

Appendix F: Guidelines for *Standard Method Performance Requirements*

Contents

| | |
|--|----|
| Introduction to <i>Standard Method Performance Requirements</i> | 1 |
| Annex A: Format of a <i>Standard Method Performance Requirement</i> | 5 |
| Annex B: Classification of Methods | 11 |
| Annex C: Understanding the POD Model | 12 |
| Annex D: Definitions and Calculations of HorRat Values from Intralaboratory Data | 13 |
| Annex E: AOAC Method Accuracy Review | 15 |
| Annex F: Development and Use of In-House Reference Materials | 16 |

Introduction to *Standard Method Performance Requirements*

Standard Method Performance Requirements (SMPRs®) are a unique and novel concept for the analytical methods community. SMPRs are voluntary consensus standards, developed by stakeholders, that prescribe the minimum analytical performance requirements for classes of analytical methods. In the past, analytical methods were evaluated and the results compared to a “gold standard” method, or if a gold standard method did not exist, then reviewers would decide retrospectively if the analytical performance was acceptable. Frequently, method developers concentrated on the process of evaluating the performance parameters of a method, and rarely set acceptance criteria. However, as the *Eurachem Guide* points out: “. . . the judgment of method suitability for its intended use is equally important . . .” (1) to the evaluation process.

International Voluntary Consensus Standards

An SMPR is a form of an international, voluntary consensus standard. A standard is an agreed, repeatable way of doing something that is published as document that contains a technical specification or other precise criteria designed to be used consistently as a rule, guideline, or definition. SMPRs are a *consensus* standards developed by stakeholders in a very controlled process that ensures that users, research organizations, government departments, and consumers work together to create a standard that meets the demands of the analytical community and technology. SMPRs are also *voluntary* standards. AOAC cannot, and does not, impose the use of SMPRs. Users are free to use SMPRs as they see fit. AOAC is very careful to include participants from as many regions of the world as possible so that SMPRs are accepted as *international* standards.

Guidance for Standard Method Performance Requirements

Commonly known as the “SMPR Guidelines.” The first version of the SMPR Guidelines were drafted in 2010 in response to the increasing use and popularity of SMPRs as a vehicle to describe

the analytical requirements of a method. Several early “acceptance criteria” documents were prepared for publication in late 2009, but the format of the acceptance criteria documents diverged significantly from one another in basic format. AOAC realized that a guidance document was needed to promote uniformity.

An early version of the SMPR Guidelines were used for a project to define the analytical requirements for endocrine disruptors in potable water. The guidelines proved to be extremely useful in guiding the work of the experts and resulted in uniform SMPRs. Subsequent versions of the SMPR Guidelines were used in the AOAC Stakeholder Panel on Infant Formula and Adult Nutritionals (SPIFAN) project with very positive results. The SMPR Guidelines are published in the *Journal of AOAC INTERNATIONAL* and *Official Methods of Analysis*.

Users of the guidelines are advised that they are: (1) a *guidance* document, not a statute that users must conform to; and (2) a “living” document that is regularly updated, so users should check the AOAC website for the latest version before using these guidelines.

The SMPR Guidelines are intended to provide basic information for working groups assigned to prepare SMPRs. The guidelines consist of the standard format of an SMPR, followed by a series of informative tables and annexes.

SMPR Format

The general format for an SMPR is provided in *Annex A*.

Each SMPR is identified by a unique SMPR number consisting of the year followed by a sequential identification number (YYYY.XXX). An SMPR number is assigned when the standard is approved. By convention, the SMPR number indicates the year a standard is approved (as opposed to the year the standard is initiated). For example, SMPR 2010.003 indicates the third SMPR adopted in 2010.

The SMPR number is followed by a method name that must include the analyte(s), matrix(es), and analytical technique (unless the SMPR is truly intended to be independent of the analytical technology). The method name may also refer to a “common” name (e.g., “Kjeldahl” method).

The SMPR number and method name are followed by the name of the stakeholder panel or expert review panel that approved the SMPR, and the approval and effective dates.

Information about method requirements is itemized into nine categories: (1) intended use; (2) applicability; (3) analytical technique; (4) definitions; (5) method performance requirements; (6) system suitability; (7) reference materials; (8) validation guidance; and (9) maximum time-to-determination.

An SMPR for qualitative and/or identification methods may include up to three additional annexes: (1) inclusivity/selectivity panel; (2) exclusivity/cross-reactivity panel; and (3) environmental material panels. These annexes not required.

Informative tables.—The SMPR Guidelines contain seven informative tables that represent the distilled knowledge of many years of method evaluation, and are intended as guidance for SMPR working groups. The informative tables are not necessarily AOAC

policy. SMPR working groups are expected to apply their expertise in the development of SMPRs.

Table A1: Performance Requirements. Provides recommended performance parameters to be included into an SMPR. Table A1 is organized by five method classifications: (1) main component quantitative methods; (2) trace or contaminant quantitative methods; (3) main component qualitative methods; (4) trace or contaminant quantitative methods; and (5) identification methods. The table is designed to accommodate both microbiological and chemical methods. Alternate microbiological/chemical terms are provided for equivalent concepts.

Table A2: Recommended Definitions. Provides definitions for standard terms in the SMPR Guidelines. AOAC relies on *The International Vocabulary of Metrology Basic and General Concepts and Associated Terms* (VIM) and the International Organization for Standardization (ISO) for definition of terms not included in Table A2.

Table A3: Recommendations for Evaluation. Provides general guidance for evaluation of performance parameters. More detailed evaluation guidance can be found in *Appendix D, Guidelines for Collaborative Study Procedures to Validate Characteristics of a Method of Analysis* (2); *Appendix I, Guidelines for Validation of Biological Threat Agent Methods and/or Procedures* (3); *Appendix K, AOAC Guidelines for Single-Laboratory Validation of Chemical Methods for Dietary Supplements and Botanicals* (4); Codex Alimentarius Codex Procedure Manual (5); and ISO Standard 5725-1-1994 (6).

Table A4: Expected Precision (Repeatability) as a Function of Analyte Concentration. The precision of a method is the closeness of agreement between independent test results obtained under stipulated conditions. Precision is usually expressed in terms

of imprecision and computed as a relative standard deviation (RSD) of the test results. The imprecision of a method increases as the concentration of the analyte decreases. This table provides target RSDs for a range of analyte concentrations.

Table A5: Expected Recovery as a Function of Analyte Concentration. Recovery is defined as the ratio of the observed mean test result to the true value. The range of the acceptable mean recovery expands as the concentration of the analyte decreases. This table provides target mean recovery ranges for analyte concentrations from 1 ppb to 100%.

Table A6: Predicted Relative Standard Deviation of Reproducibility (PRSD_R). This table provides the calculated PRSD_R using the Horwitz formula:

$$PRSD_R = 2C^{-0.15}$$

where C is expressed as a mass fraction.

Table A7: POD and Number of Test Portions. This table provides the calculated probability of detection (POD) for given sample sizes and events (detections). A method developer can use this table to determine the number of analyses required to obtain a specific POD.

Informative annexes.—The SMPR Guidelines contain informative annexes on the topics of classification of methods, POD model, HorRat values, reference materials, and method accuracy and review. As with the informative tables, these annexes are intended to provide guidance and information to the working groups.

Initiation of an SMPR

See Figure 1 for a schematic flowchart diagram of the SMPR development process.

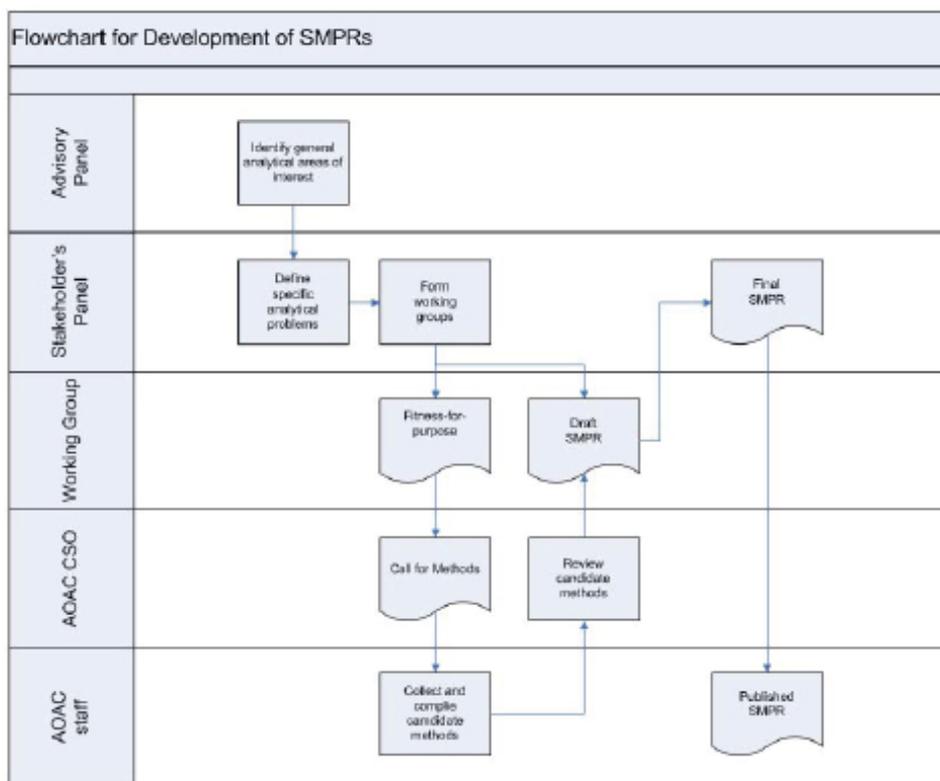


Figure 1. Schematic flowchart diagram of the SMPR development process.

Advisory panels.—Most commonly, an SMPR is created in response to an analytical need identified by an advisory panel. Advisory panels normally consist of sponsors and key stakeholders who have organized to address analytical problems. Usually, the advisory panel identifies general analytical problems, such as the need to update analytical methods for determination of nutrients in infant formula. An advisory panel, with the input of appropriate subject matter experts, also prioritizes the specific analytical problems within the general topic. This panel is critical in planning for the stakeholder panel meeting.

Stakeholder panels.—After an advisory panel has identified a general analytical problem, AOAC announces the standards development activity, identifies stakeholders, and organizes a stakeholder panel. Membership on a stakeholder panel is open to anyone materially affected by the proposed standard. AOAC recruits scientists to participate on stakeholder panels on the basis of their expertise with the analytical problem identified by the advisory panel. Experts are recruited from academia, government, nongovernmental organizations (such as ISO), industry, contract research organizations, method developers, and instrument/equipment manufacturers. AOAC employs a representative voting panel model to ensure balance with regards to stakeholder perspective, and to ensure that no particular stakeholder perspective dominates the proceedings of the stakeholder panel. All stakeholder candidates are reviewed by the AOAC Chief Scientific Officer (CSO) for relevant qualifications, and again by the Official Methods Board to ensure that the stakeholder panel is balanced and all stakeholders are fairly represented.

Stakeholder panels are extremely important as they serve several functions: (1) identify specific analytical topics within the general analytical problem described by the advisory panel; (2) form working groups to address the specific analytical topics; (3) identify additional subject matter experts needed for the working groups; (4) provide oversight of the SMPR development; and (5) formally adopt SMPRs originally drafted by working groups.

Working groups.—Working groups are formed by the stakeholder panel when a specific analytical topic has been identified. The primary purpose of a working group is to draft an SMPR. Working groups may also be formed to make general recommendations, such as developing a common definition to be used by multiple working groups. For example, SPIFAN formed a working group to create a definition for “infant formula” that could be shared and used by all of the SPIFAN working groups.

The process of drafting an SMPR usually requires several months, and several meetings and conference calls. An SMPR drafted by a working group is presented to a stakeholder panel. A stakeholder panel may revise, amend, or adopt a proposed SMPR on behalf of AOAC.

Fitness-for-Purpose Statement and Call for Methods

One of the first steps in organizing a project is creating a fitness-for-purpose statement. In AOAC, the fitness-for-purpose statement is a very general description of the methods needed. It is the responsibility of a working group chair to draft a fitness-for-purpose statement. A working group chair is also asked to prepare a presentation with background information about the analyte, matrix, and the nature of the analytical problem. A working group chair presents the background information and proposes a draft fitness-for-purpose statement to the presiding stakeholder panel. The stakeholder panel is asked to endorse the fitness-for-purpose statement.

The AOAC CSO prepares a call for methods based on the stakeholder panel-approved fitness-for-purpose statement. The call for methods is posted on the AOAC website and/or e-mailed to the AOAC membership and other known interested parties. AOAC staff collects and compiles candidate methods submitted in response to the call for methods. The CSO reviews and categorizes the methods.

Creating an SMPR

Starting the process of developing an SMPR can be a daunting challenge. In fact, drafting an SMPR should be a daunting challenge because the advisory panel has specifically identified an analytical problem that has yet to be resolved. Completing an SMPR can be a very rewarding experience because working group members will have worked with their colleagues through a tangle of problems and reached a consensus where before there were only questions.

It is advisable to have some representative candidate methods available for reference when a working group starts to develop an SMPR. These methods may have been submitted in response to the call for methods, or may be known to a working group member. In any case, whatever the origin of the method, candidate methods may assist working group members to determine reasonable performance requirements to be specified in the SMPR. The performance capabilities of existing analytical methodologies is a common question facing a working group.

Normally, a working chair and/or the AOAC CSO prepares a draft SMPR. A draft SMPR greatly facilitates the process and provides the working group with a structure from which to work.

Working group members are advised to first consider the “intended use” and “maximum time-to-determination” sections as this will greatly affect expectations for candidate methods. For example, methods intended to be used for surveillance probably need to be quick but do not require a great deal of precision, and false-positive results might be more tolerable. Whereas methods intended to be used for dispute resolution will require better accuracy, precision, and reproducibility, but time to determination is not as important.

Once a working group has agreed on the intended use of candidate methods, then it can begin to define the applicability of candidate methods. The applicability section of the SMPR is one of the most important, and sometimes most difficult, sections of the SMPR. The analyte(s) and matrix(es) must be explicitly identified. For chemical analytes, International Union of Pure and Applied Chemistry (IUPAC) nomenclature and/or Chemical Abstracts Service (CAS) registry numbers should be specified. Matrix(es) should be clearly identified including the form of the matrix such as raw, cooked, tablets, powders, etc. The nature of the matrix may affect the specific analyte. It may be advantageous to fully identify and describe the matrix before determining the specific analyte(s). It is not uncommon for working groups to revise the initial definition of the analyte(s) after the matrix(es) has been better defined.

Table 1. Example of method performance table for a single analyte

| | | |
|-----------------------------|-----------------|-----|
| Analytical range | 7.0–382.6 µg/mL | |
| Limit of quantitation (LOQ) | ≤7.0 µg/mL | |
| Repeatability (RSD,) | <10 µg/mL | ≤8% |
| | ≥10 µg/mL | ≤6% |

Table 2. Example of method performance table for multiple analytes

| | Analyte 1 | | Analyte 2 | | Analyte 3 | |
|-----------------------------|-------------|-----|---------------|-----|---------------|------|
| Analytical range | 10–20 µg/mL | | 100–200 µg/mL | | 200–500 µg/mL | |
| Limit of quantitation (LOQ) | ≤10 µg/mL | | ≤100 µg/mL | | ≤200 µg/mL | |
| Repeatability (RSD,) | <10 µg/mL | ≤8% | <10 µg/mL | ≤8% | <200 µg/mL | ≤10% |
| | ≥10 µg/mL | ≤6% | ≥10 µg/mL | ≤6% | ≥200 µg/mL | ≤8% |

For projects with multiple analytes, for example, vitamins A, D, E, and K in infant formula, it may be useful to organize a separate working group to fully describe the matrix(es) so that a common description of the matrix(es) can be applied to all of the analytes.

For single analyte SMPRs, it is most common to organize the method performance requirements into a table with 2–3 columns as illustrated in Table 1. For multiple analyte SMPRs, it is often convenient to present the requirements in an expanded table with analytes forming additional columns as illustrated in Table 2.

Once the intended use, analytical techniques, and method performance requirements have been determined, then a working group can proceed to consider the quality control parameters, such as the minimum validation requirements, system suitability procedures, and reference materials (if available). It is not uncommon that an appropriate reference material is not available. *Annex F* of the SMPR Guidelines provides comprehensive guidance for the development and use of in-house reference materials.

Most working groups are able to prepare a consensus SMPR in about 3 months.

Open Comment Period

Once a working group has produced a draft standard, AOAC opens a comment period for the standard. The comment period provides an opportunity for other stakeholders to state their perspective on the draft SMPR. All collected comments are reviewed by the AOAC CSO and the working group chair, and the comments are reconciled. If there are significant changes required to the draft standard as a result of the comments, the working group is convened to discuss and any unresolved issues will be presented for discussion at the stakeholder panel meeting.

Submission of Draft SMPRs to the Stakeholder Panel

Stakeholder panels meet several times a year at various locations. The working group chair (or designee) presents a draft SMPR to the stakeholder panel for review and discussion. A working group chair is expected to be able to explain the conclusions of the working group, discuss comments received, and to answer questions from the stakeholder panel. The members of the stakeholder panel may revise, amend, approve, or defer a decision on the proposed SMPR. A super majority of 2/3 or more of those voting is required to adopt an SMPR as an AOAC voluntary consensus standard.

Publication

Adopted SMPRs are prepared for publication by AOAC staff, and are published in the *Journal of AOAC INTERNATIONAL* and in the AOAC *Official Methods of Analysis*SM compendium. Often, the AOAC CSO and working group chair prepare a companion article to introduce an SMPR and describe the analytical issues considered and resolved by the SMPR. An SMPR is usually published within 6 months of adoption.

Conclusion

SMPRs are a unique and novel concept for the analytical methods community. SMPRs are voluntary, consensus standards developed by stakeholders that prescribe the minimum analytical performance requirements for classes of analytical methods. The SMPR Guidelines provide a structure for working groups to use as they develop an SMPR. The guidelines have been employed in several AOAC projects and have been proven to be very useful. The guidelines are not a statute that users must conform to; they are a “living” document that is regularly updated, so users should check the AOAC website for the latest version before using the guidelines.

References

- (1) Eurachem, *The Fitness for Purpose of Analytical Methods: A Laboratory Guide to Method Validation and Related Topics, Validation*, <http://www.eurachem.org/guides/pdf/valid.pdf>, posted December 1998, accessed March 2012
- (2) *Guidelines for Collaborative Study Procedures to Validate Characteristics of a Method of Analysis* (2012) *Official Methods of Analysis, Appendix D*, AOAC INTERNATIONAL, Gaithersburg, MD
- (3) *AOAC INTERNATIONAL Methods Committee Guidelines for Validation of Biological Threat Agent Methods and/or Procedures* (2012) *Official Methods of Analysis*, 19th Ed., *Appendix I, Calculation of CPOD and dCPOD Values from Qualitative Method Collaborative Study Data*, AOAC INTERNATIONAL, Gaithersburg, MD
- (4) *AOAC Guidelines for Single-Laboratory Validation of Chemical Methods for Dietary Supplements and Botanicals* (2012) *Official Methods of Analysis*, 19th Ed., *Appendix K*, AOAC INTERNATIONAL, Gaithersburg, MD
- (5) Codex Alimentarius Codex Procedure Manual
- (6) International Organization for Standardization, Geneva, Switzerland

ANNEX A
Format of a
Standard Method Performance Requirement

AOAC SMPR® YYYY.XXX
(YYYY = Year; XXX = sequential identification number)

Method Name: Must include the analyte(s), matrix(es), and analytical technique [unless the *Standard Method Performance Requirement* (SMPR®) is truly intended to be independent of the analytical technology]. The method name may refer to a “common” name (e.g., “Kjeldahl” method).

Approved By: Name of stakeholder panel or expert review panel

Final Version Date: Date

Effective Date: Date

1. Intended Use: Additional information about the method and conditions for use.

2. Applicability: List matrixes if more than one. Provide details on matrix such as specific species for biological analytes, or International Union of Pure and Applied Chemistry (IUPAC) nomenclature and Chemical Abstracts Service (CAS) registry number for chemical analytes. Specify the form of the matrix such as raw, cooked, tablets, powders, etc.

3. Analytical Technique: Provide a detailed description of the analytical technique if the SMPR is to apply to a specific analytical technique; or state that the SMPR applies to any method that meets the method performance requirements.

4. Definitions: List and define terms used in the performance parameter table (*see* Table A2 for list of standard terms).

5. Method Performance Requirements: List the performance parameters and acceptance criteria appropriate for each method/analyte/matrix. *See* Table A1 for appropriate performance requirements.

If more than one analyte/matrix, and if acceptance criteria differ for analyte/matrix combinations then organize a table listing each analyte/matrix combination and its minimum acceptance criteria for each performance criteria.

6. System Suitability Tests and/or Analytical Quality Control: Describe minimum system controls and QC procedures.

7. Reference Material(s): Identify the appropriate reference materials if they exist, or state that reference materials are not available. Refer to *Annex E (AOAC Method Accuracy Review)* for instructions on the use of reference materials in evaluations.

8. Validation Guidance: Recommendations for type of evaluation or validation program such as single-laboratory validation (SLV), *Official Methods of Analysis*SM (OMA), or *Performance Tested Methods*SM (PTM).

9. Maximum Time-to-Determination: Maximum allowable time to complete an analysis starting from the test portion preparation to final determination or measurement.

Annex I: Inclusivity/Selectivity Panel. Recommended for qualitative and identification method SMPRs.

Annex II: Exclusivity/Cross-Reactivity Panel. Recommended for qualitative and identification method SMPRs.

Annex III: Environmental Materials Panel. Recommended for qualitative and identification method SMPRs.

Table A1. Performance requirements

| Classifications of methods ^a | | | | |
|--|--|---|---|-------------------------------------|
| Quantitative method | | Qualitative method | | Identification method |
| Main component ^b | Trace or contaminant ^c | Main component ^b | Trace or contaminant ^c | |
| Parameter | | | | |
| Single-laboratory validation | | | | |
| Applicable range | Applicable range | Inclusivity/selectivity | Inclusivity/selectivity | Inclusivity/selectivity |
| Bias ^d | Bias ^d | Exclusivity/cross-reactivity | Exclusivity/cross-reactivity | Exclusivity/cross-reactivity |
| Precision | Precision | Environmental interference | Environmental interference | Environmental interference |
| Recovery | Recovery | Laboratory variance | Laboratory variance | |
| Limit of quantitation (LOQ) | LOQ | Probability of detection (POD) ^e | POD at AMDL ^f | Probability of identification (POI) |
| Reproducibility | | | | |
| RSD _r or target measurement uncertainty | RSD _r or target measurement uncertainty | POD (0) POD (c) Laboratory POD ^g | POD (0) POD (c) Laboratory POD ^g | POI (c) Laboratory POI |

^a See Annex B for additional information on classification of methods.

^b ≥100 g/kg.

^c <100 g/kg.

^d If a reference material is available.

^e At a critical level.

^f AMDL = Acceptable minimum detection level.

^g LPOD = CPOD.

Table A2. Recommended definitions

| | |
|---|--|
| Bias | Difference between the expectation of the test results and an accepted reference value. Bias is the total systematic error as contrasted to random error. There may be one or more systematic error components contributing to the bias. |
| Environmental interference | Ability of the assay to detect target organism in the presence of environmental substances and to be free of cross reaction from environmental substances. |
| Exclusivity | Strains or isolates or variants of the target agent(s) that the method must not detect. |
| Inclusivity | Strains or isolates or variants of the target agent(s) that the method can detect. |
| Laboratory probability of detection (POD) | Overall fractional response (mean POD = CPOD) for the method calculated from the pooled POD_j responses of the individual laboratories ($j = 1, 2, \dots, L$). ^a See Annex C. |
| Limit of quantitation (LOQ) | Minimum concentration or mass of analyte in a given matrix that can be reported as a quantitative result. |
| POD (0) | Probability of the method giving a (+) response when the sample is truly without analyte. |
| POD (c) | Probability of the method giving a (–) response when the sample is truly without analyte. |
| POD | Proportion of positive analytical outcomes for a qualitative method for a given matrix at a given analyte level or concentration. Consult Annex C for a full explanation. |
| Probability of identification (POI) | Expected or observed fraction of test portions at a given concentration that gives positive result when tested at a given concentration. Consult <i>Probability of Identification (POI): A Statistical Model for the Validation of Qualitative Botanical Identification Methods</i> . ^c |
| Precision | Closeness of agreement between independent test results obtained under stipulated conditions. The measure of precision is usually expressed in terms of imprecision and computed as a standard deviation of the test results. ^d |
| Recovery | Fraction or percentage of the analyte that is recovered when the test sample is analyzed using the entire method. There are two types of recovery: (1) Total recovery based on recovery of the native plus added analyte, and (2) marginal recovery based only on the added analyte (the native analyte is subtracted from both the numerator and denominator). ^e |
| Repeatability | Precision under repeatability conditions. |
| Repeatability conditions | Conditions where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time. |
| Reproducibility | Precision under reproducibility conditions. |
| Reproducibility conditions | Conditions where independent test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment. |
| Relative standard deviation (RSD) | $RSD = s_i \times 100/\bar{x}$ |
| Standard deviation (s_i) | $s_i = [\sum(x_i - \bar{x})^2/n]^{0.5}$ |

^a AOAC INTERNATIONAL Methods Committee Guidelines for Validation of Biological Threat Agent Methods and/or Procedures (Calculation of CPOD and dCPOD Values from Qualitative Method Collaborative Study Data), *J. AOAC Int.* **94**, 1359(2011) and *Official Methods of Analysis of AOAC INTERNATIONAL* (current edition), Appendix I.

^b *International Vocabulary of Metrology (VIM)—Basic and General Concepts and Associated Terms* (2008) JCGM 200:2008, Joint Committee for Guides in Metrology (JCGM), www.bipm.org.

^c LaBudde, R.A., & Harnly, J.M. (2012) *J. AOAC Int.* **95**, 273–285.

^d ISO 5725-1-1994.

^e *Official Methods of Analysis* (current edition) Appendix D (Guidelines for Collaborative Study Procedures to Validate Characteristics of a Method of Analysis), AOAC INTERNATIONAL, Rockville, MD, USA.

Table A3. Recommendations for evaluation

| | |
|---|--|
| Bias (if a reference material is available) | A minimum of five replicate analyses of a Certified Reference Material. ^a |
| Environmental interference | Analyze test portions containing a specified concentration of one environmental materials panel member. Materials may be pooled. Consult with AOAC statistician. |
| Exclusivity/cross-reactivity | Analyze one test portion containing a specified concentration of one exclusivity panel member. More replicates can be used. Consult with AOAC statistician. |
| Inclusivity/selectivity | Analyze one test portion containing a specified concentration of one inclusivity panel member. More replicates can be used. Consult with AOAC statistician. |
| Limit of quantitation (LOQ) | Estimate the LOQ = average (blank) + 10 × s ₀ (blank). Measure blank samples with analyte at the estimated LOQ. Calculate the mean average and standard deviation of the results. Guidance ^b : For ML ≥ 100 ppb (0.1 mg/kg): LOD = ML × 1/5. For ML < 100 ppb (0.1 mg/kg): LOD = ML × 2/5. |
| Measurement uncertainty | Use ISO 21748: <i>Guidance for the use of repeatability, reproducibility, and trueness estimates in measurement uncertainty estimation to analyze data collected for bias, repeatability, and intermediate precision to estimate measurement uncertainty.</i> |
| POD(0) | Use data from collaborative study. |
| POD (c) | |
| Repeatability | Prepare and homogenize three unknown samples at different concentrations to represent the full, claimed range of the method. Analyze each unknown sample by the candidate method seven times, beginning each analysis from weighing out the test portion through to final result with no additional replication (unless stated to do so in the method). All of the analyses for one unknown sample should be performed within as short a period of time as is allowed by the method. The second and third unknowns may be analyzed in another short time period. Repeat for each claimed matrix. |
| Probability of detection (POD) | Determine the desired POD at a critical concentration. Consult with Table A7 to determine the number of test portions required to demonstrate the desired POD. |
| Probability of identification (POI) | Consult <i>Probability of Identification (POI): A Statistical Model for the Validation of Qualitative Botanical Identification Methods</i> ^c . |
| Recovery | Determined from spiked blanks or samples with at least seven independent analyses per concentration level at a minimum of three concentration levels covering the analytical range. Independent means at least at different times. If no confirmed (natural) blank is available, the average inherent (naturally containing) level of the analyte should be determined on at least seven independent replicates. Marginal % recovery = $(C_f - C_u) \times 100 / C_A$ Total % recovery = $100(C_f) / (C_u + C_A)$ where C _f = concentration of fortified samples, C _u = concentration of unfortified samples, and C _A = concentration of analyte added to the test sample. ^d Usually total recovery is used unless the native analyte is present in amounts greater than about 10% of the amount added, in which case use the method of addition. ^e |
| Reproducibility (collaborative or interlaboratory study) | Quantitative methods: Recruit 10–12 collaborators; must have eight valid data sets; two blind duplicate replicates at five concentrations for each analyte/matrix combination to each collaborator. |
| | Qualitative methods: Recruit 12–15 collaborators; must have 10 valid data sets; six replicates at five concentrations for each analyte/matrix combination to each collaborator. |

^a *Guidance for Industry for Bioanalytical Method Validation* (May 2001) U.S. Department of Health and Human Services, U.S. Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Veterinary Medicine (CVM).

^b Codex Alimentarius Codex Procedure Manual.

^c LaBudde, R.A., & Harnly, J.M. (2012) *J. AOAC Int.* **95**, 273–285.

^d *Guidelines for Collaborative Study Procedures to Validate Characteristics of a Method of Analysis* (current edition) *Official Methods of Analysis, Appendix D*, AOAC INTERNATIONAL, Rockville, MD, USA.

^e *AOAC Guidelines for Single-Laboratory Validation of Chemical Methods for Dietary Supplements and Botanicals* (current edition) *Official Methods of Analysis, Appendix K*, AOAC INTERNATIONAL, Rockville, MD, USA.

Table A4. Expected precision (repeatability) as a function of analyte concentration^a

| Analyte, % | Mass fraction (C) | Unit | RSD _r , % |
|------------|-------------------|-----------------|----------------------|
| 100 | 1 | 100% | 1.3 |
| 10 | 10 ⁻¹ | 10% | 1.9 |
| 1 | 10 ⁻² | 1% | 2.7 |
| 0.1 | 10 ⁻³ | 0.1% | 3.7 |
| 0.01 | 10 ⁻⁴ | 100 ppm (mg/kg) | 5.3 |
| 0.001 | 10 ⁻⁵ | 10 ppm (mg/kg) | 7.3 |
| 0.0001 | 10 ⁻⁶ | 1 ppm (mg/kg) | 11 |
| 0.00001 | 10 ⁻⁷ | 100 ppb (µg/kg) | 15 |
| 0.000001 | 10 ⁻⁸ | 10 ppb (µg/kg) | 21 |
| 0.0000001 | 10 ⁻⁹ | 1 ppb (µg/kg) | 30 |

^a Table excerpted from AOAC Peer-Verified Methods Program, Manual on Policies and Procedures (1998) AOAC INTERNATIONAL, Rockville, MD, USA.

The precision of a method is the closeness of agreement between independent test results obtained under stipulated conditions. Precision is usually expressed in terms of imprecision and computed as a relative standard deviation of the test results. The imprecision of a method increases as the concentration of the analyte decreases. This table provides targets RSDs for a range of analyte concentrations.

Table A5. Expected recovery as a function of analyte concentration^a

| Analyte, % | Mass fraction (C) | Unit | Mean recovery, % |
|------------|-------------------|-----------------|------------------|
| 100 | 1 | 100% | 98–102 |
| 10 | 10 ⁻¹ | 10% | |
| 1 | 10 ⁻² | 1% | 97–103 |
| 0.1 | 10 ⁻³ | 0.1% | 95–105 |
| 0.01 | 10 ⁻⁴ | 100 ppm (mg/kg) | 90–107 |
| 0.001 | 10 ⁻⁵ | 10 ppm (mg/kg) | 80–110 |
| 0.0001 | 10 ⁻⁶ | 1 ppm (mg/kg) | |
| 0.00001 | 10 ⁻⁷ | 100 ppb (µg/kg) | |
| 0.000001 | 10 ⁻⁸ | 10 ppb (µg/kg) | 60–115 |
| 0.0000001 | 10 ⁻⁹ | 1 ppb (µg/kg) | 40–120 |

^a Table excerpted from AOAC Peer-Verified Methods Program, Manual on Policies and Procedures (1998) AOAC INTERNATIONAL, Rockville, MD, USA.

Recovery is defined as the ratio of the observed mean test result to the true value. The range of the acceptable mean recovery expands as the concentration of the analyte decreases. This table provides target mean recovery ranges for analyte concentrations from 100% to 1 ppb.

Table A6. Predicted relative standard deviation of reproducibility (PRSD_R)^a

| Analyte, % | Mass fraction (C) | Unit | RSD _r , % |
|------------|-------------------|-----------------|----------------------|
| 100 | 1 | 100% | 2 |
| 10 | 10 ⁻¹ | 10% | 3 |
| 1 | 10 ⁻² | 1% | 4 |
| 0.1 | 10 ⁻³ | 0.1% | 6 |
| 0.01 | 10 ⁻⁴ | 100 ppm (mg/kg) | 8 |
| 0.001 | 10 ⁻⁵ | 10 ppm (mg/kg) | 11 |
| 0.0001 | 10 ⁻⁶ | 1 ppm (mg/kg) | 16 |
| 0.00001 | 10 ⁻⁷ | 100 ppb (µg/kg) | 22 |
| 0.000001 | 10 ⁻⁸ | 10 ppb (µg/kg) | 32 |
| 0.0000001 | 10 ⁻⁹ | 1 ppb (µg/kg) | 45 |

^a Table excerpted from Definitions and Calculations of HorRat Values from Intralaboratory Data, HorRat for SLV.doc, 2004-01-18, AOAC INTERNATIONAL, Rockville, MD, USA.

Predicted relative standard deviation or reproducibility = PRSD_R.
Reproducibility relative standard deviation calculated from the Horwitz formula:

$$\text{PRSD}_R = 2C^{-0.15}$$

where C is expressed as a mass fraction.

This table provides the calculated PRSD_R for a range of concentrations. See Annex D for additional information.

Table A7. POD and number of test portions^{a,b}

| Sample size required for proportion | | | | | | | |
|-------------------------------------|--|------------------------|---------------------------|--|---|---|------------------------------------|
| Assume | 1. Binary outcome (occur/not occur). 2. Constant probability rho of event occurring. 3. Independent trials (e.g., simple random sample). 4. Fixed number of trials (N) | | | | | | |
| Inference | 95% Confidence interval lies entirely at or above specified minimum rho | | | | | | |
| Desired | Sample size N needed | | | | | | |
| Minimum probability rho, % | Sample size (N) | Minimum No. events (x) | Maximum No. nonevents (y) | 1-Sided lower confidence limit on rho ^c , % | Expected lower confidence limit on rho, % | Expected upper confidence limit on rho, % | Effective AOQL ^d rho, % |
| 50 | 3 | 3 | 0 | 52.6 | 43.8 | 100.0 | 71.9 |
| 50 | 10 | 8 | 2 | 54.1 | 49.0 | 94.3 | 71.7 |
| 50 | 20 | 14 | 6 | 51.6 | 48.1 | 85.5 | 66.8 |
| 50 | 40 | 26 | 14 | 52.0 | 49.5 | 77.9 | 63.7 |
| 50 | 80 | 48 | 32 | 50.8 | 49.0 | 70.0 | 59.5 |
| 55 | 4 | 4 | 0 | 59.7 | 51.0 | 100.0 | 75.5 |
| 55 | 10 | 9 | 1 | 65.2 | 59.6 | 100.0 | 79.8 |
| 55 | 20 | 15 | 5 | 56.8 | 53.1 | 88.8 | 71.0 |
| 55 | 40 | 28 | 12 | 57.1 | 54.6 | 81.9 | 68.2 |
| 55 | 80 | 52 | 28 | 55.9 | 54.1 | 74.5 | 64.3 |
| 60 | 5 | 5 | 0 | 64.9 | 56.5 | 100.0 | 78.3 |
| 60 | 10 | 9 | 1 | 65.2 | 59.6 | 100.0 | 79.8 |
| 60 | 20 | 16 | 4 | 62.2 | 58.4 | 91.9 | 75.2 |
| 60 | 40 | 30 | 10 | 62.4 | 59.8 | 85.8 | 72.8 |
| 60 | 80 | 56 | 24 | 61.0 | 59.2 | 78.9 | 69.1 |
| 65 | 6 | 6 | 0 | 68.9 | 61.0 | 100.0 | 80.5 |
| 65 | 10 | 9 | 1 | 65.2 | 59.6 | 100.0 | 79.8 |
| 65 | 20 | 17 | 3 | 67.8 | 64.0 | 94.8 | 79.4 |
| 65 | 40 | 31 | 9 | 65.1 | 62.5 | 87.7 | 75.1 |
| 65 | 80 | 59 | 21 | 65.0 | 63.2 | 82.1 | 72.7 |
| 70 | 7 | 7 | 0 | 72.1 | 64.6 | 100.0 | 82.3 |
| 70 | 10 | 10 | 0 | 78.7 | 72.2 | 100.0 | 86.1 |
| 70 | 20 | 18 | 2 | 73.8 | 69.9 | 97.2 | 83.6 |
| 70 | 40 | 33 | 7 | 70.7 | 68.0 | 91.3 | 79.7 |
| 70 | 80 | 63 | 17 | 70.4 | 68.6 | 86.3 | 77.4 |
| 75 | 9 | 9 | 0 | 76.9 | 70.1 | 100.0 | 85.0 |
| 75 | 10 | 10 | 0 | 78.7 | 72.2 | 100.0 | 86.1 |
| 75 | 20 | 19 | 1 | 80.4 | 76.4 | 100.0 | 88.2 |
| 75 | 40 | 35 | 5 | 76.5 | 73.9 | 94.5 | 84.2 |
| 75 | 80 | 67 | 13 | 75.9 | 74.2 | 90.3 | 82.2 |
| 80 | 11 | 11 | 0 | 80.3 | 74.1 | 100.0 | 87.1 |
| 80 | 20 | 19 | 1 | 80.4 | 76.4 | 100.0 | 88.2 |
| 80 | 40 | 37 | 3 | 82.7 | 80.1 | 97.4 | 88.8 |
| 80 | 80 | 70 | 10 | 80.2 | 78.5 | 93.1 | 85.8 |
| 85 | 20 | 20 | 0 | 88.1 | 83.9 | 100.0 | 91.9 |
| 85 | 40 | 38 | 2 | 86.0 | 83.5 | 98.6 | 91.1 |
| 85 | 80 | 74 | 6 | 86.1 | 84.6 | 96.5 | 90.6 |
| 90 | 40 | 40 | 0 | 93.7 | 91.2 | 100.0 | 95.6 |
| 90 | 60 | 58 | 2 | 90.4 | 88.6 | 99.1 | 93.9 |
| 90 | 80 | 77 | 3 | 91.0 | 89.5 | 98.7 | 94.1 |
| 95 | 60 | 60 | 0 | 95.7 | 94.0 | 100.0 | 97.0 |
| 95 | 80 | 80 | 0 | 96.7 | 95.4 | 100.0 | 97.7 |
| 95 | 90 | 89 | 1 | 95.2 | 94.0 | 100.0 | 97.0 |
| 95 | 96 | 95 | 1 | 95.5 | 94.3 | 100.0 | 97.2 |
| 98 | 130 | 130 | 0 | 98.0 | 97.1 | 100.0 | 98.6 |
| 98 | 240 | 239 | 1 | 98.2 | 97.7 | 100.0 | 98.8 |
| 99 | 280 | 280 | 0 | 99.0 | 98.6 | 100.0 | 99.3 |
| 99 | 480 | 479 | 1 | 99.1 | 98.8 | 100.0 | 99.4 |

^a Table excerpted from Technical Report TR308, *Sampling plans to verify the proportion of an event exceeds or falls below a specified value*, LaBudde, R. (June 4, 2010) (not published). The table was produced as part of an informative report for the Working Group for Validation of Identity Methods for Botanical Raw Materials commissioned by the AOAC INTERNATIONAL Presidential Task Force on Dietary Supplements. The project was funded by the Office of Dietary Supplements, National Institutes of Health.

^b Copyright 2010 by Least Cost Formulations, Ltd. All rights reserved.

^c Based on modified Wilson score 1-sided confidence interval.

^d AOQL = Average outgoing quality level.

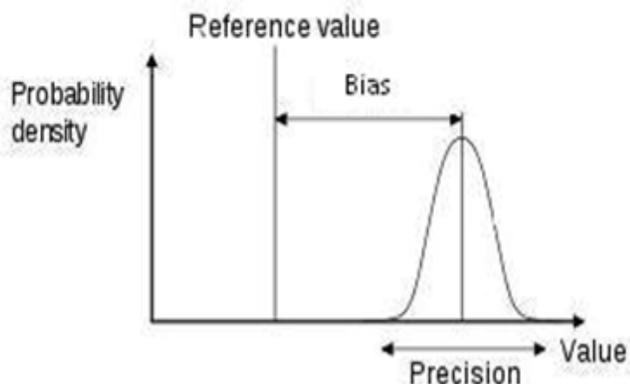


Figure A1. Relationship between precision versus bias (trueness). Trueness is reported as bias. Bias is defined as the difference between the test results and an accepted reference value.

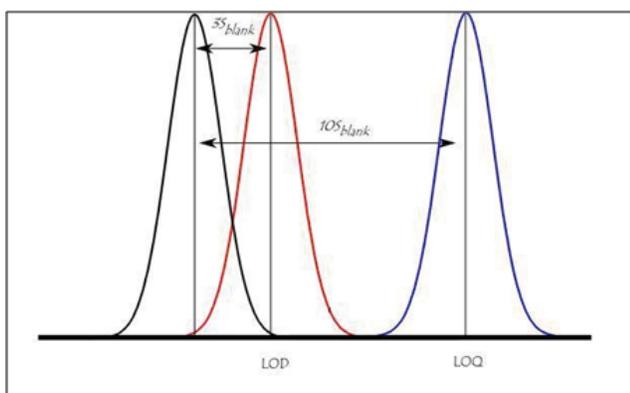


Figure A2. Relationship between LOD and LOQ. LOD is defined as the lowest quantity of a substance that can be distinguished from the absence of that substance (a blank value) within a stated confidence limit. LOQ is the level above which quantitative results may be obtained with a stated degree of confidence.

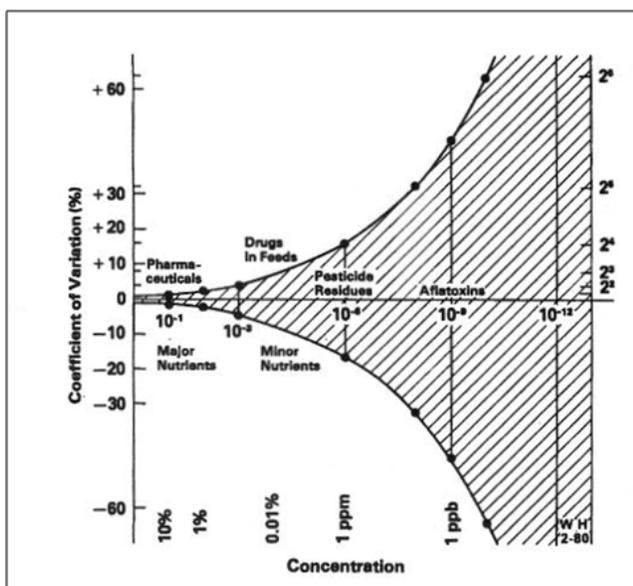


Figure A3. Horwitz Curve, illustrating the exponential increase in the coefficient of variation as the concentration of the analyte decreases [J. AOAC Int. 89, 1095(2006)].

ANNEX B Classification of Methods

The following guidance may be used to determine which performance parameters in Table A1 apply to different classifications of methods. AOAC INTERNATIONAL does not recognize the term “semiquantitative” as a method classification. Methods that have been self-identified as semiquantitative will be classified into one of the following five types:

Type I: Quantitative Methods

Characteristics: Generates a continuous number as a result.

Recommendation: Use performance requirements specified for quantitative method (main or trace component). Use recovery range and maximum precision variation in Tables A4 and A5.

In some cases and for some purposes, methods with less accuracy and precision than recommended in Tables A4 and A5 may be acceptable. Method developers should consult with the appropriate method committee to determine if the recommendations in Tables A4 and A5 do or do not apply to their method.

Type II: Methods that Report Ranges

Characteristics: Generates a “range” indicator such as 0, low, moderate, and high.

Recommendation: Use performance requirements specified for qualitative methods (main component). Specify a range of POD for each range “range” indicator.

Type III: Methods with Cutoff Values

Characteristics: Method may generate a continuous number as an interim result (such as a CT value for a PCR method), which is not reported but converted to a qualitative result (presence/ absence) with the use of a cutoff value.

Recommendation: Use performance requirements specified for qualitative methods.

Type IV: Qualitative Methods

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

Recommendation: Use performance requirements specified for qualitative methods.

Type V: Identification Methods

Characteristics: Method of analysis whose purpose is to determine the identity of an analyte.

Recommendation: Use performance requirements specified for identification methods.

ANNEX C Understanding the POD Model

Excerpted from AOAC INTERNATIONAL Methods Committee Guidelines for Validation of Biological Threat Agent Methods and/or Procedures, J. AOAC Int. 94, 1359(2011) and Official Methods of Analysis of AOAC INTERNATIONAL (current edition), Appendix I.

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, presence/absence, or +/-.

The single parameter of interest is the POD, which is defined as the probability at a given concentration of obtaining 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, a high probability of obtaining a positive response is desired. The goal of method validation is to characterize how method response transitions from low concentration/low response to high concentration/high response.

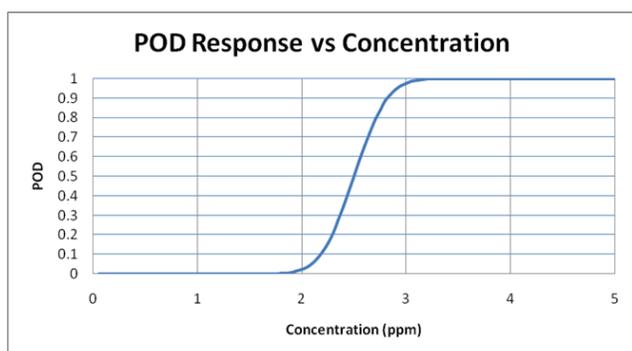


Figure C1. Theoretical POD curve for a qualitative detection method.

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 C1*).

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 incorporates 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 incorporates 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, these terms are obsolete under this model (*see Table C1*).

The terms “sensitivity,” “specificity,” “false positive,” and “false negative” are obsolete under the POD model (*see Figure C2*).

Table C1. Terminology

| Traditional terminology | Concept | POD equivalent | Comment |
|---|--|---------------------------|---|
| False positive | Probability of the method giving a (+) response when the sample is truly without analyte | POD(0) POD at conc = 0 | POD curve value at conc = 0; “Y-intercept” of the POD curve |
| Specificity | Probability of the method giving a (-) response when the sample is truly without analyte | 1-POD(0) | Distance along the POD axis from POD = 1 to the POD curve value |
| False negative (at a given concentration) | Probability of a (-) response at a given concentration | 1-POD(c) | Distance from the POD curve to the POD = 1 “top axis” in the vertical direction |
| Sensitivity (at a given concentration) | Probability of a (+) response at a given concentration | POD(c) | Value of the POD curve at any given concentration |
| True negative | A sample that contains no analyte | C = 0 | Point on concentration axis where c = 0 |
| True positive | A sample that contains analyte at some positive concentration | C > 0 | Range of concentration where c > 0 |

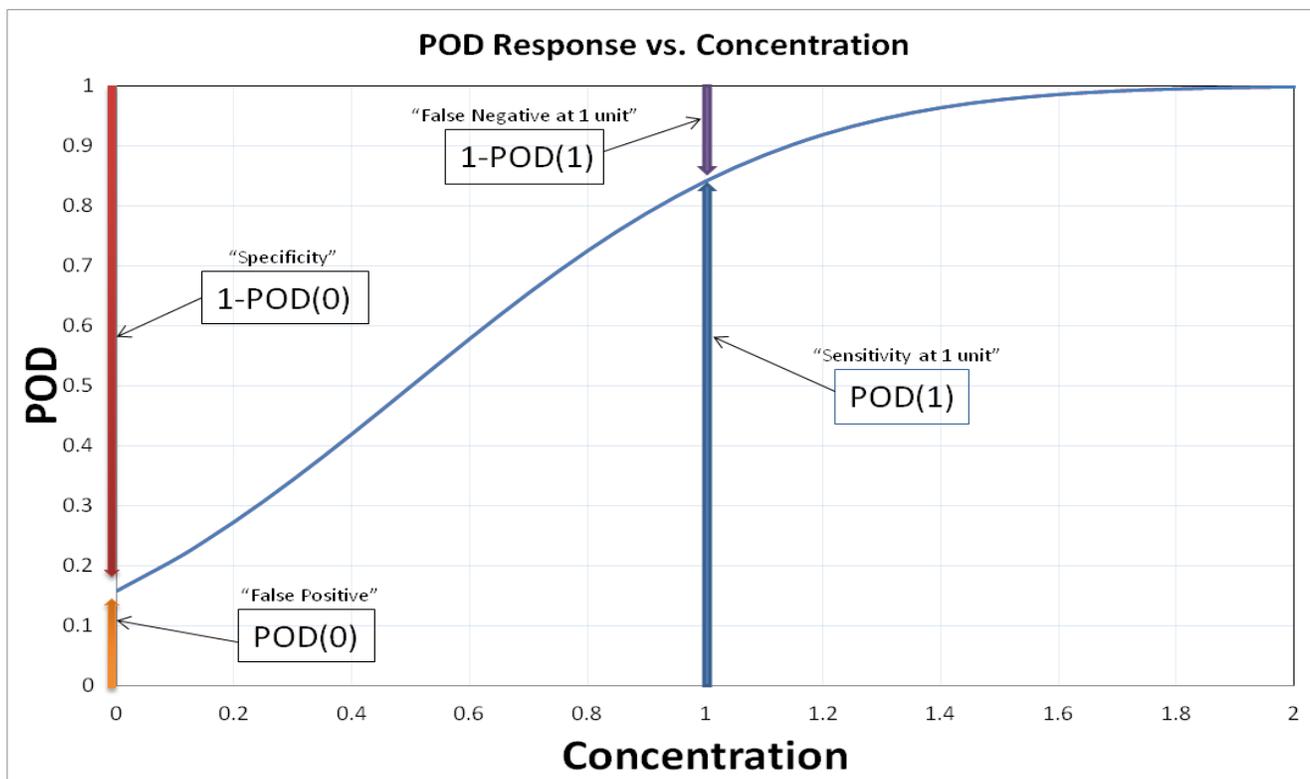


Figure C2. Comparison of POD model terminology to other obsolete terms.

ANNEX D
Definitions and Calculations
of HorRat Values from Intralaboratory Data

Excerpted from Definitions and Calculations of HorRat Values from Intralaboratory Data, AOAC INTERNATIONAL, HorRat for SLV.doc, 2004-01-18.

1. Definitions

1.1 Replicate Data

Data developed under common conditions in the same laboratory: simultaneous performance, or, if necessary to obtain sufficient values, same series, same analyst, same day. Such data provides “repeatability statistical parameters.”

1.2 Pooled Data

Replicate data developed in the same laboratory under different conditions but considered sufficiently similar that, for the purpose of statistical analysis, they may be considered together. These may include different runs, different instruments, different analysts, and different days.

1.3 Average

\bar{x} = Sum of the individual values, x_i , divided by the number of individual values, n .

$$\bar{x} = (\sum x_i)/n$$

1.4 Standard Deviation

$$s_i = [\sum(x_i - \bar{x})^2/n]^{0.5}$$

1.5 Relative Standard Deviation

$$RSD = s_i \times 100/\bar{x}$$

1.5.1 Repeatability Relative Standard Deviation [RSD(r) or RSD_r]

The relative standard deviation calculated from within-laboratory data.

1.5.2 Reproducibility Relative Standard Deviation [RSD(R) or RSD_R]

The relative standard deviation calculated from among-laboratory data.

Table D1. Predicted relative standard deviations

| Concentration (C) | Mass fraction (C) | PRSD _R , % |
|-------------------|-------------------|-----------------------|
| 100% | 1.0 | 2 |
| 1% | 0.01 | 4 |
| 0.01% | 0.0001 | 8 |
| 1 ppm | 0.000001 | 16 |
| 10 ppb | 0.00000001 | 32 |
| 1 ppb | 0.0000000001 | 45 |

1.6 Mass Fraction

Concentration, C, expressed as a decimal fraction. For calculating and reporting statistical parameters, data may be expressed in any convenient units (e.g., %, ppm, ppb, mg/g, µg/g; µg/kg; µg/L, µg/µL, etc.). For reporting HorRat values, data must be reported as a mass fraction where the units of the numerator and denominator are the same: e.g., for 100% (pure materials), the mass fraction C = 1.00; for 1 µg/g (ppm), C = 0.000001 = (E-6). See Table D1 for other examples.

1.7 Predicted Relative Standard Deviation [PRSD(R) or PRSD_r]

The reproducibility relative standard deviation calculated from the Horwitz formula:

$$\text{PRSD(R)} = 2C^{-0.15}$$

where C is expressed as a mass fraction. See Table D1.

In spreadsheet notation: PRSD(R) = 2 * C ^(-0.15).

1.8 HorRat Value

The ratio of the reproducibility relative standard deviation calculated from the data to the PRSD(R) calculated from the Horwitz formula:

$$\text{HorRat} = \text{RSD(R)}/\text{PRSD(R)}$$

To differentiate the usual HorRat value calculated from reproducibility data from the HorRat value calculated from repeatability data, attach an R for the former and an r for the latter. But note that the denominator always uses the PRSD(R) calculated from reproducibility data because this parameter is more predictable than the parameter calculated from repeatability data:

$$\text{HorRat(R)} = \text{RSD}_R/\text{PRSD(R)}$$

$$\text{HorRat(r)} = \text{RSD}_r/\text{PRSD(R)}$$

Some expected, predicted relative standard deviations are given in Table D1.

2 Acceptable HorRat Values

2.1 For Interlaboratory Studies

HorRat(R): The original data developed from interlaboratory (among-laboratory) studies assigned a HorRat value of 1.0 with limits of acceptability of 0.5 to 2.0. The corresponding within-laboratory relative standard deviations were found to be typically 1/2 to 2/3 the among-laboratory relative standard deviations.

Table D2. Predicted relative standard deviations

| Concentration (C) | PRSD _R , % | PRSD _r , % |
|-------------------|-----------------------|-----------------------|
| 100% | 2 | 1 |
| 1% | 4 | 2 |
| 0.01% | 8 | 4 |
| 1 ppm | 16 | 8 |
| 10 ppb | 32 | 16 |
| 1 ppb | 45 | 22 |

2.1.1 Limitations

HorRat values do not apply to method-defined (empirical) analytes (moisture, ash, fiber, carbohydrates by difference, etc.), physical properties or physical methods (pH, viscosity, drained weight, etc.), and ill-defined analytes (polymers, products of enzyme reactions).

2.2 For Intralaboratory Studies

2.2.1 Repeatability

Within-laboratory acceptable predicted target values for repeatability are given in Table D2 at 1/2 of PRSD(R), which represents the best case.

2.2.2 HorRat(r)

Based on experience and for the purpose of exploring the extrapolation of HorRat values to SLV studies, take as the minimum acceptability 1/2 of the lower limit ($0.5 \times 0.5 \approx 0.3$) and as the maximum acceptability 2/3 of the upper limit ($0.67 \times 2.0 \approx 1.3$).

Calculate HorRat(r) from the SLV data:

$$\text{HorRat(r)} = \text{RSD(r)}/\text{PRSD(R)}$$

Acceptable HorRat(r) values are 0.3–1.3. Values at the extremes must be interpreted with caution. With a series of low values, check for unreported averaging or prior knowledge of the analyte content; with a series of high values, check for method deficiencies such as unrestricted times, temperatures, masses, volumes, and concentrations; unrecognized impurities (detergent residues on glassware, peroxides in ether); incomplete extractions and transfers and uncontrolled parameters in specific instrumental techniques.

2.3 Other Limitations and Extrapolations

The HorRat value is a very rough but useful summary of the precision in analytical chemistry. It overestimates the precision at the extremes, predicting more variability than observed at the high end of the scale (C > ca 0.1; i.e., >10%) and at the low end of the scale (C < E-8; i.e., 10 ng/g; 10 ppb).

ANNEX E

AOAC Method Accuracy Review

Accuracy of Method Based on Reference Material

Reference material (RM) used.—The use of RMs should be seen as integral to the process of method development, validation, and performance evaluation. RMs are not the only component of a quality system, but correct use of RMs is essential to appropriate quality management. RMs with or without assigned quantity values can be used for measurement precision control, whereas only RMs with assigned quantity values can be used for calibration or measurement trueness control. Method development and validation for matrices within the scope of the method is done to characterize attributes such as recovery, selectivity, “trueness” (accuracy, bias), precision (repeatability and reproducibility), uncertainty estimation, ruggedness, LOQ or LOD, and dynamic range. RMs should be chosen that are fit-for-purpose. When certified reference materials (CRMs) are available with matrices that match the method scope, much of the work involved in method development has already been completed, and that work is documented through the certificate. RMs with analyte values in the range of test samples, as well as “blank” matrix RMs, with values below or near detection limits, are needed.

Availability of RM.—Consideration needs to be given to the future availability of the chosen RM. Well-documented methods that cannot be verified in the future due to lack of material may lose credibility or be seen as inferior.

Fit to method scope.—Natural matrix CRMs provide the greatest assurance that the method is capable of producing accurate results for that matrix. When selecting an RM to perform a method validation, analysts should consider the method to material fit. An example of a good fit would be a method for specified organic molecules in infant formula and using an infant formula or powder milk RM. A poor fit would be a method for specified organic molecules in infant formula and using a sediment material.

Stability.—Providing a stable RM can be challenging where analytes are biologically active, easily oxidized, or interactive with other components of the matrix. CRM producers provide assurance of material stability, as well as homogeneity. CRMs are accompanied by a certificate that includes the following key criteria:

- (1) Assigned values with measurement uncertainty and metrological traceability
- (2) Homogeneity
- (3) Stability, with the expiration date for the certificate
- (4) Storage requirements
- (5) Information on intended use
- (6) Identity of matrix

For some RMs, such as botanical RMs, the source and/or authenticity can be a very important piece of information that should be included with the certificate. Even under ideal storage conditions, many analytes have some rate of change. Recertification may be done by the supplier, and a certificate reissued with a different expiration date and with certain analyte data updated or removed.

Definition of CRM.—Refer to the AOAC TDRM document for definitions from ISO Guide 30, Amd. 1 (2008), <http://www.aoc.org/divisions/References.pdf>.

Information on source of RM is available.—It is the responsibility of the material producer to provide reliable authentication of the RM and make a clear statement in the accompanying documentation. This should be an as detailed listing as possible, including handling of ingredients, identification of plant materials as completely as feasible (species, type, subtype, growing region), etc. This is comparable to other required information on an RM for judging its suitability for a specific application purpose (e.g., containing how much of the targeted analyte, stabilized by adding acid—therefore not suited for certain parameters/procedures, etc.).

Separate RM used for calibration and validation.—A single RM cannot be used for both calibration and validation of results in the same measurement procedure.

Blank RM used where appropriate.—Blank matrix RMs are useful for ensuring performance at or near the detection limits. These are particularly useful for routine quality control in methods measuring, for instance, trace levels of allergens, mycotoxins, or drug residues.

Storage requirements were maintained.—Method developers should maintain good documentation showing that the RM producer’s recommended storage conditions were followed.

Cost.—The cost of ongoing method checks should be considered. Daily use of CRMs can be cost prohibitive. Monthly or quarterly analysis of these materials may be an option.

Concentration of analyte fits intended method.—Concentration of the analyte of interest is appropriate for *Standard Method Performance Requirements* (SMPRs®).

Uncertainty available.—Every measurement result has an uncertainty associated with it, and the individual contributions toward the combined uncertainty arise from multiple sources. Achieving the target measurement uncertainty set by the customer for his/her problem of interest is often one of the criteria used in selecting a method for a given application. Estimation of measurement uncertainty can be accomplished by different approaches, but the use of RMs greatly facilitates this part of a method validation.

Demonstration of Method Accuracy when No Reference Material Is Available

If an RM is not available, how is accuracy demonstrated?

There are many analytes for which a CRM with a suitable matrix is not available. This leaves the analyst with few options. For some methods, there may be proficiency testing programs that include a matrix of interest for the analyte. Proficiency testing allows an analyst to compare results with results from other laboratories, which may or may not be using similar methods. Spiking is another technique that may be used. When alternative methods are available, results may be compared between the different methods. These alternatives do not provide the same level of assurance that is gained through the use of a CRM.

Spike recovery.—In the absence of an available CRM, one technique that is sometimes used for assessing performance is the spiking of a matrix RM with a known quantity of the analyte. When this method is used, it cannot be assumed that the analyte is bound in the same way as it would be in a natural matrix. Nevertheless, a certified blank RM would be the preferred choice for constructing a spiked material.

When preparing reference solutions, the pure standards must be completely soluble in the solvent. For insoluble materials in a liquid suspension or for powdered forms of dry materials, validation is required to demonstrate that the analyte is homogeneously

The document, *AOAC Method Accuracy Review*, was prepared by the AOAC Technical Division on Reference Materials (TDRM) and approved by the AOAC Official Methods Board in June 2012.

distributed and that the response of the detection system to the analyte is not affected by the matrix or preparation technique. When a matrix material is selected for spiking, it should be reasonably characterized to determine that it is sufficiently representative of the matrix of interest. Spiked samples must be carried through all steps of the method. Many analytes are bound in a natural matrix and whether the spiked analyte will behave the same as the analyte in a natural matrix is unknown.

Other.—Use of a substitute RM involves the replacement of the CRM with an alternative matrix RM matching the matrix of interest as close as possible based on technical knowledge.

ANNEX F Development and Use of In-House Reference Materials

The use of reference materials is a vital part of any analytical quality assurance program. However, you may have questions about their creation and use. The purpose of this document is to help answer many of these questions.

- What is a reference material?
- Why use reference materials?
- What certified reference materials (CRMs) are currently available?
- Why use an in-house reference material?
- How do I create an in-house reference material?
- How do I use the data from an in-house reference material?

What Is a Reference Material?

The International Organization for Standardization (ISO) defines a reference material as a “material or substance one or more of whose property values are sufficiently homogeneous and well established to be used for the calibration of an apparatus, the assessment of a measurement method, or for assigning values to materials” (1). In plain English, natural-matrix reference materials, such as those you might prepare for use in-house, can be used to validate an analytical method or for quality assurance while you’re using your method to analyze your samples. (Natural-matrix materials are not generally used as calibrants because of the increased uncertainty that this would add to an analysis.) The assigned values for the target analytes of an in-house reference material can be used to establish the precision of your analytical method and, if used in conjunction with a CRM, to establish the accuracy of your method.

ISO defines a CRM as a “reference material, accompanied by a certificate, one or more of whose property values are certified by a procedure which establishes traceability to an accurate realization of the unit in which the property values are expressed, and for which each certified value is accompanied by an uncertainty at a stated level of confidence” (1).

Why Use Reference Materials?

CRMs can be used across the entire scope of an analytical method and can provide traceability of results to the International System of Units (SI). During method development, CRMs can be used to optimize your method. During method validation, they can be used to ensure that your method is capable of producing the “right” answer, and to determine how close your result is to that

answer. During routine use, they can be used to determine within-day and between-day repeatability, and so demonstrate that your method is in control and is producing accurate results every time it is used.

Natural-matrix reference materials should mimic the real samples that will be analyzed with a method. They should behave just as your samples would during a procedure, so if you obtain accurate and precise values for your reference material, you should obtain accurate and precise values for your samples as well.

What Certified Reference Materials Are Currently Available?

CRMs are available from a number of sources, including (but not limited to):

- American Association of Cereal Chemists (AACC)
- American Oil Chemists Society (AOCS)
- International Atomic Energy Agency (IAEA)
- Institute for Reference Materials and Measurements (IRMM)
- LGC Promochem
- National Institute of Standards and Technology (NIST)
- National Research Council Canada (NRC Canada)
- UK Food Analysis Proficiency Assessment Program (FAPAS)

A number of websites provide general overviews and catalogs of producers’ and distributors’ reference materials:

<http://www.aocs.org/tech/crm/>
<http://www.comar.bam.de>
<http://www.erm-crm.org>
<http://www.iaea.org/oregrammeslaqcs>
<http://www.aaccnet.org/checksample>
<http://www.irmm-ire.be/mrm.html>
<http://www.lgcpromochem.com>
<http://www.naweb.iaea.org/nahu/nmrm/>
<http://www.nist.gov/srm>
<http://www.fapas.com/index.cfm>
<http://www.virm.net>

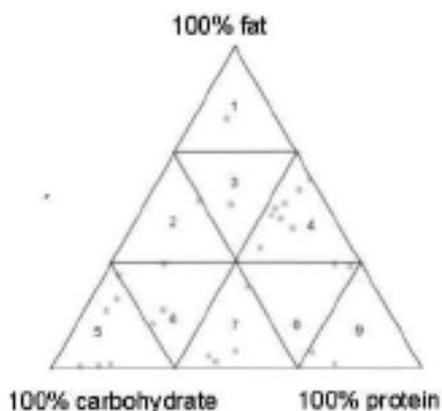
Because new reference materials are produced regularly, it is important to check these websites to determine what is currently available.

Why Use an In-House Reference Material?

There are many benefits to the use of a CRM. CRMs have been prepared to be homogeneous and, if stored under the proper conditions, stable. You are provided with a certified value as well as the statistical data for the concentration of your analyte; this is about as close as you can come to knowing the true value of the concentration of the analyte. The material has been tested by experienced analysts in leading laboratories, so you have the security of knowing that your method is generating values similar to those generated in other competent laboratories. The CRMs from the sources mentioned above are nationally and/or internationally recognized, so when you obtain acceptable results for a CRM using your analytical method, you give credibility to your methodology and traceability to your results.

But there are some drawbacks associated with CRMs. Unfortunately, many analyte/matrix combinations are not currently available. When testing food products for nutrient content, for example, a laboratory can be asked to analyze anything that might be found in a kitchen or grocery store. Reference materials that represent all of the types of foods that need to be tested are not available, and most CRMs are certified for a limited number of analytes. It is important to match the reference material matrix to your sample matrix. (Food examples dominate the discussion

Excerpted from *Development and Use of In-House Reference Materials*, Rev. 2, 2009. Copyright 2005 by the AOAC Technical Division on Reference Materials (TDRM).



below, but the same processes apply to the development of in-house reference materials in other areas of analytical chemistry.)

To demonstrate the applicability of an analytical method to a wide variety of food matrices, AOAC INTERNATIONAL's Task Force on Methods for Nutrition Labeling developed a triangle partitioned into sectors in which foods are placed based on their protein, fat, and carbohydrate content (2, 3). Since ash does not have a great impact on the performance of an analytical method for organic-material foods, and water can be added or removed, it can be assumed that the behavior of an analytical method is determined to large extent by the relative proportions of these proximates. AOAC INTERNATIONAL anticipated that one or two foods in a given sector would be representative of other foods in that sector and therefore would be useful for method assessment. Similarly, one or two reference materials in a given sector (or near each other in adjacent sectors) should be useful for quality assurance for analyses involving the other foods in the sector. The positions of many of the food-matrix CRMs from the sources listed above are shown in the triangle and are provided in the list.

These food-matrix reference materials are spread through all sectors of the triangle, thereby making it likely that you can find an appropriate CRM to match to your samples. Ultimately, however, the routine use of a CRM can be cost prohibitive, and is not really the purpose of CRMs. For example, in order to use NIST's Standard Reference Material (SRM) 2387 Peanut Butter for all mandatory nutrition labeling analyses, you could buy one sales unit (three jars, each containing 170 g material) for \$649 (2009 price). If you charge your customer about \$1000 for analysis of all mandatory nutrients in a test material, the control material would account for more than 60% of your fees. Therefore, many laboratories have found it more cost-effective to create in-house reference materials for routine quality control and characterize them in conjunction with the analysis of a CRM (4). You can prepare larger quantities of a reference material by preparing it in-house, and you have more flexibility in the types of matrices you can use. There are not many limitations on what can be purchased.

How Do I Create an In-House Reference Material?

There are basically three steps to preparing an in-house reference material: selection (including consideration of homogeneity and stability), preparation, and characterization. Additional guidance through these steps can be provided from the AOAC Technical Division on Reference Materials (TDRM), as well as in ISO Guides 34 (5) and 35 (6).

| Sector | RM No. | Matrix |
|--------|------------|----------------------------------|
| | NIST 1563 | Coconut oil |
| 1 | NIST 3274 | Fatty acids in botanical oils |
| 1 | NIST 3276 | Carrot extract in oil |
| 1 | LGC 7104 | Sterilized cream |
| 2 | NIST 2384 | Baking chocolate |
| 3 | NIST 2387 | Peanut butter |
| 4 | NIST 1546 | Meat homogenate |
| 4 | LGC 7106 | Processed cheese |
| 4 | LGC 7000 | Beef/pork meat |
| 4 | LGC 7150 | Processed meat |
| 4 | LGC 7151 | Processed meat |
| 4 | LGC 7152 | Processed meat |
| 4 | SMRD 2000 | Fresh meat |
| 4 | LGC 7101 | Mackerel paste |
| 4 | LGC QC1001 | Meat paste 1 |
| 4 | LGC QC1004 | Fish paste 1 |
| 5 | BCR-382 | Wheat flour |
| 5 | BCR-381 | Rye flour |
| 5 | LGC 7103 | Sweet digestive biscuit |
| 5 | LGC 7107 | Madeira cake |
| 5 | LGC QC1002 | Flour 1 |
| 6 | NIST 1544 | Fatty acids |
| 6 | NIST 1548a | Typical diet |
| 6 | NIST 1849 | Infant/adult nutritional formula |
| 6 | LGC 7105 | Rice pudding |
| 7 | LGC 7001 | Pork meat |
| 7 | NIST 1566b | Oyster tissue |
| 7 | NIST 1570a | Spinach leaves |
| 7 | NIST 2385 | Spinach |
| 8 | NIST 1946 | Lake trout |
| 8 | LGC 7176 | Canned pet food |
| 9 | NIST 1974a | Mussel tissue |
| 9 | NIST 3244 | Protein powder |

References

- (1) JCGM 200:2008, *International vocabulary of metrology—Basic and general concepts and associated terms (VIM)*, International Bureau of Weights and Measures (www.bipm.org)
- (2) Wolf, W.R., & Andrews, K.W. (1995) *Fresenius' J. Anal. Chem.* **352**, 73–76
- (3) Wolf, W.R. (1993) *Methods of Analysis for Nutrition Labeling*, D.R. Sullivan & D.E. Carpenter (Eds), AOAC INTERNATIONAL, Gaithersburg, MD
- (4) European Reference Materials (2005) *Comparison of a Measurement Result with the Certified Value*, Application Note 1
- (5) *ISO Guide 34 General Requirements for the Competence of Reference Material Producers* (2009) 2nd, International Organization for Standardization, Geneva, Switzerland

- (6) *Guide 35 Certification of Reference Materials—General and Statistical Principles* (2006) International Organization for Standardization, Geneva, Switzerland

For more information about TDRM, visit <http://aoac.org/divisions/tdrm>.