DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
Food and Drug Administration

Memorandum

Date: May 19, 2015

From: FDA Foods and Veterinary Medicine Science and Research Steering Committee

Subject: Guidelines for the Validation of Chemical Methods for the FDA FVM Program, 2nd Edition

To: FVM Executive Council

The FDA Foods and Veterinary Medicine (FVM) Science and Research Steering Committee (SRSC), made up of representatives from the Office of Foods and Veterinary Medicine, the Center for Food Safety and Applied Nutrition, the Center for Veterinary Medicine, the Office of Regulatory Affairs, the National Center for Toxicological Research, the Office of International Programs, and the Office of the Chief Scientist, is charged with the task of prioritizing, coordinating and integrating food- and feed-related science and research activities across the operating units of FDA's FVM Program.

As a regulatory agency tasked with ensuring the safety of the nation's food supply, it is imperative that the laboratory methods needed to support regulatory compliance, investigations and enforcement actions meet the highest analytical performance standards appropriate for their intended purposes. Development of standardized validation requirements for all regulatory methods used to detect chemical and radiological contaminants, as well as microbial pathogens, used in our laboratories is a critical step in ensuring that we continue to meet the highest standards possible.

The attached document, now formally adopted by the SRSC, re-establishes those requirements that must be fulfilled in the evaluation of chemical methods to be used in our testing laboratories and supersedes the prior guidelines approved March 22, 2012. In the near future, these updated guidelines will be posted on FDA's website and additional venues for publication and dissemination of these guidelines are being explored and will be announced when they become available. Please share this chemical methods validation standard operating procedure with anyone who may be conducting or supervising chemical methods validation projects or otherwise needs to be aware of these updated requirements.

The Chemical Methods Validation Subcommittee is charged with providing guidance and oversight to all validation studies.

Thank you,

Palmer A. Orlandi Jr -S

Digitally signed by Palmer A. Orlandi Jr -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=1300128095, cn=Palmer A. Orlandi Jr -S
Date: 2015.05.21 16:39:56 -04'00'

Palmer Orlandi, Jr., Ph.D., Chair
FDA FVM Science and Research Steering Committee
Acting OFVM Chief Science Officer/Research Director

Guidelines for the Validation of Chemical Methods for the FDA FVM Program

2nd Edition

US Food & Drug Administration
Office of Foods and Veterinary Medicine

April 2015

ACKNOWLEDGMENT

The first edition of these guidelines published in 2011 and the second edition in 2015 were developed at the request of the US FDA Office of Foods and Veterinary Medicine. In cooperation with members of the Science and Research Steering Committee, direct input, review, and consent were provided by the following FDA research and regulatory offices:

Center for Food Safety and Applied Nutrition
Office of Regulatory Science
Office of Food Safety
Office of Applied Research and Safety Assessment

Center for Veterinary Medicine
Office of Research
Office of New Animal Drug Evaluation

Office of Regulatory Affairs
Office of Regulatory Science
ORA Laboratories

APPROVAL PAGE

This document is approved by the FDA Foods and Veterinary Medicine (FVM) Science and Research Steering Committee (SRSC). The FVM SRSC Project Manager is responsible for updating the document as change requirements are met, and disseminating updates to the SRSC and other stakeholders, as required.

APPROVED BY:

Palmer A. Orlandi Jr -S
Digitally signed by Palmer A. Orlandi Jr -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=13001 28095, cn=Palmer A. Orlandi Jr -S
Date: 2015.05.21 16:41:11 -04'00'

OFVM Chief Science Officer/Research Director

William T. Flynn -A
Digitally signed by William T. Flynn -A
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=1300085840, cn=William T. Flynn -A
Date: 2015.05.28 16:44:31 -04'00'

CVM, Deputy Director for Science Policy


Digitally signed by Jeffrey L. Ward -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=1300184354, cn=Jeffrey L. Ward -S
Reason: I am approving this document
Location: Silver Spring, MD
Date: 2015.05.26 12:28:38 -04'00'

OFVM Senior Science Advisor

John Graham -S
Digitally signed by John Graham -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, cn=John Graham -S, 0.9.2342.19200300.100.1.1=2001387754
Date: 2015.06.01 09:34:29 -04'00'

CVM, Director, Office of Research

Donald L. Zink -S
Digitally signed by Donald L. Zink -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, cn=Donald L. Zink -S, 0.9.2342.19200300.100.1.1=1300188518
Date: 2015.05.26 13:33:51 -04'00'

CFSAN Senior Science Advisor

Paul E. Norris -S
Digitally signed by Paul E. Norris -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, cn=Paul E. Norris -S, 0.9.2342.19200300.100.1.1=1300234978
Date: 2015.07.01 16:30:07 -04'00'

ORA, Director Office of Regulatory Science

Vincent K. Bunning -A
Digitally signed by Vincent K. Bunning -A
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=1300010297, cn=Vincent K. Bunning -A
Date: 2015.05.26 14:09:12 -04'00'

CFSAN, Director, Office of Regulatory Science

Timothy Mcgrath -A
Digitally signed by Timothy Mcgrath -A
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=2000098076, cn=Timothy Mcgrath -A
Date: 2015.07.02 08:49:47 -04'00'

ORA, Director, Food and Feed Scientific Staff

Mary E. Torrence -S
Digitally signed by Mary E. Torrence -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=1000833936, cn=Mary E. Torrence -S
Date: 2015.05.26 14:53:51 -04'00'

CFSAN, Director Office of Applied Research & Safety Assessment


Digitally signed by William B. Martin -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=1300128280, cn=William B. Martin -S
Reason: I am approving this document
Date: 2015.07.02 10:22:16 -07'00'

ORA, Member of the ORA Scientific Advisory Council

US Food & Drug Administration
Office of Foods and Veterinary Medicine

Guidelines for the Validation of Chemical Methods for the FDA FVM Program, *2nd Edition*

TABLE OF CONTENTS

1.0 INTRODUCTION

1.1 Purpose

The U.S. Food and Drug Administration (FDA) is responsible for ensuring the safety of approximately 80% of the nation's food supply. FDA laboratories contribute to this mission through routine surveillance programs, targeted regulatory analyses, and emergency response when contaminated food or feed is detected or suspected in a public health incident. The effectiveness of these activities is highly dependent on the quality and performance of the laboratory methods needed to support regulatory compliance, investigations and enforcement actions. To ensure that the chemical methods employed for the analysis of foods and feeds meet the highest analytical performance standards appropriate for their intended purposes, the FDA Office of Foods and Veterinary Medicine (OFVM) through the Science and Research Steering Committee (SRSC) has established criteria by which all Foods and Veterinary Medicine (FVM) Program chemical methods shall be evaluated and validated. This document defines four standard levels of performance for use in the validation of analytical regulatory methods for chemical analytes in foods and feeds.

1.2 Scope

These criteria apply to FDA laboratories as they develop and participate in the validation of analytical regulatory methods for chemical analytes in anticipation of Agency-wide FVM Program implementation. These criteria do not apply to methods developed by or submitted to FDA under a codified process or official guidance (e.g., in the Code of Federal Regulations, CPGs, etc.) such as for veterinary drug approval. For such studies, the appropriate Center for Veterinary Medicine (CVM) or other Program guidance documents should be followed. This guidance is a forward-looking document; the requirements described here will only apply to ***newly***-developed methods and significant modifications to existing methods (see Requirements). Once a method has been validated at the appropriate level, it can be implemented according to OFVM document, FDA-OFVM-3, "Methods Development, Validation, and Implementation Program", which establishes a standard operating procedure for the methods development, validation and implementation process [1]. For example, for a multi-laboratory validated method to be used in a widespread regulatory application, it can be implemented by other FDA laboratories following the method verification process. However, method verification is normally part of a local laboratory's quality control procedures and is not considered within the scope of this validation document.

1.3 Administrative Authority and Responsibilities

All criteria established in this document for analytical method validation have been adopted and approved by the OFVM and the SRSC. The OFVM document, FDA-OFVM-3, establishes the standard operating procedure for the approval and tracking of method development and validation activities within the FVM Program [1]. Single laboratory validation (SLV) studies (including both Level 1 and Level 2 validations) can be managed wholly by the respective Center and Office line management structure. Oversight and coordination of multi-laboratory validation (MLV) studies (including both Level 3 and Level 4 validations) are the responsibility of the Methods Validation Subcommittees (MVS).

1.4 The Method Validation Subcommittee

Under the charge of the SRSC, the Chemistry Methods Validation Subcommittee (CMVS) will have oversight responsibility for MLV studies involving chemical methods associated

with the FVM Program which are intended for use in a regulatory context. The CMVS is a subcommittee of the Chemistry Research Coordinating Group (CRCG), which reports directly to the SRSC. The CMVS is governed by the organizational structure, roles and responsibilities as detailed in its charter [2]. Briefly, the CMVS will oversee and coordinate, in collaboration with the originating laboratory, all MLV studies for chemical methods developed within the FDA OFVM Enterprise to support regulatory analytical needs. This includes the evaluation and prioritization of proposed MLV studies as well as evaluation of completed MLV studies and reports. Submissions of chemical validation proposals, reports, questions, *etc.* can be directed to the CMVS through a central email account:

Chemistry.mvs@fda.hhs.gov

However, where possible, MLVs should be discussed in appropriate Technical Advisory Groups or with the CRCG to ensure the broadest possible consideration of factors before committing resources to an MLV.

1.5 General Responsibility of the Originating Laboratory

It is the responsibility of the originating laboratory to ensure proper adherence to all criteria described in this document. The originating laboratory should work in consultation with the CMVS and/or its designated Technical Advisory Group (TAG) throughout the multi-laboratory validation process. It will be the responsibility of the originating laboratory to include their respective QA/QC manager in all aspects of the validation process.

1.6 Overview of Method Validation

Method validation is the process of demonstrating or confirming that a method is suitable for its intended purpose. The purpose of these methods may include but is not limited to qualitative analysis, quantitative analysis, screening analysis, confirmatory analysis, limit tests, matrix extensions, platform extensions, and emergency/contingency operations. Validation includes demonstrating performance characteristics such as accuracy, precision, sensitivity, selectivity, limit of detection, limit of quantitation, linearity, range, and ruggedness, to ensure that results are meaningful and appropriate to make a decision. Method validation is a distinct phase from method development/ optimization and should be performed *subsequent* to method development. Methods may be validated for one or more analytes, one or more matrices, and one or more instruments or platforms. The method is validated by conducting experiments to determine the specific performance characteristics that serve to define and quantify method performance.

1.7 Applicability

This document establishes validation criteria for regulatory methods that are to be widely used to detect chemical analytes in food, feed and other FDA regulated products covered by the FVM Program including, but not limited to, the following:

Chemotherapeutic Residues
Color Additives
Decomposition Products
Dietary Supplement Ingredients/Adulterants
Elemental and Metals
Food and Feed Additives and Preservatives
Food Allergens
Gluten

Intentional Adulterants/Poisons
Mycotoxins
Nutrients
Persistent Organic Pollutants
Pesticides
Seafood and plant toxins
Toxic Elements
Veterinary Drug Residues

Please note that although these guidelines mainly cover multi-laboratory validations, criteria for several validation levels are discussed and are differentiated from full MLVs. There are situations where a method is being extended to handle what is likely to be a very limited (perhaps one time) use by one laboratory and is therefore not intended for Agency-wide regulatory use, thus would be validated at a lower level. For example, when a single pesticide laboratory receives several new food matrices for multi-residue analyses that were not covered in the previous validation of the method, these guidelines would not generally be required and a more abbreviated validation/verification within the pesticide program's guidelines may be acceptable.

1.8 Requirements

Method validation is required for:

- Submission of a new or original method.
- Expansion of the scope of an existing method to include additional analytes.
- Expansion of the scope of an existing method to include additional matrices.
- Changes in the intended use of an existing method (*e.g.,* screening vs. confirmatory).
- Modifications to a method that may alter its performance specifications (*e.g.*, modifications that could significantly affect the precision and accuracy, changes to the fundamental science of an existing method, significant changes to reagents, apparatus, instrumental parameters, sample preparation and/or extraction, or modification of a method's range beyond validated levels). Some examples of allowable modifications that would not require further validation are provided in the document, ORA-LAB.5.4.5 Attachment A-Modification Criteria [3].

2.0 CRITERIA AND GUIDANCE FOR THE VALIDATION OF CHEMICAL METHODS

2.1 General Validation Tools and Protocol Guidance

There are a number of excellent references and guides available providing further information on method validation for chemical methods [3-20]. 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 1 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 as described in the Performance Characteristics section below.

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 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.

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.

Matrix Spikes (Laboratory Fortified 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.

Incurred Samples: This type of sample contains (not laboratory fortified) the analyte(s) of interest (if available) 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.

2.2 Reference Method

A reference method is a method by which the performance of an alternate or new method may be measured or 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 the CMVS and the Program Office is suggested when deciding if the use of a reference method will be necessary.

2.3 Performance Characteristics

Performance characteristics that should be evaluated in order to validate a method will vary depending on the intended use of the method, the type of method (*e.g.*, quantitative vs. qualitative), and the degree to which it has been previously validated (e.g., matrix extension, analyte extension, platform extension). Although definitions of these characteristics are included in Appendix 1, this document is not meant to address the various ways of calculating characteristics such as method detection level, limit of detection or limit of quantitation.

Performance Characteristics for Validation of New Quantitative Methods: Validation of new quantitative methods should include at a minimum evaluation of the following performance characteristics: accuracy, precision, selectivity, limit of detection, limit of quantitation, linearity (or other calibration model), range, measurement uncertainty, ruggedness, confirmation of identity and spike recovery.

Performance Characteristics for Validation of New Qualitative Methods: Validation of new qualitative methods should include at a minimum evaluation of the following performance characteristics: sensitivity, selectivity, false positive rate, false negative rate, minimum detectable concentration, ruggedness, and confirmation of identity.

Performance Characteristics for Validation of Method Extensions: Validating the extension of methods 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.

2.4 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.

FDA has issued guidance documents on the development, evaluation, and application of mass spectrometric methods for confirming the identity of target analytes including: CVM Guidance for Industry 118: Mass Spectrometry for Confirmation of the Identity of Animal Drug Residues [4] and ORA-LAB.010, Guidance for the Analysis and Documentation to Support Regulatory Action on Pesticide Residues [5]. Following the CVM guidance is required for veterinary drug residue methods. The ORA-LAB.010 document was written specifically for pesticide analyses. For other types of chemical contaminants in food (e.g. food additives, mycotoxins, etc.), the CVM document should be followed because it was written as a Guidance for Industry and therefore has been more widely internally and externally reviewed and distributed. In addition, OFVM is currently drafting a supplement to CVM Guidance for Industry 118 specifically addressing the use of high resolution mass spectrometry and the evaluation of exact mass measurement data.

2.5 Method Validation Levels

The following describes the four standard levels of performance defined for method validation of analytical regulatory methods for chemical analytes in foods. This approach is based on the Food Emergency Response Network (FERN), SOP No: FERN-ADM.0008.00, FERN Validation Guidelines for FERN Chemical, Microbiological, and Radiological Methods [6], as well as AOAC guidelines for single-laboratory validation [7] and collaborative studies [8]. Key validation parameters for each level are summarized in Table 1. It is the responsibility of the originating (developing) laboratory to determine the appropriate level of validation required up to and through single laboratory validations. It is highly recommended that originating laboratories work with the appropriate Technical Advisory Group when determining the appropriate level of validation.

NOTE: *Not all methods will or should be validated to the highest level.*

Level One

This is a single laboratory validation level with the lowest level of validation requirements and is appropriate for emergency/limited use. Performance of the method at this initial level of scrutiny will determine, in part, whether further validation is useful or warranted.

Intended Use: emergency/limited use/matrix extension/analyte extension/ platform extension. Examples of where Level One validation would be acceptable include, isolated consumer complaints, single-occurrence samples, and application of a method developed for a specific analyte(s) to a matrix, not previously validated in response to a real or perceived threat to food safety or public health. Validation of method performance with a new matrix is intended to assure that the new matrix will produce accurate and reliable results for all the analytes in the scope of the method. Generally, all targeted analytes still must be included in matrix spikes at this level, if widespread use in this matrix is anticipated for regulatory purposes. As the first level of validation of methods for matrix, analyte or platform extension/emergency use, it would be expected that a

more rigorous single laboratory validation at least equivalent to Level Two below would be performed before more widespread non-emergency regulatory use.

Level Two

This is a single laboratory validation level. The originating lab has conducted a comprehensive validation study, with performance criteria similar to an AOAC Single Laboratory Validation study. If appropriate, a comparison with an existing reference method has been performed. Some of the criteria of the study may be at a lower level than the AOAC Single Laboratory Validation study, but are appropriate for the developing method at this stage.

Intended Use: Routine regulatory testing, emergency needs, minor method modifications, analyte and matrix extensions of screening methods. If a method validated at this level is expected to have use that is widespread, long term, of high public visibility or potentially involved in international trade conflicts, its validation should be extended to at least Level Three below.

Level Three

This is a multi-laboratory validation level. Level Three validation employs a minimum of one collaborating laboratory in addition to the originating laboratory. Most of the criteria followed by the originating lab are at a level similar to the AOAC full collaborative study level with comparison to an existing reference method when available and appropriate. The additional collaborating laboratories follow many of the criteria found in an AOAC collaborative study. The main differences are that Level Three validation employs at least one additional collaborating laboratory instead of the eight to ten used by AOAC and requires fewer replicates for each food matrix/spike level.

Intended Use: Methods validated to this level of scrutiny are acceptable for use in all regulatory circumstances including screening analyses, confirmatory analyses, regulatory surveys, and compliance support. If the method is expected to have use that is widespread, long term, of high public visibility or involved in international trade conflicts, it may be appropriate to have its validation extended to Level Four.

Level Four

This validation level has criteria equivalent to a full AOAC or ISO Collaborative Study. Any method reaching this level of validation should be able to be submitted for adoption by the AOAC as a fully collaborated method.

2.6 Acceptability Criteria

There are various acceptability ranges for method validation performance criteria that may be appropriate depending on the application or intended use of the methodology and especially the levels of concern, action levels or tolerance for the chemical analyte. Some examples of acceptability ranges used by various national and international organizations and their sources are provided in Appendix 2. Acceptable spike recoveries vary with analyte concentration as indicated in Appendix 2 (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. The acceptability ranges in Appendix 2 provide approximate target ranges for method developers and the MVS 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 and with emergency situations these general acceptability ranges may not be achievable or required.

Table 1. Key Validation Parameter Requirements for Chemical Methods

	Level One: Emergency/ Limited Use	Level Two: Single Laboratory Validation	Level Three: Multi-Laboratory Validation	Level Four: Full Collaborative Study
Number participating labs	1	1	≥ 2	8 (quantitative) 10 (qualitative)
Number of matrix sources per matrix*	≥1	≥3 recommended where available	≥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	≥3 spike levels + 1 matrix blank
Replicates required per matrix source at each level tested per laboratory	≥2 (quantitative) ≥2 (qualitative)	≥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)	≥3 (quantitative) ≥6 (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 sample 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).

** Number of spike levels is recommended for at least one source of matrix. Other similar sources of matrix (*e.g.,* within the same category; see Appendix 4) 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).

3.0 ADDITIONAL PROCEDURAL GUIDANCE

In addition to the criteria described above in Table 1 for standard quantitative and qualitative methods, additional guidance is provided in this section for specific types of methods or validation situations.

3.1 Platform/Instrumentation Extension

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 and approved for use; however, it may have major differences in configuration, or detection scheme.

Platform extension validation should generally be performed using Table 1, Level 2 as a guide and should compare the proposed new platform to the platform used in the reference method. In planning platform extension validation, one must determine what degree of cross-correlation between the results obtained on the two platforms will be acceptable.

Examples:
Method A is a validated method for the screening of pesticides on a gas chromatograph coupled to a single quadrupole mass spectrometer (GC-MSD). Gas chromatography coupled to a triple quadrupole mass spectrometer (GC-QQQ), offers certain advantages over the GC-MSD platform in terms of sensitivity, selectivity and scope. In this instance, a comparative method extension validation is indicated to ensure equivalent results. However, if new analytes are added to the scope of the method via the use of the new platform, a new method validation is indicated for the GC-QQQ method.

Method Z is a validated method for the screening of polycyclic aromatic hydrocarbons in seafood using liquid chromatography with a fluorescence detector (LC-FLD). A laboratory would like to transfer this method to a liquid chromatography system that utilizes only a diode-array detector (LC-DAD). In this instance, a comparative method extension validation would be indicated to ensure that the new detection system produces equivalent results to the originally validated method.

3.2 Analyte Extension

Multi-residue, multi-class methods are becoming more common. Many of these methods are semi-quantitative (limits tests) or qualitative broad band screens. Performance requirements for these types of procedures are described below. However, if a multi-residue method is meant to be used for quantitation, the same performance characteristics as required for single analyte methods should be evaluated for each analyte (accuracy, precision, selectivity, limit of detection, limit of quantitation, linearity range, uncertainty, and ruggedness). It is understood that with a large multi-residue method , not all analytes will meet the recommended acceptability ranges listed in Appendix 2, but the performance for each compound should be tested and reported so that the accuracy and precision are known for any given analyte and are sufficient for the intended purpose of the method.

When new analytes are added to a quantitative multi-residue method, tests should be performed to ensure that the addition of new compounds do not affect the performance of the instrumental conditions, e.g. duty cycle or scan rates for other eluting analytes, and that

the analytes do not present a chemical or physical interaction with the stabilities of the other tested analytes.

3.3 Food Matrix Extension

The validation of method performance with a new matrix is intended to assure that the method will continue to produce accurate and reliable results. Emergency matrix extensions (Level 1 in Table1) are intended for those instances in which a validated method is used with a matrix not previously validated in response to a real or perceived threat to food safety or public health, and in this type of urgent situation it is not expected that the MVS would be consulted. Matrix extensions of validated methods that are intended to increase the regulatory scope and applicability on a recurring basis would minimally fall under Level 2 validation in Table 1. This section provides guidance to extend validated methods to matrices in anticipation that these food commodities will be included in Agency-wide testing. Method developers may wish to consult with the appropriate Technical Advisory Group or MVS before initiating any Level 2 validation work on matrix extension.

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. It is also usually the case that the regulatory chemical methods employed by FDA are used to analyze a diversity of products representing a large spectrum of matrices. It becomes unfeasible to carry out a matrix extension validation for each single matrix in order to expand the scope of the method. A more reasonable approach to demonstrate the applicability of a method to a set of product matrices is to validate the method for different "categories" of products. For instance, a multi-residue pesticide method can be validated for "high-sugar", "high-fat", "high-water", "dry" and "high-protein" matrices. Appendix 4 provides guidance on commodity categories and gives examples of representative matrices in each category.

The number of different food categories to be validated depends on the applicability and intended use of the method. If the method is specific to only one category, only one type of food need be included. If the applicability is wider (e.g., detection of phthalates in processed foods), then an appropriate number of food categories should be included to represent all anticipated matrices. Depending on how many categories will be validated, a minimum of 1 – 3 representative matrices from each category should be selected.

3.4 Limit Tests (common semi-quantitative screening method)

One specific category of qualitative methods includes limit tests (binary or pass/fail tests) for analytes that have a defined level of concern. The purpose of these screening methods is to determine if analyte is present with a concentration near or above the level of concern. This is in contrast to screening methods whose intended purpose is to determine the presence or absence of an analyte at any level. Limit test method validations must include determination of the precision of the method for an analyte(s) at the level(s) of concern.

Limit test screening methods, in general, should avoid false negatives with false negative rates representing less than 5% of the analytical results. The occurrence of false positives is less critical since presumptive positives are further analyzed by quantitative or confirmatory methods. However, false positive rates should typically be less than 10-15% to avoid unnecessary confirmatory testing. Ideally, limit tests are capable of rapidly screening a large number of samples to minimize the need for additional analysis. A common approach used in limit test screening methods is to use a confidence interval to set a laboratory threshold or cut-off value whereby only responses above that value require further testing. For a limit

test based on an instrument response, a threshold or cut-off value can be determined by a confidence limit, based on an estimate of the standard deviation of the response or concentration of an analyte in samples fortified with the analyte at the level of concern.

Example:
Milk samples (n=21) were fortified with sulfamethazine at the level of concern (10 ng/mL). A LC-MS/MS limit test screening method was used to measure this drug in the extracted milk samples. The mean concentration found was to be 10.99 ng/mL with a standard deviation of 2.19. A threshold or cut-off value was calculated so that 95% of samples containing sulfamethazine at or above 10 ng/mL would have a response above the threshold value:

Threshold value = [mean concentration – (t * standard deviation)]
= [10.99 – (1.725 * 2.19)] = 7.21 ng/mL

Where t = one-tailed Student's t value for n-1 degrees of freedom at the 95% confidence level

This approach can also be used for immunosorbent assays such as enzyme linked immunosorbent assay (ELISA) or optical biosensor assays. These tests may be non-competitive (direct measurement of analyte response) or competitive (indirect measurement). Analysis of data from a competitive immunosorbent test should account for the fact that the observed response decreases with increasing analyte concentration; therefore, a response lower than the threshold or cut-off would be considered a presumptive positive response. For immunosorbent assays, it is also important to measure the response observed for blank matrix samples and to verify that the blank response is distinguishably (statistically) different from that of the threshold.

Performance characteristics of limit tests:
Validation of new limit tests should include, at a minimum, evaluation of the following performance characteristics: sensitivity, specificity, precision, threshold or cut-off value, false positive rate, false negative rate, minimum detectable concentration (should be lower than the threshold/cut-off value), and ruggedness/robustness.

3.5 Qualitative Broad-band Analyte Screening

Broad-band methods that can detect many compounds are being utilized more frequently as an initial screening step as part of chemical contaminant testing in FDA laboratories. These methods usually involve mass spectrometric analyses and provide qualitative information. For example, the data obtained may be compared to an established reference such as a database of compounds with exact mass and molecular formula information or spectra in a compiled library. For regulatory action, any positive findings from this screen should be confirmed by a targeted method (for example using a LC-MS/MS or GC-MS/MS platform).

Typically, initial validation of these methods is performed using a limited set of representative analytes and representative matrices. For example, sets of analytes that contain compounds from a variety of chemical classes from the area of interest (e.g. pesticides, veterinary drug residues, or common chemical toxins) are tested with the method using representative matrices. The performance characteristics that may be evaluated include: sensitivity, selectivity, false positive rate, false negative rate, minimum detectable concentration, ruggedness, and confirmation of identity. It is understood that the method

performance may vary with the different classes of compounds, but it is important to have an initial evaluation of the method's capabilities.

Laboratories continuously expand the scope of these broad-band methods by adding new analytes that come to their attention through various sources of intelligence. In addition, a new compound might be found in a sample after acquired data are compared to the reference databases. In these cases, some verification that the analyte can be detected reliably by the screening method is required. When a new compound is added to the scope of a qualitative method, it should first be determined whether this compound belongs to a class of compounds that has already been validated for the broad-band method. If the new compound shares chemical characteristics with an existing class of compounds in the scope of the method, then it may suffice to select a few representative matrices, perform a single level spike in these representative matrices in duplicate and determine that reproducible recovery is obtained in order to assess whether the analyte can be detected effectively by the method. Scenarios that may require a full validation would include a new analyte being added to the scope of the broad-band method that was not represented by any of the compound classes already in the scope. Also, if the new analyte requires modifications in the extraction protocol due to its chemical characteristics, then its inclusion in the scope should be fully validated as recommended by this guidance.

Although positive findings by the broad-band method are subjected to confirmatory testing using a targeted method, it is still important to determine, through proper validation and verification protocols, that the broad-band method does not give rise to a high number of false negative findings. False negative in this context means the method fails to detect a residue in its scope when the residue is present in the matrix at or above the level of concern or minimum detectable concentration. While the positive finding by the broad-band method is subjected to further analysis and scrutiny, negative findings are upheld as such and a regulatory decision is made based on these results, *e.g.*, to release the products into commerce.

4.0 REFERENCES AND SUPPORTING DOCUMENTS

[1] Food and Drug Administration, Office of Foods and Veterinary Medicine, "Methods Development, Validation and Implementation Program", Document #: FDA-OFVM-3. Effective Date 10/16/2014.

Methods
Development-Validatio

[2] Food and Drug Administration, "FDA Office of Foods and Veterinary Medicine Method Validation Subcommittee Charter", 3/19/2014.

SRSC Method
Validation Subcommit

[3] Food and Drug Administration, "Methods, Method Verification and Validation", Laboratory Manual, ORA Laboratory Procedure, Volume II, ORA-LAB.5.4.5. Accessed 12/12/2014. http://www.fda.gov/downloads/ScienceResearch/FieldScience/LaboratoryManual/UCM092147.pdf

[4] Food and Drug Administration, "Mass Spectrometry for Confirmation of the Identity of Animal Drug Residues", Center for Veterinary Medicine (CVM), Guidance for Industry #118, 2003. Accessed 12/12/2014. http://www.fda.gov/downloads/AnimalVeterinary/GuidanceComplianceEnforcement/GuidanceforIndustry/UCM052658.pdf

[5] Food and Drug Administration, "Guidance for the Analysis and Documentation to Support Regulatory Action on Pesticide Residues", ORA-Wide Procedure, ORA-LAB.010. Effective Date 11/06/2009. Accessed 12/12/2014. http://inside.fda.gov:9003/downloads/ORA/OfficeofRegionalOperations/DivisionofFieldScience/UCM254490.pdf

[6] Food Emergency Response Network (FERN), SOP No: FERN-ADM.0008.00, "FERN Validation Guidelines for FERN Chemical, Microbiological, and Radiological Methods", 06/22/2010. Available in eLEXNET.

[7] 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

[8] 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/imis15_prod/AOAC_Docs/StandardsDevelopment/eoma_appendix_d.pdf

[9] AOAC International, "Appendix F: Guidelines for Standard Method Performance Requirements", 2012. Accessed 12/4/2014. http://www.eoma.aoac.org/app_f.pdf

[10] Codex Alimentarius, Codex Alimentarius Commission, Procedural Manual, Twenty-second ed., "Principles for the Establishment of Codex Methods of Analysis", 2014. Accessed 12/4/2014.
http://www.fao.org/3/a-i3243e.pdf

[11] ISO, "Accuracy (Trueness and Precision) of Measurement Methods and Results", Parts 1-6, International Standard ISO 5725-1:1994, 5725-2:1994, 5725-3:1994, 5725-4:1994, 5725-5:1994, and 5725-6:1994. Downloaded 12/10/2014.

[12] 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 978-9187461-59-0. Accessed 12/9/2014.
www.eurachem.org

[13] M. Thompson, S.L.R. Ellison, R. Wood. "Harmonized Guidelines for Single-Laboratory Validation of Methods of Analysis", (IUPAC Technical Report) *Pure and Applied Chemistry*, 2002, 74, 835 - 855.

[14] European Commission, Health & Consumer Protection Directorate-General, "Guidance Document on Analytical Quality Control and Validation Procedures for Pesticide Residues Analysis in Food and Feed", SANCO/12571/2013, 19 November 2013 rev. 0. Accessed 12/9/2014.
http://ec.europa.eu/food/plant/pesticides/guidance_documents/docs/qualcontrol_en.pdf

[15] JCGM 200:2012, International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM), International Bureau of Weights and Measures. Accessed 12/9/2014.
http://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf

[16] Codex Alimentarius Commission, "Guidelines on Analytical Terminology", Standard CAC/GL 72-2009. Accessed
12/9/2014. http://www.codexalimentarius.org/input/download/standards/11357/cxg_072e.pdf

[17] ISO, "Statistics-Vocabulary and Symbols-Part 2: Applied Statistics", International Standard ISO 3534-2:2006, Downloaded 12/10/2014.

[18] Codex Alimentarius Volume 3 "Residues of Veterinary Drugs in Foods", 2nd ed. (1993), Joint FAP/WHO Food Standards Program, FAO, Rome Italy, p 59.

[19] W.R. Wolf, and K.W. Andrews, "A System for Defining Reference Materials Applicable to All Food Matrices"*, Fresenius' Journal of Analytical Chemistry,* 1998, 352, 73-76.

[20] K.E. Sharpless, R.R. Greenberg, M.M. Schantz, M.J. Welch, S.A. Wise, and M. Ihnat, "Filling the AOAC Triangle with Food-Matrix Standard Reference Materials", *Analytical and Bioanalytical Chemistry*, 2004, 378, 1161-1167.

APPENDIX 1 - Glossary of Terms

Generally, references 13-17 were utilized in preparation of this glossary.

Accuracy: The closeness of agreement between a test result and an accepted reference value. When applied to test results, accuracy includes a combination of random and systematic error. When applied to test method, accuracy refers to a combination of trueness and precision.

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

Analyte: The chemical substance measured and/or identified in a test sample by the method of analysis.

Analytical batch: An analytical batch consists of samples, standards, and blanks which are analyzed together with the same method sequence and same lots of reagents and with the manipulations common to each sample within the same time period (usually within one day) or in continuous sequential time periods.

Bias: The difference between the expectation of the test result and the true value or accepted reference value. Bias is the total systematic error, and there may be one or more systematic error components contributing to the bias.

Blank: A substance that does not contain the analytes of interest and is subjected to the usual measurement process. Blanks can be further classified as method blanks, matrix blanks, reagent blanks, instrument blanks, and field blanks.

Calibration: Determination of the relationship between the observed analyte signal generated by the measuring/detection system and the quantity of analyte present in the sample measured. Typically, this is accomplished through the use of calibration standards containing known amounts of analyte.

Calibration Standard: A known amount or concentration of analyte used to calibrate the measuring/detection system. May be matrix matched for specific sample matrices.

Carryover: Residual analyte from a previous sample or standard which is retained in the analytical system and measured in subsequent samples. Also called *memory*.

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).

Check Analysis: Result from a second independent analysis which is compared with the result from the initial analysis. Typically, check analyses are performed by a different analyst using the same method.

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). See also *Threshold Value*.

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.

Guidance Level: Level of concern or action level issued under good guidance practices that must be reliably identified or quantified in a sample.

Incurred Samples: Samples that contain the analyte(s) of interest, which were not derived from laboratory fortification but from sources such as exogenous exposure or endogenous origin. Exogenous exposure includes, for example, pesticide use, consumption by an animal, or environmental exposure.

Interference: A positive or negative response or effect on response produced by a substance other than the analyte. Includes spectral, physical, and chemical interferences which result in a less certain or accurate measurement of the analyte.

Intermediate Precision: Within-laboratory precision obtained under variable conditions, e.g., different days, different analysts, and/or different instrumentation.

Internal Standard: A chemical added to the sample, in known quantity, at a specified stage in the analysis to facilitate quantitation of the analyte. Internal standards are used to correct for matrix effects, incomplete spike recoveries, etc. Analyte concentration is deduced from its response relative to that produced by the internal standard. The internal standard should have similar physico-chemical properties to those of the analyte.

Laboratory Fortified Matrix: See *Matrix Spike*.

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 violative. This concentration can be a regulatory tolerance, safe level, action level, guidance level or a laboratory performance level.

Limit of Detection (LOD): The minimum amount or concentration of analyte that can be reliably distinguished from zero. The term is usually restricted to the response of the detection system and is often referred to as the *Detection Limit*. When applied to the complete analytical method it is often referred to as the *Method Detection Limit* (MDL).

Limit of Quantitation (LOQ): The minimum amount or concentration of analyte in the test sample that can be quantified with acceptable precision. Limit of quantitation (or quantification) is variously defined but must be a value greater than the MDL and should apply to the complete analytical method.

Limit Test: A type of semi-quantitative screening method in which analyte(s) has a defined level of concern. Also referred to as binary or pass/fail tests.

Linearity: The ability of a method, within a certain range, to provide an instrumental response or test results proportional to the quantity of analyte to be determined in the test sample.

Matrix: All the constituents of the test sample with the exception of the analyte.

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.

Matrix Effect: An influence of one or more components from the sample matrix on the measurement of the analyte concentration or mass. Matrix effects may be observed as increased or decreased detector responses, compared with those produced by simple solvent solutions of the analyte.

Matrix Source: The origin of a test matrix used in method validation. A sample matrix may have variability due to its source. Different food matrix sources can be defined as different commercial brands, matrices from different suppliers, or in some cases different matrices altogether. For example, if a variety of food matrices with differing physical and chemical properties are selected, the number of sources for each food sample matrix may be one or more.

Matrix spike: An aliquot of a sample prepared by adding a known amount of analyte(s) to a specified amount of matrix. A matrix spike is subjected to the entire analytical procedure to establish if the method is appropriate for the analysis of a specific analyte(s) in a particular matrix. Also referred to as a *Laboratory Fortified Matrix*.

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 Detection Limit (MDL): The minimum amount or concentration of analyte in the test sample that can be reliably distinguished from zero. MDL is dependent on sensitivity, instrumental noise, blank variability, sample matrix variability, and dilution factor.

Method Development: The process of design, optimization and preliminary assessment of the performance characteristics of a method.

Method Validation: The process of demonstrating or confirming that a method is suitable for its intended purpose. Validation includes demonstrating performance characteristics such as accuracy, precision, specificity, limit of detection, limit of quantitation, linearity, range, ruggedness and robustness.

Method Verification: The process of demonstrating that a laboratory is capable of replicating a validated method with an acceptable level of performance.

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.

Precision: The closeness of agreement between independent test results obtained under specified conditions. The precision is described by statistical methods such as a standard deviation or confidence limit of test results. See also *Random Error*. Precision can be further classified as *Repeatability, Intermediate Precision*, and *Reproducibility*.

Qualitative Analysis/Method: Analysis/method in which substances are identified or classified on the basis of their chemical, biological or physical properties. The test result is either the presence or absence of the analyte(s) in question.

Quantitative Analysis/Method: Analysis/method in which the amount or concentration of an analyte may be determined (or estimated) and expressed as a numerical value in appropriate units with acceptable accuracy and precision.

Random error: Component of measurement error that in replicate measurements varies in an unpredictable manner. See also *Precision*.

Range: The interval of concentration over which the method provides suitable accuracy and precision.

Reagent Blank: Reagents used in the procedure taken through the entire method. Reagent Blanks are used to determine the absence of significant interference due to reagents or equipment used in the analysis.

Recovery: The proportion of analyte (incurred or added) remaining at the point of the final determination from the analytical portion of the sample measured. Usually recovery is expressed as a percentage.

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.

Reference standard: A standard, generally having the highest metrological quality available at a given location in a given organization, from which measurements are made or derived. Note: Generally, this refers to recognized national or international traceable

standards provided by a standards producing body such as the National Institute of Standards and Technology (NIST).

Repeatability: Precision obtained under observation conditions 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.

Representative Analyte: An analyte used to assess probable analytical performance with respect to other analytes having similar physical and/or chemical characteristics. Acceptable data for a representative analyte are assumed to show that performance is satisfactory for the represented analytes. Representative analytes should include those for which the worst performance is expected. Representative analytes are used mostly for non-targeted analysis and unknown screening procedures.

Representative Matrix: Matrix used to assess probable analytical performance with respect to other matrices, or for matrix-matched calibration, in the analysis of broadly similar commodities. For food matrices, similarity is usually based on the amount of water, fats, protein, and carbohydrates. Sample pH and salt content can also have a significant effect on some analytes.

Reproducibility: Precision obtained 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.

Ruggedness/Robustness: A measure of the capacity of an analytical procedure to remain unaffected by small but deliberate variations in method parameters and provides an indication of its reliability during normal usage.

Screening Analysis/Method: An analysis/method intended to detect the presence of analyte in a sample at or above some specified concentration (action or target level). Screening methods typically attempt to use simplified methodology for decreased analysis time and increased sample throughput.

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.

Standard: A substance of known identity and purity and/or concentration.

Standard Reference Material (SRM): A certified reference material issued by the National Institutes of Standards and Technology (NIST) in the United States. (www.nist.gov/SRM).

Systematic error: Component of measurement error that in replicate measurements remains constant or varies in a predictable manner. This may also be referred to as *Bias*.

Threshold Value: 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). See also *Cut-off Concentration*.

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.

APPENDIX 2 – Examples of Acceptability Criteria for Certain Performance Characteristics

Examples of acceptability criteria are found in references 7,9,10,14 and 18. No single set of acceptability is going to be truly applicable to all methodology covered in the FVM program. However a good starting point for many methods is found in the Codex Alimentarius Commission, Procedural Manual, Twenty-second ed., 2014 [10]

A. Quantitative Method Acceptability Criteria

Table A2.1. Method Criteria for Method Levels at Increasing Orders of Magnitude
(reproduced in part from reference 10, Table 4, p. 72 and reference 7)

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 (C_{ML})	10^{-9}	10^{-8}	10^{-7}	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-2}
Minimum applicable range	From 0.0006 to 0.0014 mg/kg	From 0.006 to 0.014 mg/kg	From 0.03 to 0.17 mg/kg	From 0.52 to 1.48 mg/kg	From 6.6 to 13.3 mg/kg	From 76 to 124 mg/kg	From 0.83 to 1.2 g/kg	From 8.8 to 11 g/kg
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
RSD_r**	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 matrice(s) 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 ($HorRat_r = RSD_r$ (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. ($HorRat_R = RSD_R / PRSD_R \leq 2$)

B. Qualitative Method Acceptability Criteria

There are significantly fewer examples of acceptability criteria for qualitative methods available. AOAC is using a relatively new Probability of Detection (POD) model as a way to characterize the performance of qualitative methods [9].

As discussed above, limit test screening methods, in general, should minimize false negatives particularly at the level of concern or reporting level. The occurrence of false positives is less critical since presumptive positives are further analyzed by quantitative or confirmatory methods. However, false positive rates should typically be less than 10-15% in order to avoid unnecessary confirmatory testing (14, 18).

Table A2.2. General Method Criteria for Limit Tests/Screening Methods

False Negative Rate	≤ 5% at the level of concern[1]
False Positive Rate	≤ 10-15%

[1] Acceptable false negative rate depends significantly on the intended purpose of the method.

APPENDIX 3 - Examples of Validation Plans

A. Extension to other matrices with the same analyte(s) at Level One Validation

This scheme represents an emergency use method extension plan for Matrix Y and Analyte Z. This plan utilizes two different sources of matrix. *In cases where a representative matrix is being used to characterize a whole family of commodities, it is recommended that additional, different commodities from that family are used as "sources".* Note that this plan is for emergency use only – the new matrix (or matrices) cannot be officially included in the scope of the method until at the minimum a Level Two Validation is performed.

Table A3.1. Plan for Matrix Extension (Level One Validation, Example)

	Matrix	Samples 1 & 2	Analyte Z Fortified Samples 3 & 4	Analyte Z Fortified Samples 5 & 6	Analyte Z Fortified Samples 7 & 8
Day 1	Matrix Y (Source 1)	Blank	½X Spike Level	X Spike Level	2X Spike Level
Day 1	Matrix Y (Source 2)	Blank	½X Spike Level	X Spike Level	2X Spike Level

Notes:
i. Test portion matrices listed as Matrix Y represent 2 different commercial brands.
ii. Fortification levels: fortification will be at the level of concern or action level (X) as stated in the method and at levels corresponding to 1/2X and 2X.
iii. Fortification of each matrix can be done on the same day.
iv. Other fortification plans meeting requirements specified in Table 1 may be used.

B. Extension to similar analytes in the same matrix at Level Two Validation

A validated method can be extended to other potential analyte(s) belonging to the same chemical group. For example, a toxin method can be extended to other toxins. An example of the composition of a set of validation studies for method extension is shown in the following table for new analytes Y and Z in canned corn from 3 different sources where the method is validated originally for analyte A in corn.

Table A3.2. Plan for Extension to Similar Analytes (Level Two Validation, Example)

	Matrix	Analyte Y fortification levels	Analyte Z fortification levels
Day 1	Corn 1,2,3	0, 1/2X, X, 2X	0, 1/2X, X, 2X
Day 2	Corn 1,2,3	0, 1/2X, X, 2X	0, 1/2X, X, 2X

Day 3	Corn 1,2,3	0, 1/2X, X, 2X	0, 1/2X, X, 2X

Notes:
i. Three different commercial brands of same product will be analyzed.
ii. Fortification levels: fortification will be at the level of concern or action level (X) as stated in the method and at levels corresponding to 1/2X and 2X.
iii. Each analyte will be analyzed in blank matrix and at 1/2X, X and 2X fortification levels.
iv. Simultaneous analysis of the analytes can be undertaken if warranted.
v. Other fortification plans meeting requirements specified in Table 1 may be used.

C. Validation at Level Two for single matrix and single analyte

This plan utilizes 3 different commercial brands of one matrix. The single matrix is being validated for a single analyte.

Table A3.3. Plan for Single Matrix and Single Analyte Level Two Validation (Example)

	Matrix 1 Source 1	Matrix 1 Source 2	Matrix 1 Source 3
Day 1	Blank Fortified (X)	Fortified (X) Fortified (2X)	Blank Fortified (1/2X)
Day 2	Fortified (2X) Fortified (1/2X)	Blank Fortified (1/2X)	Blank Fortified (2X)
Day 3	Fortified (1/2X) Fortified (X)	Fortified (2X) Blank	Fortified (X) Fortified (X)
Day 4	Fortified (2X) Blank	Fortified (X) Fortified (1/2X)	Fortified (2X) Fortified (1/2X)

Notes:
i Sample matrix, represents one matrix from 3 different sources of matrix.
ii Fortification levels: fortification will be at the level of concern or action level (X) as stated in the method and at levels corresponding to 1/2X and 2X.
iii Each of 3 different sources of matrix will be analyzed 8 times (replicate analyses) over the course of experiment, two times unfortified, two times fortified at each level.
iv. The validation will take place over a period of 4 days.
v. Other fortification plans meeting requirements specified in Table 1 may be used.

APPENDIX 4 – Selection of Representative Matrices

Two tools that can aid in selection of representative matrices and CRMs when designing a validation protocol for a method intended to have applicability to a broad scope of products are shown below. Food composition varies greatly making the validation of methods intended for a wide variety of foods a difficult balance between available resources and sufficient validation with a variety of food types.

A. Commodity groups and representative commodities

Table A4.1. Vegetable and Fruits, Cereals and Food of Animal Origin (reproduced in part from reference 14)

Commodity groups	Typical commodity categories	Typical representative commodities
1. High water content	Pome fruit	Apples, pears
	Stone fruit	Apricots, cherries, peaches
	Other fruit	Bananas
	Alliums	Onions, leeks
	Fruiting vegetables/cucurbits	Tomatoes, peppers, cucumber, melon
	Brassica vegetables	Cauliflower, Brussels sprouts, cabbage, broccoli
	Leafy vegetables and fresh herbs	Lettuce, spinach, basil
	Stem and stalk vegetables	Celery, asparagus
	Forage/fodder crops	Fresh alfalfa, fodder vetch, fresh sugar beets
	Fresh legume vegetables	Fresh peas with pods, peas, mange tout, broad beans, runner beans, French beans
	Leaves of root and tuber vegetables	Sugar beet and fodder beet tops
	Fresh Fungi	Champignons, canterelles
	Root and tuber vegetables or feed	Sugar beet and fodder beet roots, carrots, potatoes, sweet potatoes
2. High acid content and high water content	Citrus fruit	Lemons, mandarins, tangerines, oranges
	Small fruit and berries	Strawberry, blueberry, raspberry, black currant, red currant, white currant, grapes
	Other	Kiwifruit, pineapple, rhubarb

Table A4.1. Vegetable and Fruits, Cereals and Food of Animal Origin (continued)

Commodity groups	Typical commodity categories	Typical representative commodities
3. High sugar and low water content	Honey, dried fruit	Honey, raisins, dried apricots, dried plums, fruit jams
4a. High oil content and very low water content	Tree nuts	Walnuts, hazelnuts
	Oil seeds	Oilseed rape, sunflower, cotton-seed, soybeans, peanuts, sesame, etc.
	Pastes of tree nuts and oil seeds	Peanut butter, tahini, hazelnut paste
	Oils from tree nuts, oil seeds and oily fruits	Olive oil, rapeseed oil, sunflower oil, pumpkin seed oil
4b. High oil content and intermediate water content	Oily fruits and products	Olives, avocados and pastes thereof
5. High starch and/or protein content and low water and fat content	Dry legume vegetables/pulses	Field bean, dried broad bean, dried haricot bean (yellow, white/navy, brown, speckled), lentils
	Cereal grain and products thereof	Wheat, rye, barley and oat grain; maize, rice, whole meal bread, white bread, crackers, breakfast cereals, pasta
6. “Difficult or unique commodities”		Hops, cocoa beans and products thereof, Coffee, tea, spices
7. Meat (muscle) and Seafood	Red muscle	Beef, pork, lamb, game, horse
	White muscle	Chicken, duck, turkey
	Offal	Liver, kidney
	Fish	Cod, haddock, salmon, trout
	Crustaceans	Shrimp, scallop, crab
8. Milk and milk products	Milk	Cow, goat and buffalo milk
	Cheese	Cow and goat cheese
	Dairy products	Yogurt, cream
9. Eggs	Eggs	Chicken, duck, quail, and goose eggs
10. Fat from food of animal origin	Fat from meat	Kidney fat, lard
	Milk fat	Butter
	Fish oil	Cod liver oil

B. AOAC Food Matrix Triangle

The AOAC Food Matrix Triangle (Figure A4.1) can be used to categorize foods and food matrix reference materials into nine sectors based on relative fat, protein and carbohydrate content [9, 19, 20]. This tool can be useful in the validation of methods intended for a wide variety of food matrices and to help in categorizing similar food matrices for methods intended for more limited applicability.

Figure A4.1. Foods Partitioned into Sectors Based on Their Protein, Fat, and Carbohydrate Content