET AL. (2010) JOURNAL OF AOAC INTERNATIONAL 93, 442–450
Soybean	Adzuki bean, Almond Barley, Beans, Beef meat, Brazil nut, Buckwheat, Carrot, Cashew, Chicken meat, Chickpea, Cocoa, Coconut, Corn, Crustacean, Egg, Hake, Hazelnut, Kiwi, Lentils, Linseed, Lupine, Macadamia nut, Milk, Oat, Pea, Peanut, Pecan nut, Pine kernel, Poppy seed, Pork meat, Pumpkin seed seed, Red beans, Rice, Rye, Sesame seed, Sunflower seed, Tuna, Walnut, Wheat, White beans I. Segura-Gil et al. (2022) Journal of Food Composition and Analysis 106, 104303-104311
Gluten	Almond flour, Amaranth flour, Arrowroot, Black bean flour, Brown rice flour, Buckwheat flour, Chestnut flour, Coconut flour, Coffee powder, Corn starch/meal, Dried fruits, Egg powder, Fava bean, flour, Flax seed flour/meal, Garfava flour, Green pea flour, Guar gum, Hazelnut flour, Lentil flour, Lima bean flour, Meats, Milk powder, Millet flour, Oat flour, Potato flour/starch, Quinoa flour, Romano bean flour, Sesame flour, Sorghum flour, Soya flour, Spices, Sweet rice flour, Tapioca flour/starch, Tea, White bean flour, White rice flour, Xanthan gum, Yellow pea flour Saito et al.: Journal of AOAC International Vol. 102, No. 4, 2019
Shell fish toxins	Okadaic acid, domoic acid, tetrodotoxin (TTX),
Crustaceans and molluscs	House dust mite, insects, tree nuts (almond, cashew etc) legumes, vegetables, seeds, flours, soybean, egg, milk protein, whey, peanut, mustard, spices, rice/rice flour, wheat, chicken, pork, goat, buffalo

5.2 Applicability (Food Matrix Study)

The matrix study is a SLV/ILV study designed to demonstrate that the LFIA/kit can detect the target compound in the claimed food matrices. Analyze test portions of the claimed matrices containing the target analyte/allergen at various concentrations. The number of different matrices to be tested depends on the claims and intended use of the method. The applicability consists in demonstrating during the full validation that the method is applicable to different claimed food matrices with the same detection capabilities.

In general, a minimum of five concentrations per target compound should be evaluated for each food matrix at the claimed 95% Probability of Detection (POD see Box 1). The number of replicates at each of the five concentrations should be a minimum of 10 replicates.

The method developer should clearly identify which matrixes the method is applicable for, on the basis of their in-house data, recognizing the variability of specific formulations. The developer should also identify any matrixes that the method is known to have difficulty with, and identify clearly which states of the food allergen (raw, cooked, or both) is detected.

Both processed (e.g., such as cooked, roasted, extruded fermented, etc.) and raw samples should be represented if the assay claims to detect the target analyte (s) in such foods e.g., allergens. If the LFIA detects more than one analyte simultaneously in the same test portion, the study should be designed so that the target compounds are fortified together into some of the test portions.

A list of foods that may be used based for applicability are:

a. **Meats*:*** Fresh meat, Frozen meat, Raw marinated/minced/comminuted meat, Semi-cooked /Smoked Meat, partially heat treated and/ or smoked meat and meat product, Canned/Retorted meat product, Chilled meat, Cooked Meat/meat product, Cured/pickled meat products, Dried/Dehydrated meat/meat products, Fermented meat products sausage, lunch meat, meat substitutes etc ("meat" means all edible parts (including edible offal) of any food animal slaughtered in an abattoir that are intended for or have been judged as safe and suitable for, human consumption; "meat food products" means any product prepared from meat and other ingredients through various processing methods in which meat should be the major ingredient of all the essential ingredients
b. **Seafood**: Chilled/Frozen Finfish, crustaceans, cephalopods, molluscs, bivalves, dried or Salted and dried fish products, thermally processed, fermented, smoked, canned fish products. Fish sticks, surimi, raw fish filet, raw oysters, raw mussels, raw clams, cooked crawfish, crabmeat (fresh or pasteurized), battered and breaded fish products,
c. **Fruits, Vegetables, and Nuts**: Fresh / frozen /dehydrated or dried fruits and vegetables, fresh fruit juice, apple cider, tomato juice, fruit cubes, berries, peanut butter, coconut, fruit powders almonds, minimally processed lettuce, spinach, kale, collard greens, cabbage, bean sprouts, seed sprouts, peas, mushroom, green beans and other minimally processed fruit and vegetable products

d. **Dairy:** Dahi, Yogurt, Paneer, Khoa, Channa, hard and soft cheeses, raw or pasteurized liquid milk, infant formula, coffee creamer, ice cream, milk powders, casein, whey, non-fat dry milk/dry whole milk,
e. **Chocolate / bakery**: Frosting and topping mixes, candy and candy coating, milk, chocolate, cake mixes,
f. **Egg and egg products**: Shell eggs, liquid whole eggs, dried whole egg or dried egg yolk, dried egg whites, salad dressing.
g. **Nuts**: All nuts such as cashew nut, walnut, peanut, almond, pistachios
h. **Herbs and spices**: Oregano, pepper, paprika, black pepper, white pepper, celery seed or flakes, chili powder, cumin, parsley flakes, rosemary, sesame seed, thyme, vegetable flakes, onion flakes, onion powder, garlic flakes, allspice,
i. **Processed grains and legumes**: Flours, grits, rice corn meal, soy flour, dried yeast, cereal based complementary food, Uncooked noodles, macaroni, spaghetti, soygurt, tofu, soy beverage

5.2.1 Incurred or Fortified

The food matrix may be either an incurred or fortified analyte in the food matrix. Incurred target analyte(s) are preferred. If not available, matrix fortified with the target compound(s) (Reference material) may be used. If a matrix with incurred target analyte(s) is used, then matrix that is known to be free of the target can be used to 'dilute' it to the desired concentration. Evidence supporting homogeneity must be provided.

5.3 Sensitivity

5.3.1 Limit of detection (LOD)

LOD (CCβ) is "the smallest amount of a substance that can be reliably detected, identified and/or quantified in a sample with a statistical certainty of 1 – β. β is the probability of error usually set at 5% (significance level 0.05). LOD (CCβ) is the concentration limit at which the LFIA detects the analyte with a 5% of error. It is the concentration of an analyte which provides a signal that can be statistically distinguished from the mean signal of relevant blank samples.

The LOD is determined by applying the LFIA procedure to solutions containing progressively smaller levels of the analyte/allergen until the likelihood of producing false results (visual using naked eye or instrument-based test line readout) reaches a pre-established criterion of <5%.

To test for matrix LOD, the extract from the reference material (e.g total hazelnut protein) is spiked into a blank matrix extract (e.g., cookie) in the range of 100 ppm to 0.5 ppm (100 ppm, 50 ppm, 25 ppm, 10 ppm, 5ppm, 2.5 ppm, 1 ppm, 0.5 ppm, 0 ppm). Subsequently, the visual/instrument LOD of the LFIA in the spiked commodity is determined.

5.3.2 Probability of detection limit (POD)

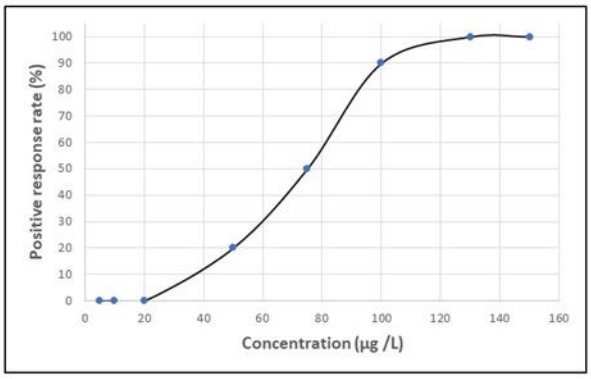
POD (or cut-off concentration) is the proportion of reliable positive analytical outcomes for a qualitative method for a given matrix at a given analyte level. POD from qualitative LFIA is calculated as the

number of positive results divided by the total number of tests at each level of added analyte (AOAC 2013). POD is concentration dependent, where false negative rates for concentrations above the limit are low – with a stated probability. Analyses to obtain POD is performed by 3 different analysts in different days and a minimum of 6 concentrations of the test are included in the study.

The POD can be determined by establishing the false positive and negative rates at a number of levels below and above the expected POD. The POD is where false negative rates for concentrations above the limit are low – with a stated probability, e.g., 95 %. See Box 1 for example.

Box 1 Determination of POD

Concentration (µg/L)	No. of positive/ negative results	% Positive
150	10/0	100
130	10/0	100
100	9/1	90
75	5/5	50
50	1/5	20
20	0/10	0
10	0/10	0
5	0/10	0



1. Measure, in random order, sample blanks spiked with the analyte at a range of concentration levels.
2. Positive or negative observations for ten replicates are recorded at each concentration level.
3. A response curve with fraction (in %) of positive results versus concentration is constructed.
4. With a criterion of < 5 % false negative results (95% positive), the POD is determined between 100 and 110 µg /L.

Adapted from "B. Magnusson and U. Ornemark (eds.) Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics, (2nd ed. 2014). ISBN 978-91-87461-59-0. Available from www.eurachem.org."

Procedure

Collect enough of each matrix to prepare more than the required number of test portions for each concentration. Prepare the required number of test portions of the matrix with the target compound(s) at the specified concentration. Blind-code, randomize, and analyze the prepared test portions.

The analyst performing the analyses should not have knowledge of the study design or the blind codes of the test portions. The analyst should be informed that the design of the study does include a certain number of "blank" samples and that both positive and negative outcomes should be expected.

Plot the positive response rate vs. concentration of target compound (see Box 1).

In the example the POD lies between 100-130 µg/L. It is possible to obtain a higher degree of reliability by performing a new set of experiments using 20 replicates of blank samples spiked with concentrations of analyte between 100 and 130 µg/L.

5.3.3 False Positive and False Negative Rate

A false positive arises when the signal from a sample containing the target analyte at a level below the reference concentration yields a "yes" response. A false positive also arises when the signal belonging to a compound different from the analyte is wrongly assigned to the target compound. On the other hand, a false negative result from the signal for a sample containing the analyte at a level above the reference concentration provides a "no" response, a mistake that its concentration is lower than the limit established. The relative concentration of the analyte in relation of the reference concentration affects the type of error made; if the concentration is slightly higher than the reference, then a false negative is expected while for concentrations lower but close to this reference value, false positives are to be expected.

Procedure

The experimental procedure to estimate the qualitative error rates at different concentration levels of analyte consists of the analysis of 10 replicates each of 6 blank samples and at 6 concentration levels of analyte (3 higher and 3 lower than the POD). These samples should be also be analyzed by a confirmatory method if it is available. The determined parameters are:

$$(\%)\ \text{False positive} = \frac{\text{Number of false positives}}{\text{total number of known negative samples}} \times 100$$

$$(\%)\ \text{False negative} = \frac{\text{Number of false negative}}{\text{total number of known positive samples}} \times 100$$

$$\text{Reliability } (\%) = 100 - (\%)\ \text{False positives} - (\%)\ \text{False negatives}$$

It is important to note that the concept of confidence intervals and statistical uncertainty needs to be applied to the risk of false positive and/or false negative results. The desired level of confidence determines the size and number of pools that need to be tested. The statistical approach is to confirm false negative and false positive rates are <5%.

The minimum number of samples that must be tested depends on the criteria for the defect rate and the level of statistical confidence is calculated using the formula

$$n = \frac{\text{Log}(\propto)}{\text{Log}\,(1-\rho)}$$

where 1-α is the confidence level and ρ is the maximum acceptable FN or FP rate. Sample sizes to assess selected criteria for FN or FP rates with varying levels of confidence are listed in Table 5.3.

Table 5.3 Sample sizes recommended for assessing FN or FP rates

False Positive/ Negative rate	Confidence level			
	80%	90%	95%	99%
< 1%	161	230	299	459
<2%	80	114	149	228
<5%	32	45	59	59
<10%	16	22	29	44

For example, if the goal is to have 95% confidence that the FN rate is <5% then carry out the test with 59 samples containing the analyte at the concentration of interest, typically the LOD or threshold, in a range of matrices. The LFIA kit is acceptable if all 59 test results are positive for the target analyte.
This sample size formula is related to the Clopper-Pearson confidence interval for Binomial proportions and frequently used for zero defect acceptance sampling plans for commodity lots. The rationale for the sample size is that when the probability of a FP/FN response is ρ for each sample then $(1 - \rho)$ n is the probability that n samples will have the correct response

5.3.4 Ruggedness

Ruggedness is the ability of a method to remain unaffected by small variations in method parameters The robustness of the LFIA method should be investigated by introducing changes in the procedure and evaluating the effects on the results. In the statistical manual AOAC describes an established technique for how to carry out such a test using a Plackett-Burman experimental design (Youden and Steiner,1975).

Procedure

1. Identify critical parameters in the procedure such as deviations in incubation times (±5% or more), and temperature (±5% or more), reagent volumes (±5% or more), extraction conditions time and temperature (±5% or more), variations on assay time and volume of sample, volume of dilution buffer. The main aim of the study is to select those variables producing greater differences in the qualitative response.
2. The most effective way to experimentally approach such a study is by using experimental design e.g., 7 parameters can be studied in 8 experiments using a Plackett-Burman experimental design (See Box 2).

Box 2 Simple procedure for the establishment of the ruggedness for a LFS method

1. Selection of 7 *a priori* critical variables (A-G) of the test method.
2. Select A high (A-G) and low (a-g) value of each variable is selected.
3. Preparation of 8 (7+1) identical aliquots of blank samples spiked with analyte at a concentration close to the threshold limit.
4. Use the LFS under different experimental conditions (1-8) according to the following scheme:

Exp No	Values of selected variables							% Reliability	Rates	
									% FP	% FN
1	A	B	C	D	E	F	G	R1	FP-1	FN-1
2	A	B	c	D	e	f	g	R2	FP-2	FN-2
3	A	b	C	d	E	F	g	R3	FP-3	FN-3
4	A	b	c	d	e	F	G	R4	FP-4	FN-4
5	a	B	C	d	e	F	g	R5	FP-5	FN-5
6	a	B	c	d	E	f	G	R6	FP-6	FN-6
7	a	b	C	D	e	f	G	R7	FP-7	FN-7
8	a	b	c	D	E	f	g	R8	FP-8	FN-9

Wherein 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, their alternative values are denoted by a, b, c, d, e, f and g.
5. Each experiment provides % reliability (R), % of false positives (FP), and % of false negatives (FN).
6. Establish the influence of each experimental factor on the variability of the LFS response

Adapted from EUR 20605 — Metrology of Qualitative Chemical Analysis BCR information series ISBN 92-894-5194-7

3. Set up experiments (analysing RM/CRMs or test samples) to monitor the effect on measurement results of systematically changing the variables. Optimally, the target analyte concentration to be tested should be at the POD/LOD.
4. Perform the LFIA assay using the same set samples at each occasion.
5. Determine the effect of each change of condition on the measurement results. Data such as % reliability, false positive and false negative rates should be used for evaluation
6. Rank the variables in order of the greatest effect on method performance.
7. Carry out significance tests to determine whether observed effects are statistically significant.
8. The results should be reflected in the assay protocol before other validation parameters are investigated. The reason for this is that a validation is linked to an assay protocol and changes in the latter might demand a new validation to be performed.
9. The ruggedness should be evaluated by the analysis of at least 10 different blank materials (i.e. specificity) and 10 different materials spiked (or incurred) at the level of interest (i.e. POD). It is recommended to perform the ruggedness studies as a blind test (unknown samples).

Acceptance criteria: The method should provide the expected results (False negative and False negative should be <5%) irrespective of these minor alterations.

5.3.5 Stability

Stability is defined as the ability of a substance to remain unchanged over time under stated or reasonably expected conditions of storage and use. Stability of LFIA kit indicates its ability to maintain expected, and consistent, performance over time without degradation. It relates to performance of a product within a specified time period.

5.3.5.1 Real-time stability testing

Real-time stability testing must include long-term stability covering the shelf-life, and in-use stability. At the time of submitting an application for FSSAI approval, all stability claims shall be justified by adequate data. Long-term stability and in-use stability studies must be carried out

During long-term stability testing, the LFIA kit shall be stored under the storage conditions recommended by the manufacturer (e. g. temperature, humidity). Examinations shall be undertaken at specified time intervals for (%) False positive, (%) False negative and reliability using concentrations of the target analyte set at LOD or threshold level. The time intervals chosen should encompass at least the whole of the target or declared shelf life and, if appropriate, continue until significant degradation in the performance of the kit can be determined. The number of time intervals shall be appropriately chosen so that trends may be discerned from variability of the data.

The minimum number of batches to be evaluated for real time stability

– 3 batches LFIA shelf-life (long-term stability);

– 1 batch LFIA for in-use stability of reagents, for example buffers etc which are to be reconstituted or were initially packed under vacuum.

5.3.5.2 Accelerated stability testing

Typically, the LFIA kit is exposed to different elevated temperatures. The manufacturer shall specify these stress conditions and the testing intervals. The test conditions chosen should, where appropriate, demonstrate significant deterioration of the kit over the testing period in order to allow for a mathematical extrapolation.

Procedure

1. Unopened LFIA test kits from the same production lot must be used. The kit in the final configuration (unopened package), in particular in terms of volume and material of the primary container must be used.
2. The kits under investigation shall be stored under the conditions defined by manufacturer (e.g., 4 °C) until the accelerated testing programme starts.
3. The kits at day zero, shall then be shifted and stored to the defined stress conditions (e.g 21/ 37 °C). LFIA shall be removed at the specified time intervals (1,2, 3….n days) and analysed for (%) False positive and (%) False negative and reliability of using concentrations of the target analyte set at LOD or threshold level. The kit is acceptable when false positive and false negative rate is <5% or same as the kit stored at 4 °C.
4. Continue the test till you obtain unacceptable results
5. One day test kit storage at 21 °C and 37°C is equivalent to 5 days and 35 days storage at 4 °C, respectively (Deshpande, 1996)

6.0 VALIDATION PARAMETERS AND ACCEPTANCE CRITERIA FOR QUANTITATIVE IMMUNOASSAY METHODS (ELISA KITS)

6.1 Selectivity

Selectivity can be defined as “the ability of the method to measure and differentiate the analytes in the presence of components that may be expected to be present in the sample”. The terms “selectivity” and “specificity” are often used interchangeably while they are different. Selectivity is something that can be graded while specificity is an absolute characteristic. Selectivity is the preferred terminology. A selectivity study requires knowledge about the analyte in terms of its structure, modifications, derivates etc. Do structurally similar metabolites of the analyte or other protein similar to the allergen interfere with the assay?

Selectivity of the method is routinely demonstrated by analyzing blank samples of the food from multiple sources. Depending on the intended use of the assay, different analytes/food selectivity assessment (See Table 5.2). Potential interfering substances in the food include endogenous matrix components such as metabolites, decomposition products During validation, it must be proven that the assay is free of potential interfering substances

Investigating exogenous interference involves determining the cross-reactivity of molecules (see Table 5.2) that could potentially interfere with the binding interaction (antigen-antibody), including molecules structurally related to the analyte/allergen.

Assay specificity is defined as the ability of an antibody to produce a measurable response only for the analyte of interest. Cross-reactivity is a measurement of antibody response to substances other than the analyte. A wide selection of foods and ingredients should be tested for cross-reactivity with the ELISA (Table 5. 2). Use the following procedures for cross-reactivity testing

Procedure for selectivity

1. Identify substances that are structurally/physiochemically similar to the one that the assay is developed for.
2. Conduct an analysis of blank samples in the matrix from ≥ 10 individual sources
3. Investigate to what degree the measurements are interfered by spiking blank food samples with substances identified above and in combination with the analyte of interest to determine its ability to cause interference.

4. The spiking concentration should be at least two times the reference limit/LOD.
5. For ≥ 90% of sources, unspiked matrix should be BQL, and spiked samples should be ± 25% at LLOQ, and ± 20% at highest QC spike.

Procedure for cross-reactivity

1. Extract food samples in the same extraction buffer and conditions specified.
2. Add the full-strength extracts (i.e., without dilution) to ELISA.
3. If a positive result is observed, serially dilute the extracts to characterize the extent of the cross-reactivity.
4. The greater the number of items tested for cross-reactivity the better.
5. Food items should be raw and processed as they would normally be consumed

6.2 Sensitivity

6.2.1 Limit of Detection (LOD) and Limit of Quantification (LOQ)

The LOD is defined as the minimum concentration of analyte that can be reliably distinguished from the background. The LOD should be estimated by a statistical analysis of the calibration data according to the ISO standard ISO 11843-2 for linear data, or ISO 11843-5 (7) for linear

and nonlinear data, using as default probabilities a = b = 0.05, where a and b represent the probability of a false positive and false negative, respectively.

Two methods have been commonly used for determining the LOD:

$$\mathrm{LOD1} = 3.3 \times \frac{\mathrm{S1}}{\mathrm{b}}$$

where S1 is the standard deviation of the blank absorbance, b is the slope of the linear calibration curve

$$\mathrm{LOD2} = \mathrm{Xb} + 3.3 \times \mathrm{S2}$$

Xb is the mean of the blank concentration, and S2 is the standard deviation of the blank concentration.

The coefficient 3.3 is based on false-positive and false negative risks of 5% each (α = β = 0.05) (Abbot et al., 2010, ISO 2008)

The LOQ is defined as the minimum concentration of analyte that can be quantitatively determined with suitable precision and accuracy. The LOQ of the ELISA can be calculated using either of the following formulas:

$$\mathrm{LOQ1} = 10 \times \frac{\mathrm{S1}}{\mathrm{b}}$$

where S1 is the standard deviation of the blank absorbance, b is the slope of the linear calibration curve

$$\mathrm{LOQ2} = \mathrm{Xb} + 10 \times \mathrm{S2}$$

Xb is the mean of the blank concentration, and S2 is the standard deviation of the blank concentration

Procedure

LOD and LOQ using a calibration curve:

1. Dilute the standard to appropriate concentrations in extraction buffer to achieve a linear standard curve (minimum five concentrations in triplicate).
2. Follow the ELISA procedures
3. Plot the absorbance against concentration and carry out regression analysis to obtain linear regression equation y = a + bx

where y is the absorbance, x is the concentration, a is the y intercept of the standard curve, and b is the slope of the standard curve.

4. Take the mean of the slope of three independent calibration curve
5. Use the formulas described above calculate LOD and LOQ

Use the following steps to determine LOD2 and LOQ2:

1. Dilute the standard to appropriate concentrations in extraction buffer (minimum 5 concentrations.
2. Follow the ELISA procedure.
3. Plot the absorbance against concentration and find the linear regression equation y = a + bx
where y is the absorbance, x is the concentration, a is the y intercept of the standard curve, and b is the slope of the standard curve.
4. Calculate the mean and standard deviation of the blank concentration using the regression equation.
5. Use the formulas described above in this section to calculate LOD2 and LOQ2

Calibration data from at least three analysts over a minimum of three different runs should be included, preferably using different instruments.

Note: the method of curve fitting as well as the number of calibration values used and the true concentrations of the standards determine the accuracy of the calibration curve. The four-parameter-logistic model and the approximation of a cubic spline which are more accurate may be used for enzyme immunoassays

6.2.2 Working Range

The working range for a method is defined by the lower and upper limits of quantification (LLOQ and ULOQ, respectively).

Upper limit of quantification (ULOQ): is defined as a mean value of 10 duplicates of maximally achieved absorbances in the linear part of the standard curve, from which three standard deviations have been subtracted. Subtraction of the multiplied standard deviations achieves accuracy in this range from 80–120%.

Lower limit of quantification (LLOQ): is the lowest concentration of analyte in a sample which can be quantified reliably, with an acceptable accuracy and precision. In practice this is a mean value of the smallest result measured in the linear part of the curve to which three standard deviations have been added.

6.2.3 Linearity of Dilution

Dilution linearity should be assessed to confirm: (i) that measured concentrations are not affected by dilution within the calibration range and (ii) that sample concentrations above the ULOQ of a calibration curve are not impacted by hook effect (i.e., a signal suppression caused by high concentrations of the analyte), whereby yielding an erroneous result. Dilution linearity is performed to demonstrate samples with high concentrations of the analyte can be diluted and still give a reliable result. It determines to what extent the dose–response of the analyte is linear in a particular diluent within the range of the standard curve. Thereby dilution of samples should not affect the accuracy and precision. At the same time, the presence of a hook effect, i.e., suppression of signal at concentrations above the ULOQ, is investigated.

Note: Dilution linearity should not be confused the linearity of the calibration curve.

Procedure

1. Spike three samples (undiluted) with calibrator stock solution, as high as possible.

Note: if possible, spike (undiluted) samples with 100- to 1000-fold the concentration at ULOQ using the calibrator stock solution.

2. Make serial dilutions of the spiked samples, using sample diluent in small vials until the theoretical concentration is below LLOQ.

Note: the dilution should be performed using vials and not directly in the wells of the ELISA plate.

3. Analyze the serial dilutions to compensate for the dilution factor. At least 5 replicates per dilution factor should be tested in one run to determine if concentrations are accurately and precisely measured within the calibration range.

4. Calculate the mean concentration of analyte for the dilutions that fall into the range of LLOQ and ULOQ. Also calculate for each sample the % change in concentration from previous dilution. See example below

$$\%\ change\ in\ concentration\ from\ previous\ dilution = \frac{c\ (concentration\ in\ 1{:}4\ dilution)}{c\ (concentration\ in\ 1{:}8\ dilution)} \times 100$$

$$\%\ change\ in\ concentration\ from\ previous\ dilution = \frac{c\ (concentration\ in\ 1{:}8\ dilution)}{c\ (concentration\ in\ 1{:}16\ dilution)} \times 100$$

Note: The analyte concentration at each dilution should be 100% (+/- 20%) of the concentration measured at the previous dilution, which then demonstrates dilutional linearity. The calculated concentration for a dilution that falls into the range of LLOQ and ULOQ should be within the acceptance criteria for the precision. The precision (%CV) of the concentrations determined at each level should not exceed 15%, except at the LLOQ, where it should not exceed 20%.

5. Also, plot the signal (absorbance) against the dilution factor to investigate if the signal is suppressed at much higher concentrations than the ULOQ of the measurand (“hook effect”).

6.2.4 Accuracy

Accuracy (Trueness) is defined as “The closeness of agreement between the average value obtained from a large series of test results and an accepted reference value”. Accuracy is typically verified through studies using one of the following

1. Reference material comparison: analytical procedure is applied to an analyte of known purity (e.g., a RM/CRM, and the measured versus theoretically expected result is evaluated. The reference value is derived directly from a CRM or from materials that can be traced to the CRM.

2. Spiking study: The analytical procedure is applied to a matrix of all components except the analyte where a known amount of the analyte of interest has been added. The results from measurements on unspiked and spiked samples are evaluated.

3. Orthogonal procedure comparison: The results of the proposed analytical procedure are compared with those of a second well characterized procedure that ideally applies a different measurement principle (HPLC/LC-MS/MS etc). The accuracy of this second procedure should be reported.

It should be demonstrated under regular test conditions of the analytical procedure (e.g., in the presence of sample matrix and using described sample preparation steps).

Procedure

Prepare the CRM/spiked material at a minimum of three different concentration levels. Minimal requirements for this are one close to ULOQ, one close to LLOQ and one in the middle of the range. Each of these concentrations must be measured in five replicates in the same plate.

Accuracy should be reported as the percent difference between the measured mean ($\bar{X}$ is calculated from five replicates) and the accepted true (X_{ref} assigned value) together with the confidence intervals.

Accuracy at each concentration level should be within ±20% of the reference values, except for the LLOQ and ULOQ, which should be within ±25%.

6.2.5 Precision

Precision is defined as “The closeness of agreement between independent test results obtained under stipulated conditions. These are repeatability (r), intermediate precision, and reproducibility (R). Repeatability is the variability observed when as many factors as possible, e.g., laboratory, analyst, days, instrument, reagent lot, are held constant and the time between the measurements is kept to a minimum Repeatability is sometimes called within-run or within-day precision while intermediate precision is also known as between-run or between day repeatability as opposed to reproducibility conditions where all factors are varied and measurements are carried out over several days.

6.2.5.1 Repeatability

Repeatability should be assessed using: a) a minimum of 9 determinations covering the reportable range for the procedure (e.g., 3 concentrations /5 replicates each); or b) a minimum of 6 determinations at 100% of the test concentration. Three samples with different levels have been suggested as a general rule to cover a wide measuring range

Procedure

1. Collect samples with known three concentrations (high, middle and low concentrations of the analyte.
2. Make 25 aliquots of each concentration and store at −80°C pending analysis.
3. At day 1 to 5 measure 5 replicates on each sample. Note: the days need not to be consecutive, only different.
4. Calculate the mean value, SD, %CV for both the repeatability and intermediate precision.

Within-run and between-run precision of the sample concentrations determined at each level should not exceed ±20%, except at the LLOQ and ULOQ, where it should not exceed ±25% and Intra-Assay: CV<10%

Within-run precision data should be reported for each run. If the within-run precision criteria are not met in all runs, an overall estimate of within-run precision for each QC level should be calculated. Between-run (intermediate) precision should be calculated by combining the data from all runs.

6.2.6 Recovery

ELISA is susceptible to matrix effects that affect sample extraction. Recovery from different food matrices can be tested using spiked or incurred samples. Recovery is reported as a percentage of the known amount of an analyte carried through the sample extraction and processing steps of the method. Recovery is determined by comparing the analyte response in a food matrix that is spiked with the analyte and processed, with the response in a blank sample that is processed and then spiked with the analyte. Recovery of the analyte does not need to be 100%, but the extent of recovery of an analyte and of the incurred sample (if used) should be consistent. Recovery experiments are recommended to be performed by comparing the analytical results for extracted samples at multiple concentrations, typically three concentrations (LLOQ, medium and ULOQ).

Procedure

1. Collect five food matrices where the concentrations of the analyte have previously been determined
2. Spike each food, using calibrator (analyte e.g aflatoxin or peanut protein) stock solution, to expected concentrations that are evenly distributed (low, medium, high) concentration.
3. Extract the analytes from the unspiked and spiked samples and carry out ELISA in the same plate.
4. The recovery (*R*) is calculated from the difference between the results obtained before and after spiking as a fraction of the added amount. The recovery (*R*) is calculated from the difference between the results obtained before and after spiking as a fraction of the added amount.

$$\% \text{ Recovery} = \frac{\text{c1} - \text{c2}}{\text{c3}} \times 100$$

Where:

*c*1 = mean of measured concentration in spiked sample
*c*2 = mean of measured concentration in unspiked sample
*c*3 = concentration of spiking

The accepted recovery range of 80–120% for spiked samples. For incurred samples, a recovery range of 50–150% will be considered acceptable so long as they can be shown to be consistent

6.2.7 Ruggedness and Lot-to-Lot variability

Ruggedness refers to the ability of a method to resist changes in the final results when minor deviations are made in the experimental conditions described in the procedure. The ruggedness of the method should be investigated by performing experiments in which specific parameters are changed to determine the impact on the experimental result. In particular, the effect of deviations in incubation times, reagent volumes, extraction conditions (time and temperature) should be investigated. The following parameters that are crucial for quantitative ELISA are studies,

- buffer temperature,
- volume of buffer
- incubation temperature,
- sample incubation time,
- primary antibody incubation time
- secondary antibody incubation time,
- number of washes before colour development,
- colour development time,
- enzyme substrate volume and the like.

It is recommended that deviations for time and volume be investigated at ±5% or more, and incubation temperatures tried at ±3°C or more.
The most effective way to experimentally approach such a study is by using a Plackett-Burman experimental design (e.g., 7 parameters can be studied in 8 experiments See Box 2).
For design procedure see Section 4.2.5 and Box 2
Acceptance criteria: The kit should provide the expected accuracy results irrespective of these minor alterations.

6.2.8 Stability

The shelf life should include the stability of all the reagents provided with the ELISA test kit, ideally through real time testing of reagents under normal storage conditions. Accelerated stability testing at higher-than-normal storage temperatures can also be used to estimate stability (see Section 4.2.6).

An expiration date for each test kit should be clearly indicated, along with appropriate conditions for storage before use

Procedure

1. Unopened quantitative ELISA test kits from the same production lot are grouped into three sets and stored at the storage temperature usually recommended by the manufacturer viz 4 °C.
2. Analyse one kit in the batch for accuracy, linearity and dilution linearity. This is the reference value
3. Store the required number of kits at 4°C, 21 °C and 37°C.

4. One kit from each temperature point is used to test concentrations of the target analyte set at the LLOQ and ULOQ.
5. Determine the 1) accuracy (five replicates of LLOQ), 2) linearity (minimum six calibration standards) and 3) dilution linearity (minimum of three dilutions of ULOQ) the sample at day zero and at regular time intervals. Suggested time intervals are every day for kits stored at 21 °C and 37°C.
6. The accuracy, linearity and dilution linearity must be the same as the reference value i.e., day 0.
7. Continue testing until the results are not acceptable
8. One day test kit storage at 21 °C and 37°C is equivalent to 5 days and 35 days storage at 4 °C, respectively (Deshpande, 1996)

6.2.8.1 Long term stability test

The long-term stability test, also called the stability test within the period of validity, is performed at normal storage condition (4 °C) to validate the efficiency of kits stored at normal condition. It is usually used to reflect the stability of the kit directly since the result of the long-term stability test is the final basis to determine the shelf life and storage condition. Carry out the recovery test, linearity test and precision test of the kits within 30 days before the expiration date and within 30 days after the expiration date to validate the kit.

6.2.8.2 Lot-to-Lot variability

A small number of test kits from each lot should be set aside for comparison with previous or future lots. When a new lot of test kits is produced, it should be tested against the previous lot. New lots should have characteristics similar to those of the previous lots. For example, a positive control sample, such as an incurred test sample or spiked sample, should be analyzed with each new lot to be sure that consistent results are achieved. Information on lot-to-lot variability should be provided by the kit manufacturer as part of the data submission package.

7.0 INFORMATION TO BE PROVIDED WITH METHOD VALIDATIONDOCUMENTS

A listing of information that should be provided from the validating laboratory when the results for a SLV/ILV are prepared for review.
This list of information is in addition to all the results obtained for the validation criteria/insert accompanying the kit.

For Qualitative and Quantitative Assays:

A. Assay Design

- Type of assay: ELISA/Sandwich ELISA/Competitive ELISA/LFS
- Name of analyte (Protein/Aflatoxin)
- Nature of antibody (Monoclonal/polyclonal)
- Reporter molecule: Enzyme/substrate characteristics
- Detection method: UV/Vis/Fluorescence etc.
- Any other reporter molecules.

B. Sample Preparation and Protein Extraction

- Form and quantity of sample required. Include information on subsampling or sample compositing as well as relevant aspects of handling and storage.

C. Method or kit used for extraction. Include any relevant modifications

D. ELISA/LFIA Conditions

- Reaction: reaction volume; identities and concentrations of all reaction components, including buffer/diluent/additives.
- Platform: State make and model of instrument used for detection as well as name and version of accompanying software. Include brief descriptions of physical format (e.g., 96/384 well or other) and optical system.
- Assay conditions and detailed methodology. Include washing and incubation steps

E. Data Analysis

- Specify which software program and version was used for data analysis.
- Report and explain any adjustments or other software default analysis parameters.
- Include all data: tables, calibration curves, recovery data etc

8.0 CONCLUSION

Based on verification of the submitted documents comprising of data on the above validation parameters meeting the acceptance criteria as per the procedure, the method will be declared as fit for intended use.

9.0 GLOSSARY

Accuracy: Closeness of agreement between a quantity value obtained by measurement and the true value of the measurand.

Analytical sample: Sample prepared from the laboratory sample by grinding, if necessary, and homogenization.

Certified reference material: Use of known materials can be used to assess the accuracy of the method, as well as obtaining information on interferences.

Cross-Reactivity: Degree to which binding occurs between an antibody and antigenic determinants, and a target sequence, which are not the analyte of primary interest.

ELISA- For the purposes of this document, ELISA is defined as “an analytical procedure characterized by the recognition and binding of specific antigens by antibodies”. This definition is not meant to be restrictive and encompasses other related binding-based technologies.

False Positive and False Negative: The false positive and false negative are the probability that respectively a negative sample would be classified positive or a positive one would be regarded as negative.

Lateral flow Immunoassays- The lateral flow immunoassay (LFIA) is a paper/membrane-based platform for the detection and quantification of analytes in complex mixtures, where the sample is placed on a test device and the results are displayed within 5–30 min.

Limit of Application (LLA): Manufacturers or method developers are free to define an LLA at whatever level of confidence they choose. This value may be higher than the LOQ and represents a level below which the method developer does not support or recommend use of the method.

Limit of Detection (LOD): LOD is the lowest amount or concentration of analyte in a sample, which can be reliably detected but not necessarily quantified. Experimentally, methods should detect the presence of the analyte at least 95% of the times at the LOD, ensuring ≤5% false negative results.

Limit of Quantification (LOQ): LOQ is the lowest amount or concentration of analyte in a sample, which can be reliably quantified with an acceptable level of precision and trueness.

Linearity of Dilution: Dilutional linearity is performed to demonstrate that a sample with a spiked concentration above the ULOQ can be diluted to a concentration within the working range and still give a reliable result.

*Matrix***:** Totality of components of a material system except the analyte.

Precision: Closeness of agreement between quantity values obtained by replicate measurements of a quantity, under specified conditions. Precision is usually expressed as the standard deviation or relative standard deviation.

Relative Repeatability Standard Deviation (RSDr): The relative standard deviation of test results obtained under repeatability conditions.

Probability of detection (POD): The probability of a positive (i.e., presence detected) analytical outcome for a qualitative method for a given matrix at a given concentration. It is estimated by the expected ratio of positive to negative results for the given matrix at the given analyte concentration.

*Qualitative ELISA***:** can be used to achieve a yes or no answer indicating whether a particular antigen is present in a sample, as compared to a blank well containing no antigen or an unrelated control antigen.

Quantitative ELISA: data can be interpreted in comparison to a standard curve (a serial dilution of a known, purified antigen) in order to precisely calculate the concentrations of antigen in various samples.

Recovery: The fraction or percentage of added analyte that is recovered when the test sample is analysed using the entire method.

Repeatability: Variation arising when all efforts are made to keep conditions constant by using the same instrument and operator (in the same laboratory) and repeating during a short time period. Expressed as the repeatability standard deviation (SDr); or % repeatability relative standard deviation (%RSDr)

Reproducibility: Variation arising when identical test materials are analyzed in different laboratory by different operators on different instruments. The standard deviation or relative standard deviation calculated from among-laboratory data. Expressed as the reproducibility standard deviation (SDR); or % reproducibility relative standard deviation (%RSDR).

Robustness: The robustness of a method is a measure of its capacity to remain unaffected by small, but deliberate deviation from the experimental conditions described in the procedure.

Semi-Quantitative ELISA: can be used to compare the relative levels of antigen in assay samples, since the intensity of signal will vary directly with antigen concentration.

Selectivity: The ability of the bioanalytical method to measure and differentiate the analytes in the presence of components that may be expected to be present

Specificity: The property of the method to respond exclusively to the characteristic or the analyte of interest.

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

Verification: Verification is the confirmation, through the provision of objective evidence, that specified test characteristics have been fulfilled.

11.0 REFERENCES

Abbott M, Hayward S, Ross W et al (2010) Validation procedures for quantitative food allergen ELISA methods: community guidance and best practices. J AOAC Int 93:442–450

Appendix D Guidelines for Collaborative Study Proceduresto Validate Characteristics of a Method of Analysis (2023) https://doi.org/10.1093/9780197610145.005.004 Pages AD-1–AD-12

Appendix M Validation Procedures for Quantitative Food Allergen ELISA Methods: Community Guidance and Best Practices https://doi.org/10.1093/9780197610145.005.013 Pages AM-1–AM-8

Appendix M Validation Procedures for Quantitative Food Allergen ELISA Methods: Community Guidance and Best Practices (2023)

CAC/GL 74-2010, Guidelines on performance criteria and validation of methods for detection, identification and quantification of specific DNA sequences and specific proteins in foods.

CLSI EP25-A (2009) Evaluation of Stability of In-Vitro Diagnostic Reagents, Approved Guideline

Deshpande, S.S.: Enzyme Immunoassays From Concept to Product Development, pp.397(1996), Chapman and Hall, International Thomson Publishing, New York

Guidance for Industry. Bioanalytical method validation. *Fed Regist* (2001) 66:28526–7.

ISO (1994) ISO 5725-2 Accuracy (trueness and precision) of measurement methods and results. Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method. International Organization for Standardization

ISO (2008) ISO 11843-5:2008 Capability of detection. Part 5: Methodology in the linear and non-linear calibration cases. International Organization for Standardization

ISO 5725-2. Accuracy (Trueness and Precision) of Measurement Methods and Results – Part 2: Basic Method for the Determination of Repeatability and Reproducibility of a Standard Measurement Method (2019)

ISO/IEC 17025. *General Requirements for the Competence of Testing and Calibration Laboratories* (2017).

Koerner TB, Abbott M, Godefroy SB et al (2013) Validation procedures for quantitative gluten ELISA methods: AOAC allergen ELISA Validation for Food Allergen Quantification 303 community guidance and best practices. J AOAC Int 96:1033–1040

Official Methods of Analysis (2002) AOAC INTERNATIONAL, Gaithersburg, MD, Appendix M, https://doi.org/10.1093/9780197610145.005.013 P

Plackett RL, Burman JP. The design of optimum multifactorial experiments. *Biometrika* (1946) **33**:305–25. doi:10.1093/biomet/33.4.305

Thompson, M., Ellison, S., & Wood, R. (2002) *Pure Appl.Chem*. **74**, 835–855

W. J. Youden, E. H. Steiner, Statistical Manual of the AOAC, AOAC International, 1975, ISBN 0-935584-15-3.

Chapter-6

GUIDELINES FOR MULTI-LABORATORY VALIDATION (COLLABORATIVE STUDY)

1.0 INTRODUCTION

A multi-laboratory validation (MLV) also known as a collaborative study, is a test of how a method/kit will perform in in actual practical applications in real life settings. This study also provides confidence in the performance statistics and demonstrates the reproducibility of the results and provides opportunities for improvements of the method/kit. The study is carried out to validate the data provided by a single laboratory validation exercise. Validation through a collaborative study is expensive to undertake and usually follows only after the method has shown acceptable performance in a single laboratory validation study. The user of a method /kit, which has been found to be fit for the intended purpose through a MLV, needs only to demonstrate the performance characteristics stated in the method/kit are achieved. Such verification for the correct use of a method is essential to meet regulatory compliance.

2.0 PURPOSE

The purpose of a MLV is to determine estimates of the attributes of a method, particularly the "precision" of the method that may be expected when the method is used in actual practice. Precision refers to repeatability and reproducibility. The reproducibility (between-laboratory variation) represents a systematic error that reflects variation arising from varying environmental conditions (e.g., condition of reagent and instruments, variation in calibration factors, and interpretations of the steps of the method) associated with the laboratories used in the study. Therefore, it is important to identify the causes of the differences among laboratories so that they may be controlled.

3.0 SCOPE

The scope of this document is to provide comprehensive technical guidelines for conducting MLV studies of methods/rapid kits submitted for FSSAI approval. Both the collaborative study design and data are subject to scrutiny before acceptance by FSSAI.

These guidelines are applicable to any candidate method/rapid kit whether proprietary or non-proprietary, that is submitted to FSSAI for approval. The approval process requires an Independent Laboratory Validation (ILV) and an MLV (collaborative study.).

The guidelines described here are based on Appendix D: AOAC Guidelines for collaborative study procedures to validate characteristics of a method of analysis and Appendix M Validation procedures for quantitative food Allergen ELISA Methods: Community Guidance and Best Practices (2023).

Note: In this document the Multiple Laboratory Validation will be referred to as the 'Collaborative Study'

4.0 PRELIMINARY WORK TO BE COMPLETED BY METHOD/KIT DEVELOPER/ LABORATORY PRIOR TO A COLLABORATIVE STUDY

4.1 Determination of purpose and scope of the method/kit

i. Determine purpose of the study (e.g., to determine attributes of method,),
ii. Define the type of method (qualitative/semiquantitative/quantitative)
iii. Determine type of analysis (chemical/microbiology/immuno-assay/DNA-based)
iv. The probable application of the method (enforcement, surveillance, monitoring, acceptance).
v. Also, on the basis of the relative importance of the various method attributes (bias, precision, specificity, limit of determination), select the design of the collaborative study. Primarily the study should be designed to determine the precision (repeatability and reproducibility)

4.2 Optimization of method/kit performance and within-laboratory attributes

Collaborative studies require considerable effort and should be conducted only on kits/methods that have undergone a single laboratory validation. Do not conduct collaborative study with an unoptimized method/kit. Conduct as much experimentation within your laboratory which should include, as applicable, information on the following:

i. Preliminary estimates of precision—estimates of the total within-laboratory standard deviation of the analytical results over the concentration range of interest; as a minimum at the upper and lower limits of the concentration range, with particular emphasis on any standard or specification value
ii. Test the variability at different days and with different calibration curves, by the same or different analysts within a single laboratory. This total within-laboratory estimate reflects both between run (between-batch) and within-run (within-batch) variability.
iii. Systematic error (bias)—estimates of systematic error of the analytical results over the concentration range and in the food categories of interest; as a minimum at the upper and lower limits of the concentration range, with particular emphasis on any standard or specification value. The results obtained by applying the method to relevant reference materials should be noted.
iv. Determine analytical function (response vs concentration in matrix, including blank) to determine applicability to food categories of interest.
v. Recoveries—the recovery of "spikes" added to real materials and to extracts, digests or other treated solutions thereof.
vi. Applicability—the ability of the method to identify and measure the physical and chemical forms of the analyte likely to be present in the materials. Delineate the range of applicability to the food matrices or commodities of interest

vii. Test for interferences (specificity): Interference is the effect of other substances that are likely to be present at appreciable concentrations in matrices of interest and that may interfere in the determination.
   a. Test for effects of impurities, ubiquitous contaminants, flavors, additives, and other components expected to be present and at usual concentrations.
   b. Test for nonspecific effects of matrices.
   c. Test effects of transformation products, if method is to indicate stability, and metabolic products, if tissue residues are involved.

viii. Method comparison—the results of comparison of the application of the method with existing tested or reference or FSSAI -methods intended for similar purposes.

ix. Calibration procedures—the procedures specified for calibration and for blank correction must not introduce important bias into the results. Determine calibration function (response vs concentration in pure or defined solvent) to determine useful measurement range of method. For some techniques, e.g., immunoassay, linearity is not a prerequisite. Indicate any mathematical transformations needed.

x. Develop performance specifications for instruments and suitability tests for systems (which utilize columns or adsorbents) to ensure satisfactory performance of critical steps (columns, instruments, etc.) in method.

xi. Have method tried by analysts not involved in its development

4.3 Written description of the method/instructions for use of kit

The method or steps involved in use of kit must be clearly and unambiguously written so that it can be replicated. A collaborative study involves practical testing of the written version of the method, in its specific style and format, by a number of laboratories on identical materials. Ensure that the following are covered in the method/instruction sent to participating laboratories:

i. Always express reagent concentrations in terms of mass (or volume) per volume (or mass); never in terms requiring the analyst to recalculate or look up formula weights, e.g., Molar in place of moles. Moles may be used, particularly with volumetric standards, but only in addition to mass and volume. Many errors are caused by incorrect recalculation of formula weights.

ii. Clearly specify requirements for chromatographic materials, enzymes, antibodies, and other performance related reagents.

iii. Clearly describe and explain every step in the analytical method so as to discourage deviations e.g., exact time and temperature of incubation. Whether a timer should be used.

iv. Use imperative directions; avoid subjunctive (uncertain) and conditional expressions (e.g. you may add) as options as far as possible.

v. Clearly describe all safety precautions needed.

vi. Edit method for completeness, credibility (e.g., buffer pH consistent with specified chemicals, volumes not greater than capacity of container), continuity, and clarity.

vii. Check for inclusion of performance specifications and system suitability tests, defined critical points, and convenient stopping points.

viii. Incorporate physical or chemical constants of working standards solutions, e.g., absorptivity, recoveries, etc., or properties of operating solutions and chromatographic materials, e.g., pH,

volumes, resolution, etc., and any other indicators (e.g., sum equals 100%) that suggest analysis is proceeding properly.

ix. If possible, conduct a pilot study involving 2–3 laboratories to see if the written method works.

4.4 Conducting the collaborative study

4.4.1 Study director

The laboratory conducting the study selects a Study Director who oversees the entire study. The Study Director has overall responsibility for the technical conduct of the study, as well as for the interpretation, analysis, documentation and reporting of results, and represents the single point of control for a collaborative study. The flow diagram shown in Figure 6.1 summarizes the process of a collaborative study.

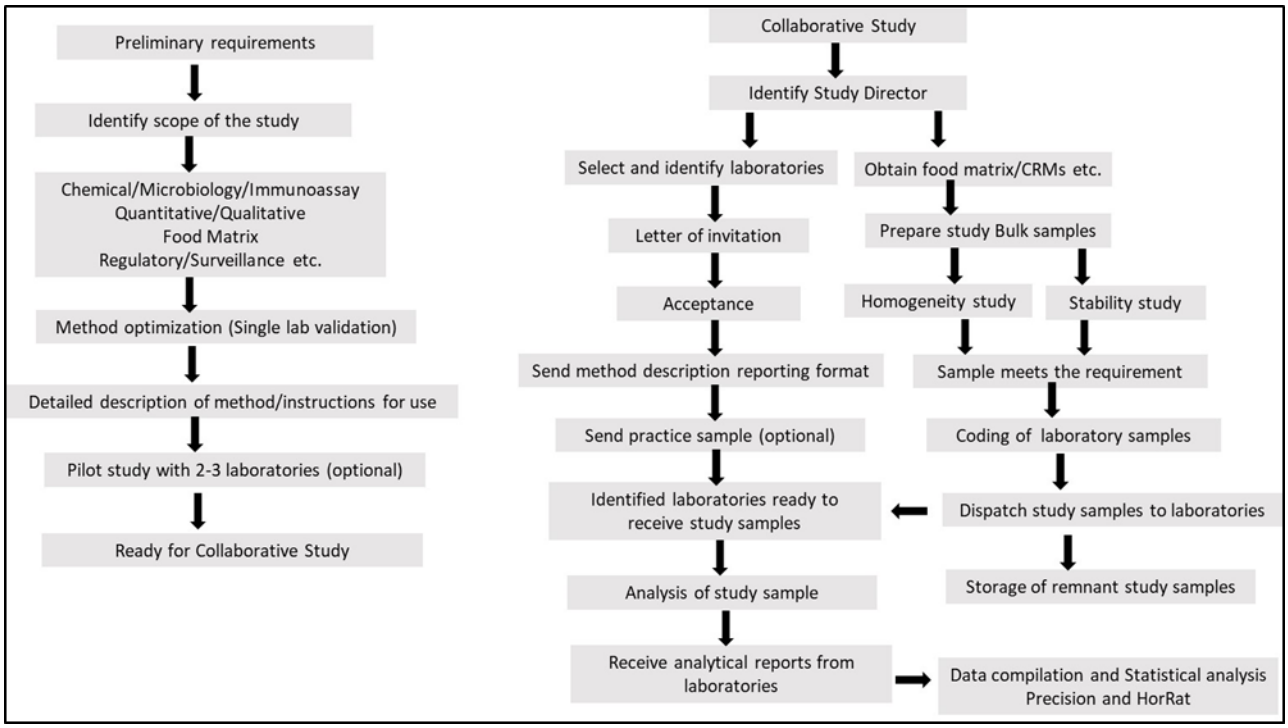


Figure 6.1: Flow diagram summarizing the process of a collaborative study.

4.4.2 Selection of laboratories

Laboratories invited to participate must have personnel with sufficient experience in the basic techniques employed (e.g. for a PCR method analyst must be experience in DNA extraction and basic knowledge of PCR set-up). Experience with the method itself is not a prerequisite for selection.

Collaborators are chosen by the organizing laboratory of the collaborative study from a diversity of laboratories with interest in the method, including regulatory agencies, industry, and universities, through literature search, NABL website, FSSAI notified labs.

4.4.3 Letter of invitation

Prepare a formal letter to the individual responsible for the participant laboratory (Head/Quality Manager/Technical lead). State your reason for selecting that laboratory (e.g., as a volunteer or has responsibility or familiarity with the problem or method), estimated number of man-hours required for performance, number of test samples that will be sent, number of analyses to be carried out, expected date for test sample distribution, and target date for completion of the study. Emphasize the importance of management support in assigning the necessary time for the project. Enclose a copy of the method and a return response form or card (with postage affixed, if appropriate), requiring only a check mark for acceptance or refusal of the invitation, a signature, space for address corrections, telephone and fax numbers, e-mail, and date.

Acceptance of the invitation should be followed by a letter suggesting that a Laboratory Coordinator be appointed who would be responsible for receiving and storing the study materials, assigning the work, dispensing study materials and information related to the study, seeing that the method is followed as written, collating the data, assuring that the data are correctly reported, and submitting the data within the deadline.

4.5 Instructions and report forms

Carefully design and prepare instructions and forms, and scrutinize them before sending to participant laboratory. A pilot study by an analysis in organising laboratory is useful for uncovering problems in these documents.

Send instructions and report forms immediately on receipt of acceptance, independent of study materials

The instructions should include in bold face or capital letters a statement:

'THIS IS A STUDY OF THE METHOD, NOT OF THE LABORATORY. THE METHOD MUST BE FOLLOWED AS CLOSELY AS PRACTICABLE, AND ANY DEVIATIONS FROM THE METHOD AS DESCRIBED, NO MATTER HOW TRIVIAL THEY MAY SEEM, MUST BE NOTED ON THE REPORT FORM'.

Include instructions on

a. storage and handling,
b. markings, and identification to be noted,
c. any special preparation for analysis,
d. criteria for use of practice or familiarization samples, if included.

Pre-code the form for each laboratory and provide sufficient space for as much sequential data as may be required for proper evaluation of the results, including a check of the calculations.

The organizing laboratory should indicate the number of significant figures to be reported, (see Box 1).

When recorder tracing reproductions are required (e.g., HPLC-chromatogram/LC-MS/MS scan/RT-PCR amplification profile) to evaluate method performance, request their submission both in the instructions and as a check item on the report form. Provide instructions how these should be labelled, axes titles, date, name of submitter, experimental conditions, and instrument settings.

Familiarization (practice) samples—with new or unfamiliar techniques, materials of stated composition for practice should be provided to collaborators to demonstrate that the stated value can be reproduced prior to analysis of collaborative test samples.

When recorder tracing reproductions are required to evaluate method performance, request their submission both in the instructions and as a check item on the form. Provide instructions with regard

to labelling of recorder tracings, such as identification with respect to item analyzed, axes, date, submitter, experimental conditions, and instrument settings.

Box 1 Significant Figures

In making statistical calculations from the reported data, the full power of the calculator or computer is to be used, with no rounding or truncating until the final reported mean and standard deviations are achieved.

At this point, the standard deviations are rounded to two (2) significant figures, and the mean and relative standard deviations are rounded to accommodate the significant figures of the standard deviations.

For example, if $S_R = 0.102$, the mean (x) is reported as 0.147, not as 0.1473 or 0.15, and RSD_R is reported as 8.2%.

Include in the report form 'a signature line for the analyst' and lines for a printed or typed version of the name and address for correct acknowledgement.

Provide for a review by the laboratory supervisor. An example of a completed form is helpful. A questionnaire may be included or sent after completion of the analyses in which the questions can be designed to reveal if any modifications have been made at critical steps in the method.

Request a copy of the calibration curve or other relationship between response and concentration or amount of analyte so that if discrepancies become apparent after examining all of the data, it can be determined whether the problem is in the calibration or in the analysis.

5.0 COLLABORATING LABORATORIES

5.1 Type of laboratories

All participating Laboratories must realize the importance of the study. A large investment is being made in studying the method and this probably will be the only collaborative study of the method that will be performed. Therefore, it is important to have a fair and thorough evaluation of the method.

The most appropriate laboratory is one with a responsibility related to the analytical problem. Laboratories with experience in the general subject matter of the method should be selected. Laboratory types should be representative (selection of laboratories that will be using the method in practice), reference/referral (assumed to be "best"). Final selection should be based on each laboratory's capabilities and past performance in collaborative studies, followed up, if possible, by telephone conversations or by personal visits. If the study is intended for international consideration, laboratories from different countries should be invited to participate.

5.2 Number of laboratories

AOAC Appendix D guidelines, for quantitative chemical analysis currently sets a minimum of **eight laboratories** contributing usable data for statistical analysis at the end of the study. It is recommended that the study is started with ≥ 12 labs. Leave room for error, non-participation, or unforeseen difficulties. Only in special cases of very expensive equipment (NMR/IRMS) or specialized laboratories (radioactivity testing) may the study be conducted with a minimum of usable data from five laboratories.

The labs should be representative of the labs where the kit/method will be used. Participation from as diverse a group of laboratories as possible is recommended. However, use of larger numbers is encouraged.
For qualitative (Binary method) analysis AOAC recommends valid data of a minimum of 10 laboratories to be used. Ideally it would be best to have ≥ 12laboratories participating.

5.3 Analysts

Most designs require only one qualified analyst per laboratory. Because the objective of the study is to evaluate the method/kit and not the analyst, all analysts must be instructed to follow the method exactly as written and provided. If analyst-within-laboratory variability is a desired variance component, multiple analysts should be requested from all participating laboratories. Ordinarily two analysts from the same laboratory cannot be substituted for different laboratories, unless they work independently, preparing standard solutions, reagents, chromatographic columns and/or materials, instrument calibrations, standard curves, etc., and no consultation between analysts is permitted during the work. Analyst must have required experience in the method that is being validated (e.g LC-MS/MS, ELISA/LFIA, RT-PCR).

6.0 TEST MATERIALS

6.1 Homogeneity

All test materials must he homogeneous and the analyte stable. This is critical so that the sampling error is only a negligible fraction of the expected analytical error. Non-homogeneity can cause outliers and will increase the variance estimates that is not due to the intrinsic method variability. Establish homogeneity by testing a representative number of laboratory samples taken at random before shipment of samples to participants. The penalty for inhomogeneity is an increased variance in the analytical results.

6.1.1 Procedure for homogeneity testing

1. Comminute and mix bulk material.
2. Split into distribution units.
3. Select m>10 distribution units at random.
4. Homogenise each one.
5. Analyse 2 test portions from each in random order, with high precision, and conduct one-way analysis of variance (ANOVA) on results.

6.2 Coding of test samples

It is very important that the samples and test portions provided to participants be blinded, so the collaborators cannot determine the expected outcome of any individual analysis. Code test samples at random so that there is no preselection from order of presentation or spiking. Code the test samples so it is not obvious to the analyst, which samples are blind duplicates and so they will not be analyzed in a set order.
For example, if a study has three levels and eight replicates per testing site per level, the testing site would need to receive 24 test vials and be asked to analyze each vial independently, and the vials should be randomly coded e so that the operator cannot distinguish the sample replication scheme.

If necessary to conserve analyst time, an indication of the potential range of concentration or amount of analyte may be provided.

6.3 Concentration range

Choose analyte levels to cover concentration range of test method/kit. A sufficient number of test samples must be prepared to cover typical matrices and the concentration range of interest. If concentration range of interest is a tolerance limit or a specification level (maximum or minimum specifications), bracket it and include it with materials of appropriate concentration. If, design includes the determination of presence/absence of analyte, include blank (not detectable) materials as part of range of interest.

6.4 Number of test materials

6.4.1 Chemical methods

Number of test samples for quantitative method/kit: A minimum of five materials (as a pair or duplicates) must be tested in the collaborative study for a quantitative method. Three materials are allowed but only when a single specification is involved for a single matrix. This parameter determines the applicability of the method (Table 6.1). Materials should be representative of commodities usually analyzed with customary and extreme values for the analyte. For an example, see Box 2.

Note: A material is an analyte (or test component)/matrix/concentration combination to which the method-performance parameters apply. This parameter determines the applicability of the method.

The duplicates (Youdan pair) statistically analyzed as a pair is considered a single material. If statistically analyzed and reported as single test samples, they are two materials.

Number of test samples for qualitative (binary) method/kit: The minimum number of concentration levels to study is three: 1) Low/blank, 2) high 3) Marginal or intermediate with six replicates per level (Figure 6.2)

1. Blank/Zero level: There should be a very low concentration where the expected probability of detection (POD) is close to zero, and if it is possible to obtain a sample with no analyte, then even better. This will demonstrate the method will not give a positive response at low, near-zero concentrations.
2. High concentration, where the method is expected to give a very high percentage of positive responses. This will demonstrate that there is a concentration where the method responds to the target compound(s).
3. Intermediate concentration where the POD is expected to be in a marginal range (0.25 to 0.75) is the transition concentration from low POD to high POD can be identified.

More levels may be added in the marginal range to increase the confidence in estimation of the detection limit of the method.

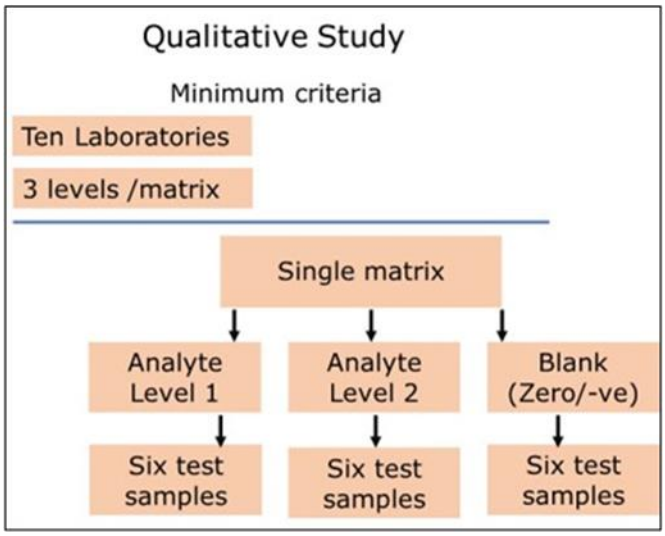


Figure 6.2: Schematic showing the minimum number of samples each laboratory participating in a collaborative study for qualitative analysis must receive

Box 2 Example of samples supplied in a collaborative study

Samples supplied to each laboratory in a collaborative study for determination of Aflatoxin in oils by HPLC and fluorescence detection

Olive oil (4 pairs)

One pair of blank olive oil sample (2 test samples, Afs <0.1 µg/kg).

Three pairs of spiked olive oil samples (6 test samples, 2 test samples/level) at 2, 4, and 20 µg/kg total AFs

Peanut oil (5 pairs)

One pair of blank peanut oil samples (2 test samples, AFs <0.1 µg/kg).

One pair of naturally contaminated peanut oil samples (2 test samples).

Three pairs of spiked peanut oil samples (total 6 test samples, 2 test samples/level) at 2, 4, and 20 µg/kg total AFs (AFB1:AFB2:AFG1:AFG2= 4:1:2:1)

Sesame oil (3 pairs)

One pair of blank sesame oil samples (AFs <0.1 µg/kg).

Two pairs of spiked sesame oil samples (total 4 test samples, 2 test samples/level) at 2 and 4 µg/kg total

(Adapted from Bao et al. 2012, Journal of AOA C International ,95, **1689-1700)**

Table 1 Minimum Criteria for Quantitative study (Chemical analysis)	
Minimum number of materials	Five When a single level specification is involved for a single matrix may this minimum be reduced to three).
Minimum number of laboratories	Eight reporting valid data for each material (only in special cases involving very expensive equipment or specialized techniques (NMR/IRMS etc.,) may the study be conducted with a minimum of five laboratories, with the resulting expansion in the confidence interval for the statistical estimates of the method characteristics)
Minimum number of replicates	Two replicates. Replication should ordinarily be attained by blind replicates or split levels (Youden pairs) collaborative study must be designed to include two analyte levels per matrix, six test samples per level, and six negative controls per matrix.
Minimum Criteria for Qualitative study (Chemical analysis)	
Minimum number of materials	Three analyte concentrations per matrix
Minimum number of laboratories	Ten reporting valid data
Minimum number of replicates	Six test samples per concentration level, and six negative controls per matrix

Box 3 Example of samples to be supplied in a collaborative ELISA/LFIA

Quantitative ELISA assay

Samples supplied to each laboratory for determination of Gluten by ELISA

1. One pair -Wheat flour in oat flour, (Two test samples, 10 mg/kg gluten)
2. One pair - Rye flour in oat flour, (Two test samples, 10 mg/kg gluten)
3. One pair - Barley flour in oat flour, (Two test samples, 0 mg/kg gluten)
4. One pair - Oat flour, unspiked ; (Two test samples, 0 mg/kg gluten)
5. One pair - Wheat flour contaminated corn-based processed snack (Two test samples: 82 mg/kg gluten)
6. One pair - Rice flour, naturally contaminated at a very low gluten level (Two test samples)

(Adapted from Lacorn et al., 2019: Journal of AOAC International. 102, 1535-1543)

Qualitative assay

Samples supplied to each laboratory for detection of peanut protein in cookies by LFIA

1. Cookie flour (Six test samples, 0 mg /kg peanut flour).
2. Cookie flour (Six test samples, 5 mg/kg peanut flour)
3. Cookie flour (Six test samples, 8 mg/kg peanut flour)
4. Cookie flour (Six test samples, 14 mg/kg peanut flour)

Each laboratory should receive 24 test samples. Eight laboratories participating 192 data points

6.4.2 ELISA/LFIA methods

Number of test samples for quantitative method/kit: A minimum of two matrixes, five concentration levels per matrix, and two replicate samples of each concentration per matrix in each laboratory. This is in compliance with AOAC Appendix D requirements for a minimum of five materials. For the concentration levels, one of the levels must be the zero level or blank (See Box 3 for example). As an example, for a study using the minimum five concentration levels, two replicates and two matrixes, each participating laboratory would receive 20 samples for analysis.

F*or qualitative binary methods*: The minimum number of concentration levels to study should be three. There should be a very low concentration where the expected POD is close to zero, and if it is possible to obtain a sample with no analyte, then even better. This will demonstrate the method **will not give a positive response at low, near-zero concentrations.**

6.4.3 Nucleic acid methods

The following materials can be used as test samples for the determination of the false-positive and false-negative rates of a qualitative real-time PCR method:

i. DNA solutions (in general extracted from reference material)
ii. Sample materials (reference material or samples from which the DNA must be extracted)

Each laboratory receives positive and negative samples. The positive samples contain definite quantities of the target sequence. The negative samples only contain non-genetically modified taxon-specific plant DNA or DNA from a non-target species.

The positive samples should contain at least twice the number of copies corresponding to the limit of detection as determined in the course of the in-house validation (but no fewer than 20 copies of the target sequence) per reaction (See Box 4)

Number of test samples:

Qualitative: Minimum Twelve DNA samples

Six vials of target DNA solution, (20 copies/µl).

Six vials of non-target DNA solution, (20 copies/µl).

Three PCR replicates each DNA sample

Quantitative: Minimum of one pair each of five concentrations per target DNA

One pair of each target DNA concentration

One pair of blank (no target DNA) concentration

The copy numbers or % should be determined using the real-time PCR method and serial dilutions of reference plasmid or CRM DNA containing the target sequence. The concentration of the plasmid-DNA (copies/µl) are generally measured by digital PCR.

Participating laboratories may also receive a PCR master mix and the oligonucleotides (primers and probes) from the collaborative test organizer to conduct the PCR experiments.

6.5 Size of test sample

The test sample furnished should be enough to provide the number of test portions specified in the instructions. If additional test portions are required, the collaborator must request them, with an explanation.

6.6 Familiarization (practice) samples

With new, complex, or unfamiliar techniques, provide material(s) of stated composition for practice, on different days, if possible. The valuable collaborative materials should not be used until the analyst can reproduce the stated value of the familiarization samples within a given range. However, it should be pointed out that one of the assumptions of analysis of variance is that the underlying distribution of results is independent of time (i.e., there is no drift). The Study Director must be satisfied that this assumption is met.

Box 4: Example of sample materials for Nucleic acid based analysis

Qualitative analysis

Type of sample	Quantity of target sequence in PCR	Examples for the number of copies of the target sequence	Number of samples per participant
Target DNA	2 x $LOD_{95\%}$	10 copies/µl.	6
Target DNA	3 x $LOD_{95\%}$	20 copies/µl.	6
Non target DNA	2 x $LOD_{95\%}$	0	6
Non target DNA	3 x $LOD_{95\%}$	0	6

Quantitative analysis of GMOs

GMO	Number of samples	Quantity of Target GMO	Number of Samples per participant
Bt-11 Maize	One pair each	0, 0.10, 0.50, 1.0, 5.0, and 10.0 %	12
GA-21 Maize	One pair each	0, 0.10, 0.50, 1.0, 5.0, and 10.0 %	12
T-25 Maize	One pair each	0, 0.10, 0.50, 1.0, 5.0, and 10.0 %	12
Event 176	One pair each	0, 0.10, 0.50, 1.0, 5.0, and 10.0 %	12
MON810 Maize	One pair each	0, 0.10, 0.50, 1.0, 5.0, and 10.0 %	12

Adapted from Shindo et al.,2002 Validation of Real-Time PCR Analyses for Line-Specific Quantitation of Genetically Modified Maize and Soybean Using New Reference Molecules, JOURNAL OF AOAC INTERNATIONAL. 85, 1119-1126

7.0 PREPARATION OF STUDY MATERIALS

Heterogeneity (Compositional and distributional) between test samples from a single test material must be negligible compared to analytical variability, as measured within the laboratory.

7.1 Sample container

The containers used to send samples must not contribute any extraneous analytes to the contents, and they must not adsorb or absorb analytes or other components from the matrix, e.g., water.

If necessary, the materials may be stabilized, preferably by physical means (freezing, dehydrating), or by chemical means (preservatives, antioxidants) which do not affect the performance of the method.

Composition changes must be avoided, where necessary, by the use of vapor-tight containers, refrigeration, flushing with an inert gas, or other protective packaging.

7.2 Suitable materials

7.2.1 Materials:

Ensure analyte and matrix stability over projected transport time and projected length of study.

*Single batch of homogenou*s, stable products such as milk powder, peanut butter, vegetable oil, starch, etc., is the best type of material.

Reference materials supplied by standards organizations or Certified reference materials (CRMs) that have certificates

Synthetic materials may be especially formulated with known amounts of analytes by actual preparation for the study.

Spiked materials consisting of normal or blank materials to which a known amount of analyte has been added may be used. The amount of analyte added should not be excessive in relation to the amount present (e.g., about 2x), and the analyte added should be in the same chemical form as present in the commodities to be analyzed subsequently.

7.2.2 Sample preparation

Preparation in bulk: This requires thorough and uniform incorporation of analyte, often by serial dilution of solids. The danger of segregation (distributional heterogeneity) due to differences in densities always exists. Fluid materials susceptible to segregation should be prepared under constant agitation. Uniformity should be checked by direct analysis, with an internal standard, or by a marker compound (dye or radioactive label).

Test samples, individually prepared: A known amount of analyte is either weighed directly or added as an aliquot of a prepared solution to premeasured portions of the matrix in individual containers. The collaborator is instructed to use each entire portion for the analysis, transferring the contents of the container quantitatively or a substantial weighed fraction of the portion.

Concentrated unknown solutions for direct addition to commodities by participating laboratories: Should be used only as a last resort when instability of the analyte precludes distribution from a central point. Supply the spiking solution as coded solutions to be added in their entirety to portions of the matrix for single analyses by each laboratory. All solutions should have the same volume and appearance. This type of material is analogous to that of test samples except for the source of matrix. This case should be used only for perishable commodities that are altered by all available preservation techniques.

7.2.3 Preparation of blank samples

When the absence of a component is as important as its presence, when determinations must be corrected for the amount of the component or the presence of background in the matrix, or when recovery data are required, blank materials containing “none” (not detected) of

the analyte should be supplied.

Two types of blanks: matrix blanks and reagent blanks. Each laboratory should perform their own reagent blanks. Matrix blanks are supplied by the Central Laboratory and the number of blanks needed depends on the combined variance of the material (sM) and of the blank (sB).

7.2.4 Number of material samples to be prepared

Prepare bulk sample and more sets (e.g., 25% more) of laboratory samples and blanks than there are collaborators. Some packages may never arrive, some materials may spoil, and some may be lost or the container broken. New laboratories may have to be substituted for those which are unable to complete. Some sets may have to be analyzed at a later time for different purposes, such as to verify stability on storage.

8.0 DISPATCH OF STUDY MATERIAL TO COLLABORATORS

i. Send prior information to collaborators of shipping arrangements, including waybill numbers, arrival time, and required storage conditions.
ii. Label test samples legibly and without ambiguity.
iii. Pack shipping cartons well and label properly to avoid transportation delays. If the containers are breakable, pack well to minimize possibility of breakage. If material is perishable, ship frozen with sufficient solid CO2, and use special transportation services.
iv. Hazardous materials must be packed and labelled as required by transportation regulations.
v. Include a return slip, with each package for confirmation. If not sent previously, include copy of method, instructions, and report forms.

9.0 OBLIGATION AND RESPONSIBILITIES OF COLLABORATING LABORATORY

i. Analyze test samples at times indicated, according to submitted protocol. With unstable materials (e.g., with microbial or decomposition problems), analyses must be started at specified times.
ii. FOLLOW METHOD EXACTLY (this is critical). If method is unclear, contact Central Laboratory. Any deviation, such as the necessity to substitute reagents, columns, apparatus, or instruments, must be recorded at the time and reported. If the collaborator has no intention of following the submitted method, he or she should not participate in the study.
iii. Conduct exactly the number of determinations as stated in the instructions received. Any other number complicates the statistical analysis.
iv. Report individual values, including blanks. Do not average or do other data manipulations unless required by the instructions.
v. Report the negative value; do not equate negative values to zero. Follow or request instructions with regard to reporting “traces” or “less than.”
vi. When results are below the limit of determination, report actual calculated result, regardless of its value not as ‘Below detection limit’.
vii. Include all raw data, graphs, recorder tracings, photographs, or other documentation as requested in the instructions.
viii. Ensure the transposition of numbers, the decimal point, or use of the correct units.
ix. Ensure use of correct standards, dilutions, calibrated clean glassware and equipment.
x. Ensure no contamination of samples, reagents etc

10.0 STATISTICAL ANALYSIS OF STUDY DATA

The statistical analysis of the collaborative study data, the required statistical procedures described below must be performed and the results reported.

10.1 Initial review

Valid data: Only valid data should be reported and subjected to statistical treatment. Valid data are values that the Study director has no reason to suspect as being wrong. Invalid data may result when: (1) the method is not followed; (2) a nonlinear calibration curve is found although a linear curve is expected; (3) system suitability specifications were not met; (4) resolution is inadequate; (5) distorted absorption curves arise; (6) unexpected reactions occur; or (7) other atypical phenomena materialize. Other potential causes of invalid data are noted such as arithmetic, clerical, or typographical errors.

One-way analysis of variance (ANOVA)—one-way analysis of variance and outlier treatments must be applied separately to each material to estimate the components of variance and repeatability and reproducibility parameters.

Initial estimate: Calculate the mean $\bar{X}$ (average of laboratory averages), repeatability relative standard deviation (RSDr) and reproducibility relative standard deviation (RSD_R), with no outliers removed, but using only data that has been determined to be valid.

10.2 Outlier treatment

Collaborative studies will have an inherent level of outliers. Rejection of more than 22.2% (2/9) of the data from each material in a study, without an explanation (e.g., failure to follow the method), is ordinarily considered excessive. A study must maintain valid data from a minimum of eight labs.

The estimated precision parameters that are reported are based on the initial valid data with removal of all outliers flagged by the IUPAC harmonized outlier removal procedure. This procedure essentially consists of sequential application of the Cochran and Grubbs tests (at 2.5% probability (*P*) level, 1-tail for Cochran, 2-tail for single Grubbs) until no further outliers are flagged, or until a drop of more than 22.2% (2 of 9 laboratories) in the original number of laboratories providing valid data would occur.

Dialogue and communication with a participating laboratory reporting suspect values can result in correction of mistakes or discovering conditions that lead to invalid data

Note: Recognizing mistakes and invalid data is much preferred to relying upon statistical tests to remove deviate values.

Step 1: Cochran Test

Cochran test for removal of laboratories (or indirectly for removal of extreme individual values from a set of laboratory values) showing significantly greater variability among replicate (within laboratory) analyses than the other laboratories for a given material. Apply as a 1-tail test at a probability value of 2.5%.

To calculate the Cochran test statistic:

1. Compute the within-laboratory variance for each laboratory
2. Divide the largest of these by the sum of all of these variances.
3. The resulting quotient is the Cochran statistic which indicates the presence of a removable outlier. If this quotient exceeds the critical value listed in the Cochran table below for P = 2.5% (1-tail) and L (number of laboratories).

For worked example Chemical Analysis see Appendix A. For DNA-based analysis refer to collaborative studies described in Detection of Animal-Derived Materials in Foodstuffs and Feedstuffs by Real Time PCR (ISO/TS 20224-Part 1to 9)

Critical values for the Cochran maximum variance ratio at the 2.5% (1-tail) rejection level, expressed as the percentage the highest variance from an individual laboratory is of the total variance (sum of variances from all laboratories).

L = Number of laboratories at a given level (concentration)	r = number of replicates per laboratory				
	$r = 2$	$r = 3$	$r = 4$	$r = 5$	$r = 6$
4	94.3	81.0	72.5	65.4	62.5
5	88.6	72.6	64.6	58.1	53.9
6	83.2	65.8	58.3	52.2	47.3
7	78.2	60.2	52.2	47.3	42.3
8	73.6	55.6	47.4	43.0	38.5
9	69.3	51.8	43.3	39.3	35.3
10	65.5	48.6	39.9	36.2	32.6
11	62.2	45.8	37.2	33.6	30.3
12	59.2	43.1	35.0	31.3	28.3
13	56.4	40.5	33.2	29.2	26.5
14	53.8	38.3	31.5	27.3	25.0
15	51.5	36.4	29.9	25.7	23.7
16	49.5	34.7	28.4	24.4	22.0
17	47.8	33.2	27.1	23.3	21.2
18	46.0	31.8	25.9	22.4	20.4
19	44.3	30.5	24.8	21.5	19.5
20	42.8	29.3	23.8	20.7	18.7
21	41.5	28.2	22.9	19.9	18.0
22	40.3	27.2	22.0	19.2	17.3
23	39.1	26.3	21.2	18.5	16.6
24	37.9	25.5	20.5	17.8	16.0
25	36.7	24.8	19.9	17.2	15.5
26	35.5	24.1	19.3	16.6	15.0
27	34.5	23.4	18.7	16.1	14.5
28	33.7	22.7	18.1	15.7	14.1
29	33.1	22.1	17.5	15.3	13.7
30	32.5	21.6	16.9	14.9	13.3
35	29.3	19.5	15.3	12.9	11.6
40	26.0	17.0	13.5	11.6	10.2
50	21.6	14.3	11.4	9.7	8.6

Cochran statistic = (largest individual within laboratory variance)/(sum of all the within-laboratory variances)

Step 2: Grubbs Tests for removal of laboratories with extreme averages

Apply in the following order: 1) Single value test (2-tail; P = 2.5%); then if no outlier is found, apply 2) pair value test (two values at the highest end, two values at the lowest end, and two values, one at each end, at an overall P = 2.5%).

A. Single value test

To calculate the Single Grubbs test statistic:

1. Compute the average for each laboratory
2. Calculate the standard deviation (SD) of these L averages (designate as the original s).
3. Calculate the SD of the set of averages with the highest average removed (sH);
4. Calculate the SD of the set averages with the lowest average removed (sL).
5. Then calculate the percentage decrease in SD as follows:

$$100 \times [1 - (s_L/s)] \text{ and } 100 \times [1 - (s_H/s)]$$

The higher of these two percentage decreases is the single 'Grubbs statistic', which signals the presence of an outlier to be omitted if the value exceeds the critical value listed in the single Grubbs table at the P = 2.5% level, 2-tail, for L laboratories.

Critical values for the Grubbs extreme deviation outlier tests at the 2.5% (2-tail), 1.25% (1-tail) rejection level, expressed as the percent reduction in the standard deviations caused by removal of the suspect value(s) (see text for calculating Grubbs statistics)

L = Number of laboratories at a given level (concentration)	One highest or lowest	Two highest or two lowest	One highest and one lowest
4	86.1	98.9	99.1
5	73.5	90.3	92.7
6	64.0	81.3	84.0
7	57.0	73.1	76.2
8	51.4	66.5	69.6
9	46.8	61.0	64.1
10	42.8	56.4	59.5
11	39.3	52.5	55.5
13	33.8	46.1	49.1
14	31.7	43.5	46.5
15	29.9	41.2	44.1
16	28.3	39.2	42.0
17	26.9	37.4	40.1
18	25.7	35.9	38.4
19	24.6	34.5	36.9
20	23.6	33.2	35.4
21	22.7	31.9	34.0
22	21.9	30.7	32.8
23	21.2	29.7	31.8
24	20.5	28.8	30.8
25	19.8	28.0	29.8
26	19.1	27.1	28.9
27	18.4	26.2	28.1
28	17.8	25.4	27.3
29	17.4	24.7	26.6
30	17.1	24.1	26.0
40	13.3	19.1	20.5
50	11.1	16.2	17.3

The critical values in Table 1 were calculated by R. Albert (October 1993) by computer simulation involving several runs of approximately 7000 cycles each for each value, and then smoothed. Although the table of critical values for the Cochran maximum variance ratio strictly applicable only to a balanced design (same number of replicates from all laboratories) it can be applied to an unbalanced design without too much error, if there are only a few deviations.

B Pair value test

To calculate the Grubbs pair statistic, proceed as described below:

1. Calculate the standard deviations
    a) s_{2L} (standard deviation after removal of the two lowest averages from original set of averages)
    b) s_{2H} (standard deviation after removal of the two highest averages from original set of averages)
    c) s_{HL}, (standard deviation after removal of highest and the lowest averages from original set of averages),
2. Take the smallest from these three SD values (s_{2L}, s_{2H}, s_{HL}) and calculate the corresponding percentage decrease in SD from the original s using formula shown above
3. A Grubbs outlier pair is present if the selected value for the percentage decrease from the original s exceeds the critical value listed in the Grubbs pair value table at the P = 2.5% level, for L laboratories.

Step 3: If the single value Grubbs test signals the need for outlier removal, remove the single Grubbs outlier and recycle back to the Cochran test as shown in the flow chart (Figure 6.3).

If the single value Grubbs test is negative, check for masking by performing the pair value Grubbs test. If this second test is positive, remove the two values responsible for activating the test and recycle back to the Cochran test as shown in the flow chart, and repeat the sequence of Cochran, single value Grubbs, and pair value Grubbs. Note, however, that outlier removal should stop before more than 2/9 laboratories are removed.

Step 4: If no outliers are removed for a given cycle (Cochran, single Grubbs, pair Grubbs), outlier removal is complete. Also, stop outlier removal whenever more than 2/9 of the laboratories are flagged for removal.

11.0 PRECISION

The precision of analytical methods is usually characterized for two circumstances of replication: within laboratory or repeatability and among laboratories or reproducibility. Repeatability is a measure of how well an analyst in a given laboratory can check himself using the same analytical method to analyze the same test sample at the same time. Reproducibility is a measure of how well an analyst in one laboratory can check the results of another analyst in another laboratory using the same analytical method to analyze the same test sample at the same or different time. Given that test samples meet the criteria for a single material, the repeatability standard deviation (s_r) is:

$$s_r = (\Sigma d_i^2/2L)^{1/2}$$

where di is the difference between the individual values for the pair in laboratory i and L is the number of laboratories or number of pairs.

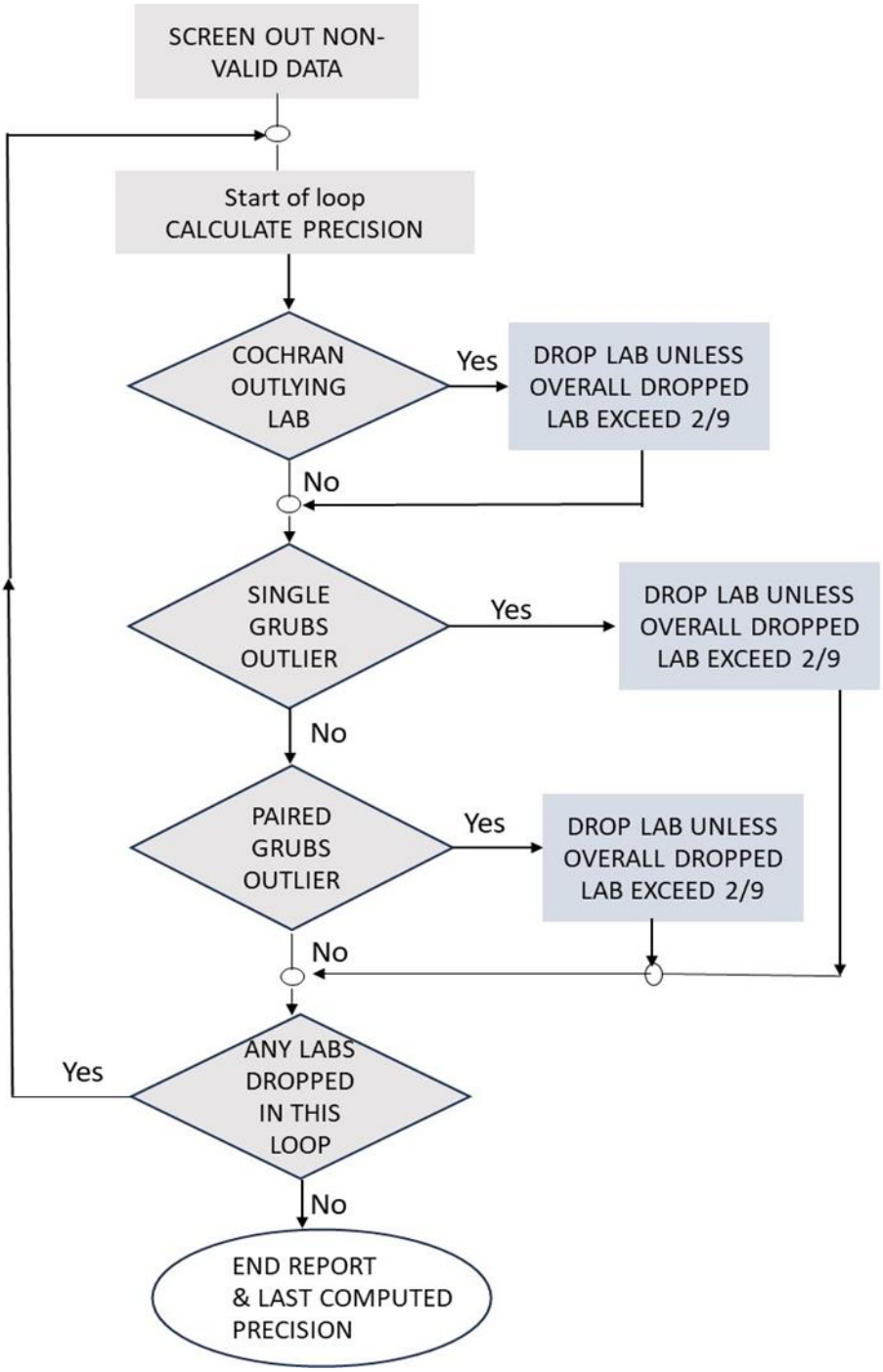


Figure 6.3: Flowchart based on IUPAC Harmonised Guidelines. Adapted from Appendix D Guidelines for Collaborative Study Procedures to Validate Characteristics of a Method of Analysis. https://doi.org/10.1093/9780197610145.005.004 Pages AD-1–AD-12 **Published:** January 2023

The reproducibility standard deviation (s_R) is computed using the formula:

$$s_R = [1/2\ (s_d^2 + s_r^2)]^{1/2}$$

Where

$s_d = \Sigma(T_i - T)\ /(2(L-1))$

$s_r = (\Sigma d_i^2/2L)^{1/2}$

T_i is the sum of the individual values for the pair in laboratory i

T is the mean of the T_i across all laboratories or pairs,

L is the number of laboratories or pairs,

When the pairs of test samples meet the criteria for Youden matched pairs, i.e., when:

$[(xc - yc)/xc\ ] \leq 0.05$ or $yc \geq (xc - 0.05xc)$

Then s_r, a practical approximation for repeatability standard deviation, is calculated as:

$$s_r = \left[\Sigma\ (d_i - d)^2/(2(L-1))\right]^{1/2}$$

where

d_i is the difference between the individual values for the pair in laboratory i,

d is the mean of the d_i across all laboratories or pairs,

L is the number of laboratories or pairs.

12.0 HORRAT

HorRat value is the ratio of the reproducibility relative standard deviation, expressed as a percent (RSD_R, %) to the predicted reproducibility relative standard deviation, expressed as a percent ($PRSD_R$, %),

$$\text{HorRAT} = \frac{RSD_R(\%)}{PRSD_R(\%)}$$

where

$PRSD_R, \% = 2C^{-0.1505}$

C = the estimated mean concentration expressed as a decimal fraction (i.e., 100% = 1; 1% = 0.01; 1 ppm = 0.000001).

HorRat values between 0.5 to 1.5 may be taken to indicate that the performance value for the method corresponds to an acceptable performance. The limits for performance acceptability are 0.5–2.

The precision of a method must be presented in the collaborative study manuscript. The HorRat will be used as a guide to determine the acceptability of the precision of a method.

The following guidelines will be used to evaluate the assay precision:

- HorRat ≤ 0.5—Method reproducibility may be in question due to lack of study independence, unreported averaging, or consultations.
- 0.5 < HorRat ≤ 1.5—Method reproducibility normally would be expected and precision accepted.
- HorRat > 1.5—Method reproducibility higher than normally expected: you should critically look into possible reasons for a "high" HorRat (e.g., were test samples sufficiently homogeneous, indefinite analyte or property).
- HorRat > 2.0—Method reproducibility is problematic. Precision is unacceptable and method/kit is rejected for regulatory use. is because it may indicate unacceptable weaknesses in the method or the study.

HorRat is applicable to most chemical methods. HorRat is not applicable to physical parameters such as viscosity, refractive index, density, pH, absorbance, etc. and empirical methods [e.g., fibre, enzymes, moisture) and "quality" measurements, e.g., drained weight.

Deviations may also occur at both extremes of the concentration scale (near 100% and ≤ 10). In area where there is a question if the HorRat is applicable, the General Referee will be the determining judge.

13.0 GLOSSARY

Method performance study: a method performance (collaborative) study is an interlaboratory study in which each laboratory uses the defined method of analysis to analyze identical portions of homogeneous materials to assess the performance characteristics obtained for that method of analysis.

Repeatability value (r)—when the mean of the values obtained from two single determinations, performed simultaneously or in rapid succession by the same operator, using the same apparatus under the same conditions for the analysis of the same test sample, lies within the range of the mean values cited in the final report, the difference between the two values obtained should not be greater than the repeatability value (r), which can generally be inferred by linear interpolation of Sr from the report.

Reproducibility value (R)—when the values for the final result, obtained by operators in different laboratories using different apparatus under different conditions for the analysis of the same laboratory sample, lie within the range of the mean values cited in the final report, the difference between the values for the final result obtained by those operators should not be greater than the reproducibility value (R), which can generally be inferred by linear interpolation of SR from the report

One-way analysis of variance—one-way analysis of variance is the statistical procedure for obtaining the estimates of within-laboratory and between-laboratory variability on a material-by-material basis. Examples of the calculations for the single-level and single-split level designs can be found in ISO 5725:1994.

14.0 REFERENCES

AOAC International ***interlaboratory*** study ***workbook blind*** (unpaired) replicates, version 2.0. Published online at www.aoac.org/stats/AOAC_BlindDup_v2-0.xls. The Association, Gaithersburg.MD. 2006.

AOCS Procedure M 4-86 Reapproved 2009 Collaborative Study Procedures

Appendix D Guidelines for Collaborative Study Procedures to Validate Characteristics of a Method of Analysis https://doi.org/10.1093/9780197610145.005.004 Pages AD-1–AD-12

Appendix M Validation Procedures for Quantitative Food Allergen ELISA Methods: Community Guidance and Best Practices https://doi.org/10.1093/9780197610145.005.013 Pages AM-1–AM-8

Collaborative Study Guidelines, *J. AOAC International 78:*143A (1995).

Horwitz, W., *Pure Appl. Chem. 60:*855 (1988).

Horwitz, W., *Pure Appl. Chem. 67:*331 (1995).

ISO/TS 20224'Molecular biomarker analysis Detection of animal-derived materials in foodstuffs and feedstuffs by real-time PCR Part 1: Bovine DNA detection method

Lacorn et al.: Journal of AOAC International Vol. 99, No. 3, 2016

Youden, W. J., *J. Assoc. Off. Agric. Chem. 46:*55 (1963).

APPENDIX A

Name of study:		Collaborative study for quantitative estimation of Aflatoxin B1 in peanut butter						
Date of study:		10-20, January, 2024						
Sample ID:		PB02						
Factor for Units of measurement		ppm						
Number of Labs		8						
Number of replicates per lab		6						
Replicate No	Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7	Lab 8
1	407.1	388.4	405	357	607.7	361.5	370.2	369.8
2	452.4	396.8	400.7	365.5	621.2	343.5	319.6	367.7
3	443.2	393.3	410.3	422.9	669.2	357.7	443.4	419.9
4	419.6	411.2	449.3	434.5	635.2	359.9	376.8	348.8
5	354.7	465.5	405.7	428.1	647.4	382.7	347	447.7
6	450	404.2	410.5	371.6	607.4	377.6	423.4	414.7
Mean	421.1667	409.9	413.583	396.6	631.35	363.81	380.066	394.766
SD (s)	33.93743	25.93248	16.31854	32.35408	22.13969	13.02209	42.32126	34.88035
(%) RSD_R	8.057956	6.326537	3.945647	8.157863	3.506722	3.579299	11.13522	8.835688
Variance=s^2	1151.749	672.4933	266.2947	1046.787	490.1658	169.5747	1791.089	1216.639
COCHRAN'S TEST								
Sum of variance		6804.792						
Largest variance		1791.089						
Cochran statistic		(1791.089/6804.792) *100= 26.32099						
Conclusion		From the table the Cochran Maximum Variance Ratio at 2.5% (1-tail) rejection level for 8 (number of laboratories) and 6 replicates critical value						

	is 38.8 %. Calculated Cochran statistic does not exceed the critical value listed. Absence of an outlier. **Outlier removal is complete**					
SINGLE GRUBBS TEST						
Lab No	Lab average	sH (highest average removed)	S_L (Lowest average removed)			
1	421.1667	421.1667	421.1667			
2	409.9	409.9	409.9			
3	413.5833	413.5833	413.5833			
4	396.6	396.6	396.6			
5	631.35	631.35	631.35			
6	363.81	363.81	363.81			
7	380.07	380.07	380.07			
8	394.77	394.77	394.77			
Mean	426.4063					
s	84.87039	sH= 20.07761	sL= 87.50508			
$\%RSD_R$	19.90365					
100× (1-SL/s)		-3.31759				
100 ×(1-SH/s)		76.34029				
Single Grubbs statistic= 76.34029						
Conclusion		76.34029 exceeds the critical value listed in the Single Grubbs table at the P = 2.5% level, 2-tail, for 8 laboratories. Therefore, highest mean (Lab 5) as outlier is to be removed and repeat single Grubbs Test				
REPEAT OF SINGLE GRUBBS TEST after highest outlier (Lab 5) is removed						
Lab No	Lab mean	sH	S_L (Lowest average			

		(highest average removed)	removed)			
1	421.1667	421.1667	421.1667			
2	409.9	409.9	409.9			
3	413.5833	413.5833	413.5833			
4	396.6	396.6	396.6			
6	363.81	363.81	363.81			
7	380.07	380.07	380.07			
8	394.77	394.77	394.77			
SD	20.07761	18.67901	14.991			
	s=20.07761	sH=18.67901	sL=14.991			
100× (1-SL/s)		25.33474				
100 ×(1-SH/s)		6.965969				
Single Grubbs statistic= 25.33474						

25.33474 is below the critical value of 57.0 listed in the Single Grubbs table at the P = 2.5% level, 2-tail, for 7 laboratories.

Outlier removal is complete

Therefore, the data of these 7 laboratories used for final calculation

Final Statistical analysis with data of 7 laboratories

Lab No	Mean					
1	421.1667					
2	409.9					
3	413.5833					
4	396.6					
6	363.81					
7	380.07					
8	394.77					
Mean	397.1286					

SD	20.07761					
%) RSD_R	5.055696					
(%) $PRSD_R$	6.393611					
HorRat=%RSDR /%PRSDR	0.790742					
Final Conclusion: HorRat 0.79 is >0.5 and ≤ 1.5. Method reproducibility is accepted						

Chapter-7

BASIC STATISTICAL TOOLS FOR DATA ANALYSIS

1.0 INTRODUCTION

Statistical analysis of analytical data obtained during a method validation should be performed to demonstrate validity of the analytical method. The primary parameters used for the interpretation of analytical method validation results are the calculation of the mean (or average), standard deviation, relative standard deviation, confidence intervals, and regression analysis. These calculations are characteristically performed using statistical software packages and Microsoft Excel. However. it is important to understand the basics of statistical analysis.

2.0 STATISTICAL TOOLS

2.1 Mean

The mean (arithmetic) or average of set of data points is the basic and most common statistics used. The mean is calculated by adding all data points and dividing the sum by the number of samples (n). It is typically denoted by $\bar{X}$(X bar) and is computed using the following formula:

$$\bar{\mathbf{X}} = \sum \frac{\mathbf{X_i}}{\mathbf{n}} = \frac{\mathbf{X_1 + X_2 + \cdots \ldots \ldots \ldots \ldots . . X_n}}{\mathbf{n}}$$

where X_i are individual values and n is the total number of individual data points.

2.2 Standard Deviation

Standard Deviation is commonly abbreviated as SD and denoted by the symbol 'σ' and it indicates how the data values deviate from the mean value. A low standard deviation indicates that the values tend to be close to the mean whereas a high standard deviation indicates values are far from the mean value. It is denoted by the symbol 'σ' and is computed using the following formula:

$$\boldsymbol{\sigma} = \sqrt{\sum_{\mathbf{i}} \frac{(\mathbf{X_i} - \bar{\mathbf{X}})^2}{\mathbf{n-1}}}$$

where Xi is individual value, $\bar{X}$ is the sample mean, and n is the number of individual data points.

2.3 Relative Standard Deviation

Relative Standard Deviation (RSD) measures the deviation of a set of numbers disseminated around the mean. One may calculate it as the ratio of standard deviation to the mean for a set of numbers. The higher the deviation, the further the numbers are from the mean. The lower the deviation, the closer the numbers are to the mean.

The RSD is computed by taking the standard deviation (σ) of the sample set and dividing it by the sample set average and multiplied by 100. The relative standard deviation is expressed as percent:

$$\%\mathbf{RSD} = \frac{\boldsymbol{\sigma}}{\bar{\mathbf{X}}} \times \mathbf{100}$$

where σ is the standard deviation and $\bar{X}$ is the mean.

Coefficient of Variation (CV)

Calculating the coefficient of variation involves a simple ratio. Simply take the standard deviation and divide it by the mean.

$$CV = \frac{Standard\ deviation}{Mean}$$

$$\mathbf{CV} = \frac{\boldsymbol{\sigma}}{\bar{\mathbf{X}}}$$

Where σ is the standard deviation and $\bar{X}$ is the mean.

2.3.1 Variance

Variance is a statistical measurement that is used to determine the spread of numbers in a data set with respect to the average value or the mean. The standard deviation squared will give us the variance. Using variance, we can evaluate how stretched or squeezed a distribution is

When needed (e.g. for the F-test,) the variance is calculated by squaring the standard deviation:

$$\mathbf{V} = \boldsymbol{\sigma}^{\mathbf{2}}$$

2.3.2 Repeatability Relative Standard Deviation [RSD(r) or RSDr]

It is the relative standard deviation calculated from within(intra) laboratory data.

2.3.3. Reproducibility Relative Standard Deviation [RSD(R) or RSDR]

The relative standard deviation calculated from among(inter) laboratory data.

2.4 Coefficient of Variation (CV)

Calculating the coefficient of variation involves a simple ratio. Simply take the standard deviation and divide it by the mean.

$$CV = \frac{Standard\ deviation}{Mean} \quad \mathbf{CV} = \frac{\boldsymbol{\sigma}}{\bar{\mathbf{X}}}$$

Where σ is the standard deviation and $\bar{X}$ is the mean.

2.5 Mass Fraction

Concentration, C, expressed as a decimal fraction. The data may be expressed in any convenient units (e.g., %, ppm, ppb, mg/g, µg/g; µg/kg; µg/L, µg/µL, etc.) when calculating and reporting statistical parameters. When reporting HorRat values, data must be reported as a mass fraction where the units of the numerator and denominator are the same See Table 7.1 for examples.

Table 7.1 Mass fraction, Predicted Relative standard deviation (Repeatability and Reproducibility)

Concentration (C)	Mass fraction (C)	$PRSD_R$, %	PRSDr, %
100%	1.0	1	1
1%	0.01	4	2
0.01%	0.0001	8	4

1 ppm	0.000001	16	8
10 ppb	0.00000001	32	16
1 ppb	0.000000001	45	22

2.6 Predicted Relative Standard Deviation [PRSD(R) or PRSDR]

The reproducibility relative standard deviation calculated from the Horwitz formula:

$PRSD(R) = 2C^{-0.15}$

where C is expressed as a mass fraction. See Table 7.1.

In Excel: PRSD (R) = 2 * C ^ (–0.15).

2.7 HorRat Value

The ratio of the reproducibility relative standard deviation calculated from the data to the PRSD(R) calculated from the Horwitz formula:

$$HorRat = \frac{\mathbf{RSD(R)}}{\mathbf{PRSD(R)}}$$

HorRat value calculated from reproducibility data is represented as HorRat (R)

$$HorRat(R) = \frac{\mathbf{RSD(R)}}{\mathbf{PRSD(R)}}$$

HorRat value calculated from repeatability data is HorRat

$$HorRat(r) = \frac{\mathbf{RSD(r)}}{\mathbf{PRSD(R)}}$$

Some expected, predicted relative standard deviations are given in Table 7.1.

2.8 One-way ANOVA ("ANalysis Of VAriance")

One-Way ANOVA compares the means of two or more independent groups in order to determine whether there is statistical evidence that the associated population means are significantly different. One-Way ANOVA is a parametric test. One-Way ANOVA can compare the means across three or more groups (e.g., in homogeneity tests). SPSS software is used. A one-way ANOVA uses the following the null and alternative hypotheses:

H_0 (null hypothesis): μ1 = μ2 = μ3 = … = μk (all the population (group) means are equal) i.e. there is no difference between group means

H_1 (alternative hypothesis): at least one population mean is different from the rest. The alternative hypothesis (H_1) is that at least one group differs significantly from the overall mean of the dependent variable.

ANOVA uses the *F* test for statistical significance. This allows for comparison of multiple means at once, because the error is calculated for the whole set of comparisons rather than for each individual two-way comparison (which would happen with a *t* test). The *F* test compares the 'variance' in each group mean from the overall group variance. If the variance 'within group' is smaller than the variance 'between groups', the *F* test will find a higher *F* value, and therefore a higher likelihood that the difference observed is real and not due to chance.

Typically, a statistical software (such as R, Excel, Stata, SPSS, etc.) is used to perform a one-way ANOVA since it's cumbersome to perform by hand. No matter which software is used the following table is obtained as output:

Sources of variation	**Sum squares (SS)**	**Degrees of freedom (df)**	**Mean Square (MS)**	**The F test statistic (F)**	**p** $F_{(k-1),\ (n-k)}$
Between groups (Treatment)	SSB	k-1	MSB	MSB/MSW	
Within group (Error)	SSW	n-k	MSW		
Total	SST	n-1			
Wherein: SSB= 'regression' sum of squares between the group means SSW= 'error' sum of squares SST=total sum of squares (SST = SSR + SSE) k: total number of groups n: total observations MSB: Between group mean square (MSB = SSB/(k-1)). It is the mean of the sum of squares, which is calculated by dividing the sum of squares by the degrees of freedom MSW: Within group mean square (MSW = SSW/(n-k) F= MSB/MSW p: The p-value that corresponds to $F_{dfr,\ dfe}$ can be calculated using a F distribution calculator					

From this table two values that we immediately analyze are the 1) F-statistic and the 2) corresponding p-value.

The F-statistic

The F-statistic is the ratio of the 'mean squares' treatment to the mean squares same as

F-statistic: Variation between sample means / Variation within samples

The larger the F-statistic, the greater the variation between sample means relative to the variation within the samples. Thus, the larger the F-statistic, the greater the evidence that there is a difference between the group means.

p value

To determine if the difference between group means is statistically significant, the p-value that corresponds to the F-statistic are examined. To find the p-value that corresponds to an F-value, the 'F Distribution Calculator with numerator = degrees of freedom (df)Treatment (Between group) and denominator= degrees of freedom (df) Error (within group).

If this p-value is less than the chosen significance level (e.g. $\alpha = .05$), we reject the null hypothesis of the ANOVA and conclude that there is a statistically significant difference, i.e. at least one of the group means is different from the others. Otherwise, if the p-value is not less than $\alpha = .05$ then we fail to reject the null hypothesis and conclude that we do not have sufficient evidence to say that there is a statistically significant difference between the means of the three groups. The ANOVA table doesn't specify *which* means are different.

Sources of variation	**Sum squares (SS)**	**Degrees of freedom (df)**	**Mean Square (MS)**	**The F test statistic (F)**	**p**
Between batch data	192.2	2	96.1	2.538	0.11385

Within each batch	1100.6	27	40.8		
Total	1292.6	29			

The above is the ANOVA table is obtained for a homogeneity testing and stability between three batches of samples and each batch is analysed ten times. The p-value that corresponds to an F-value of 2.358, numerator df = 2, and denominator df = 27 the calculator gives a value of 0.11382.

The p-value is not less than 0.05, we fail to reject the null hypothesis. This means we don't have there is not sufficient evidence to say that there is a statistically significant difference between the mean exam values of the three batches. These batches can be used for interlaboratory trials.

2.9 Student's t Test

The T-Statistic was introduced by W.S. Gossett under the pen name "student". Therefore, the T-test is also known as the "student T-test". The T-test is a commonly used statistical analysis for testing hypothesis. The T-test is applied, if you have a limited sample, size. Student's t-test is mainly used for comparison of two independent sets of data with very similar standard deviations usually and sample size in each group is less than 30.

$$\mathbf{t_{cal}} = \frac{[\overline{\mathbf{X_1}} - \overline{\mathbf{X}}_2]}{\mathbf{S_\sigma}} \sqrt{\frac{\mathbf{n_1\, n_2}}{\mathbf{n_1 + n_2}}}$$

Where

$\overline{\mathbf{X_1}}$ is mean of data set 1

$\overline{\boldsymbol{X}}_2$ is mean of data set 2

Sσ= pooled σ of the sets

n_1= number of data points in set 1

n2= number of data points in set 2

$$\mathbf{S_\sigma} = \sqrt{\frac{(\mathbf{n_1 - 1})\mathbf{S_1}^2 + (\mathbf{n_2 - 1})\mathbf{S_2}^2}{\mathbf{n_1 + n_2 - 2}}}$$

To perform the t-test, the critical t_{cal} has to be found in the 'Students-t-table (two-tailed table) shown below'. The applicable number of degrees of freedom (df) given by: df = n_1+n_2-2 and significance level of α=0.5 (95% confidence)

When the t_{cal}, is lower than the critical value t_{tab}, the null hypothesis (no difference) is accepted and the two data sets are assumed to belong to the same set meaning there is no significant difference between the mean results of the two data sets (e.g.two analysts) (with 95% confidence).

Example: Two-sided Student's t test for Protein analysis carried out by two analysts from the same laboratory

Analyst 1	Analyst 2
Protein content (g%)	
10.2	9.7
10.7	9.0
10.5	10.2
9.9	10.3
9.0	10.8

11.2	11.1
11.5	94
10.9	9.2
8.9	9.8
10.6	10.2
$\bar{X}$ =10.34	$\bar{X}$ = 9.97
S (σ) =0.819	S (σ) =0.644
n=10	n=10
t_{cal} = 1.12	
T_{tab} = 2.10 at 95% confidence (from table below using df = 18	

Two data sets were used for the calculation: tcal, is calculated as 1.12 which is lower than the critical value t_{tab} of 2.10 (df = 18, two-sided), hence the null hypothesis (no difference) is accepted and the two data sets are assumed to belong to the same population: there is no significant difference between the mean results of the two analysts (with 95% confidence).

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

एफएसएसएआई
fssai
भारतीय खाद्य सुरक्षा और मानक प्राधिकरण
Food Safety and Standards Authority of India

सत्यमेव जयते

स्वास्थ्य एवं
परिवार कल्याण मंत्रालय
MINISTRY OF
HEALTH AND
FAMILY WELFARE

fssai
RAFT
RAPID ANALYTICAL FOOD TESTING

FDA Bhawan, Kotla Marg, Near Bal Bhavan, Mandi House, New Delhi, Delhi - 110002 India
Ph.: 011-23667417, E-Mail: raft-approval@fssai.gov.in, Website: www.fssai.gov.in/cms/raft.php