

TECHNICAL COMMUNICATIONS

AOAC INTERNATIONAL Methods Committee Guidelines for Validation of Qualitative and Quantitative Food Microbiological Official Methods of Analysis

PHILIP FELDSINE

BioControl Systems, Inc.
12822 SE 32nd St
Bellevue, WA 98005

CARLOS ABEYTA

U.S. Food and Drug Administration
Seafood Products Research Center
22201 23rd Dr, SE
Bothell, WA 98021-4421

WALLACE H. ANDREWS

U.S. Food and Drug Administration
Division of Microbiological Studies
5100 Paint Branch Pkwy
College Park, MD 20740-3835

Responding to a need for a guide for conducting Official Method validation studies of microbiological methods, AOAC utilized the experience of three microbiologists who have been active in the field of method validation. In collaboration, a document was prepared which covered the following areas: terms and their definitions associated with the Official Methods program (e.g., reference methods, alternative methods, and ruggedness testing), protocols and validation requirements for qualitative methods versus those for quantitative methods, the concept of the precollaborative study, ruggedness testing, tests for significant differences, performance indicators, and the approval process. After its preparation, this document was reviewed by the members of the Methods Committee on Microbiology and Extraneous Materials and by members of the Official Methods Board. Herein is presented the approved version of that document.

1 Scope

The purpose of this document is to provide comprehensive guidelines for conducting AOAC INTERNATIONAL (AOAC) food microbiological Methods Comparison or Precollaborative Studies and Interlaboratory Collaborative Studies for those methods submitted for AOAC® *Official Methods*SM Program (OMA) status. The guidelines define steps involved in the validation process including selection of the Study Director (SD), ruggedness testing, the Methods Comparison or Precollaborative Study that includes inclusivity/exclusivity testing, the Interlaboratory Collaborative Study, and the approval process by AOAC.

These microbiological guidelines were provided by the Methods Committee on Microbiology and Extraneous Materials as part of an initiative to specify validation criteria for Methods Comparison/Precollaborative Studies and Interlaboratory Collaborative Studies and to harmonize validation methodology with ISO 16140. They have been adopted by the AOAC Board of Directors. AOAC INTERNATIONAL thanks the Methods Committee on Microbiology and Extraneous Materials for their contribution in the development of these guidelines.

2 Applicability

These guidelines are applicable to the validation of any alternative method, whether proprietary or nonproprietary, that is submitted to AOAC for OMA status recognition. It is the intent of this guideline to harmonize validation procedures with ISO standard 16140, *Protocol for the Validation of Alternative Methods*. Data produced for an alternative method that satisfies the protocol requirements and acceptance criteria contained in ISO 16140 may be reciprocally recognized

and may apply toward validation requirements specified in these guidelines; however, some additional data could be required.

3 Terms and Definitions

3.1 Alternative Method

Method of analysis that demonstrates or estimates, for a given category of products, the same analyte (3.4) as is measured by using the corresponding reference method (3.2). The method can be proprietary or noncommercial and does not need to cover an entire analysis procedure, that is, from the preparation of samples to the test results.

3.2 Reference Method

Taken in order of priority herein, the appropriate AOAC, FDA/BAM or USDA reference culture procedure that is applicable to the analyte and sample type that the alternative method is intended to detect. Other internationally recognized methods may also be appropriate reference methods and will be considered on a case-by-case basis.

3.3 Validation of an Alternative Method

Demonstration that adequate confidence is provided when the results obtained by the alternative method are comparable to those obtained using the reference method using the statistical criteria contained in the approved validation protocol.

3.4 Analyte

Component measured by the method of analysis. In the case of microbiological methods, it is the microorganism or associated by-products (e.g., enzymes or toxins).

3.5 Qualitative Method

Method of analysis whose response is either the presence or absence of the analyte (3.4) detected either directly or indirectly in a certain amount of sample.

3.6 Quantitative Method

Method of analysis whose response is the amount of the analyte (3.4) measured either directly (e.g., enumeration in a mass or a volume), or indirectly (e.g., color absorbance, impedance, etc.) in a certain amount of sample.

3.7 Methods Comparison or Precollaborative Study

A study, performed by the organizing laboratory or Study Director (SD) of the alternative method against the reference method.

3.8 Interlaboratory Collaborative Study

Study of the alternative method's performance using common samples in numerous laboratories and under the control of the organizing laboratory or SD.

3.9 Fractional Recovery

Validation criterion that is satisfied when a common set of samples (e.g., inoculation level), yields a partial number of positive determinations and a partial number of negative determinations on that same set of samples. The proportion of positive samples should approximate 50% of the total number of samples in the set.

3.10 Ruggedness Testing (RT)

Subjecting the proposed alternative method to small procedural changes or environmental factors to determine what, if any, influence they have on method performance. It is not a formal part of the validation protocol and is not a submission requirement. It may provide useful information to the alternative methods sponsor. Analytical results from various collaborating laboratories may vary. Thus ruggedness testing is conducted in a single laboratory, generally the SD's organizing laboratory. This testing can be used to measure analytical differences in the same laboratory that may result from changes in operational or environmental conditions.

4 Validation Protocol

4.1 Validation Options

There are 3 validation options available through of AOAC and its subsidiary AOAC Research Institute (AOAC-RI). They are the AOAC[®] *Official Methods*SM Program (OMA), AOAC[®] *Peer-Verified Methods*SM Program (PVM) and AOAC[®] *Performance Tested Methods*SM Program (PTM). This document will only address the *Official Methods*SM Program (OMA). For information regarding the other 2 validation schemes, see References, Section 7, (5) and (7).

4.2 The OMA Method

AOAC OMA methods are validated by the interlaboratory Collaborative Study (CS), in which experienced, competent analysts (collaborators) work independently in different laboratories under the direction of a SD using a specific method to analyze replicated test samples for a particular analyte. The SD, who is supervised by the General Referee (GR), man-

ages the CS. The selected method may be subjected to ruggedness testing prior to the Methods Comparison or Collaborative Study to determine its behavior under various in-house operating conditions. (See Section 3.10.)

4.3 Participants of the *Official Methods*SM Process

There are many technical experts who participate in the *Official Methods*SM process. Each serves a unique, defined role and their functions are described below and in the *Official Methods*SM Program Manual.

4.3.1 The Official Methods Board (OMB)

OMB consists of a Chair and the chairs of the 11 Methods Committees. The OMB approves Final Action status for First Action methods.

4.3.2 Methods Committee (MC)

Reviews protocols for Methods Comparison and Interlaboratory Collaborative Studies, reviews completed studies, grants First Action status to new and revised methods, and recommends revisions or repeals, Final Action and surplus status to the Official Methods Board. Recommend scientists for appointment or termination as General Referees. Evaluate performances of General Referees; recommend re-appointments as appropriate.

4.3.3 Methods Committee Statistician (MCS)

Provides statistical review of protocols and completed Methods Comparison and Interlaboratory Collaborative Studies. Gives comments on precision and accuracy parameters of studies submitted. Works with the Study Directors in designing Interlaboratory Collaborative Studies to incorporate the correct statistical procedures. Identifies statistical problems and/or shortcomings in the evaluation of method performance needing research.

4.3.4 Methods Committee Safety Advisor (MCSA)

Reviews protocols and Collaborative Studies submitted to the Methods Committee. Examines the proposed method and study design for possible hazards and ensures that appropriate safety precaution statements are listed when necessary.

4.3.5 General Referee (GR)

A senior scientist in Microbiology with experience using the Collaborative Study process. Recommends scientists for appointment as Study Directors and advises them on method development concepts. Reviews the studies and recommends appropriate action on proposed methods to the Methods Committee.

4.3.6 Study Director (SD) or Organizing Laboratory Responsibilities

A volunteer required to be an expert in the particular area of study involved. Responsible for administration of the validation process and conducting the Interlaboratory study. Proposes the reference methodology, develops in-house validation data and develops a protocol for the Interlaboratory Collaborative validation of the method following AOAC guidelines.

The responsibilities of the SD/Organizing Laboratory are summarized as follows:

- selects a method to be validated
- conducts in-house ruggedness testing (optional procedure);
- conducts a Methods Comparison/Precollaborative Study and submits to AOAC for Methods Committee approval;
- prepares a Collaborative Study protocol and submits to AOAC for Methods Committee approval;
- prepares samples, ships samples to participants, conducts study and analyzes results;
- prepares manuscripts containing results and methods in AOAC format and submits to GR, MCS, MCSA, MC, and OMB for review and approval upon completion of the PCS and CS.

4.3.7 Collaborators

Responsible for performing the test analyses as directed by the SD/Organizing Laboratory, following the proposed or reference methods precisely and reporting all results as required. Any experienced scientist in microbiology may be chosen from a diversity of laboratories that include regulatory agencies, industry and academia.

5 Qualitative Methods—Technical Protocol for Validation

5.1 Methods Comparison or Precollaborative Study (PCS)

5.1.1 Measurement Protocol

The Precollaborative Study (PCS), also known as the Methods Comparison Study, is a formal submission requirement for OMA microbiology methods and is normally conducted in the organizing laboratory. It precedes the Interlaboratory Collaborative Study. The purpose of a Precollaborative Study is to define the applicability claims of a proposed OMA method by demonstrating the applicability of the method to various food categories. For OMA methods, the applicability statement immediately follows the method title. The applicability statement for microbiological methods is generally concerned with target analyte and food type coverage.

5.1.1.1 Reference Procedure

Alternative methods are always compared to a reference culture method. Reference culture methods contained in AOAC OMA, FDA/BAM or USDA (for meat and poultry products) are examples of reference methods. (See 3.2.)

5.1.1.2 Number of Laboratories

The PCS is normally conducted in the organizing laboratory.

5.1.1.3 Food Categories

Refer to the Recommended Food Categories and Food Types for Microbiological Methods contained in Annex A. Analyze at least twenty (20) diverse food types selected from the food categories contained in Annex A, if the applicability claim is for all or most foods.

Note: AOAC INTERNATIONAL has recognized *Salmonella* methods for ALL FOODS and has recognized other target microorganisms, including *E. coli* O157, *Listeria*, and *Campylobacter* methods for only the range of food categories or specific food types included in the PCS and CS.

If it is sought to validate the method for all or most foods, study at least 6 categories of food and 20 food types. This number may be reduced to 1, 2, 3, 4, or 5 categories if the validation of alternative method is restricted to these stated categories. The recommended categories are listed in Annex A.

If it is not possible to acquire a sufficient number of naturally contaminated food types for each of the categories, artificial contamination of food samples is permissible.

Food selection must cover as broad a range as is practical and should include foods that have been implicated in food poisoning or recall situations.

5.1.1.4 Inoculating Cultures and Conditions of Microorganisms

Foods are commonly inoculated with the organism of interest since naturally contaminated products are infrequently available. For a genus level method, such as *Salmonella* detection, numerous strains representing different serovars are required.

Typically a different isolate, strain, serotype or species is used for each food type. The product inoculation should be conducted with a pure culture of one strain. Mixed cultures are not recommended.

Microorganisms in processed foods are typically stressed, thus the contaminating microorganisms are also stressed for these types of foods. Microorganism stress may occur at the time of inoculation or during preparation of the food. Raw, unprocessed foods may be inoculated with unstressed organisms. Lyophilized inocula are generally used for dry powder/granulated foods and wet inocula are used for wet foods. Inoculated samples of solid food types, if included, are held at appropriate storage conditions to stabilize the population prior to analysis.

5.1.1.5 Levels of Inoculum and Controls

Each food type is divided into at least 2 portions. One portion serves as the negative control, one portion is inoculated at a level that will produce fractional recovery (see 3.9). Control and inoculated test samples should be prepared at the same time. Depending on the organizing laboratory's confidence in satisfying this validation criterion, it may be advisable to prepare a third portion that has a high inoculum level. This level is optional as the only acceptance requirement is fractional recovery at one level. The target for the low inoculum level is typically set at the lowest detection limit of the test method, e.g., 1–5 cfu/25 g test portion. The high inoculum is set at 10–50 cfu/25 g test portion. Additional inoculum levels may be added as necessary.

More than one inoculum level may be used to increase the probability that one level will exhibit fractional recovery. Inoculum levels that test totally positive or negative are not useful in determining the lowest detection limit of the method and do not satisfy validation requirements.

Most probable number (MPN) quantification of contamination levels is conducted in the organizing laboratory on the day that the analysis of test samples is initiated [see References, Section 7 (12)].

5.1.1.6 Number of Test Samples

The number of test portions per inoculum level is 20.

5.1.1.7 Naturally Contaminated Test Samples

At least 2 lots of each naturally contaminated food type are required. However, naturally contaminated products are infrequently available to most analysts. An effort should be made to obtain them as they are most representative of the method usage environment. For these products, there is no negative control. Twenty replicates are analyzed per lot. If all test portions are positive, dilute the test sample to obtain fractional positives and repeat analysis of the lot.

5.1.1.8 Need for Competitive Microflora

It is more realistic and challenging to include microorganisms that act as competitors to the analyte microorganisms. The purpose of including these organisms is to more closely simulate conditions found in nature. It is sufficient to demonstrate this recovery in one food type. This requirement may be satisfied in the Methods Comparison Study. The competitor contamination levels, which may be naturally occurring or artificially introduced, should be at least one logarithm higher than the target microorganism.

5.1.1.9 Inclusivity and Exclusivity

Inclusivity or sensitivity is the ability of the alternative method to detect the target analyte from a wide range of strains. Exclusivity or specificity is the lack of interference in the alternative method from a relevant range of nontarget strains, which are potentially cross-reactive.

Generate inclusivity and exclusivity data to substantiate that the method is reactive for the major serotypes of the specified microorganism and is nonreactive to other related genera and/or species. Select at least 50 pure strains of the specific microorganism and select at least 30 strains of potentially competitive strains to be analyzed as pure culture preparations. For *Salmonella* methods, this number of target analyte strains is increased to at least 100 strains that are selected to represent the majority of known somatic groups of *Salmonella*.

5.1.1.10 Contaminated Controls

Inoculated test samples and uninoculated controls are prepared at the same time. If any uninoculated control test portion is positive for the inoculated microorganism, the results are invalid and the run is repeated because it is assumed that cross contamination has occurred. Control samples are not included with naturally contaminated food types.

5.2 Interlaboratory Collaborative Study (CS)

5.2.1 Measurement Protocol

The Collaborative Study (CS) also known as the Interlaboratory Collaborative Study, is a formal submission requirement for OMA methods. It follows the Methods Comparison or Precollaborative Study. The purpose of the Collaborative Study is to provide a realistic estimate of the attributes of a method, particularly systematic and random deviations, to be expected when the method is used in actual practice.

5.2.1.1 Reference Procedure

Alternative methods are always compared to a reference culture method. Reference culture methods contained in AOAC OMA, FDA/BAM or USDA (for meat and poultry products) are examples of reference methods.

5.2.1.2 Number of Laboratories

A minimum of 10 valid laboratories data sets per food type is needed. The SD should plan on including at least 12–15 laboratories per food because some laboratories are generally eliminated for various reasons. No guidelines have been established for qualitative methods involving very expensive equipment. It is advisable to consult with General Referee and Committee Statistician.

5.2.1.3 Food Categories

The number of different food categories depends on the applicability of the method. If the method is specific to only one category (e.g., detection of *Campylobacter* in oysters), only one type of food need be included. If the applicability is wider (e.g., detection of *Salmonella* in all foods), then 6 food categories shall be included in the CS. As mentioned previously, the data from both the PCS and CS studies form the basis for defining the method applicability statement.

5.2.1.4 Inoculating Cultures

The inoculating microorganisms must represent different genera, species and/or toxin-producing microorganisms that are intended to be included in the method applicability statement. The choice of cultures should be broad enough to represent the inherent variation in the microorganisms of interest.

Typically a different isolate, strain, serotype or species is used for each food type. The product inoculation should be conducted with a pure culture of one strain. Mixed cultures are not recommended.

Microorganisms in processed foods are typically stressed. Thus, the contaminating microorganisms are also stressed for these types of foods. Microorganism stress may occur at the time of inoculation or during preparation of the food. Raw, unprocessed foods may be inoculated with unstressed organisms. Lyophilized inocula are generally used for dry powder/granulated foods and wet inocula are used for wet foods. Inoculated samples of solid food types, if included, are held at appropriate storage conditions to stabilize the population prior to analysis. Products must be inoculated to assure homogeneity of the analyte in each food type.

5.2.1.5 Levels of Inoculum and Controls

Each food type is divided into 3 portions. One portion serves as the negative control, one portion is inoculated at a level (usually the low inoculum level) that will produce fractional recovery (*see* 3.9), and a third portion is inoculated at a high inoculum level. Control and inoculated test samples should be prepared at the same time. The target for the low inoculum level is typically set at the lowest detection limit of the test method, e.g., 1–5 cfu/25 g test portion. The high inoculum is set at 10–50 cfu/25 g test portion.

Most probable number (MPN) quantification of contamination levels is conducted in the organizing laboratory on the day that the analysis of test samples is initiated. MPNs are conducted according to the reference culture method specified in the protocol.

5.2.1.6 Number of Test Samples

Six test portions per analyte level for each food type and 6 negative (uninoculated) control portions for each food type are required. Test portions are to be blind coded when sent to participating laboratories for analysis.

5.2.1.7 Use of Artificially and Naturally Contaminated Test Samples

The use of both naturally and artificially contaminated test samples is strongly encouraged. Naturally contaminated foods are not always available. Therefore, artificially contaminated test samples are acceptable.

5.3 Test for Significant Difference and Performance Indicators

These are applicable to both the Precollaborative and Collaborative Studies. Examine data to determine whether any laboratory shows consistently aberrant results, which differ from the determinations from the other laboratories to a greater degree than could be reasonably expected or found by chance alone. Perform outlier tests in order to discard the outlying values [*see* References, Section 7 (8)].

5.3.1 Test for Significant Difference (χ^2)

The proportion confirmed positive for the alternative method must not be statistically different from the proportion confirmed positive to the reference method for each food type and each inoculation level. McNemar's test [a Chi square (χ^2) test] is used to compare the proportions for the methods.

A Chi square value <3.84 indicates that the proportions positive for the alternative and reference methods are not statistically different at the 5% level of significance. This criterion must be satisfied for each level of each food type. However, a significant difference between the proportions positive for the 2 methods is acceptable provided that the alternative method demonstrates superior recovery to the reference method.

Chi square, as defined by McNemar, is:

$$\chi^2 = \frac{(|a-b|-1)^2}{a+b}$$

where a = test samples confirmed positive by the alternative method but tested negative by the reference method, b = test samples that tested negative by the alternative method but are confirmed positive by the reference method. Refer to Table 1 for method of calculation from study data.

A Chi square value ≥ 3.84 indicates that the proportions confirmed positive for the alternative and reference methods differ significantly at $P \leq 0.05$. If the McNemar test indicates statistical significance when applied to the analytical results from the analysis of a food at an inoculum level, that food must be removed from the applicability statement or the method must be modified and additional testing performed to demonstrate that the results are now acceptable.

If no assignable cause for the failure can be found, the failure may be assumed to be attributable to random error. In this case, the specific food category can be re-evaluated by testing 3 lots of the food product according to the above protocol

5.3.2 Performance Indicators

The 4 performance indicators for qualitative methods are sensitivity, specificity, false negative rate and false positive rate. Together with the test for significant difference (5.3.1), these performance indicators provide a total assessment of the test method. Table 1 provides the calculations of these 4 performance indicators. Enter method results in the Table 1 format and calculate the performance indicators according to the formulae provided.

5.3.2.1 Sensitivity Rate (p+) for a Food Type and Inoculation Level

The probability that the method, alternative or reference, will classify a test sample as positive, given that a test sample is a known positive. A known positive refers to the confirmation of inoculated analyte.

Sensitivity rate is defined as: Total number of confirmed positive test portions by the method divided by total number of confirmed positive test portions by both the alternative and reference methods.

5.3.2.2 Specificity Rate (p-) for a Food Type and Inoculation Level

The probability that the method will classify the test sample as negative, given that the test sample is a known negative. A known negative refers to a confirmed negative test portion.

Specificity rate is defined as: Total number of analyzed negative test portions by the method divided by total number of confirmed negative test portions by both the alternative and reference methods.

For microbiological methods involving a confirmation step, a presumptive positive result is taken through the cultural procedure and confirmed to be a positive or determined to be a negative. In other words, the confirmation procedure allows the sample to be reclassified as a known positive or a known negative. As such, the specificity rate of results after confirmation is always 100%.

5.3.2.3 False Negative Rate (pf-) for a Food Type and Inoculation Level

The probability that a test sample is a known positive, given that the test sample has been classified as negative by the method. $pf-$ is the number of misclassified known positives divided by the total number of positive test samples (misclassified positives plus the number of correctly classified known positives) obtained with the method.

Incidence of false negatives equals 100 minus the sensitivity rate.

5.3.2.4 False Positive Rate (pf+) for a Food Type and Inoculation Level

The probability that a test sample is a known negative, given that the test sample has been classified as positive by the method. $pf+$ is the number of misclassified known negatives divided by the total number of negative test samples (misclassified negatives plus the number of correctly classified known negatives) obtained with the method.

Incidence of false positives equals 100 minus the specificity rate.

6 Quantitative Methods—Technical Protocol for Validation

6.1 Methods Comparison or Precollaborative Study

6.1.1 Measurement Protocol

The Precollaborative Study (PCS), also known as the Methods Comparison Study, is a formal submission requirement for OMA microbiology methods and is normally conducted in the organizing laboratory. It precedes the Interlaboratory Collaborative Study. The purpose of a Precollaborative Study is to define the applicability claims of a proposed OMA microbiology method by demonstrating the applicability of the method to various food categories. For OMA methods, the applicability statement immediately follows the method title. The applicability statement for microbiological methods is generally concerned with target analyte and food type coverage.

6.1.1.1 Reference Procedure

Alternative methods are always compared to a reference culture method. Reference culture methods contained in AOAC OMA, FDA/BAM or USDA (for meat and poultry products) are examples of reference methods.

6.1.1.2 Number of Laboratories

The PCS is normally conducted in the organizing laboratory.

6.1.1.3 Food Categories

Refer to the Recommended Food Categories and Food Types for Microbiological Methods contained in Annex A. Analyze at least twenty (20) diverse food types selected from the food categories contained in Annex A, if the applicability claim is for all or most foods.

Note: AOAC INTERNATIONAL has recognized *Salmonella* methods for ALL FOODS and has recognized other target microorganisms, including *E. coli* O157, *Listeria*, and *Campylobacter* methods for only the range of food categories or specific food types included in the PCS and CS.

6.1.1.4 Inoculating Cultures

Inoculating cultures are used only if the method is for a specific target analyte which may not routinely be found in all food types (e.g., enumeration of *Listeria* spp.) or a certain type has been referenced and the subject flora (e.g., yeast) has not been found in measurable levels.

6.1.1.5 Levels of Inoculum and Controls and Number of Test Samples

For the artificially contaminated food types, 3 inoculated levels (high, medium, and low) and one uninoculated control are required. For each of these 3 levels and for the controls, test 5 samples by the alternative method and 5 samples by the reference method. The low level should be at the limit of detection, and the medium and high levels may be approximately one and 2 log units higher, respectively. Intermediate levels may be added to improve precision but they are not required.

6.1.1.6 Use of Artificially and Naturally Contaminated Test Samples

Approximately 50% of the food types should be naturally contaminated unless the method is for a specific microorganism that may not be naturally occurring in that number of food types. For the food types that are naturally contaminated, 3 different lots are required per food type. There are no uninoculated controls for the food types that are naturally contaminated.

The balance of the food types may be either naturally contaminated or artificially contaminated.

6.1.1.7 Need for Competitive Flora

For those alternative methods that are specific for target organisms, it is more realistic and challenging to include microorganisms that act as competitors to the analyte microorganisms. The purpose of including these organisms is to more closely simulate conditions found in nature. It is sufficient to demonstrate this recovery in one food type. This requirement may be satisfied in the Methods Comparison Study. The competitor contamination levels, which may be naturally occurring or artificially introduced, should be at least one logarithm higher than the target microorganism.

6.1.2 Inclusivity and Exclusivity

This requirement is not applicable to total viable count, yeast and mold count or similar total enumeration methods that are not directed at specific microorganisms.

Inclusivity is the ability of the alternative method to detect the target analyte from a wide range of strains. Exclusivity is the lack of interference in the alternative method from a relevant range of nontarget strains, which are potentially cross-reactive.

Generate inclusivity and exclusivity data to substantiate that the method is reactive for the major serotypes of the specified microorganism and is nonreactive to other related genera and/or species.

In microbiology, except in the case of total counting methods, the inclusivity and exclusivity is established by the analysis of:

- at least 30 pure strains of the specific microorganism being studied.
- at least 20 pure strains of potentially competitive strains, taken from strains known to commonly cause interference with the target analyte.

6.2 Interlaboratory Collaborative Study (CS) Protocol Design—Quantitative Methods**6.2.1 Measurement Protocol**

The Collaborative Study (CS), also known as the Interlaboratory Collaborative Study, is a formal submission requirement for OMA methods. It follows the Methods Comparison or Precollaborative Study. The purpose of the Collaborative Study is to provide a realistic estimate of the attributes of a method, particularly systematic and random deviations, to be expected when the method is used in actual practice.

6.2.1.1 Reference Procedure

Alternative methods are always compared to a reference culture method. Reference culture methods contained in AOAC OMA, FDA/BAM or USDA (for meat and poultry products) are examples of reference methods.

6.2.1.2 Number of Laboratories

A minimum of eight laboratories reporting valid data for each food type is required. It is suggested that at least 10–14 laboratories begin the analysis. In special cases involving very expensive equipment or specialized laboratories, the study may be conducted with a minimum of 5 laboratories.

6.2.1.3 Food Categories

The number of different food categories depends on the applicability of the method. If the method is specific to only one category (e.g., enumeration of *Listeria* spp. in fresh, unpasteurized cheese), only one type of food type needs to be included. If the applicability is wider (e.g., enumeration of total plate count in foods), then 6 food categories shall be included in the CS. As mentioned previously, the data from both the PCS and CS form the basis for defining the method applicability statement.

6.2.1.4 Enumeration of Specific Microorganisms

If the alternative method is for enumeration of a specific microorganism, it may be necessary to include certain food types known to support the growth of such analytes. The inoculating microorganisms must represent different genera, species and/or toxin-producing microorganisms that are intended to be included in the method applicability statement. The choice of strains should be broad enough to represent the inherent variation in the microorganisms of interest.

6.2.1.5 Inoculating Cultures

Inoculating cultures are used only if the method is for a specific target analyte which may not routinely be found in all food types (e.g., enumeration of *Listeria* spp.) or lots of a certain type that have been analyzed and the subject flora (e.g., yeast) have not been found in measurable levels.

6.2.1.6 Levels of Inoculum and Controls

For each food type, prepare 3 contamination levels (high, medium and low) plus an uninoculated control level. The lowest level should approximate the limit of detection of the alternative method and the medium and high levels may be approximately one and 2 logarithms higher, respectively. Uninoculated control samples are also included for each inoculated food type.

6.2.1.7 Number of Test Samples per Level

For each food type (3 contamination levels and the uninoculated control) 2 test portions are analyzed by the alternative method and 2 test portions are analyzed by the reference method.

6.2.1.8 Use of Artificially and Naturally Contaminated Test Samples

The use of both naturally and artificially contaminated test samples is strongly encouraged. Because naturally contaminated foods are not always available particularly for methods applicable to specific microorganisms, artificially contaminated test samples may be used.

6.3 Quantitative Method Calculations—PCS and CS

For detailed explanation of the quantitative method calculations to be performed, refer to Youden and Steiner, References, Section 7 (10).

6.3.1 General Considerations

In microbiology, the data often does not show a normal statistical distribution. In order to get a more symmetric distribution, counts should be transformed into logarithms. Other more complicated procedures can also be used.

Data from study results should first be plotted. The vertical y-axis (dependent variable) is used for the alternative method and the horizontal x-axis (independent variable) for the reference method. This independent variable x is considered to be accurate and have known values. Usually major discrepancies will be apparent.

6.3.1.1 Initial Review of Data

The Study Director may first plot the Collaborative Study results, material by material (or one value against the other) for a split level [Youden pair], value vs laboratory, preferably in ascending or descending order of reported average concentration. Usually major discrepancies will be apparent: displaced means, unduly spread replicates, outlying values, differences between methods, consistently high or low laboratory rankings, etc.

Only valid data should be included in the statistical analysis. 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 previously.

6.3.2 Outliers

It is often difficult to make reliable estimations (average, standard deviation, etc.) with a small bias and in presence of outliers. Data should be examined to determine whether any laboratory shows consistently high or low values or an occasional result, which differs from the rest of the data by a greater amount than could be reasonably expected or found by

chance alone. Perform outlier tests (Cochran, Dixon, Grubbs) in order to discard the outlying values and to obtain a better estimate [see References, Section 7 (10) and (11)].

6.3.3 Performance Indicators

Performance indicators for quantitative methods include repeatability, reproducibility and relative standard deviations.

6.3.3.1 Analysis of Variance

To determine if the alternative method mean is not statistically different from the reference method mean, a one way analysis of variance or a paired *t*-test by food type and inoculation level is performed [see References, Section 7 (10)].

6.3.3.2 Repeatability (s_r)

The repeatability is within laboratory precision, designated s_r or the closeness of agreement between successive and independent results obtained by the same method on identical test material, under the same conditions (e.g., apparatus, operator, laboratory and incubation time).

6.3.3.3 Repeatability Value (r)

The repeatability value is the value below which the absolute difference between 2 single test results obtained under repeatability conditions may be expected to lie within 95% probability.

6.3.3.4 Reproducibility (s_R)

The reproducibility is among laboratories precision, designated s_R , or the closeness of agreement between single test results on identical test material using the same method and obtained by operators in different laboratories using different equipment.

6.3.3.5 Reproducibility Value (R)

The reproducibility value is the value below which the absolute difference between single test results under reproducibility conditions may be expected to lie within 95% probability.

6.3.3.6 Relative Standard Deviation (RSD)

Relative standard deviation (RSD) is a useful measure of precision in quantitative studies. RSD is computed by dividing s_R and s_r by the mean. This is done so that one can compare variability of sets with different means. RSD values are independent of the amount of analyte over a reasonable range and facilitate comparison of variabilities at different concentrations. The result of a collaborative test may be summarized by giving the RSD for repeatability (RSD_r) and RSD for reproducibility (RSD_R).

6.3.3.7 Calculations

For details, refer to Youden and Steiner, References, Section 7 (10).

7 References

- (1) AOAC INTERNATIONAL (1989) *Handbook for AOAC Members*, 6th Ed., AOAC INTERNATIONAL, Gaithersburg, MD 20877-2417 USA
- (2) Andrews, W.H. (1987) *J. AOAC Int.* **70**, 931–936
- (3) Andrews, W.H. (1996) *Food Control* **7**, 19–29
- (4) Andrews, W.H. (1996) *Trends in Food & Tech.* **7**, 147–151
- (5) AOAC[®] *Peer-Verified MethodsSM* Program, AOAC INTERNATIONAL, Gaithersburg, MD 20877-2417 USA
- (6) AOAC INTERNATIONAL (1995) *Quick and Easy*, AOAC INTERNATIONAL, Gaithersburg, MD 20877-2417 USA
- (7) AOAC Research Institute. AOAC[®] *Performance Tested MethodsSM* Program, AOAC INTERNATIONAL, Gaithersburg, MD 20877-2417 USA
- (8) McClure, F.D. (1990) *J. AOAC Int.* **73**, 953
- (9) ISO 16140 (2000) *Microbiology of food and animal feeding stuffs—Protocol for the validation of alternative methods*, ISO TC 34/SC 9
- (10) Youden, W.J., & Steiner, E.H. (1987) *Statistical Manual of the Association of Official Analytical Chemists Qualitative and Quantitative Checklist*
- (11) Dixon, W.J., & Massey, F.J. (1969) *Introduction to Statistical Analysis*, 3rd Ed., McGraw-Hill Book Co., New York, NY
- (12) *Bacteriological Analytical Manual* (1998) 8th Ed., Rev. A, AOAC INTERNATIONAL, Gaithersburg, MD

Table A.2 Food categories relevant to nonpathogenic microorganisms

Type of product	Yeasts and molds	Lactic acid bacteria	Total viable counts	Coliforms	<i>Escherichia coli</i>
A. Meat products					
Raw	x	x	x	x	x
Heat processed		x	x	x	
Frozen	x		x	x	x
Fermented	x	x	x		
Cured		x	x		
B. Poultry products					
Raw	x	x	x	x	x
Heat processed		x	x	x	
Frozen	x		x	x	x
Others					
C. Fish and seafood products					
Raw	x	x	x	x	x
Heat processed		x	x	x	
Frozen	x		x	x	x
Smoked	x	x	x	x	
D. Fruits and vegetable based products					
Raw	x	x	x	x	x
Heat processed			x	x	
Frozen	x		x	x	
Dry	x		x	x	
Fermented	x		x		
Cured/salted	x		x		
Juice/concentrates	x	x	x		
Low moisture	x				
E. Dairy products					
Raw	x	x	x	x	x
Heat processed			x	x	
Frozen	x		x	x	x
Fermented	x				x
Dry			x	x	
F. Chocolate/bakery products					
Low moisture/IMF	x		x	x	
Dry			x	x	
Milk chocolate	x		x	x	
G. Animal feeds					
Low moisture	x		x	x	
Dry pet food	x			x	x
H. Pasta					
Uncooked noodles	x		x	x	
I. Miscellaneous					
Dressings	x	x	x	x	x
Spices			x		x
Mayonnaise	x	x	x		x
Egg and derivatives			x	x	
Cereals/rice			x	x	

Notes: Examples of representative food products contained in food categories listed in Annex A, Tables A.1 and A.2

Food category

A. Meat products	Ground beef, ground pork, meat by products, glandular products, frog legs, rabbit carcasses, lamb, sausage, frankfurters, lunch meat, beef jerkies, meat substitutes
B. Poultry	Ground chicken, ground turkey, cooked chicken, raw chicken parts
C. Fish and seafood products	Raw shrimp, fish sticks, surimi, raw fish filet, raw oysters, raw mussels, raw clams, cooked crawfish, smoked fish, pasteurized crab meat
D. Fruits and vegetable based products	Fresh/frozen fruits or dried fruits, orange juice, apple juice, apple cider, tomato juice, melon cubes, berries Pecans, walnuts, peanut butter, coconut, almonds Lettuce, spinach, kale, collard greens, cabbage, bean sprouts, seed sprouts, spent water from bean sprouts and seed sprouts, peas, mushrooms, green beans
E. Dairy products	Yogurt, cottage cheese, hard and soft cheeses, raw or pasteurized liquid milk (skim, 2% fat, whole, buttermilk), infant formula, coffee creamer, ice cream, nonfat dry milk/dry whole milk, dried buttermilk, dried cheese spray
F. Chocolate/bakery products	Frosting and topping mixes, candy and candy coating, milk chocolate
G. Animal feed	Dry pet food, meat and bone meal, chicken and feather meal
H. Uncooked pasta	Uncooked noodles, macaroni, spaghetti
I. Miscellaneous	Shell eggs, liquid whole eggs, oral or tube feedings containing egg, dried whole egg or dried egg yolk, dried egg whites 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 Wheat flour, casein, cake mixes, whey, nonfat dry milk/dry whole milk, corn meal, dried whole egg or dried egg yolk, dried egg whites, soy flour, dried yeast, cereals, dried buttermilk, dry cheese spray

Enter study data in Table 1 and calculate the test for significant difference and performance indicators as shown below. This calculation should be done at each inoculation level.

Table 1. Calculation of performance indicators after generalized categorization of test samples (8)

Status of test samples ^a	Test result ^b		
	Positive	Negative	Total
Positive	N ₁₁	N ₁₂	N _{1•}
Negative	N ₂₁	N ₂₂	N _{2•}
Total	N _{•1}	N _{•2}	N = N _{1•} + N _{2•} or N _{•1} + N _{•2}

^a Status defined by reference method.

^b Test result is defined as the alternative method result. For sensitivity calculation, use method results after confirmation.

N = Number of results in any particular cell. First subscript is the row and second subscript is the column. Examples: N₁₁: Row 1, Column 1; N_{1•}: Row 1 total; N_{2•}: Column 2 total; N₁₂: Row 1, Column 2.

Chi square (χ^2) = $((|N_{12} - N_{21}|) - 1)^2 / (N_{12} + N_{21})$, degrees of freedom (df) = 1.

Sensitivity rate = $p+ = N_{11}/N_{1•} = P(T^+/S^+)$ where $P(T^+/S^+)$ is the probability that the test is positive given that a randomly chosen test sample is a true positive. Refer to Table 1, define test result as results of completed method (screen assay with confirmation) of positive assay.

Specificity rate = $p- = N_{22}/N_{2•} = P(T^-/S^-)$ where $P(T^-/S^-)$ is the probability that the test is negative given that a randomly chosen test sample is a true negative. Refer to Table 1, define test result as results of completed method (screen assay with confirmation) of positive assay. Specificity should be recalculated with the negative control results included.

False negative rate = $pf- = N_{12}/N_{1•} = 1 - \text{sensitivity rate} = P(T^-/S^+)$ where $P(T^-/S^+)$ is the probability that the test is negative for samples that contain the analyte.

False positive rate = $pf+ = N_{21}/N_{2•} = 1 - \text{specificity rate} = P(T^+/S^-)$ where $P(T^+/S^-)$ is the probability that the test is positive for samples that do not contain the analyte.