International Stakeholder Panel on Alternative Methods

(ISPAM)

AOAC INTERNATIONAL ANNUAL MEETING
Friday, September 16, 2011
8:30am – 5:30am

Sheraton New Orleans Hotel/New Orleans, Louisiana
Meeting of the International Stakeholder Panel on Alternative Methods (ISPAM) and Working Groups

DRAFT AGENDA
Friday, September 16, 2011
8:30 am – 5:30 pm
Sheraton New Orleans Hotel, New Orleans, Louisiana

SESSION I:
ISPAM JOINT SESSION (Bayside C) 8:30 am – 9:45 am

A. Welcome and Introductions, and overview of ISPAM goals and activities to date – Russ Flowers, Mérieux NutriSciences and Chair, ISPAM

B. Summary of activities and overview of Meeting Agenda of the Working Group on Microbiological Validation Guidelines for Alternative Methods – Russ Flowers, Mérieux NutriSciences and WG Chair

C. Summary of work to date and Meeting Agenda of the Working Group on Qualitative Chemistry Guidelines for Alternative Methods – Bert Pöpping, Eurofins and WG Chair

D. Overview of different statistical models, the strengths and weaknesses of each model and the appropriate application of each – Paul Wehling, General Mills and Chair WG on Statistics

E. Standard Method Performance Requirements (SMPRs) and proposal to develop SMPRs to complement general validation guidelines – Scott Coates, AOAC Chief Scientific Officer

SESSION II:
ISPAM WORKING GROUP ON MICROBIOLOGICAL GUIDELINES (Bayside C)
10:00 am – 3:00 pm

A. Discussion of proposal from the Sub-Group (SG) to harmonize the use of reference methods – Phil Feldsine, BioControl Systems

B. Discussion of proposal from the Sub-Group (SG) to harmonize the number of levels/samples/fractional positives – Phil Feldsine, BioControl Systems

C. Discussion of proposal from the Sub-Group (SG) to harmonize the selection of food/category (sample matrix) – Ron Johnson, BioMérieux
LUNCH 12:00 pm – 1:00 pm (Grand Chenier, 5th Floor)

D. Discussion of proposal from the Sub-Group (SG) to harmonize results analysis & criteria/statistical analysis – Morgan Wallace, DuPont Qualicon

E. Discussion of proposal from the Sub-Group (SG) to harmonize number of samples/replicates/method comparison & collaborative – Wendy Lauer, Bio-Rad

F. Summary of recommendations that will be proposed to ISPAM for vote and discussion of next steps members – Russ Flowers, Mérieux NutriSciences and Chair, ISPAM

SESSION III:
ISPAM WORKING GROUP ON QUALITATIVE CHEMISTRY GUIDELINES FOR ALTERNATIVE METHODS (Oak Alley) 10:00 am – 3:00 pm

A. Overview and Comparison of Current Qualitative Chemistry Method Validation Guidance – Bert Pöpping, Eurofins and WG Chair

B. Presentation of Draft Guidelines for Validation of Qualitative Chemistry alternative methods – Scott Coates, AOAC Chief Scientific Officer

LUNCH 12:00 pm – 1:00 pm (Grand Chenier, 5th Floor)

C. Development of Standard Method Performance Requirements for Gluten as an example – Scott Coates, AOAC Chief Scientific Officer

SESSION IV:
ISPAM JOINT SESSION (Bayside C) 3:15 pm – 5:30 pm

A. Presentation of recommendations from WG on Microbiological Guidelines for discussion and vote by ISPAM voting members – Russ Flowers, Mérieux NutriSciences and Chair, ISPAM

B. Presentation of recommendations from WG on Qualitative Chemistry Validation Guidelines for vote by ISPAM voting members – Bert Pöpping, Eurofins and WG Chair

C. Next Steps - Russ Flowers, Mérieux NutriSciences and Chair, ISPAM
   i. Implementation of recommendations – how will recommendations be implemented by agencies and organizations
   ii. Should WGs develop a standard validation study design template
   iii. Should the goal be adoption of a standard validation study design by international agencies and organizations
   iv. Should WGs develop standard method performance requirements for new and updated study protocols
AOAC INTERNATIONAL
ANTITRUST POLICY
STATEMENT AND GUIDELINES

**Introduction**

It is the policy of AOAC INTERNATIONAL (AOAC) and its members to comply strictly with all laws applicable to AOAC activities. Because AOAC activities frequently involve cooperative undertakings and meetings where competitors may be present, it is important to emphasize the ongoing commitment of our members and the Association to full compliance with national and other antitrust laws. This statement is a reminder of that commitment and should be used as a general guide for AOAC and related individual activities and meetings.

**Responsibility for Antitrust Compliance**

The Association's structure is fashioned and its programs are carried out in conformance with antitrust standards. However, an equal responsibility for antitrust compliance — which includes avoidance of even an appearance of improper activity — belongs to the individual. Even the appearance of improper activity must be avoided because the courts have taken the position that actual proof of misconduct is not required under the law. All that is required is whether misconduct can be inferred from the individual's activities.

Employers and AOAC depend on individual good judgment to avoid all discussions and activities which may involve improper subject matter and improper procedures. AOAC staff members work conscientiously to avoid subject matter or discussion which may have unintended implications, and counsel for the Association can provide guidance with regard to these matters. It is important for the individual to realize, however, that the competitive significance of a particular conduct or communication probably is evident only to the individual who is directly involved in such matters.

**Antitrust Guidelines**

In general, the U.S. antitrust laws seek to preserve a free, competitive economy and trade in the United States and in commerce with foreign countries. Laws in other countries have similar objectives. Competitors (including individuals) may not restrain competition among themselves with reference to the price, quality, or distribution of their products, and they may not act in concert to restrict the competitive capabilities or opportunities of competitors, suppliers, or customers.

Although the Justice Department and Federal Trade Commission generally enforce the U.S. antitrust laws, private parties can bring their own lawsuits.
Penalties for violating the U.S. and other antitrust laws are severe: corporations are subject to heavy fines and injunctive decrees, and may have to pay substantial damage judgments to injured competitors, suppliers, or customers. Individuals are subject to criminal prosecution, and will be punished by fines and imprisonment. Under current U.S. federal sentencing guidelines, individuals found guilty of bid rigging, price fixing, or market allocation must be sent to jail for at least 4 to 10 months and must pay substantial minimum fines.

Since the individual has an important responsibility in ensuring antitrust compliance in AOAC activities, everyone should read and heed the following guidelines.

1. Don't make any effort to bring about or prevent the standardization of any method or product for the purpose or intent of preventing the manufacture or sale of any method or product not conforming to a specified standard.

2. Don't discuss with competitors your own or the competitors' prices, or anything that might affect prices such as costs, discounts, terms of sale, distribution, volume of production, profit margins, territories, or customers.

3. Don't make announcements or statements at AOAC functions, outside leased exhibit space, about your own prices or those of competitors.

4. Don't disclose to others at meetings or otherwise any competitively sensitive information.

5. Don't attempt to use the Association to restrict the economic activities of any firm or any individual.

6. Don't stay at a meeting where any such price or anticompetitive talk occurs.

7. Do conduct all AOAC business meetings in accordance with AOAC rules. These rules require that an AOAC staff member be present or available, the meeting be conducted by a knowledgeable chair, the agenda be followed, and minutes be kept.

8. Do confer with counsel before raising any topic or making any statement with competitive ramifications.

9. Do send copies of meeting minutes and all AOAC-related correspondence to the staff member involved in the activity.

10. Do alert the AOAC staff to any inaccuracies in proposed or existing methods and statements issued, or to be issued, by AOAC and to any conduct not in conformance with these guidelines.

**Conclusion**

Compliance with these guidelines involves not only avoidance of antitrust violations, but avoidance of any behavior which might be so construed. Bear in mind, however, that the above antitrust laws are stated in general terms, and that this statement is not a summary of applicable laws. It is intended only to highlight and emphasize the principal antitrust standards which are relevant to AOAC programs. You
must, therefore, seek the guidance of either AOAC counsel or your own counsel if antitrust questions arise.

* * * * *

Adopted by the AOAC Board of Directors: September 24, 1989
Revised: March 11, 1991
Revised October 1996
Meeting of the International Stakeholder Panel on Alternative Methods (ISPAM) and Working Groups

DRAFT AGENDA
Friday, September 16, 2011
8:30 am – 5:30 pm
Sheraton New Orleans Hotel, New Orleans, Louisiana

SESSION I:
ISPAM JOINT SESSION (Bayside C)
8:30 am – 9:45 am

A. Welcome and Introductions, and overview of ISPAM goals and activities to date – Russ Flowers, Mérieux NutriSciences and Chair, ISPAM
Page: 1

B. Summary of activities and overview of Meeting Agenda of the Working Group on Microbiological Validation Guidelines for Alternative Methods – Russ Flowers, Mérieux NutriSciences and WG Chair
Page: 5

C. Summary of work to date and Meeting Agenda of the Working Group on Qualitative Chemistry Guidelines for Alternative Methods – Bert Pöpping, Eurofins and WG Chair

D. Overview of different statistical models, the strengths and weaknesses of each model and the appropriate application of each – Paul Wehling, General Mills and Chair WG on Statistics
Page: 7

E. Standard Method Performance Requirements (SMPRs) and proposal to develop SMPRs to complement general validation guidelines – Scott Coates, AOAC Chief Scientific Officer
Page: 10
INTERNATIONAL Stakeholder Panel on Alternative Methods

ISPAM

- Purpose and Scope
  - Develop harmonized, internationally accepted standard validation guidelines for alternative chemical and microbiological methods

- Objectives:
  - Improve efficiency and minimize economic burden by leveraging global networks of experts to reach consensus on analytical validation protocols
Working Groups

- *Smaller Working Groups of subject matter experts have been formed to identify potential conflicts and propose solutions.*
  - AOAC has learned that this upfront work saves a tremendous amount of time and focuses the larger stakeholder panel on consensus decision making.
  - Three working groups were established: Microbiology Work Group, Chemistry Work Group, and Statistics Work Group.

ISPAM History

- November 29, 2010
  - AOAC Research Board endorsed supporting a stakeholder panel to work on harmonizing validation guidelines for alternative methods.
- December, 2010
  - AOAC Board of Directors approves establishing an Advisory Panel to define the project.
- March 15, 2011
  - Advisory Panel meets to define the project and recommends experts for Stakeholder Panel and Working Groups.
- April 12, 2011
  - Review of potential Panel and Working Group members.
ISPAM History

- April 21, 2011
  - Teleconference to accept nominations for Stakeholder Panel and Working Groups
- April 22, 2011
  - AOAC staff meeting to organize to do the impossible!
- April 27 & 28, 2011
  - First of numerous conference calls of the working groups
- June 30
  - Mid-Year Stakeholder meeting

ISPAM History

- August 16 – 17
  - Sub-Group on Microbiological Guidelines meets to lay the groundwork for recommendations to harmonize top five priority areas
- September 16
  - Meeting of the International Stakeholder Panel on Alternative Meeting (ISPAM)
Objectives for Today

› Microbiology and Chemistry Working Groups will develop recommendations to present to the Stakeholder Panel
› Stakeholder Panel will discuss/consider the recommendation
› Voting members of Stakeholder Panel will vote on recommendations
   › Voting members - balance from the different sectors and countries, so that no one group dominates
Summary of Activities and Meeting Agenda of the Working Group on Microbiological Validation Guidelines for Alternative Methods

Microbiology Work Group

- The Microbiology WG developed a spread sheet including the pertinent validations schemes that identified the main requirements and criteria for each scheme; e.g., AOAC, ISO, NordVal, FDA, USDA, Health Canada, CFIA.
- Microbiology WG – Focus has been on validation schemes for qualitative microbiology analysis of foods and food environments, as there were more discrepancies compared to quantitative methods.
- Five issues were identified as the most critical to harmonize, and a sub-group of the Micro WG, headed by Phil Feldsine, was formed to develop ideas for harmonization.
Microbiology Work Group

- The top 5 priority areas decided in the teleconference meetings and confirmed at the June mid-year meeting were:
  - Reference Methods
  - Selection of Food/Category (sample matrix)
  - Number of Levels/Samples/Fractional Positive
  - Results analysis & Criteria/Statistical Analysis
  - Number of Samples/Replicates/Method Comparison & Collaborative

Work continued following the June meeting by teleconference.

- August 16 and 17
  - Meeting of a sub-group consisting of regulators and organizational affiliates at AOAC offices
  - Proposals for harmonization of the 5 most critical issues were developed and agreed upon

- September 16 ISPAM meeting
  - Proposals for harmonization of this issues will be presented and discussed
  - Work Group will make recommendations to the Stakeholder Panel for acceptance.
Working Group on Statistics

- Develop scientific consensus on the best statistical techniques to use for validating qualitative methods.
- Provide guidance to Micro and Chem Working groups on aspects of statistical methodologies.
- Advise on strengths, weaknesses and applicability of various models.
- Look for potential areas of agreement and encourage flow of ideas across Chem/Micro working groups.

Presentations

- Dan Tholen – Chi Square
- Paul Wehling – POD
- Robert LaBudde – LOD, etc.
Summary – X²

- Advantages
  - Fairly simple
  - Historical
- Disadvantages
  - Only works for paired studies
  - Lab Effect not accounted for
  - Only a hypothesis test for difference
  - Power of test not inherently obvious

Summary - POD

- Advantages
  - Easy to calculate stats with existing software
  - Graphical representation
  - Confidence intervals
- Disadvantages
  - No comparison across levels – some loss of power
Regression Strategies

- Advantages
  - Power of experiments can be increased
  - LOD-like parameters can be estimated
- Disadvantages
  - Assumptions needed for how to model levels
  - Requires much better concentration estimates
Validation Guidelines and SMPRs

Scott Coates
Chief Scientific Officer
AOAC INTERNATIONAL

Outline

• Terminology
• SMPR
• SMPR Development Process
• SMPR Examples
• How SMPR could be used

Terminology

• validation guidelines
• standard validation study design template
• validation protocols
Terminology

• Fitness-for-purpose
• Call for Methods
• First Action Official Methods of Analysis
• Final Action Official Methods of Analysis

Standard Method Performance Requirement (SMPR)

• a very detailed description of the analytical requirements.
• includes acceptance requirements.
• used to qualify methods for AOAC approval.
• published as a standard.
• may include study designs for measuring analytical performance.
For the chemists:

<table>
<thead>
<tr>
<th>Range (ppm)</th>
<th>LOQ (LOD)</th>
<th>LOD</th>
<th>Limit of quantitation (LOQ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nitrogen</td>
<td>0.12-0.11</td>
<td>0.1</td>
<td>0.11</td>
</tr>
<tr>
<td>Manganese</td>
<td>0.12-0.11</td>
<td>0.1</td>
<td>0.11</td>
</tr>
</tbody>
</table>

For the microbiologists:

Note to self: standardized panels make fairer comparison.
SMPR Guideline and Education Session

Unique document:
- Chemistry & microbiology
- Quantitative & qualitative
- Definitions
- Evaluation recommendations
- Expected results
- Informative sections

2010 AOAC SMPRs

<table>
<thead>
<tr>
<th>Number</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010.001</td>
<td>PCR Methods for Detection of Francisella tularensis</td>
</tr>
<tr>
<td>2010.002</td>
<td>PCR Methods for Detection of Yersinia pestis</td>
</tr>
<tr>
<td>2010.003</td>
<td>PCR Methods for Detection of Bacillus anthracis</td>
</tr>
<tr>
<td>2010.004</td>
<td>HHAs for Detection of Bacillus anthracis</td>
</tr>
<tr>
<td>2010.005</td>
<td>HHAs for Detection of Ricin</td>
</tr>
<tr>
<td>2010.006</td>
<td>Portable PCR Methods for Detection of Bacillus anthracis in visible powders</td>
</tr>
</tbody>
</table>

2011 AOAC SMPRs

<table>
<thead>
<tr>
<th>Number</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011.001</td>
<td>PCR Methods for Detection of Burkholderia mallei [pending]</td>
</tr>
<tr>
<td>2011.002</td>
<td>PCR Methods for Detection of Burkholderia pseudomallei [pending]</td>
</tr>
<tr>
<td>2011.003</td>
<td>Vitamin A in Infant Formula and Adult/Pediatric Nutritional Formula</td>
</tr>
<tr>
<td>2011.004</td>
<td>Vitamin D in Infant Formula and Adult/Pediatric Nutritional Formula</td>
</tr>
<tr>
<td>2011.005</td>
<td>Vitamin B12 in Infant Formula and Adult/Pediatric Nutritional Formula</td>
</tr>
<tr>
<td>2011.006</td>
<td>Folate in Infant Formula and Adult/Pediatric Nutritional Formula</td>
</tr>
<tr>
<td>2011.007</td>
<td>Myo-Inositol in Infant Formula and Adult/Pediatric Nutritional Formula</td>
</tr>
<tr>
<td>2011.008</td>
<td>Vitamin E in Infant Formula and Adult/Pediatric Nutritional Formula</td>
</tr>
</tbody>
</table>
SMPR Development Process

- Advisory Panel
- Stakeholders Panel
- Working Groups
- Expert Review Panel

Advisory Panel

- Thought leaders who propose ideas.
- Topic areas are usually broad.
- Recommend stakeholders.

Stakeholder Panels

- Identify and prioritize specific topics.
- Identify experts.
- Create working groups.
- Review and approve (or not) SMPRs.
Working Groups

- Draft Fitness-for-Purpose statements.
- Call for methods.
- Draft SMPRs.
- Review collected methods.
- Recommend methods for OMA.

Expert Review Panels

- Review methods recommended for OMA by WG.
- Determine criteria that method needs to meet for the ERP to recommend Final Action status.
- Monitor method between First Action status and Final Action status.
- Recommend First Action methods for Final Action to Official Methods Board.

SMPR:

- Can be used as the basis for method approval.
- Provide guidance to method developers for the development of new methods.
- Can be used to thoughtfully advance the state-of-the-art in a particular direction.
SMPR:

- Can be used to address specific analytical needs.
- Allow AOAC to reach a broader community of method developers and users.
- Can be used to replace reference methods with objective criteria.

SMPRs and Guidelines

- Guidelines are very general, and a little abstract, by their nature.
- It helps to apply the guidelines to a real-life issue as they are being developed.
- SMPRs connect the guidelines to reality.

SMPRs and Guidelines

- SMPRs are the wave of the future
- Development of a couple of SMPRs as a part of the guideline development process ensures a better product in the end.
Conclusion

SMPRs have many flexible uses.
Development time = 6 months.
Transparent.
The kinks have been hammered out.

Scott Coates
Chief Scientific Officer
AOAC INTERNATIONAL
301-924-7090
scoates@aoac.org
Meeting of the International Stakeholder Panel on Alternative Methods (ISPAM) and Working Groups

DRAFT AGENDA
Friday, September 16, 2011
8:30 am – 5:30 pm
Sheraton New Orleans Hotel, New Orleans, Louisiana

SESSION II:
ISPAM WORKING GROUP ON MICROBIOLOGICAL GUIDELINES
(Bayside C) 10:00 am – 3:00 pm

A. Discussion of proposal from the Sub-Group (SG) to harmonize the use of reference methods – Phil Feldsine, BioControl Systems
Page: 18

B. Discussion of proposal from the Sub-Group (SG) to harmonize the number of levels/samples/fractional positives – Phil Feldsine, BioControl Systems
Page: 23

C. Discussion of proposal from the Sub-Group (SG) to harmonize the selection of food/category (sample matrix) – Ron Johnson, BioMérieux
Page: 24

LUNCH 12:00 pm – 1:00 pm (Grand Chenier, 5th Floor)

D. Discussion of proposal from the Sub-Group (SG) to harmonize results analysis & criteria/statistical analysis – Morgan Wallace, DuPont Qualicon
Page: 26

E. Discussion of proposal from the Sub-Group (SG) to harmonize number of samples/replicates/method comparison & collaborative – Wendy Lauer, Bio-Rad
Page: 28

F. Summary of recommendations that will be proposed to ISPAM for vote and discussion of next steps members – Russ Flowers, Mérieux NutriSciences and Chair, ISPAM
1. **Use of Reference methods**

Many microbiological validation studies have been conducted according to multiple validation schemes using a number of different international reference methods. This provides an existing, extensive database upon which a statistical analysis can be conducted to demonstrate the equivalence of reference methods.

The project group proposes that it may be possible to conduct such an analysis to show equivalence of different reference methods by analyzing the existing data from alternative method validation studies that have been previously conducted.

It is proposed that to reach consensus on the equivalence of reference methods, a retrospective data analysis be conducted in two stages.

Proposal:

**Stage 1:**

A. Data will be requested from multi-national food test kit manufacturers. Data will be kept entirely confidential, and must fulfill all of the following requirements:

   i. Raw data.

   ii. Data from a comparative validation study where the test kit was compared to at least one international reference method.

   iii. Data from food samples that were artificially contaminated

   iv. Data from sample sets that produced fractional positive results

   v. Data where all positive and negative results have been confirmed

   vi. Data for which all methodology is clearly described, including revision date

   vii. Data produced from 2001 – 2011

   viii. Data generated in testing for *Salmonella*, *Listeria* or *E. coli O157*

   ix. Data from any food type, where the food type is clearly indicated

   x. Data where the testing laboratory has been clearly identified

B. The overall charge for the statistical consultant at Stage 1 will be to categorize the available data and to deliver an opinion on the scope and the impact of analysis that is possible, if the project group agrees
to proceed to Stage 2. The data must be “blind” coded in this process, so that the method manufacturers cannot be identified.

In particular, to fulfill the contract, the consultant will provide satisfactory answers to each of the following questions:

i. What is the structure of the dataset? The answer should make particular reference to the number of testing laboratories, the number of reference methods used, and the degree of overlap between different test methods in different laboratories.

ii. Can the analysis described in Stage 2 be performed? For which organisms? For which food categories? For which international reference methods?

iii. Can inter-laboratory variability be modeled using the existing dataset? How would you recommend this be performed?

iv. Can the inter-laboratory variability be considered to be random?

v. What additional assumptions would be required to perform the analysis as described in Stage 2?

vi. What groupings of food types or food categories could be analyzed? How would changing the food category groupings impact the statistical validity of the analysis?

vii. What is your recommended statistical procedure for analysis as described in Stage 2? What are the advantages and disadvantages of this protocol?

viii. What alternative statistical methods could be used to perform the same analysis? What are their advantages and disadvantages?

Stage 2:

A. Perform an analysis of existing validation studies of multiple alternative methods that have been compared to a variety of reference methods. For clarity, an individual alternate method study may have employed only one reference method or may have included more than one reference method. Both studies are acceptable for purposes of this analysis. These studies may have employed a variety of reference methodology, i.e., AOAC, BAM, USDA/FSIS, ISO and Health Canada.

Using assumptions as indicated by the project group after their analysis of the Stage 1 report, the overall charge for the statistical consultant will be to use the alternative methods as the anchor methods (or normalizer), and to use the dataset as a whole to provide a comparison of the equivalence of the international reference methods for *Salmonella*, *Listeria* and *E. coli* O157:H7 in food categories as defined by the project group after their review of the Stage 1 report.

Note that the detailed contract requirements for Stage 2 will be further developed by the project group after their review of the Stage 1 report.

The objective of this analysis is to provide supporting data to substantiate that various reference methods are acceptable for use in alternative method validation studies.
### Instructions for Completing Data Tab:

**Sample ID (column C):**
The Sample IDs are generated using the following format:

```
Method-Pathogen-Matrix-Random Name-number
```

- **Method:** For the method, please use the following letters:
  - EU or ISO methods – ISO
  - AOAC OMA – AOAC
  - US FDA – BAM
  - USDA / FSIS – FSIS
  - Health Canada – HC

- **Pathogen:** For the pathogens, please use the following letters:
  - Salmonella – Sal
  - Listeria monocytogenes – Lm
  - Listeria species – Ls
  - E. coli O157:H7 – EC

- **Matrix:** For the Matrix, please use the following letters:
  - Raw Ground Poultry – GP
  - Raw Ground Beef – GB
  - Non-Fat Dry Milk – DM
  - Ice Cream – IC
  - Chocolate – CH

- **Random Name:** For this field, please use the initial of someone in the lab, in the office or any other two letters.

**Number:** Numerical value starting from 1 and ending with the number of samples that are in the set.

**Food Item (column D):**
This Column will be one of the following food matrix or food item:
- Raw Ground Poultry
- Raw Ground Beef
- Non-Fat Dry Milk
- Ice Cream
- Chocolate

**Spike level (column E):**
Specify the spike level of the target pathogen in Colony Forming Units (CFU) per 25 grams of sample size (CFU / 25g).

**Reference Method (column F):**
This is the actual reference method that is used. In the name only an acronym is used; however, in this field please specify the exact method that has been used in the study.

**Alternate Results (column H, I):**
- **Screen (column H):** The results of the rapid (PCR, ELISA, etc.) test. Place a + or a – in this field.
- **Confirm (column I):** The result of the cultural confirmation of the rapid test. Place a + or a – in this field.

**Status (column J):**
The status of the test is filled in automatically with a formula. The status can be True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

**Reference Results (column L):**
The results of the reference method. Place a + or a – in this field.

**Paired / Unpaired (column M):**
Results are considered to be paired when the reference and alternate method started from the same primary enrichment. Results are considered to be unpaired when the reference method and the alternate method start from different primary enrichments.

Please note that in case of unpaired samples, the reference method results and the confirmation results of the alternate method will be identical unless sheer persistence is used to prove the alternate method as more sensitive than the reference method.
## CFIA Performance Parameters – Example
### Data Sheet

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Food Item</th>
<th>Spike Level</th>
<th>Reference Method</th>
<th>Alternate Results Screen</th>
<th>Confirm</th>
<th>Status</th>
<th>Reference Results</th>
<th>Paired / Unpaired</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISO-Sal-GB-SM-05</td>
<td>Ground Beef</td>
<td>2 CFU / 25g</td>
<td>ISO 6579:2006</td>
<td>+</td>
<td>+</td>
<td>TP</td>
<td>-</td>
<td>Unpaired</td>
</tr>
<tr>
<td>ISO-Sal-GT-SM-01</td>
<td>Ground Turkey</td>
<td>5 CFU / 25g</td>
<td>ISO 6579:2010</td>
<td>+</td>
<td>+</td>
<td>TP</td>
<td>-</td>
<td>Unpaired</td>
</tr>
<tr>
<td>ISO-Sal-GT-SM-02</td>
<td>Ground Turkey</td>
<td>5 CFU / 25g</td>
<td>ISO 6579:2011</td>
<td>+</td>
<td>+</td>
<td>TP</td>
<td>+</td>
<td>Unpaired</td>
</tr>
<tr>
<td>ISO-Sal-GT-SM-03</td>
<td>Ground Turkey</td>
<td>5 CFU / 25g</td>
<td>ISO 6579:2012</td>
<td>+</td>
<td>-</td>
<td>TP</td>
<td>+</td>
<td>Unpaired</td>
</tr>
<tr>
<td>ISO-Sal-GT-SM-05</td>
<td>Ground Turkey</td>
<td>5 CFU / 25g</td>
<td>ISO 6579:2014</td>
<td>-</td>
<td>+</td>
<td>TP</td>
<td>+</td>
<td>Unpaired</td>
</tr>
<tr>
<td>ISO-Sal-GT-SM-06</td>
<td>Ground Turkey</td>
<td>5 CFU / 25g</td>
<td>ISO 6579:2015</td>
<td>-</td>
<td>-</td>
<td>TN</td>
<td>-</td>
<td>Unpaired</td>
</tr>
<tr>
<td>ISO-Sal-GT-SM-07</td>
<td>Ground Turkey</td>
<td>5 CFU / 25g</td>
<td>ISO 6579:2016</td>
<td>+</td>
<td>+</td>
<td>TP</td>
<td>+</td>
<td>Unpaired</td>
</tr>
<tr>
<td>ISO-Sal-IC-SM-01</td>
<td>Ice Cream</td>
<td>8 CFU / 25g</td>
<td>ISO 6579:2017</td>
<td>+</td>
<td>+</td>
<td>TP</td>
<td>-</td>
<td>Unpaired</td>
</tr>
<tr>
<td>ISO-Sal-IC-SM-02</td>
<td>Ice Cream</td>
<td>8 CFU / 25g</td>
<td>ISO 6579:2018</td>
<td>+</td>
<td>+</td>
<td>TP</td>
<td>-</td>
<td>Unpaired</td>
</tr>
<tr>
<td>ISO-Sal-IC-SM-03</td>
<td>Ice Cream</td>
<td>8 CFU / 25g</td>
<td>ISO 6579:2019</td>
<td>+</td>
<td>+</td>
<td>TP</td>
<td>+</td>
<td>Unpaired</td>
</tr>
<tr>
<td>ISO-Sal-IC-SM-04</td>
<td>Ice Cream</td>
<td>8 CFU / 25g</td>
<td>ISO 6579:2020</td>
<td>+</td>
<td>+</td>
<td>TP</td>
<td>+</td>
<td>Unpaired</td>
</tr>
<tr>
<td>ISO-Sal-IC-SM-05</td>
<td>Ice Cream</td>
<td>8 CFU / 25g</td>
<td>ISO 6579:2021</td>
<td>+</td>
<td>+</td>
<td>TP</td>
<td>-</td>
<td>Unpaired</td>
</tr>
</tbody>
</table>
APPENDIX I Sample Analysis Guidelines

Supplemental Validation Parameters

In the event that supplemental test method procedure validation is required, it is important to apply a standardized approach to validation design. Specifically, there are key validation parameters that must be consistently applied in order to demonstrate a method is fit for intended use.

Matrix:

1. Sample matrix must be representative of product to be tested. Assumptions drawn from a more complex matrix could be applied to less complex matrices.

   **Logic:** Properties inherent to the product type (e.g., pH, ingredients and additives) may impact test method performance. Variations on lean point and ingredient components are inherent to all product types, and each lean point or ingredient variation does not require validation, however, a range from high fat to low fat samples could be considered as a part of matrix definition in validations. **Industry experience lend to a focus on high fat as a worst case scenario for most methods.**

   Significant differences or changes in matrix must be considered. Validation should occur when appropriate, based on questions that may be raised in the product matrix that cannot be answered by looking to existing science.

   Existing science, to understand the potential impact of an individual component on method performance, does not necessarily require validation in the meat matrix. Typically, as a part of the methods approvals multiple product types and matrixes are evaluated.

   Validation of a more complex matrix may be applied to a matrix which is considered less complex

   The intent of this section is not to require validation for every minor matrix difference that may occur, nor should every minor matrix difference be validated. However, it is intended that as a part of method selection matrix and the impact of matrix components be considered.

2. Sample matrix temperature must be defined within typical product temperature range at the time of analysis, or worst case scenario (frozen). Validation should be completed at lower range temperature for fresh products.

   **Logic:** Initial sample temperature should be a consideration in the validation study design. The sample temperature can impact incubation time requirements and time required to get sample to optimum temperature (fresh 35°F to 40°F; frozen less than 28°F).
Sub-Group on Microbiological Guidelines for Alternative Methods
Harmonization Proposals for Five Top Priority Areas

Number of Levels/Samples/Fractional Positive

The sub-group discussed fractional level recovery and recommends the following:

Fractional Level Recovery

Target is 50% with a range of ± 25%. Provided, however, that in the context of the entire study data set, values outside of the fractional range may be considered. For example, for studies where a larger number of test portions were analyzed, (i.e., 60), a larger fractional range may be acceptable. Other parameters may be considered on an individual basis.
### Table A. Food categories and suggested relevance to major foodborne pathogens

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Raw Meat Products</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw unprocessed</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw processed</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frozen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carcass punch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>X</td>
</tr>
<tr>
<td><strong>Raw Poultry Products</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw unprocessed</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw processed</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frozen</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rinsates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ready-to-eat Meat and Poultry Products</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heat processed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fermented</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fish and Seafood Products</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heat processed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frozen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoked</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fruit and Vegetable Based Products</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heat processed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frozen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Juices/concentrates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low moisture</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nuts/Nut meats</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Eggs and Egg Products</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shell eggs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heat processed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dairy Products</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heat processed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frozen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fermented</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------------</td>
<td>------------------------</td>
<td>------------------</td>
<td>-----------------------</td>
<td>-------------------</td>
<td>----------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Personnel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipment</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Air</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Process Water / Brine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary production samples</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Animal feces</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Animal non-feces</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environmental</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carcass drag swabs/dry swabs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table B. Environmental categories and suggested relevance to major foodborne pathogens

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Personnel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipment</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Air</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Process Water / Brine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary production samples</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Animal feces</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Animal non-feces</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environmental</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carcass drag swabs/dry swabs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Statistical Treatment of Data

Assumptions

- Statistical interpretation of data will not be fully harmonized across Governments and Validating Organizations
- Multiple statistical approaches can be justified to evaluate the data sets obtained using the validation approach described in the previous presentations
Conclusion

- The ISPAM group will not proscribe the statistical tests to be used as long as the statistical treatments used are appropriate for the evaluation of the experimental design described above

Proposal for Adoption

- “The statistical analysis on data generated from harmonized study designs will be performed by the RLOD and/or Chi-square models. As appropriate, other statistical methodologies that have been adopted by certification bodies or regulatory agencies may be applied”
## Top 5 Priorities: Comparison Table

<table>
<thead>
<tr>
<th>QUALITATIVE Method</th>
<th>ISO 16140</th>
<th>AOAC OMA</th>
<th>Health Canada</th>
<th>NordVal</th>
<th>FDA Guidelines</th>
<th>Draft USDA/FSIS Guidelines</th>
</tr>
</thead>
</table>

*References:*

| 1st priority is ISO method, 2nd priority is CEN method, if neither exists, then 3rd priority is other recognized methods |
| ISO, CEN, NMKL, BAM, etc. It is up to the applicant; however, as the EU regulation in EC 2073/2005 Microbiological criteria states EN ISO methods, these are most frequently used. |
| Must be BAM, unless there is no BAM reference method. |
| For FSIS regulated products, the current FSIS method, which is found in the Microbiology Laboratory Guidebook (MLG), is the most appropriate reference cultural method for validating methods used by FSIS-regulated establishments. FDA BAM, or methods referenced by ISO or Codex Alimentarius may be appropriate. Non-cultural methods applicable in some circumstances. |

*Pre-collaborative phase(s):*

- Can be various pre-existing recognized analytical methods e.g. AOAC OMA, ISO, FDA BAM, FSIS MLG and Health Canada |
- Acceptable Ref published by HC (Part 1) |
- May include any methods from methods organizations, such as AOAC, BAM, APHA, ICMSF, IDF, ISO etc. |
- Where no Ref exists, MMC assess on case by case basis |
- ISO, CEN, NMKL, BAM, etc. It is up to the applicant; however, as the EU regulation in EC 2073/2005 Microbiological criteria states EN ISO methods, these are most frequently used. |
- Must be BAM, unless there is no BAM reference method. |
- If these is no BAM reference method, but if there is a nationally/internationally recognized reference method, then FSIS MLG, AOAC, ISO, and Health Canada are all potential reference methods. APHA, ICMSF, and IDF methods also may be used as reference methods. |
| summary tables for POD |
| -If no appropriate Ref can indicate “NA” in summary tables for POD |
| -If no appropriate Ref can indicate “NA” in summary tables for POD |
| Note: definition still under discussion at ISO level to open up for non ISO/CEN methods (PIV) |

**SESSION II ISPAM**
<table>
<thead>
<tr>
<th>ISO 16140</th>
<th>AOAC OMA</th>
<th>Health Canada</th>
<th>NordVal</th>
<th>FDA</th>
<th>Draft USDA/FSIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection of foods</td>
<td>RA</td>
<td>-5 categories for all foods applications, 3 food types per category (see below) -Feed, environmental samples and primary production samples (PIV) are additional categories</td>
<td>SLV</td>
<td>-5 categories for all foods applications, 3 food types per category (Table 1). Environmental samples is additional category</td>
<td>RA, SE, SP, Kappa</td>
</tr>
<tr>
<td></td>
<td>RLOD</td>
<td>Same, except 1 food type per category (if possible) a different food type</td>
<td></td>
<td>Internal</td>
<td>LOD</td>
</tr>
<tr>
<td></td>
<td>Internal</td>
<td>At least 1 matrix that was tested in the SLV. For every 5 foods claimed, 1 food matrix must be included</td>
<td></td>
<td>Internal</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>Internal</td>
<td>The selection of foods is determined by FDA’s regulatory needs.</td>
<td></td>
<td>Internal</td>
<td>Internal</td>
</tr>
<tr>
<td></td>
<td>Internal</td>
<td>All claimed matrices must be included in the study. Contains proposal to create matrix categories based on intrinsic properties. “All Foods” claim not applicable</td>
<td></td>
<td>Internal</td>
<td>Internal</td>
</tr>
<tr>
<td>ISO 16140</td>
<td>AOAC OMA</td>
<td>Health Canada</td>
<td>NordVal</td>
<td>FDA</td>
<td>Draft USDA/FSIS</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>---------------</td>
<td>---------</td>
<td>-----</td>
<td>-----------------</td>
</tr>
<tr>
<td>- Food category/ type/ item</td>
<td>Each Food type can be made of various relevant food items. Annex B provides guidance (not mandatory). These are then grouped together to meet the sample number requirement of a food type, i.e. 20 samples. <em>This is allow for the use of naturally contaminated samples (BL)</em></td>
<td>Only one single food item is accepted to meet the sample size requirement of a food type, i.e. 20 replicates.</td>
<td>Each food type can be made of various relevant food items. Table 1. <strong>CLARIFICATION NEEDED</strong> Can these be grouped together to meet the sample number requirement of a food type, in this case 20 samples? Yes, they can be grouped together to meet the sample number requirement. This notion has been introduce to allow for heterogeneity within a food type. Products in a type may vary greatly in origin, composition, preparation processes, natural background; all those small variabilities could have an influence on the detectability of the target organism. (I.I.)</td>
<td>Each Food type can be made of various relevant food items. At NordVal’s homepage (<a href="http://www.nmkl.org">www.nmkl.org</a>) provides a list of food categories. These are then grouped together to meet the sample number requirement of a food type, i.e. 20 samples.</td>
<td>Currently, foods are validated individually and there are no category claims. There are no “All Foods” claims.</td>
</tr>
<tr>
<td>ISO 16140</td>
<td>AOAC OMA</td>
<td>Health Canada</td>
<td>NordVal</td>
<td>FDA</td>
<td>Draft USDA/FSIS</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>---------------</td>
<td>---------</td>
<td>-----</td>
<td>-----------------</td>
</tr>
<tr>
<td><strong>-No. of levels/samples</strong>&lt;br&gt;RA</td>
<td>20 samples per food type or 60 samples per category&lt;br&gt;RLOD</td>
<td>3 levels&lt;br&gt;-negative controls =5 samples&lt;br&gt;-1 level (theoretical LOD, with fractional positive results (BL)) = 20 samples&lt;br&gt;-Another level = at least 5 samples</td>
<td>3 levels:&lt;br&gt;-negative controls =5 samples&lt;br&gt;-1 level with fractional positive results = 20 samples&lt;br&gt;-Another level up to 1 log higher = 20 samples</td>
<td>RA</td>
<td>20 samples per food type or 60 samples per category&lt;br&gt;LLOD</td>
</tr>
<tr>
<td><strong>-Sample size</strong>&lt;br&gt;MicroVal: Is specified in the reference method, other (larger) samples size is allowed but specified in the certificate. (PIV)</td>
<td>Undefined</td>
<td>Standard is 25 g or 25 mL, unless Ref method specified larger sample size</td>
<td>Undefined</td>
<td>25 g unless otherwise specified.</td>
<td>Application dependent. Portions should not be made larger without validation. Validation study conclusions from larger portions applicable to smaller portions.</td>
</tr>
<tr>
<td><strong>-Fractional positive</strong>&lt;br&gt;Can be achieved by either alternate or Ref. - All samples should not be all positive or all negative.&lt;br&gt;-Ideal is 10 positive and 10 negative (50%) but any fractional results is acceptable</td>
<td>Can be achieved by either alternate or Ref. -proportion of positives 25% to 75%, ideal is approx 50% (10% to 90% is under review (RF))</td>
<td>Can be achieved by either alternate or Ref. -proportion of positives 25% to 75%,</td>
<td>Can be achieved by either alternate or Ref. - All samples should not be all positive or all negative.&lt;br&gt;-Ideal is 10 positive and 10 negative (50%) but any fractional results is acceptable</td>
<td>Yes, one or both methods must give 40 – 90% positive results. It is proposed that the percentage positive results be changed to 25 – 75%.</td>
<td>defined as a range of 20-80% confirmed positive results using reference method</td>
</tr>
<tr>
<td>ISO 16140</td>
<td>AOAC OMA</td>
<td>Health Canada</td>
<td>NordVal</td>
<td>FDA</td>
<td>Draft USDA/FSIS</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>---------------</td>
<td>---------</td>
<td>-----</td>
<td>-----------------</td>
</tr>
<tr>
<td><strong>Results analysis &amp; criteria</strong></td>
<td><strong>RA</strong> - By type and by category - Relative accuracy AC, relative specificity SP, relative sensitivity SE - First by unconfirmed results, again by confirmed results - McNemar test as criteria, (for paired and unpaired) with caveats i.e. really not suitable for unpaired and “never be interpreted by only the McNemar test” <strong>RLOD</strong> - by category - LOD of alternate method divided by LOD of Ref For paired, no lower limit, but LOD alternate might not be &gt; 2 times the LOD Ref For unpaired samples, no lower limit, the LOD alternate might not be &gt; 3 times the LOD Ref <strong>(In the ISO/CD 16140-2 version, I don’t find any acceptability limit settled for unpaired samples, only specified for paired samples BL)</strong> The values of 2 (paired) and 3 (unpaired) are still tentative values!!!!!! (PIV)</td>
<td>Method Equivalence: - <strong>POD</strong> - one-tailed POD 95% confidence interval (I.I.) Performance parameters: - by level and by food, but only calculated for those that passed POD successfully For Unpaired: - Performance parameters is the comparison of presumptive vs. confirmed results of the alternate method (not the Ref method results) - Specificity is based on presumptive results - Sensitivity is based on final (confirmed) results - Equivalence of alternate method and Ref can only be determined by the number of true positives in both sets, done by POD method For Paired: Use “absolute” results where Ref can have FN Criteria: Sensitivity ≥98% Specificity ≥90.4% False negative rate &lt; 2% False positive rate ≤ 9.6% Efficacy ≥94% LOD must be comparable or exceed the lower LOD of the Ref</td>
<td><strong>RLOD</strong> - By type and by category - Relative accuracy AC, - Relative specificity SP, - Relative sensitivity SE - Kappa - First by unconfirmed results, again by confirmed results Criteria: SE ≥ 95% Kappa ≤ 0.80 LOD: fit for purpose</td>
<td>By level/individual experiment for each matrix. Per AOAC Microbiology guidelines, McNemar Chi Square statistics are used.</td>
<td>Performed for each matrix. Unpaired study: One sided chi-square test with alpha = 0.05. Criterion: indistinguishable or better performance than reference method. Paired study: Evaluate sensitivity with minimum 29 confirmed positive results. Zero false negative results from 29 confirmed positives would be consistent with a test having a sensitivity that met or exceeded 90% and zero negative results from 50 confirmed positives would be consistent with a test with a sensitivity that met or exceeded 94%. Criterion: none proposed</td>
</tr>
<tr>
<td>ISO 16140</td>
<td>AOAC OMA</td>
<td>Health Canada</td>
<td>NordVal</td>
<td>FDA</td>
<td>Draft USDA/FSIS</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>---------------</td>
<td>---------</td>
<td>-----</td>
<td>----------------</td>
</tr>
<tr>
<td><strong>Inter-laboratory Study</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Applicable to alternative methods with a major modification, defined as any significant change in the design or the component reagents for a screening test, for example, the introduction of a new antibody or oligonucleotide primer. Follow guidance provided by the AOAC International Official Methods of Analysis Program</td>
</tr>
<tr>
<td>- minimum no of valid data sets/ collaborators</td>
<td>-10; defined as individuals working independently using different sets of samples; from a min. of 5 different organizations, including organizing lab and different locations from same company</td>
<td>-10 valid lab data sets required - Specifies that 12 labs should start</td>
<td>Minimum of 8 labs reporting valid data, labs should be accredited per 17025 or demonstrate is functioning under equivalent quality system</td>
<td>-10; defined as individuals working independently using different sets of samples; from a min. of 5 different organizations, including organizing lab and different locations from same company</td>
<td>2 for a Level 2 study, 3 for a Level 3 study and 10 for a level 4 study.</td>
</tr>
<tr>
<td>-Sample size</td>
<td>NA</td>
<td>Standard is 25 g or 25 mL, unless Ref method specified larger sample size</td>
<td>CLARIFICATION NEEDED Consistent with Pre-collaborative? Sample size is 25g unless otherwise specified by the method or need for larger size (to achieve enhance detectability, regulatory purpose or compositing) (I.I.)</td>
<td>NA</td>
<td>25 g unless otherwise specified.</td>
</tr>
</tbody>
</table>

*SESSION II ISPAM*
<table>
<thead>
<tr>
<th>ISO 16140</th>
<th>AOAC OMA</th>
<th>Health Canada</th>
<th>NordVal</th>
<th>FDA</th>
<th>Draft USDA/FSIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>- number of foods</td>
<td>1; relevant food item, inoculated with target, using a challenging enrichment protocol</td>
<td>1</td>
<td>At least 1</td>
<td>1; relevant food item, inoculated with target, using a challenging enrichment protocol</td>
<td>One or more.</td>
</tr>
<tr>
<td>- number of levels</td>
<td>3; negative control, one level which produce fractional positive and another level</td>
<td>3; negative control, one level which produce fractional positive and another level</td>
<td>3; negative control, one level which produce fractional positive and another level about 10 times greater than the detection level</td>
<td>3; negative control, one level which produce fractional positive and another level</td>
<td>2 for a Level 2 study (1 inoculated and 1 uninoculated. 3 for Levels 3 &amp; 4 (high, low, and uninoculated.</td>
</tr>
<tr>
<td>- number of replicates</td>
<td>8; per level of contamination -minimum of 48 results per collaborator = 8 replicates x 3 levels x 2 methods -minimum of 480 results (48 from each collaborator) = (240 per method) for statistical analysis</td>
<td>12 per level of contamination - 72 results per collaborator = 12 replicates x 3 levels x 2 methods = 72 - minimum of 720 results (360 per method) for statistical analysis</td>
<td>8 per level - min of 24 results per collaborator (8 x3 levels) per method</td>
<td>8 laboratories; - 3 levels in duplicates</td>
<td>6</td>
</tr>
<tr>
<td>- Confirmation</td>
<td>for Paired, only confirm the + Alt/- Ref, for Unpaired, confirm all enrichments</td>
<td>Matched or unmatched, confirm all samples</td>
<td>Confirm all samples</td>
<td>for Paired, only confirm the + Alt/- Ref, for Unpaired, confirm all enrichments</td>
<td>Yes.</td>
</tr>
<tr>
<td>ISO 16140</td>
<td>AOAC OMA</td>
<td>Health Canada</td>
<td>NordVal</td>
<td>FDA</td>
<td>Draft USDA/FSIS</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>---------------</td>
<td>---------</td>
<td>-----</td>
<td>-----------------</td>
</tr>
<tr>
<td>- Comparisons</td>
<td>Analyzed two ways: 1. Unconfirmed Alternate method results vs. confirmed Ref 2. Confirmed Alternate method results vs. confirmed Ref</td>
<td>By level and by matrix analyzed and reported separately</td>
<td>CLARIFICATION NEEDED Consistent with Pre-collaborative? By level and by matrix, all result confirmed Confirmed alternate method results vs reference (I.I.)</td>
<td>Alternative to reference method (if available).</td>
<td></td>
</tr>
<tr>
<td>- Parameters Calculated</td>
<td>Specificity (only for Neg controls) Sensitivity (only for inoculated levels) Relative Accuracy (% of agreements) RLOD of the different participants (BL)</td>
<td>Cross Lab Probability of Detection (LPOD) Difference between Alternate LPOD and Ref LPOD</td>
<td>CLARIFICATION NEEDED Consistent with Pre-collaborative? Yes, POD, dPOD determined for each matrix-level. All dPOD data is then used to assess the comparative performance of both methods All 5 method parameter (specificity, selectivity, FP, FN and method efficacy) calculated in one of two ways, depending if sample is paired or unpaired. (I.I.)</td>
<td>Rel Specificity Rel Sensitivity Rel Accuracy Kappa</td>
<td>Per AOAC guidelines, Sensitivity, Specificity, False Negative, and False Positive Rates.</td>
</tr>
<tr>
<td>- Interpretation</td>
<td>McNemar test (chi square) RLOD is for information only: analysis of deviance test to assess the laboratory effect on RLOD then acceptability of RLOD global value (BL)</td>
<td>If confidence interval of dLPOD does not contain zero, then the diff is statistically significant</td>
<td>CLARIFICATION NEEDED Consistent with Pre-collaborative? Yes, dPOD one-tailed and method parameter requirement must be met. (I.I.)</td>
<td>Criteria: SE ≥ 95% Kappa ≤ 0.80 [LOD: fit for its purpose]</td>
<td>Per AOAC guidelines, McNemar Chi Square statistics.</td>
</tr>
</tbody>
</table>
Meeting Attendees:
Yi Chen, FDA, AOAC General Referee, Microbiology
Peter Evans, USDA FSIS
Phil Feldsine, BioControl Systems
Russ Flowers, Silliker, ISPAM Chair and AOAC President
Ron Johnson, BioMerieux
Wendy Lauer, Bio-Rad Laboratories
Harry Marks, USDA, FSIS (Tuesday only)
Kirsten Mattison, Health Canada
Mark Mozola, Neogen
Sam Mohajer, Canadian Food Inspection Agency
Palmer Orlandi, US FDA
Morgan Wallace, DuPont Qualicon

AOAC Staff:
Zerlinde Johnson
Krystyna McIver
Nora Marshall

I. Introductions
Russ Flowers, Chair, ISPAM provided a review of project scope and the five areas of primary concern to the harmonization effort for discussion during the meeting (see Attachments 1 & 2.

II. Harmonization of Reference Methods
Discussion ensued on the proposal for collecting, organizing and analyzing existing method comparison data to ascertain if the data can be analyzed to demonstrate method equivalency between reference methods (see Attachment 3 and 4).

The group decided to limit the data collection to *Salmonella* in five food types. The question of food type and what food type to start with and ask for data sets to be submitted. The following food types were selected with the caveat that there are sufficient data available for these categories from studies using the ISO reference method to do the analysis:
1. Raw ground poultry*
2. Raw ground beef
3. Non-fat dry milk
4. Ice cream
5. Chocolate
*homogenized, naturally contaminated data sets will be accepted.

There was discussion of the fact that some reference methods contain options written into the method. If there are options within the method, then the option used must be specified in the submission of the data set. Submitted data will be blind coded for before statistical analysis.

Action Items:
1. CFIA will provide a spreadsheet for data entry and a draft call for data memo with instructions to method developers on how to populate the data sheet with the requested data sets (reference method vs. alternative method). An example of the spreadsheet is under Attachment 5.

There was discussion of the concept and food types selected with Bertrand Lombard (ISO) in a teleconference on August 17. Dr. Lombard agreed with the concept and food types and:
   - Recommends gaining input from the ISO TC 34/SC 9/WG 2, on statistics on what is the most appropriate statistical model(s) that should be used in the analysis of the data submitted.
   - Ideally everyone will agree on the statistical model used to analyze the data collected.
   - The goal is to agree that the data supports the acceptance of the equivalency of the reference methods, and mutual acceptance.

III. Number of Levels/Samples/Fractional Positive

The group discussed fractional positives (see Attachment 3) and what is the appropriate number of fractional positives necessary for method validation. The following is a proposal for the definition of acceptable fraction positive results:

Target is 50% with a range of ± 25%, based upon 20 replicates. Provided, however, that in the context of the entire study data set, values outside of the fractional range may be considered. For example, for studies where a larger number of test portions were analyzed, (i.e., 60), a larger fractional range may be acceptable. Other parameters may be considered on an individual basis.

The sub-group agreed upon the following proposal.

For inoculated samples, for a food matrix (item) –
   - Minimum 20 replicates with fractional positive
   - At least 5 replicates for high level
   - 5 uninoculated replicates

The option of pooling data from multiple day trials to fill the 60 replicates per category criteria of some regulatory agencies was discussed.
It was agreed on a minimum of 60 replicates per category, subject to review by the General Referee

Possible language for inclusion. Group has not agreed on the inclusion of this language or any language to this affect:
   For certain commodities a higher number of replicates may be needed for specific regulatory agency acceptance.

IV. **Selection of Food/Category (sample matrix)**

The group decided on the following proposal:

1. There is a need for category list for AOAC-claimed matrices, since currently no category claim exists.
2. There was discussion of claims for “all foods” versus “variety of food”. The group recommended proposing to HC and ISO to remove the “all foods” claim.
3. Group would like AOAC to move to 3 foods per category, with 5 categories needing to be tested for “variety of foods” claim.

However, because there is no common use of the definitions for category of foods, versus type of food for example, the group agreed to propose to harmonize terminology and agree on the use of common definitions using the draft ISO 16140 Part 1: Terminology as a starting point (see Attachment 7).

The group discussed and proposed the following food categories, considering the AOAC, ISO and Health Canada (HC) food categories. The HC food category table under Attachment 6 was also used as a reference:

**Ready-To-Eat category (RTE):**
FSIS – proposed three RTE meat/poultry categories (cooked, acidified, and dried/cured). And a minimum of 60 replicates of a single food item to represent each category.

Discussion of “RTE meat” claim and the need for RTE category under the ISO food categories.

USDA FSIS requires 60 reps per category and has RTE meats broken down into four food types based upon intrinsic/extrinsic factors (acidic, cured, etc…). For FSIS validation guidance it would be acceptable for 30 reps for typical matrix and 30 reps for challenge matrix = 60 reps for total RTE class.

Since ISO 16140 does not break out RTEs as a separate category, it was proposed that SC 9 consider this possibility.

**Fish and Seafood Category**
HC same as AOAC and ISO. Therefore, ISO 16140 list can be used.

**Fruit and Vegetable-based product category**
It was recommended that nuts/nut butter/nut meat should be added to ISO 16140 as a category.

**Egg and Egg Derivatives**
Add egg category to ISO and AOAC.

**Dairy category**
No differences between AOAC, HC and ISO.

**Chocolate and Baked Goods**
No differences between AOAC, HC and ISO.

**Animal feeds**
Animal feeds should be a separate category for HC, like it is for AOAC and ISO.

**Miscellaneous**
No differences between AOAC, HC and ISO.

**Multi-ingredient foods**
HC and ISO already have this category. Group recommended that AOAC incorporate this category.

Group recommended that environmental categories be done separately.

The group asked Bertrand Lombard (ISO) whether the removal of “all foods” claim from the ISO standard is possible. Lombard agreed with the idea of removing the “all foods” claim, but indicated that the issue should be brought before Subcommittee 9 for discussion and decision on the issue.

**Action Items:**
1. HC/CFIA will provide a draft category and food items table as they relate to micro validations by September 1st.
2. HC/CFIA will provide a draft terms and definitions for category, food item and type. There will be a draft food categories and environmental categories to also be submitted by September 1st.
3. HC will also propose a draft proposal for harmonizing the use of definitions using ISO 16140 Part 1 as the starting point for the common validation terminology (see Attachment 7). This will be submitted by September 1 for discussion by the sub-group and further discussion by the larger Working Group at the September 16 meeting in New Orleans.

V. **Results analysis & Criteria/Statistical Analysis**

The group discussed the issue and decided on the following proposal to bring forth to ISPAM:

“The statistical analysis on data generated from harmonized study designs will be performed by the RLOD and/or Chi-square models. As appropriate, other
statistical methodologies that have been adopted by certification bodies or regulatory agencies may be applied.

Health Canada and CFIA provided documents that explain their criteria for statistical analysis. These documents were provided by email to those who were not in attendance.

**VI. Number of Data Sets for Collaborative Studies/Sample Size**

The group has decided on the following proposal:

Minimum number of collaborative data sets required is 10 valid data sets.

All samples will be confirmed and therefore meet ISO/AOAC/HC requirements.

The group agreed on the following statement:
Analytical test portion size is 25g unless otherwise specified by the methods (alternative or reference). The reference method and alternative method may have different analytical test portion sizes.

**VII. Summary and next steps for September 16, 2011 meeting.**

A follow-up conference call has been scheduled for Sep. 1st at 11:00 am US Eastern Daylight Time to discuss progress on the action items listed above with the entire sub-group and to discuss the plan for the ISPAM stakeholder meeting on September 16th in conjunction with the AOAC Annual Meeting in New Orleans, LA, USA.

Proposals for resolving the top five priority areas of harmonization will be presented to the larger ISPAM stakeholder group for discussion, amendment, and approval at the AOAC Annual Meeting.

**Attachments:**

1. Sub-Group meeting agenda
2. Comparison chart top 5 priorities table
3. Sub-Group harmonization proposals for top priority areas v2
4. Sub-Group harmonization proposal comments from Sam Mohajer
5. CFIA performance parameters - example data sheet
6. Health Canada Health Product and Food Branch Part 4 Table 4.1 – Classification of food categories and suggested food type-pathogen combinations for validation studies
7. ISO 16140 Part 1 – Terminology

Note: Other Documents used for reference during the sub-groups discussion can be found on the ISPAM website: [http://www.aoac.org/ISPAM/guidelines.htm](http://www.aoac.org/ISPAM/guidelines.htm)
Discussion of Proposal from the Sub-Group (SG) to Harmonize the use of Reference Methods

(Phil Feldsine, BioControl Systems)
EQUIVALENCE OF REFERENCE METHODS

- Sub Group Recommendation:

“The Sub Group recognized the essential need to determine the equivalence of reference methods and recommends that a statistical analysis be performed according to a scheme to be defined by ISO TC34 SC9 WG2. If reference methods are determined to be equivalent by statistical analysis for defined matrices, then studies conducted using any of the equivalent reference methods should be acceptable for alternate method validation. Agencies will provide guidance on acceptable reference methods for this analysis.”
EQUIVALENCE OF REFERENCE METHODS

ISPAM Micro Working Group Vote

Motion: Ron Johnson & Morgan Wallace

Unanimous
EQUIVALENCE OF REFERENCE METHODS

Joint ISPAM Vote

Motion:
Phil Feldsine & Pam Wilger

Vote: Unanimous
Pass
Discussion of Proposal from the Sub-Group (SG) to Harmonize the number of Levels/Samples/Fractional Positives

*(Phil Feldsine, BioControl Systems)*
Definition of Fractional Positives

**Sub Group Recommendation:** The Sub Group proposes to define Fractional Recovery as:

“The target recovery is 50% with a range of 25%-75% based on a sample size of 20. However, in the context of the entire study, values outside of the fractional range may be considered. For example, for studies where a larger number of test portions were analyzed, (i.e., 60), a larger fractional range may be acceptable. At least one of the methods should meet this recommendation. Percent recovery outside the target range may be considered on an individual basis.”
TOP 5 ISSUES:
Definition of Fractional Positives

ISPAM Micro Working Group Vote

Motion:
Michael Brodsky & Jim Agin

Unanimous
TOP 5 ISSUES:
Definition of Fractional Positives

Joint ISPAM Vote

Motion:
Ron Johnson & Patrice Arbault

Vote: Unanimous - Pass
Discussion of Proposal from the Sub-Group (SG) to Harmonize the Selection of Food/Category (Sample Matrix)

(Ron Johnson, BioMérieux)
TOP 5 ISSUES: Selection of Food/Category (Sample Matrix)

- ISPAM Micro Working Group Recommendation

- Discussion referred to the Sub-Group (SG):
  - Daniele Sohier
  - Ron Johnson
  - Nuri Gras-Rebolledo
  - Pamela Wilger
  - Irene Iugovaz
  - Tom Hammack
  - Bala Jagadeesan/Julie Moulin
TOP 5 ISSUES:
Selection of Food/Category (Sample Matrix)

Joint ISPAM Vote

Motion:
Bert Popping & Jim Agin

Discussion: Group should look at the BPMM document.

Vote: (23) Pass - Unanimous
Discussion of Proposal from the Sub-Group (SG) to Harmonize Results Analysis & Criteria/Statistical Analysis

(Morgan Wallace, DuPont Qualicon)
TOP 5 ISSUES: Criteria/Statistical Analysis

- Proposal for Adoption

“The statistical analysis on data generated from harmonized study designs will be performed by the RLOD, Chi-square, and/or POD models. As appropriate, other statistical methodologies that have been adopted by certification bodies or regulatory agencies may be applied”
TOP 5 ISSUES:
Criteria/Statistical Analysis

ISPAM Micro Working Group Vote

Motion:

Paul Feldsine & Michael Brodsky

Unanimous
TOP 5 ISSUES:
Criteria/Statistical Analysis

Joint ISPAM Vote

- Motion:
  
  Paul In’t Veld & Morgan Wallace

Discussion: False positives/negatives (method comparison). Move topic to another item.

Vote: (22) Pass / (1) abstain
Discussion of Proposal from the Sub-Group (SG) to Harmonize Number of Samples/Replicates/Method Comparison, Confirmation & Collaborative

(Wendy Lauer, Bio-Rad)
TOP 5 ISSUES:
Minimum number of valid data sets/collaborators

- Methods mostly harmonized at 10 valid data sets with the exception of Health Canada (minimum of 8) and FDA (varies with different study levels).

- Proposal for Adoption
  
  Minimum number of collaborative data sets required is 10 valid data sets.
TOP 5 ISSUES:
Sample size (Test Portion)

- Mostly 25g or as specified by reference method.
- This can differ from industry requirements (i.e., 375g sample size).

Proposal for Adoption

“Analytical test portion size is 25g unless otherwise specified by the methods (alternative or reference). The reference method and alternative method may have different analytical test portion sizes. The alternative test portion size can not be smaller than the reference method test portion size.”
TOP 5 ISSUES:
Number of foods/levels

- Proposal for Adoption

  “Currently harmonized at 3 levels for each matrix – one negative control (n=5), one which produces fractional positive results (n=20) and one higher level (n=5). For certain commodities, a higher number of replicates may be required for specific regulatory agency acceptance.”
TOP 5 ISSUES:
Number of Collaborative Study Replicates

- Methods mostly harmonized at minimum of 8 per level with the exception of AOAC (12 per level).

- Proposal for Adoption
  “A minimum of 8 replicates will be required per level.”
TOP 5 ISSUES: Confirmation

- Mostly harmonized that all samples will be confirmed with the exception of ISO 16140 (for paired samples only confirm + Alternate /- Reference).

- Proposal for Adoption
  “All samples presumptively positive or negative will be confirmed.”
Top 5 Issues: Number of Samples/Replicates/Method Comparison, Confirmation & Collaborative

ISPAM Micro Working Group Vote

• Motion:

  Ron Johnson & Mark Mozola

  Unanimous
Top 5 Issues: Number of Samples/Replicates/Method Comparison, Confirmation & Collaborative

Joint ISPAM Vote
• Motion:
  Bertrand Lombard & Jim Agin

Discussion: replicates of 12 for AOAC, will we change the replicates for AOAC? Why 12?
• Statistically valid, the number of foods are reduced.
• Need to minimize the requirement.

Vote: (23) Pass
Meeting of the International Stakeholder Panel on Alternative Methods (ISPAM) and Working Groups

DRAFT AGENDA
Friday, September 16, 2011
8:30 am – 5:30 pm
Sheraton New Orleans Hotel, New Orleans, Louisiana

SESSION III:
ISPAM WORKING GROUP ON QUALITATIVE CHEMISTRY GUIDELINES FOR ALTERNATIVE METHODS
(Oak Alley) 10:00 am – 3:00 pm

A. Overview and Comparison of Current Qualitative Chemistry Method Validation Guidance – Bert Pöpping, Eurofins and WG Chair

B. Presentation of Draft Guidelines for Validation of Qualitative Chemistry alternative methods – Scott Coates, AOAC Chief Scientific Officer
Page: 41

LUNCH 12:00 pm – 1:00 pm (Grand Chenier, 5th Floor)

C. Development of Standard Method Performance Requirements for Gluten as an example – Scott Coates, AOAC Chief Scientific Officer
Page: 65
Guidelines for Validation of

Qualitative Chemistry Methods
# Table of Contents

1.0 Scope  
2.0 Terms and Definitions  
3.0 Analyte Variants / Specificity Study  
4.0 Sensitivity / Matrix Study  
5.0 Analytical Range  
6.0 Collaborative Study  

Appendix A. Understanding the POD Model  
Appendix B. Raw Format Data Table Template and Example for Qualitative Method: Method Developer Studies, Independent Studies and Collaborative Studies  
Appendix C. Calculation of POD and dPOD Values from Qualitative Method: Method Developer and Independent Validation Data  
Appendix D. Summary Data Table for Qualitative Method: Method Developer and Independent Validation Studies  
Appendix E. Calculation of CPOD and dCPOD Values from Qualitative Method Collaborative Study Data  
Appendix F. Data Summary Table Template and Example for Qualitative Method Collaborative Studies
1 Scope
The purpose of this document is to provide comprehensive guideline for the validation of qualitative methods intended to detect chemical analytes. Qualitative methods are defined as those methods of analysis whose response is either the presence or absence of the analyte detected either directly or indirectly in a specified test portion.

2 Terms and Definitions
Where appropriate, definitions have been taken from international standards and the source is noted. Sources of definitions include the following:
ISO 14971:2007, Medical devices – Application of risk management to medical devices
ISO 17511:2003, In vitro diagnostic medical devices – Measurement of quantities in biological samples – Metrological traceability of values assigned to calibrators and control materials
ISO 5725-1: 1994, Accuracy (trueness and precision) of measurement methods and results – Part 1: General principles and definitions

Acceptable Minimum Detection Level (AMDL)
A concentration that a method must detect the analyte(s) with a specified POD. Both the AMDL and POD are specified by an expert review panel.

Candidate Method
The method submitted for validation.

Collaborator
An intended end user who participates in the collaborative study.

Collaborator Data Set
Results from a combination of an analytical instrument, device, or equipment and an operator, technician, or analyst. Results from a single test site may serve as multiple collaborator data sets dependent on non-redundant operators and acceptable separation of effort.
Intended Use
Use for which a product, process, or service is intended according to the specifications, instructions, and information provided by the manufacturer (ISO 14971).

Matrix
Totality of components of a material system except the analyte (ISO 17511).

Method
A procedure that includes sample processing, assay, and data interpretation.

Probability of Detection (POD)
The proportion of positive analytical outcomes for a qualitative method for a given matrix at a given analyte level or concentration. POD is concentration dependent. Several POD measures can be calculated, e.g., POD_{R} (for the reference method), POD_{C} (for the confirmed candidate method), POD_{CP} (for the candidate presumptive method) and POD_{CC} (for the candidate confirmation method).

Qualitative Method
Method of analysis whose response is either the presence or absence of the analyte detected either directly or indirectly in a specified test portion.

Reproducibility
Precision under reproducibility conditions (ISO 5725-1).

Reproducibility Conditions
Conditions where independent test results are obtained with the same methods on equivalent test items in different laboratories with different operators using separate instruments.

Sample
The batch of material from which replicate test portions are removed for analysis.

Sensitivity
The probability of a (+) response at a given concentration denoted as POD(c).

Specificity
The probability of the method giving a (-) response when the sample is truly without analyte denoted as 1-POD(0).

Test Portion
A quantity of sub sample or member of a sample set that is taken for analysis by the method.
3.0 Analyte Variants / Specificity Study

The analyte variants study is designed to demonstrate that a candidate method can detect the different varieties of the claimed analyte, especially when the claim is for a class of compounds such as "beta-lactam antibiotics". A specificity study demonstrates that a candidate method does not cross react with similar compounds, and therefore the candidate method is sufficiently specific. It is recommended to conduct the analyte variants study and specificity study together since the study designs for both are similar and the two studies serve as countercheck to each other.

3.1 Test Panel

Collect samples representing the variation present in the target analyte population and organize into a 'primary' test panel. For example, if the target analyte is "beta-lactam antibiotics", then a primary test panel of a variety of beta-lactam antibiotics samples must be collected. Document the source and origin of each primary test panel sample. All documentation of analyte identity must be on file and available for review.

Collect samples chosen to adequately cover potentially cross reactive substances into a 'specificity' test panel.

3.2 Study Design

The target concentration for testing primary test panel samples is the pre-determined acceptable minimum detection level (AMDL). Prepare and test one replicate per primary test panel sample using the candidate method.

Similarly, the target concentration for testing specificity test panel samples is 10 times (10X) the pre-determined acceptable minimum detection level (AMDL). Prepare and test one replicate per specificity test panel sample using the candidate method.

Primary and specificity test panel samples evaluations shall be performed together as one study. Primary and specificity test panel samples must be blind coded and intermingled to minimize the possibility that the analysts can know the identity or concentration of the test samples.

If an individual primary test panel sample tests negative, it may be retested in 96 replicates with no failures allowed to demonstrate a 5% lower confidence limit on the probability of detection (POD) of 0.95 or higher for that panel member. If an individual specificity test panel sample tests positive, it may be retested in 96 replicates with no failures allowed to demonstrate a 95% upper confidence limit on the probability of detection (POD) of 0.05 or lower for that panel member. All test and retest data must be reported.
4.0 **Sensitivity / Matrix Study**

The matrix study is designed to demonstrate that a candidate method can detect the target analyte in the claimed matrices. Test samples of the claimed matrix with the target analyte at the AMDL are tested. This guideline assumes a 95% POD. Other POD levels may be used.

4.1 **Matrix Categories**

AOAC INTERNATIONAL recognizes claims for only the range of matrix categories or specific matrix types included in the Method Developer Study or Collaborative Study. The number of different matrices to be tested depends on the claims and intended use of the method.

4.2 **Study Design**

Obtain enough a master sample of each matrix to prepare at least 192 samples (it is recommended to a lot an additional 10% of material to compensate for loss). Test portions are to be taken from each sub-sample.

If the method is intended to detect more than one target analyte simultaneously from the same test portion, the validation study should be designed so that target analytes are spiked into a common sample and the validation tests are performed in a simultaneous manner.

**Incurred or Spiking:** A master sample may be incurred (the analyte is present in the sample) or spiked. If master sample with naturally occurring analyte is used, then additional material that is known to be free of the analyte can be used to ‘dilute’ the incurred sample to the AMDL.

If spiked material is used, then: a) the base material can spiked and then sub-divided, or b) the base material can sub-divided and then spiked.

In either case, prepare 96 test portions of the matrix with the target analyte at the AMDL (a lower concentration for each matrix may be used.) The method developer may choose the specific analyte of more than one variety exists. For example, penicillin may be selected for this study as a representative for beta-lactam antibiotics. Code the prepared samples.

Similarly, prepare 96 test portions of each matrix with one of the specificity test panel compound/substances at 10 times (10X) the AMDL. Code the prepared samples.

Randomly mix the blinded coded samples.

Use the candidate method to evaluate the blinded samples.
5.0 Analytical Range

This study is designed to determine the response of a candidate method over a range of concentrations above and below the AMDL. Most (all?) qualitative methods demonstrate a response curve to the analyte as opposed to step function response. Therefore, the response curve of a candidate method below the AMDL may be an important characteristic of the candidate method, especially if there is a regulatory level above zero established for the target analyte. For example, the safe maximum level for penicillin is 5 ppm. It is important to understand how a candidate method responds at 1, 2, 3 and 4 ppm as well as at 5ppm.

5.1 Study Design

Select and obtain analyte-free material for testing. The study director may select one matrix, or several matrices, for testing.

The study director may select one analyte, or several analytes, for testing.

Select at least four concentrations bracketing the ADML: two concentrations above and below the AMDL. The highest concentration will be used to validate the upper limit of the analytical range.

Prepare five replicates at each concentration, plus five unspiked samples. Blind code replicates. Mix the blinded coded samples so that the analyst cannot know the concentration of the analyte.

An analysts not involved in the preparation of the samples should analyze the samples using the candidate method.

Prepare a ‘dose response’ graph.

6.0 Collaborative Study

A Collaborative Study characterizes the performance parameters (e.g., POD, repeatability, reproducibility) of the candidate method across collaborators.

Methods shall be validated under simulated conditions of intended use. For example, a bacterial threat agent method intended for use by trained emergency response personnel in Personal Protective Equipment at an outdoor mobile site must be validated under conditions that simulate the outdoor mobile site and include representative personnel as collaborators.

A trial run should be conducted prior to the initiation of the validation study. The purpose of the trial run is to ensure that logistics, sample handling and data reporting processes are worked out and understood by all of the collaborators. The trial run should be conducted under the same conditions as the validation study. A reasonable amount of time should be allotted for troubleshooting after the completion of the trial run, including a discussion with all of the collaborators to address issues and answer questions. The data should not be analyzed or included in the validation report.
6.1 Study Design

A collaborative study must include a minimum of 12 collaborators. After eliminating data sets for assignable cause, the study must yield at least 10 valid data sets. The study must include a minimum of 3 test sites with no more than 4 collaborators at any one test site. Care must be taken to ensure the independence of collaborators at each test site.

Divide each matrix into at least 2 samples. Inoculate one sample with the target analyte at the AMDL. Inoculate a second sample with a specificity test panel compound at 10 times the AMDL of the target to serve as the negative control. Analyze test portions from each sample.

If the method is intended to detect more than one target analyte simultaneously from the same test portion, the validation study should be designed so that target analytes are inoculated into a common sample and the validation tests are performed in a simultaneous manner.

The number of test portions per sample per collaborator is 12. Test portions are to be randomized and blind-coded when sent to participating collaborators for analysis.

Each concentration level of each matrix must be analyzed and reported separately. Data sets may be excluded from analysis only for assignable cause.

6.2 Data Analysis

6.2.1 Raw Data Tables
For each matrix and concentration level, report each result from each test portion separately. See Appendix B for raw data table in a software-compatible format.

6.2.2 between Collaborator Standard Deviation
For each matrix, level, and collaborator calculate the POD. For each matrix and level, calculate the standard deviation of collaborator PODs ($s_{POD}$).

6.2.3 Between Collaborator Probability of Detection (CPOD)
Report the CPOD estimates with 95% confidence intervals for the candidate method and, if included, the presumptive and confirmed results. See Appendix E for details.
6.2.4 Summary Data Tables
For all matrices and levels, use the summary table from Appendix F.

6.2.5 Collaborator Comments
Comments on the candidate method should be collected from all collaborators and reported in the Collaborative Study report.
Appendix A. Understanding the POD Model

The Probability of Detection (POD) model is a way of characterizing the performance of a qualitative (binary) method. A binary qualitative method is one that gives a result as one of two possible outcomes, either positive or negative, or presence/absence or +/-.

The single parameter of interest is the Probability of Detection (POD), which is defined as the probability at a given concentration of getting a positive response by the detection method. POD is assumed to be dependent on concentration, and generally, the probability of a positive response will increase as concentration increases.

For example, at very low concentration, the expectation is that the method will not be sensitive to the analyte, and at very high concentration, we desire a high probability of getting a positive response. The goal of method validation is to characterize how method response transitions from low concentration/low response to high concentration/high response.

POD is always considered to be dependent upon analyte concentration. The POD curve is a graphical representation of method performance where the probability is plotted as a function of concentration (see, for example, Figure 1).

Figure 1: Theoretical POD Curve for a qualitative detection method

The POD Model is designed to allow an objective description of method response without consideration to an a priori expectation of the probabilities at given concentrations. The model is general enough to allow comparisons to any theoretical probability function.

The POD Model is also designed to allow for an independent description of method response without consideration to the response of a reference method. The model is general enough to allow for comparisons between reference and candidate method responses, if desired.

Older validation models have used the terms Sensitivity, Specificity, False Positive and False Negative to describe method performance. The POD model has incorporated all of the performance concepts of these systems into a single parameter, POD.
For example, False Positive has been defined by some models as the probability of a positive response, given the sample is truly negative (concentration = 0). The equivalent point on the POD curve for this performance characteristic is the value of the curve at Conc = 0.

Similarly, False Negative has sometimes been defined as the probability of a negative response when the sample is truly positive (concentration >0). In the POD curve, this would always be specific to a given sample concentration, but would be represented as the distance from the POD curve to the POD = 1 horizontal top axis at all concentrations except C=0.

The POD model has incorporated all these method characteristics into a single parameter, which is always assumed to vary by concentration. In other models, the terms “False Positive”, “False Negative”, “sensitivity”, and “specificity” have been defined in a variety of ways, usually not conditional on concentration. For these reasons, their use is denigrated under this model.

<table>
<thead>
<tr>
<th>Denigrated Terminology</th>
<th>Concept</th>
<th>POD Equivalent</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>False Positive</td>
<td>The probability of the method giving a (+) response when the sample is truly without analyte</td>
<td>POD(0) POD at Conc = 0</td>
<td>The POD curve value at Conc = 0 – The “Y-intercept” of the POD curve</td>
</tr>
<tr>
<td>Specificity</td>
<td>The probability of the method giving a (-) response when the sample is truly without analyte</td>
<td>1-POD(0)</td>
<td>The distance along the POD axis from POD=1 to the POD curve value.</td>
</tr>
<tr>
<td>False Negative (at a given concentration)</td>
<td>The probability of a (-) response at a given concentration</td>
<td>1-POD(c)</td>
<td>The distance from the POD curve to the POD = 1 “top axis” in the vertical direction</td>
</tr>
<tr>
<td>Sensitivity (at a given concentration)</td>
<td>The probability of a (+) response at a given concentration</td>
<td>POD(c)</td>
<td>The value of the POD curve at any given concentration</td>
</tr>
<tr>
<td>True Negative</td>
<td>A sample that contains no analyte</td>
<td>C = 0</td>
<td>Point on concentration axis where c = 0</td>
</tr>
<tr>
<td>True Positive</td>
<td>A sample that contains analyte at some positive concentration</td>
<td>C &gt; 0</td>
<td>Range of concentration where c &gt; 0</td>
</tr>
</tbody>
</table>
Use of the terms “Sensitivity”, “Specificity”, “False Positive” and “False Negative” is denigrated under the POD model.
Appendix B. Raw Format Data Table Template and Example for Qualitative Method: Method Developer Studies, Independent Studies and Collaborative Studies

The purpose of the Raw Format Data Table is to document in a software-friendly dataset comprising all of the factors, variables and measurements in the experiment in a standardized format. By matrix and concentration level, report each result from each method for each test portion separately.

Each row (record) in the Raw Format Data Table should contain the following columns (fields):

1. **Matrix type**: An identifier indicating the matrix involved, such as “FILTERS”. The same exact identifier must be used for the same matrix.
2. **Concentration level**: The concentration/test portion for the level.
3. **Test Site**: An identifier uniquely indicating the test site involved, such as “S1”.
4. **Collaborator Team**: An identifier uniquely specifying the collaborator team across test sites, e.g. “C01”.
5. **Instrument**: An identifier uniquely specifying the apparatus used in testing, across test sites and collaborator teams, e.g., “I01”.
6. **Method**: An identifier indicating the test method used, such as “REF” for the reference method, “C-P” for the candidate presumptive method or “C-C” for the candidate confirmation method.
7. **Replicate**: A unique identifier for the test portion involved. If this identifier is common to two rows in the Table, this implies the results are matched by test portion. Example identifiers might be “01” or “001” or “A1”.
8. **Result**: “0” for absence or “1” for presence (detection).

In computer format, the Raw Format Data Table should be given either as: 1) a “fixed-format” file with fixed column widths and blanks or tabs as separators and a file extension of “.txt” or “.xls”; or 2) a “comma-separated value” file with commas as separators between columns and identifiers within quotes, and a file extension of “.csv”.

It is desirable to include a “header” record as the first record in the file with identifiers for each column.

An example file named “banthracis.csv” might be:

```
matrix", "level", "site", "collab", "instrument", "method", "replicate", "result"
"filter", "2.20", "S1", "C01", "I01", "CP", "001", 0
"filter", "2.20", "S1", "C01", "I01", "CC", "002", 1
"filter", "2.20", "S1", "C01", "I01", "R", "003", 1
"filter", "2.20", "S1", "C01", "I01", "CP", "004", 1
"filter", "2.20", "S1", "C01", "I01", "CC", "005", 1
```

etc.
Appendix C. Calculation of POD and dPOD Values from Qualitative Method: Method Developer and Independent Validation Data

In general, four different probabilities detected (PODs) are to be calculated: POD_R (for the reference method), POD_C (for the confirmed candidate method), POD_CP (for the candidate presumptive method) and POD_CC (for the candidate confirmation method).

For each of these four cases, calculate the POD as the ratio of the number positive (x) to total number tested (N):

$$POD = \frac{x}{N}, \quad \text{Where POD is POD}_C, \text{ POD}_R, \text{ etc.}$$

The POD estimates and 95% confidence interval (LCL, UCL) estimates are given by:

1. For the case where $x = 0$,

   POD = 0
   LCL = 0
   $$UCL = \frac{3.8415}{N + 3.8415}$$

2. For the case where $x = N$,

   POD = 1
   $LCL = \frac{N}{N + 3.8415}$
   $UCL = 1$

3. For the case where $0 < x < N$,

   $$POD = \frac{x}{N}$$
   $$LCL = \frac{x + 1.9207 - 1.9600 \sqrt{x - \frac{x^2}{N} + 0.9604}}{N + 3.8415}$$
   $$UCL = \frac{x + 1.9207 + 1.9600 \sqrt{x - \frac{x^2}{N} + 0.9604}}{N + 3.8415}$$

   where $1.9600 = z$, the Gaussian quantile for probability 0.975, $1.9207 = z^2 / 2$, $0.9604 = z^2 / 4$ and $3.8415 = z^2$.

Finally, if $x \leq 1$, set $LCL = 0$. If $x \geq N-1$, set $UCL = 1$.

The confidence interval corresponds to the uncorrected Wilson-score method, modified for $x = 1$ and...
\( x = N - 1 \) to improve coverage accuracy on the boundary (9).

The differences in proportions detected are estimated by (10):

\[
\begin{align*}
\delta_{POD} & = POD_c - POD_R \\
\delta_{CP} & = POD_{CP} - POD_{CC}
\end{align*}
\]

The associated 95% confidence interval (LCL, UCL) for the expected value of \( \delta_{POD} = POD_1 - POD_2 \) is estimated by:

\[
\begin{align*}
LCL & = \delta_{POD} - \sqrt{\frac{POD_1 - LCL_1}{N} + \frac{POD_2 - UCL_2}{N}} \\
UCL & = \delta_{POD} + \sqrt{\frac{POD_1 - UCL_1}{N} + \frac{POD_2 - LCL_2}{N}}
\end{align*}
\]

where \((LCL_1, UCL_1)\) is a 95% confidence interval for \( POD_1 \) and \((LCL_2, UCL_2)\) is a 95% confidence interval for \( POD_2 \), as determined above.

**Example:**

Suppose a single laboratory does a study with \( N = 96 \) replicates and finds all replicates detected \( (x = 96) \). Then \( x = N, \) so

\[
\begin{align*}
POD & = 1.00 \\
LCL & = \frac{N}{N + 3.8415} \\
& = \frac{96}{99.8415} \\
& = 0.962 \\
UCL & = 1.000
\end{align*}
\]

Suppose instead that one of the replicates is not detected, so \( x = 95 = N - 1 \). Then

\[
\begin{align*}
POD & = \frac{95}{96} \\
& = 0.990 \\
LCL & = \frac{x + 1.9207 - 1.9600 \sqrt{[x - x^2 / N + 0.9604]}}{N + 3.8415} \\
& = \frac{96.9207 - 1.9600 \sqrt{1.9500}}{99.8415} \\
& = 0.943
\end{align*}
\]
\[ UCL = 1.000 \quad (\text{because} \ x = N - 1) \]

Suppose finally that \( x = 94 = N - 2 \). Then

\[
POD = \frac{94}{96} = 0.979
\]

\[
LCL = \{ x + 1.9207 - 1.9600 \sqrt{x - x^2 / N + 0.9604} \} / [N + 3.8415] = 0.927
\]

\[
UCL = \{ x + 1.9207 + 1.9600 \sqrt{x - x^2 / N + 0.9604} \} / [N + 3.8415] = 0.994
\]
Appendix D. Summary Data Table for Qualitative Method: Method Developer and Independent Validation Studies

Prepare one table per matrix and indicate laboratory or intended use conditions.

Example: Detection of *Bacillus anthracis* Spores on Nitrocellulose Aerosol Collection Filters by PCR

<table>
<thead>
<tr>
<th>Organism</th>
<th>Concentration</th>
<th>N</th>
<th>x</th>
<th>POD_C</th>
<th>95% LCL</th>
<th>95% UCL</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>B. anthracis</em> Ames</td>
<td>20,000 spores/filter</td>
<td>96</td>
<td>95</td>
<td>0.9896</td>
<td>0.9433</td>
<td>1.0000</td>
</tr>
<tr>
<td><em>B. cereus</em> E33L</td>
<td>200,000 spores/filter</td>
<td>96</td>
<td>1</td>
<td>0.0104</td>
<td>0.000</td>
<td>0.0567</td>
</tr>
</tbody>
</table>

Where POD_C is the POD for the confirmed candidate method, 

x is the number of positive outcomes, 

And N is the number of replicates tested.
Appendix E. Calculation of CPOD and dCPOD Values from Qualitative Method Collaborative Study Data

Calculations are done in 4 steps.

Step 1. The overall fractional response (mean POD = CPOD) for the method is calculated from the pooled POD responses of the individual collaborators (j = 1, 2, ..., L). Then the standard deviation of an individual x value, called the repeatability standard deviation $s_r$, is estimated. Next a 95% confidence interval on $s_r$ is obtained.

Step 2. The standard deviation $s_C$ of the collaborator variance component due to differences in detection among collaborators is estimated. $s_C$ is called the "collaborator effect" standard deviation. A 95% confidence interval for $s_C$ is obtained.

Step 3. Both $s_C$ and $s_r$ are used to estimate a 95% confidence interval for the expected value of CPOD. Also estimated is the reproducibility standard deviation $s_R$, which is the standard deviation of measurement of a single x from a single collaborator, including both the collaborator-collaborator and repeatability error sources. A 95% confidence interval is estimated for the expected value of $s_R$. Also estimated is the "intraclass correlation coefficient" (or "ICC") $I_r$ for the repeatability effect with an associated 95% confidence interval. The $I_r$ indicates the proportion of the total variance that is due to repeatability variance.

Step 4. A T statistic based on the $\chi^2$ distribution or Fisher randomization test of homogeneity of POD across laboratories is done to see if the observed collaborator effect is detectably greater than zero.

**CPOD**

Report the CPOD estimates with 95% confidence intervals for the candidate method and, if included, the presumptive and confirmed results. CPOD estimates are determined as for the single-laboratory case, but are based on the composite data across collaborators.

$$\text{CPOD} = \frac{\sum x_j}{\sum N_j} = \frac{x}{N} \quad j = 1, 2, 3, ..., C$$

where $N = \sum N_j$ is the total number of data, $x = \sum x_j$ is the total number of detections, and $C$ is the number of collaborators.

The CPOD is the proportion detected as positive in the entire set of data across collaborators for the concentration level for the particular method. CPOD is distinguished from POD in that it includes between-collaborator variation as well as within-collaborator variation.

**Repeatability Standard Deviation ($s_r$)**

Estimate the repeatability standard deviation:
\[
S_t^2 = \frac{\sum \left( x_j - x_j' \right)^2 / N_j}{N - C}
\]

\[s_t = \sqrt{s_t^2}\]

where \(N\) is the total number of data and \(C\) is the number of collaborators.

**Collaborator effect standard deviation (\(s_C\))**

Estimate the collaborator effect standard deviation:

\[s_C^2 = \max \left\{ 0, \frac{s_{POD}^2 - s_t^2}{n} \right\}\]

where

\[n = \left( N - \sum \frac{N_j^2}{N} \right) \left( \frac{1}{C - 1} \right)\]

is the weighted average number of replicates per collaborator and

\[s_{POD}^2 = \frac{\sum (POD_j - CPOD)^2}{C - 1}\]

is the observed variance in POD values across collaborators. The degrees of freedom, df, for \(s_{POD}\) are calculated as follows:

\[df = \frac{\left[ \frac{s(L)^2}{L} + \frac{s_r^2}{N} \right]^2}{\left[ \frac{s(L)^2}{L} \right]^2 + \left[ \frac{s_r^2}{N} \right]^2} + \frac{\left[ \frac{s(L)^2}{L} \right]^2}{L - 1} + \frac{\left[ \frac{s_r^2}{N} \right]^2}{N - L}\]

**Reproducibility Standard Deviation (\(s_R\))**

Calculate the reproducibility standard deviation:

\[s_R^2 = s_t^2 + s_C^2\]

If the collaborator effect \(s_C\) is near zero, then \(s_t \approx s_R \approx \sqrt{CPOD (1 - CPOD)}\)

**Intraclass correlation coefficient for repeatability (\(I_0\))**
The "intraclass correlation coefficient" (or "ICC") $I_r$, measures the fraction of the variance in POD values due to repeatability $s_r^2$, and is estimated by:

$$I_r = \frac{s_r^2}{s_r^2 + s_C^2} = \frac{s_r^2}{s_R^2}$$

**Confidence Intervals for CPOD, $s_r$, $s_C$ and $s_R$:**

Methods for calculating confidence intervals for $s_r$, $s_C$ and $s_R$ will be developed by the AOAC Committee on Statistics and incorporated into the AOAC Qualitative Collaborative Study software.

For CPOD, calculate 95% confidence limits as follows:

If $0.15 \leq \text{CPOD} \leq 0.85$:

$$LCL = \max \left\{ 0, \text{CPOD} - \frac{t_{0.975, df} \cdot s(POD)}{\sqrt{C}} \right\},$$

$$UCL = \min \left\{ 1, \text{CPOD} + \frac{t_{0.975, df} \cdot s(POD)}{\sqrt{C}} \right\},$$

If CPOD <0.15 or CPOD > 0.85:

$$LCL = \frac{x + 1.9207 - 1.9600 \sqrt{x - \frac{x^2}{N} + 0.9604}}{N + 3.8415},$$

$$UCL = \frac{x + 1.9207 + 1.9600 \sqrt{x - \frac{x^2}{N} + 0.9604}}{N + 3.8415}$$

Where $x$ is the number of observed positive outcomes and $N$ is the total number of trials.

If CPOD = 0:

$$LCL = 0$$

$$UCL = 3.8415/ (N + 3.8415)$$

If CPOD = 1:

$$LCL = N/(N + 3.8415)$$

$$UCL = 1$$
**dCPOD Estimates and Confidence Intervals**

The dCPOD statistics are differences between the CPOD estimates:

\[
\begin{align*}
\text{dCPOD}_C &= \text{CPOD}_C - \text{CPOD}_R \\
\text{dCPOD}_{CP} &= \text{CPOD}_{CP} - \text{CPOD}_{CC}
\end{align*}
\]

Confidence intervals for the dCPOD values may be obtained from the standard deviations of the paired differences by collaborator:

\[
s_{\text{dCPOD}} = \sqrt{\frac{\sum (d\text{POD} - d\text{CPOD})^2}{C - 1}}
\]

Methods for calculating confidence intervals for dCPOD will be developed by the AOAC Committee on Statistics and incorporated into the Qualitative Collaborative Study software.

**Test of intercollaborator variability**

An intercollaborator effect causing POD variation will almost always be present at some level. The minimum size of this effect can be judged by the estimate \(s_C\) above and how much its LCL exceeds zero. However, "Heywood cases" where \(s_C = 0\) make LCL = 0, which is then uninformative. To test the specific question of whether or not an intercollaborator effect in POD is detectable in the study, a more direct test of this is by, e.g., computing the usual test of homogeneity statistic

\[
T = \sum \left[ \frac{x_i - N_i \cdot \text{CPOD}}{N_i} \cdot \left( \frac{x_i - N_i \cdot \text{CPOD}}{N} \right)^2 \right] + \left( \frac{N_i - x_i - \frac{N_i \cdot N - x}{N}}{N} \cdot \frac{N_i \cdot N - x}{N} \right)^2
\]

\[
= \sum \frac{x_i - N_i \cdot \text{CPOD}^2}{N_i \cdot \text{CPOD} \cdot 1 - \text{CPOD}}
\]

which is approximately distributed as \(\chi^2\) for C-1 degrees of freedom. Alternatively, a Fisher "exact" randomization test can be used on the 2 x C contingency table counts. The size of \(s_C\) should be assessed on its own merit, and the test of T should be used solely to judge whether or not the study is large enough to isolate the effect clearly (e.g., to choose number of replicates or collaborators in a study design). A significance level of \(\alpha = 0.10\) is recommended for the statistic T above, or \(\alpha = 0.05\) if a Fisher "exact" test is used.

**Example:**

Suppose the reference method in an interlaboratory study gave the following results when 12 replicate test portions were tested by each of 10 collaborators:
<table>
<thead>
<tr>
<th>Method Lab</th>
<th>Method R Pos</th>
<th>Method R Neg</th>
<th>Method R Total</th>
<th>Method R POD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>5</td>
<td>12</td>
<td>0.5833</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>3</td>
<td>12</td>
<td>0.7500</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>6</td>
<td>12</td>
<td>0.5000</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>2</td>
<td>12</td>
<td>0.8333</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>7</td>
<td>12</td>
<td>0.4167</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>5</td>
<td>12</td>
<td>0.5833</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>7</td>
<td>12</td>
<td>0.4167</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>5</td>
<td>12</td>
<td>0.5833</td>
</tr>
<tr>
<td>9</td>
<td>11</td>
<td>1</td>
<td>12</td>
<td>0.9167</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>3</td>
<td>12</td>
<td>0.7500</td>
</tr>
<tr>
<td>All</td>
<td>76</td>
<td>44</td>
<td>120</td>
<td></td>
</tr>
</tbody>
</table>

Here, $x = 76$, $N = 120$ and $CPOD = 0.6333 (= 76/120)$.

The repeatability standard deviation

$$s_r^2 = \frac{\sum \left[ x_j - \bar{x}_j / \sqrt{N_j} \right]^2}{N - C} = \left[ \frac{7 - \frac{49}{120} + 9 - \frac{81}{120} + \ldots + 9 - \frac{81}{120}}{120 - 10} \right]$$

$$s_r = \sqrt{s_r^2} = \sqrt{0.2242} = 0.4735$$

And $\sqrt{CPOD (1 - CPOD)} = 0.4819$, suggesting $s_c$ will be small compared to $s_r$.

The among collaborator standard deviation is

$$s_c^2 = \max \left\{ 0, \frac{\sum \text{POD}_i - CPOD^2}{C-1} - \frac{s_r^2}{n} \right\}$$

$$= \max \left\{ 0, \left[ \frac{0.5833 - 0.6333^2 + \ldots + 0.75 - 0.6333^2}{10 - 1} \right] - \frac{0.2242}{12} \right\}$$

$$= \max 0, 0.02963 - 0.0187$$

$$= 0.01093$$

and $s_c = \sqrt{0.01093} = 0.1045$, which is noticeably less than $s_r$, as expected.
The reproducibility standard deviation is

\[ s_R^2 = s_r^2 + s_C^2 \]
\[ = 0.01093 + 0.2242 \]
\[ = 0.2351 \]

so \( s_R = \sqrt{0.2351} = 0.4849 \approx s_r \)

The results are summarized here:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPOD</td>
<td>0.6333</td>
</tr>
<tr>
<td>s_r</td>
<td>0.4735</td>
</tr>
<tr>
<td>s_C</td>
<td>0.1046</td>
</tr>
<tr>
<td>s_R</td>
<td>0.4850</td>
</tr>
<tr>
<td>p-value for T test</td>
<td>0.1703</td>
</tr>
</tbody>
</table>

The “homogeneity test” reported above is the T statistic based on the \( \chi^2 \) distribution, so the p-value of 0.1703 should be compared to 0.10 to see if an intercollaborator effect was detectable. The test indicates the observed value of \( s_C = 0.1046 \) is not statistically significant, so the study was not large enough to reliably detect an intercollaborator effect of this size.
Appendix F. Data Summary Table Template and Example for Qualitative Method Collaborative Studies

Prepare one table per matrix.

Table 1: Comparative results for the detection of *B. anthracis* Ames on aerosol collection filters by the Candidate and Reference methods in a Collaborative Study.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Concentration</th>
<th>Site</th>
<th>Collab or</th>
<th>Instrument</th>
<th>Candidate result $(C)$</th>
<th>Reference method $(R)$</th>
<th>C vs R</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CFU/mL</td>
<td>CFU/filter</td>
<td>N</td>
<td>x</td>
<td>POD$_C$</td>
<td>N</td>
<td>x</td>
</tr>
<tr>
<td>Estimate</td>
<td>0.00</td>
<td>0.00</td>
<td>All</td>
<td>12</td>
<td>0</td>
<td>0%</td>
<td>12</td>
</tr>
<tr>
<td>LCL</td>
<td>0.00</td>
<td>0.00</td>
<td>All</td>
<td>0</td>
<td>0</td>
<td>0%</td>
<td>0</td>
</tr>
<tr>
<td>UCL</td>
<td>0.00</td>
<td>0.01</td>
<td>All</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

| Estimate  | 1830 | 18300 | All | 12 | 11 | % | 12 | 12 | % | -1% |
| LCL       | 400   | 4000  | All | 0 | 9 | % | 0 | 0 | % | -5% |
| UCL       | 9150 | 91500 | All | 95 | 100 | % | 100 | 100 | % | 2% |

| Estimate  | 2120 | 21200 | All | 12 | 12 | % | 12 | 12 | % | -3% |
| LCL       | 4240 | 42400 | All | 97 | 97% | % | 97% | % | 3% |
| UCL       | 1060 | 10600 | All | 100 | 100 | % | 100 | % | 3% |
DETECTION OF GLUTEN IN FOODS

Approved by: International Stakeholder Panel for Alternative Methods (ISPAM)
Final version date:
Effective date:

Intended Use:

1. Applicability:
   Detection of gluten (CAS no.8002-80-0) in raw ingredients, foods, and cooked foods
   that have been formulated, processed or prepared to meet the special dietary needs
   of gluten intolerant people.

2. Analytical Technique:
   Enzyme-linked Immunoassay (ELISA)

3. Definitions:
   Acceptable Minimum Detection Level (AMDL)
   A concentration that a method must detect the analyte(s) with a specified POD.
   Both the AMDL and POD are specified by an expert review panel.

   Specificity
   The probability of the method giving a (-) response when the sample is truly without
   analyte denoted as 1-POD(0).

   Sensitivity
   The probability of a (+) response at a given concentration denoted as POD(c).

   Probability of Detection (POD)
   The probability at a given concentration of getting a positive response by the
   detection method

4. Method Performance Requirements:

<table>
<thead>
<tr>
<th>Analytical range</th>
<th>1 to 200 ppm</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMDL</td>
<td>≤ 20 ppm</td>
</tr>
<tr>
<td>POD at AMDL</td>
<td>≥ 95%</td>
</tr>
<tr>
<td>Specificity (1-POD(0))</td>
<td>≥ 95%</td>
</tr>
<tr>
<td>POD (c)</td>
<td>5 20%</td>
</tr>
<tr>
<td></td>
<td>10 50%</td>
</tr>
<tr>
<td></td>
<td>20 95%</td>
</tr>
<tr>
<td></td>
<td>200 95%</td>
</tr>
<tr>
<td></td>
<td>400 95%</td>
</tr>
<tr>
<td>Laboratory Probability of Detection @ AMDL</td>
<td>≥ 95%</td>
</tr>
<tr>
<td>Laboratory Probability of Detection @ 0 concentration</td>
<td>≥ 95%</td>
</tr>
</tbody>
</table>
5. **System suitability tests and/or analytical quality control:**
   Suitable methods will include blank check samples, and check standards at the lowest point and midrange point of the applicability range.

6. **Reference Material(s):** NIST Standard Reference Material 8418: Wheat Gluten. The source of material for Reference Material 8418 was food-grade-purity Whetpro-80 vita wheat gluten from Canadian western spring wheat flour. NIST Wheat Flour SRM 1567a (protein content 12% by Kjeldahl; assumed gluten content of 10.8%)

7. **Validation Guidance:**
   Recommended level of validation: *Performance Tested Methods*\textsuperscript{SM}

9. **Maximum Time-To-Result:** 2 hours.
2 main documents:

Guidelines for validation of qualitative methods

SMPR for Gluten
Comparison Microbiology – Qual Chem Methods

• Micro: One area with a limited group of methods

• Qual Chem: ideally covers all areas for qual chem detection ranging from ELISA, LC-MS/MS, PCR, SPR etc.
Guidance Document

• General enough to cover all areas
• Specific enough to still give general guidance
• Ideally in line with ISO/CODEX etc
GLUTEN

• Applicability of procedure
• Method Performance Parameters
• Underlying model: POD
TIMELINE

• First draft by Oct 15 (Scott Coates + Group)
• Revision by group
• Ready by Dec/Jan
Meeting of the International Stakeholder Panel on Alternative Methods (ISPAM) and Working Groups

DRAFT AGENDA
Friday, September 16, 2011
8:30 am – 5:30 pm
Sheraton New Orleans Hotel, New Orleans, Louisiana

SESSION IV:
ISPAM JOINT SESSION
(Bayside C) 3:15 pm – 5:30 pm

A. Presentation of recommendations from WG on Microbiological Guidelines for discussion and vote by ISPAM voting members – Russ Flowers, Mérieux NutriSciences and Chair, ISPAM

B. Presentation of recommendations from WG on Qualitative Chemistry Validation Guidelines for vote by ISPAM voting members – Bert Pöpping, Eurofins and WG Chair

C. Next Steps - Russ Flowers, Mérieux NutriSciences and Chair, ISPAM
   i. Implementation of recommendations – how will recommendations be implemented by agencies and organizations
   ii. Should WGs develop a standard validation study design template
   iii. Should the goal be adoption of a standard validation study design by international agencies and organizations
   iv. Should WGs develop standard method performance requirements for new and updated study protocols
INTERNATIONAL Stakeholder Panel on Alternative Methods
Purpose and Scope

- Develop harmonized, internationally accepted standard validation guidelines for alternative chemical and microbiological methods

Objectives:

- Improve efficiency and minimize economic burden by leveraging global networks of experts to reach consensus on analytical validation protocols
Working Groups

- **Smaller Working Groups of subject matter experts have been formed to identify potential conflicts and propose solutions.**
  - AOAC has learned that this upfront work saves a tremendous amount of time and focuses the larger stakeholder panel on consensus decision making.
  - Three working groups were established: Microbiology Work Group, Chemistry Work Group, and Statistics Work Group.
ISPAM History

- November 29, 2010
  - AOAC Research Board endorsed supporting a stakeholder panel to work on harmonizing validation guidelines for alternative methods.

- December, 2010
  - AOAC Board of Directors approves establishing an Advisory Panel to define the project

- March 15, 2011
  - Advisory Panel meets to define the project and recommends experts for Stakeholder Panel and Working Groups

- April 12, 2011
  - Review of potential Panel and Working Group members
ISPAM History

- April 21, 2011
  - Teleconference to accept nominations for Stakeholder Panel and Working Groups

- April 22, 2011
  - AOAC staff meeting to organizes to do the impossible!

- April 27 & 28, 2011
  - First of numerous conference calls of the working groups

- June 30
  - Mid-Year Stakeholder meeting

- September 16
  - Meeting of the International Stakeholder Panel on Alternative Methods (ISPAM) Meeting
Discussion of Proposal from the Sub-Group (SG) to Harmonize the use of Reference Methods

(Phil Feldsine, BioControl Systems)
EQUIVALENCE OF REFERENCE METHODS

Sub Group Recommendation:

“The Sub Group recognized the essential need to determine the equivalence of reference methods and recommends that a statistical analysis be performed according to a scheme to be defined by ISO TC34 SC9 WG2. If reference methods are determined to be equivalent by statistical analysis for defined matrices, then studies conducted using any of the equivalent reference methods should be acceptable for alternate method validation. Agencies will provide guidance on acceptable reference methods for this analysis.”
EQUIVALENCE OF REFERENCE METHODS

ISPAM Micro Working Group Vote

Motion:
Ron Johnson & Morgan Wallace

Unanimous
Joint ISPAM Vote

Motion:
Phil Feldsine & Pam Wilger

Vote: Unanimous
Pass
Discussion of Proposal from the Sub-Group (SG) to Harmonize the number of Levels/Samples/Fractional Positives

(Phil Feldsine, BioControl Systems)
TOP 5 ISSUES:
Definition of Fractional Positives

- **Sub Group Recommendation:** The Sub Group proposes to define Fractional Recovery as:

  “The target recovery is 50% with a range of 25%-75% based on a sample size of 20. However, in the context of the entire study, values outside of the fractional range may be considered. For example, for studies where a larger number of test portions were analyzed, (i.e., 60), a larger fractional range may be acceptable. At least one of the methods should meet this recommendation. Percent recovery outside the target range may be considered on an individual basis.”
TOP 5 ISSUES:
Definition of Fractional Positives

ISPAM Micro Working Group Vote

Motion:
Michael Brodsky & Jim Agin

Unanimous
TOP 5 ISSUES:
Definition of Fractional Positives

Joint ISPAM Vote

Motion:
Ron Johnson & Patrice Arbault

Vote: Unanimous - Pass
Discussion of Proposal from the Sub-Group (SG) to Harmonize the Selection of Food/Category (Sample Matrix)

(Ron Johnson, BioMérieux)
TOP 5 ISSUES: Selection of Food/Category (Sample Matrix)

- ISPAM Micro Working Group Recommendation

- Further discussion was referred to a Sub-Group (SG):
  - Daniele Sohier
  - Ron Johnson
  - Nuri Gras-Rebolledo
  - Pamela Wilger
  - Irene Iugovaz
  - Tom Hammack
  - Bala Jagadeesan/Julie Moulin
TOP 5 ISSUES:
Selection of Food/Category (Sample Matrix)

Joint ISPAM Vote to defer food categories to a sub-group

Motion:
Bert Popping & Jim Agin

Discussion: Group should look at the BPMM document.

Vote: (23) Pass - Unanimous
Discussion of Proposal from the Sub-Group (SG) to Harmonize Results Analysis & Criteria/Statistical Analysis

(Morgan Wallace, DuPont Qualicon)
TOP 5 ISSUES:
Criteria/Statistical Analysis

- Proposal for Adoption

“The statistical analysis on data generated from harmonized study designs will be performed by the RLOD, Chi-square, and/or POD models. As appropriate, other statistical methodologies that have been adopted by certification bodies or regulatory agencies may be applied”
TOP 5 ISSUES:
Criteria/Statistical Analysis

ISPAM Micro Working Group Vote

• Motion:

  Paul Feldsine & Michael Brodsky

  Unanimous
TOP 5 ISSUES:
Criteria/Statistical Analysis

Joint ISPAM Vote

- Motion:

  Paul In’t Veld & Morgan Wallace

  Discussion: False positives/negatives (method comparison). Move topic to another item.

Vote: (22) Pass / (1) abstain
Discussion of Proposal from the Sub-Group (SG) to Harmonize Number of Samples/Replicates/Method Comparison, Confirmation & Collaborative

(Wendy Lauer, Bio-Rad)
TOP 5 ISSUES:
Minimum number of valid data sets/collaborators

- Methods mostly harmonized at 10 valid data sets with the exception of Health Canada (minimum of 8) and FDA (varies with different study levels).

- Proposal for Adoption
  Minimum number of collaborative data sets required is 10 valid data sets.
TOP 5 ISSUES:
Sample size (Test Portion)

- Mostly 25g or as specified by reference method.
- This can differ from industry requirements (ie, 375g sample size).

Proposal for Adoption

“Analytical test portion size is 25g unless otherwise specified by the methods (alternative or reference). The reference method and alternative method may have different analytical test portion sizes. The alternative test portion size can not be smaller than the reference method test portion size.”
TOP 5 ISSUES:
Number of foods/levels

- Proposal for Adoption
  
  “Currently harmonized at 3 levels for each matrix – one negative control (n=5), one which produces fractional positive results (n=20) and one higher level (n=5). For certain commodities, a higher number of replicates may be required for specific regulatory agency acceptance.”
TOP 5 ISSUES:
Number of Collaborative Study Replicates

- Methods mostly harmonized at minimum of 8 per level with the exception of AOAC (12 per level).

- Proposal for Adoption
  “A minimum of 8 replicates will be required per level.”
TOP 5 ISSUES: Confirmation

- Mostly harmonized that all samples will be confirmed with the exception of ISO 16140 (for paired samples only confirm + Alternate /- Reference).

- Proposal for Adoption
  “All samples presumptively positive or negative will be confirmed.”
Top 5 Issues: Number of Samples/Replicates/Method Comparison, Confirmation & Collaborative

ISPAM Micro Working Group Vote

- Motion:

  Ron Johnson & Mark Mozola

  Unanimous
Top 5 Issues: Number of Samples/Replicates/Method Comparison, Confirmation & Collaborative

Joint ISPAM Vote

Motion:
Bertrand Lombard & Jim Agin

Discussion: replicates of 12 for AOAC, will we change the replicates for AOAC? Why 12?
- Statistically valid, the number of foods are reduced.
- Need to minimize the requirement.

Vote: (23) Pass
Meeting of the International Stakeholder Panel on Alternative Methods (ISPAM) and Working Groups

DRAFT AGENDA
Friday, September 16, 2011
8:30 am – 5:30 pm
Sheraton New Orleans Hotel, New Orleans, Louisiana

V. MINUTES:

V.1 ISPAM MICROBIOLOGICAL WORKING GROUP
   a) Minutes Microbiological Working Group August 16-17, 2011
      Page: 67
   b) Minutes Microbiological Working Group June 30, 2011
      Page: 72

V.2 ISPAM QUALITATIVE CHEMISTRY WORKING GROUP
   a) Minutes Chemistry Working Group June 30, 2011
      Page: 75

V.3 ISPAM STATISTICAL WORKING GROUP
   a) Minutes Statistics Working Group June 29, 2011
      Page: 79
Meeting of Sub-Group to Harmonize Microbiological Guidelines
AOAC Headquarters
481 N. Frederick Ave., Gaithersburg, MD 20877
Meeting Minutes
August 16-17, 2011

Meeting Attendees:
Yi Chen, FDA, AOAC General Referee, Microbiology
Peter Evans, USDA FSIS
Phil Feldsine, BioControl Systems
Russ Flowers, Silliker, ISPAM Chair and AOAC President
Ron Johnson, BioMerieux
Wendy Lauer, Bio-Rad Laboratories
Harry Marks, USDA, FSIS (Tuesday only)
Kirsten Mattison, Health Canada
Mark Mozola, Neogen
Sam Mohajer, Canadian Food Inspection Agency
Palmer Orlandi, US FDA
Morgan Wallace, DuPont Qualicon

AOAC Staff:
Zerlinde Johnson
Krystyna McIver
Nora Marshall

I. Introductions

Russ Flowers, Chair, ISPAM provided a review of project scope and the five areas of primary concern to the harmonization effort for discussion during the meeting (see Attachments 1 & 2).

II. Harmonization of Reference Methods

Discussion ensued on the proposal for collecting, organizing and analyzing existing method comparison data to ascertain if the data can be analyzed to demonstrate method equivalency between reference methods (see Attachment 3 and 4).

The group decided to limit the data collection to Salmonella in five food types. The question of food type and what food type to start with and ask for data sets to be submitted. The following food types were selected with the caveat that there are sufficient data available for these categories from studies using the ISO reference method to do the analysis:

1. Raw ground poultry*
2. Raw ground beef
3. Non-fat dry milk
4. Ice cream
5. Chocolate
*homogenized, naturally contaminated data sets will be accepted.

There was discussion of the fact that some reference methods contain options written into the method. If there are options within the method, then the option used must be specified in the submission of the data set. Submitted data will be blind coded for before statistical analysis.

Action Items:
1. CFIA will provide a spreadsheet for data entry and a draft call for data memo with instructions to method developers on how to populate the data sheet with the requested data sets (reference method vs. alternative method). An example of the spreadsheet is under Attachment 5.

There was discussion of the concept and food types selected with Bertrand Lombard (ISO) in a teleconference on August 17. Dr. Lombard agreed with the concept and food types and:
   - Recommends gaining input from the ISO TC 34/SC 9/WG 2, on statistics on what is the most appropriate statistical model(s) that should be used in the analysis of the data submitted.
   - Ideally everyone will agree on the statistical model used to analyze the data collected.
   - The goal is to agree that the data supports the acceptance of the equivalency of the reference methods, and mutual acceptance.

III. Number of Levels/Samples/Fractional Positive

The group discussed fractional positives (see Attachment 3) and what is the appropriate number of fractional positives necessary for method validation. The following is a proposal for the definition of acceptable fraction positive results:

Target is 50% with a range of ± 25%, based upon 20 replicates. Provided, however, that in the context of the entire study data set, values outside of the fractional range may be considered. For example, for studies where a larger number of test portions were analyzed, (i.e., 60), a larger fractional range may be acceptable. Other parameters may be considered on an individual basis.

The sub-group agreed upon the following proposal.

For inoculated samples, for a food matrix (item) –
   - Minimum 20 replicates with fractional positive
   - At least 5 replicates for high level
   - 5 uninoculated replicates

The option of pooling data from multiple day trials to fill the 60 replicates per category criteria of some regulatory agencies was discussed.
It was agreed on a minimum of 60 replicates per category, subject to review by the General Referee.

Possible language for inclusion. Group has not agreed on the inclusion of this language or any language to this affect:
For certain commodities a higher number of replicates may be needed for specific regulatory agency acceptance.

IV. Selection of Food/Category (sample matrix)

The group decided on the following proposal:

1. There is a need for category list for AOAC-claimed matrices, since currently no category claim exists.
2. There was discussion of claims for “all foods” versus “variety of food”. The group recommended proposing to HC and ISO to remove the “all foods” claim.
3. Group would like AOAC to move to 3 foods per category, with 5 categories needing to be tested for “variety of foods” claim.

However, because there is no common use of the definitions for category of foods, versus type of food for example, the group agreed to propose to harmonize terminology and agree on the use of common definitions using the draft ISO 16140 Part 1: Terminology as a starting point (see Attachment 7).

The group discussed and proposed the following food categories, considering the AOAC, ISO and Health Canada (HC) food categories. The HC food category table under Attachment 6 was also used as a reference:

**Ready-To-Eat category (RTE):**
FSIS – proposed three RTE meat/poultry categories (cooked, acidified, and dried/cured). And a minimum of 60 replicates of a single food item to represent each category.

Discussion of “RTE meat” claim and the need for RTE category under the ISO food categories.

USDA FSIS requires 60 reps per category and has RTE meats broken down into four food types based upon intrinsic/extrinsic factors (acidic, cured, etc…). For FSIS validation guidance it would be acceptable for 30 reps for typical matrix and 30 reps for challenge matrix = 60 reps for total RTE class.

Since ISO 16140 does not break out RTEs as a separate category, it was proposed that SC 9 consider this possibility.

**Fish and Seafood Category**
HC same as AOAC and ISO. Therefore, ISO 16140 list can be used.

**Fruit and Vegetable-based product category**
It was recommended that nuts/nut butter/nut meat should be added to ISO 16140 as a category.

**Egg and Egg Derivatives**
Add egg category to ISO and AOAC.

**Dairy category**
No differences between AOAC, HC and ISO.

**Chocolate and Baked Goods**
No differences between AOAC, HC and ISO.

**Animal feeds**
Animal feeds should be a separate category for HC, like it is for AOAC and ISO.

**Miscellaneous**
No differences between AOAC, HC and ISO.

**Multi-ingredient foods**
HC and ISO already have this category. Group recommended that AOAC incorporate this category.

Group recommended that environmental categories be done separately.

The group asked Bertrand Lombard (ISO) whether the removal of “all foods” claim from the ISO standard is possible. Lombard agreed with the idea of removing the “all foods” claim, but indicated that the issue should be brought before Subcommittee 9 for discussion and decision on the issue.

**Action Items:**
1. HC/CFIA will provide a draft category and food items table as they relate to micro validations by September 1st.
2. HC/CFIA will provide a draft terms and definitions for category, food item and type. There will be a draft food categories and environmental categories to also be submitted by September 1st.
3. HC will also propose a draft proposal for harmonizing the use of definitions using ISO 16140 Part 1 as the starting point for the common validation terminology (see Attachment 7). This will be submitted by September 1 for discussion by the subgroup and further discussion by the larger Working Group at the September 16 meeting in New Orleans.

**V. Results analysis & Criteria/Statistical Analysis**

The group discussed the issue and decided on the following proposal to bring forth to ISPAM:

“The statistical analysis on data generated from harmonized study designs will be performed by the RLOD and/or Chi-square models. As appropriate, other
statistical methodologies that have been adopted by certification bodies or regulatory agencies may be applied.

Health Canada and CFIA provided documents that explain their criteria for statistical analysis. These documents were provided by email to those who were not in attendance.

VI. Number of Data Sets for Collaborative Studies/Sample Size

The group has decided on the following proposal:

Minimum number of collaborative data sets required is 10 valid data sets.

All samples will be confirmed and therefore meet ISO/AOAC/HC requirements.

The group agreed on the following statement:
Analytical test portion size is 25g unless otherwise specified by the methods (alternative or reference). The reference method and alternative method may have different analytical test portion sizes.

VII. Summary and next steps for September 16, 2011 meeting.

A follow-up conference call has been scheduled for Sep. 1st at 11:00 am US Eastern Daylight Time to discuss progress on the action items listed above with the entire sub-group and to discuss the plan for the ISPAM stakeholder meeting on September 16th in conjunction with the AOAC Annual Meeting in New Orleans, LA, USA.

Proposals for resolving the top five priority areas of harmonization will be presented to the larger ISPAM stakeholder group for discussion, amendment, and approval at the AOAC Annual Meeting.

Attachments:

1. Sub-Group meeting agenda
2. Comparison chart top 5 priorities table
3. Sub-Group harmonization proposals for top priority areas v2
4. Sub-Group harmonization proposal comments from Sam Mohajer
5. CFIA performance parameters - example data sheet
6. Health Canada Health Product and Food Branch Part 4 Table 4.1 – Classification of food categories and suggested food type-pathogen combinations for validation studies
7. ISO 16140 Part 1 – Terminology

Note: Other Documents used for reference during the sub-groups discussion can be found on the ISPAM website: http://www.aoac.org/ISPAM/guidelines.htm
Russ Flowers, Silliker, AOAC President
Jim Agin, Q Labs
Patrice Arbault, Bioadvantage Consulting
Marcia Armstrong, Qiagen
Stan Bailey, bioMerieux
DeAnn Benesh, 3M
Reginald Bennet, FDA
Peter Bodnaruk, Ecolab
Yan Cao, Life Technologies
Yi Chen, U.S. FDA CFSAN
Benjamin Diep, Nestle´
Peter Evans, USDA FSIS
Phil Feldsine, BioControl
Peter Feng, U.S. FDA CFSAN
Nuri Gras, Labser, Ltda.
Qian Graves, U.S. FDA CFSAN
Brad Goskowicz, Microbiologics, Inc.
Bob Jechorek, 3M
Ron Johnson, bioMerieux
Tom Hammack, U.S. FDA CFSAN
Melinda Haymen, FSNS
Robert LaBudde, Least Cost Formulation
Roxane Lafortune, bioMerieux Canada
Wendy Lauer, Bio-Rad
Zhiyong Li, Guangdong Entry-Exit Inspection and Quarantine Bureau
Tony Lupo, Neogen
Harry Marks, USDA FSIS
Kirsten Mattison, Health Canada
Sam Mohajer, CFIA
Dan Morse, 3M Infection Prevention
Julie Moulin, Nestle
Mark Mozola, Neogen
Palmer Orlandi, U.S. FDA FERN
Wendy Ortman, Decagon Devices
Linda Peng, DuPont
Yvonne Salfinger, Florida State Department Agriculture
Sharan Shea, APHL
Appavu Kalyana Sundaram, Engineering Services, Inc.
Dan Tholen, Dan Tholen Statistical Consulting
Morgan Wallace, DuPont
Pamela Wigler, Cargill

AOAC Staff:
Arlene Fox
Krystyna McIver
Zerlinde Johnson
Working Group Chair Russ Flowers welcomed attendees and indicated that there will be no voting in this meeting, just discussions and consensus on the direction from this ISPAM Micro Working Group as to how to attain the goal of harmonizing the various microbiological guidelines.

A small sub-group of the Microbiology Working Group, key representatives from regulatory agencies and standard setting bodies who developed microbiological method validation guidelines, has been tasked to work on the five top priority areas of divergence and propose a harmonized plan. Sub-group Chair, Phil Fieldsine, indicated that the group reached consensus on two of the areas and needed more time to present their harmonization plan to the entire group.

The top five priority areas as decided in teleconference meetings and confirmed at this WG meeting are as follows:

1. Reference Methods
2. Selection of Food/Category (sample matrix)
3. Number of Levels/Samples/Fractional Positive
4. Results analysis & Criteria/Statistical Analysis
5. Number of Samples/Replicates/Method Comparison & Collaborative

More representation from AFNOR and ISO was requested by working group members. Chair Russ Flowers indicated that European representatives, Paul In’t Veld and Bertrand Lombard, two key ISO representatives who are currently working on ISO 16140 guidelines, will attend the ISPAM meeting at the AOAC Annual Meeting in New Orleans, LA, but acknowledged that more ISO representation is needed at the working group level.

Regulatory agency representatives were asked to provide their thoughts on the harmonization efforts and those issues of importance to their agencies.

U.S. FDA CFSAN comments were provided by Thomas Hammack:

FDA is interested in the harmonization effort as it impacts international trade and is participating in the discussions to try and harmonize the top priority issues. Of the top five areas of divergence, they are most interested in the use of reference methods and fractional positives. FDA would like to see comparison data on the reference methods with fractional positives in order to consider accepting reference methods other than BAM. It was suggested that proficiency testing data might serve as a tool for comparison of reference method performance. The matrices of particular interest to the FDA include shrimp, leafy green produce, sprouts, and spices.

USDA/FSIS comments were provided by Peter Evans:

USDA/FSIS uses methods in the Microbiology Laboratory Guidelines (MLG). FSIS does not endorse or approve methods for industry. Industry may use a variety of methods to verify the effectiveness of a HACCP plan. FSIS finds AOAC, ISO, and AFNOR methods acceptable for this purpose. USDA/FSIS has the greatest concern for methods that produce false negatives. They would like validation studies to prove that alternative methods do not provide false negative results when compared to the reference method. The USDA FSIS utilizes a one-sided chi square test to analyze unpaired data sets. USDA FSIS performs more replicates to increase the odds of detecting a method that does not perform adequately. USDA FSIS also would like to see a
standard matrix and common strains found in outbreaks used in validations to test the efficacy of the test method to detect these strains. The impact on the assay of larger sample size is particularly of concern to FSIS during validation.

Health Canada (HC) comments were provided by Kirsten Mattison:
Health Canada has been in transition during the revision of the HC guidelines for alternate method validation. Now that the new guidelines are available, it is important to engage the industry and try to harmonize validation schemes for the alternative methods. Harmonization of validation efforts would reduce the time it takes HC to review method submissions. Prevention of false negatives is of greatest concern. Detection of *Listeria* has been the focus of analysis following an outbreak. A main concern of HC is the use of reference methods; ideally HC would like Health Protection Branch (HPB) to be used, or equivalent. Equivalency is decided on a case by case basis. Showing reference methods are equivalent is becoming increasingly important to HC since many alternative methods are not validated in comparison to a HC HPB reference method. Validation of HPB methods must include a minimum of eight government laboratories.

The next steps for the ISPAM Working Group will be for the sub-working group to come to consensus on the five areas of major divergence and provide a proposal for review to the working group to discuss and provide comments. This would ideally happen before the AOAC Annual Meeting. A stakeholder meeting of the ISPAM group will take place on September 16, 2011 in conjunction with the AOAC Annual Meeting in New Orleans, LA, and it is anticipated that the group would vote on the proposals to move forward in the five areas of divergence.
1. **Introductions:**

   Bert Pöpping, Eurofins and WG Chair provided an overview of the working group activities and discussed the objectives of the meeting. With that he introduced the day’s presenters.

2. **Presentations:**

   **Jim Harnly, USDA-ARS BHNRC**

   Presented a power point on the “Draft AOAC Single Laboratory Validation of the Identification of [botanical or specific botanical material]”. The presentation gave an overview of what distinguishes the SLV Botanical Identification Methods (BIM) from other methods of analysis. The (BIM) analyzes the whole material, not specific individual components, markers or analytes. Identification is based on comparisons of an unknown botanical material to the target material and utilizes inclusivity (non-target material) and exclusivity (target material) for measurement criteria and analytical measurement. The SLV method use a POD concept which was adapted to become a POI, Point of Identification and log-log calculation. Superior and inferior materials (acceptable/un-acceptable limits) need to be established and if there is a unknown non-target material, the BIM inclusivity panel determines mean spectra and variance.
Roy Macarthur, Food Environmental Research Agency, UK

Overview of why it is important to validate methods. The validation gives a prediction of the performance a method will possess when in use. Testing performed at “zero” and at least one other level is done to achieve target LOD and use of “positive” and “negative” results. The POD concept assists the validation method with the probability of “false” negatives and “false” positives. The Protocol takes into consideration inter-laboratory and single laboratory testing. Whereby the inter-laboratory tests at least 10 replicates under repeatability conditions and a prediction of the interval within which we can expect the probability of a false positive result to lie when the method is applied in a new laboratory. A prediction of the interval within expected false negative probability to lie at concentrations of interest. Hence, a prediction interval for limit of detection when the method is applied in a new laboratory. The Single Laboratory is all of the above, but on a new day.

Paul Wehling, General Mills
“Validation Scheme for Qualitative Analytical Method”, ISO Technical Committee 34, Standing Committee 16, Horizontal methods for molecular biomarker analysis, 2011 and “Probability of Detection (POD) as a Statistical Model for the Validation of Qualitative Methods.”

Discussion of the POD concept and its intimate correlation of assessing the parameters of probability of detection. Its versatility as part of any qualitative (binary) method. Probability will change with the concentration and as such the POD concept assists laboratories assess the suitability of a method of intended purpose by predicting the probability of a positive result in any given concentration. The POD concept is a simple statistic inherent in all other systems, such as Chi-Square, LOD, RLOD. The “POD Concept” is only new in that it recognizes the POD as a key parameter and plots a graph of POD vs concentration. The POD curve can be an indicator of the “usefulness” of the method. If POD were constant across all concentrations, the method would not be useful. Works for single lab and Multilab experiments. Works for paired and unpaired designs. Provides harmonization across qualitative/quantitative methodologies. Does comparisons and hypothesis tests via confidence interval analysis – equivalent to chi-squared tests. POD Curve plots mean response and uncertainty on the same graph.

Scott Coates, Chief Scientific Officer AOAC INTERNATIONAL

Coates reviewed the SMPR guideline that was developed in response to the number of SMPRs being written for a number of different projects. The guideline is a compilation of a number of existing AOAC guidelines and documents. The SMPR guideline covers chemistry and microbiology methods; and qualitative, quantitative, and identification
methods. The guideline provides guidance on the types of data required to validate methods. The guideline also include definitions of and recommendations for validation parameters such as reference method comparison, inclusivity/selectivity, exclusivity /specificity, environmental interference , laboratory variance, bias , and probability of detection (pod).

Scott Coates, Chief Scientific Officer AOAC INTERNATIONAL
“AOAC INTERNATIONAL Guidelines for Validation of Biological Threat Agent Methods and/or Procedures” (BTAM)
The Biological Threat Agent Methods and/or Procedures (BTAM) validation guideline was prepared in response to the development of a series of SMPRs for biological threat agents. The guideline was published in March 2011. The project to write the BTAM document was funded by the US Department of Defense and US Department of Homeland Security. The BTAM guideline was the first guideline to incorporate: 1) intended user testing; 2) test site definitions (as opposed to the traditional laboratory testing site); 3) instrument variation testing; and 4) the POD concept. The guideline is also the first document to allow for a reduction of collaborative study test sites.

Deborah McKenzie, Senior Director, AOAC Methods Development & Approval Processes
Hilde Skaar Norli, Nordic Committee on Food Analysis - Authored
Deborah McKenzie presented on behalf of Hilde Skaar Norli to offer contrasts and similarities between NordVal protocol for validations and the previously reviewed presentations. NordVal resonates AOAC’s attitude of being a practical guide, a procedure that can be used by everyone, without requiring profound statistical knowledge. The NordVal protocol is based on a NMKL guide, NMKL Procedure No 20, 2007: "Evaluation of results from qualitative methods". The NMKL procedure has also an on-line excel spreadsheet to perform test calculations. Deviations from AOAC style validations include the use of kappa instead of POD concept to estimate agreement between the candidate and reference methods (or between the expected/true results and the obtained results). It also requires the use of 2 independent laboratories for confirmation instead of only 1 required by AOAC INTERNATIONAL. NordVal differs from previous protocols whereby it lends a definition of comprises a proprietary method.

Scott Coates, Chief Scientific Officer AOAC INTERNATIONAL
Comparison of Current Qualitative Chemistry Method Validation Guidance
Six guidelines were reviewed that specifically referenced recommendations for qualitative chemistry. The six documents are as follows:

MoniQA: Draft – “A Protocol for the Validation of Qualitative Methods of Detection”
Roy Macarthur (Fera) & Christoph von Holst (IRMM), Joint IUPAC/MoniQA protocol for validation of qualitative methods, Monitoring & Quality Assurance (MoniQA) 2011
ISO: Draft- “Validation Scheme for Qualitative Analytical Methods, ISO Technical Committee 34, Standing Committee 16


BTAM: “AOAC INTERNATIONAL Methods Committee Guidelines for Validation of Biological Threat Agent Methods and/or Procedures (BTAM), Journal of AOAC INTERNATIONAL Vol. 94, NO. 4, 2011

The guidelines contain many similar concepts with differences specific to the application. The POD concept is prevalent throughout the guidelines.

Bert Pöpping, ISPAM WG Chair, Eurofins
Definition of Qualitative Chemistry Methods
Qualitative methods test for either “positive” or “negative”, not corresponding concentration. Any method that gives a binomial response can be considered a qualitative method. A good example of qualitative method would be real-time PCR. The PCR results can be both, qualitative and quantitative with its binary evaluation. It was proposed during the Working Group meeting that perhaps the PCR results be taken in to groups. One using the binary results, and the second group use the results in a further detailed analysis. This discussion will be continued at future meetings.

All of the presentations listed above are available on line at the ISPAM page. While on our website, please take a moment to review the up-coming AOAC INTERNATIONAL Annual Meeting agenda, being hosted in New Orleans, Louisiana, USA on September 18-21st, 2011.
Committee on Statistics Attendees:
Paul Wehling, Chair, General Mills/Medallion Labs
Joshua Hicks, Bruker-BioSpin
Robert LaBudde, Least Cost Formulations Ltd.
Roy MacArthur, Food Environmental Research Institute
Julie Moulin, Nestlé S.A.
Dan Tholen, Statistical Consultant
Caryn Thompson, Elanco Animal Health

Observers:
Jim Agin, Q Laboratories, AOAC OMB Member
Stan Bacler, Health Canada, AOAC President-Elect
DeAnn Benesh, 3M Food Safety
Yi Chen, FDA, AOAC General Referee for AOAC Committee on Microbiology
Yan Cao, Life Technologies
Ron Johnson, bioMerieux, Member of AOAC Board of Directors
Peter Feng, FDA
Bob Jechorek, 3M Food Safety
Prabakar Katsuri, PepsiCo
Wendy Lauer, Bio-Rad Laboratories
Markus Lipp, USP
Tony Lupo, Neogen
Jeff Moore, USP
Dan Morse, 3M
Mark Mozola, Neogen
Linda Peng, DuPont Qualicon
Yvonne Salfinger, Florida State Dept Agriculture, Chair of AOAC Committee on Microbiology
Lindell Ward, Covance Laboratories
Morgan Wallace, DuPont Qualicon
Pamela Wilger, Cargill

AOAC Staff:
Arlene Fox
Zerlinde Johnson

Overall goal of the harmonization meeting is to have draft validation guidelines for microbiological and chemical qualitative methods by AOAC Annual Meeting. End goal of this meeting would be a plan on how to help the working groups get the guidance together by September. Assigning of statisticians to the other two working groups to provide input on statistical design for microbiological and chemical qualitative methods.

Statistician Assignments are as follows:
Microbiology WG:                   Chemistry WG:
Robert LaBudde          Joshua Hicks
Julie Moulin            Caryn Thompson
Dan Tholen              Roy MacArthur
Paul Wehling            Paul Wehling
It was noted that Max Fienberg is a member of the AOAC Committee on Statistics and will also be involved in this effort. Statistical input from Health Canada and Canadian Food Inspection Agency is welcomed.

The primary responsibilities for this group will be to come to consensus on the parameters and estimators for parameters with advice from the microbiology & chemistry working groups. The working groups will decide on size of study with advice from statisticians. This division of responsibility will help the process and ensure consensus on studies.

Even though there are issues specific to microbiological methods we would like same statistical models for microbiological and chemical qualitative methods.

ISO TC 68 revising standard 5725 standard dealing with quantitative chemical methods
ISO/IUPAC/AOAC – addition of qualitative binary methods. This standard will move forward soon.

Qualitative also called binary methods provide a yes/no; +/-; 0/1, presence/absence result, but method could have thresholds with yes/no output at each threshold.

Dan Tholen spoke about several validation study conditions and parameters that require statistical procedures for analyzing data (Appendix 1).

**Appendix 1: Validation study conditions and parameters that require statistical procedures**
(Microbiological and chemical methods)

List by Dan Tholen

1. Reference method exists.
2. No reference method exists.
3. Reference method exists, but not in the range of interest.
4. Single lab study.
5. Mult-lab study (reproducibility).
8. Can quantify concentration.
9. Can not quantify concentration.
10. Can confirm.
11. Can not confirm.
12. Paired with common enrichment or extraction procedures.
13. Unpaired without common enrichment.
14. Unpaired with independent samples (not allowed).

Parameters:

1. LOD
2. Accuracy (trueness and precision)
3. Compare with reference method
4. Agreement among labs