

**Standard Format and Guidance for
AOAC Standard Method Performance Requirement (SMPR) Documents
(Version 12.1; 31-Jan-11)**

AOAC SMPR 2010.XXX

Method Name: 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 refer to a “common” name (e.g. “Kjeldahl” method).

Approved by: Stakeholder Panel or Expert Review Panel name

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 IUPAC¹ nomenclature and CAS² 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 2 for list of standard terms)
- 5. Method Performance Requirements:** List the performance parameters and acceptance criteria appropriate for each method/analyte/matrix. See table 1 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 Appendix D *Development and Use of In-House Reference Materials* if a reference material cannot be identified
- 8. Validation Guidance:** Qualify the method into one of the method classifications in Table 1. which provides general recommendations regarding validation of the method. Validation study protocols should be provided as an annex to the SMPR. Identify which studies should be carried out at the method developers site (single laboratory validation); and the independent laboratory (for the *Performance Tested Methods* program).
- 9. Maximum Time-To-Determination:** Maximum allowable time to complete an analysis starting from the test portion preparation to final determination or measurement.

¹ International Union of Pure and Applied Chemistry (IUPAC)

² Chemical Abstracts Service

Annex I: Validation Study Protocols

(Required for all SMPRs)

Introduction: Provide basic information about the type of methods that are applicable to the validation outline; and general information about the levels of validation required for the SMPR method.

Level 1: Method Developer Validation Study Protocol. Describe the studies that a method developer must carry out to demonstrate that a candidate method meets the performance parameters specified in an SMPR. Refer to Table 3 for the recommended studies.

Level 2: Independent Laboratory Validation Study Protocol. Describe the studies that an independent laboratory must carry out to demonstrate that a candidate method meets the performance parameters specified in an SMPR. Refer to Table 3 for the recommended studies

Level 3: Collaborative Study Protocol. Describe the collaborative study that a Method Developer / Study Director must carry out to demonstrate that a candidate method meets the performance parameters specified in an SMPR.

ALL VALIDATION PROTOCOLS MUST INLCUDE THE FOLLOWING STATEMENTS:

Method developers seeking to submit data to the *Official Methods* or *Performance Tested Methods* programs must prepare, submit, and obtain approval of their individualized validation study protocol.

An approved protocol is valid for one year from its approval date. Validation protocols with approval dates more than one year must be submitted for re-approval.

Annex II: Inclusivity/Selectivity Panel

(Required for qualitative and identification method SMPRs)

Annex III: Exclusivity/Specificity Panel

(Required for qualitative and identification method SMPRs)

Annex IV: Environmental Materials Panel

(Required for qualitative and identification method SMPRs)

Table 1: Performance Requirements

		Classifications of Methods ⁹				
		Quantitative Method (main component ¹)	Quantitative Method (trace or contaminant ²)	Qualitative Method (main component ¹)	Qualitative Method (trace or contaminant ²)	Identification Method
Parameters	Single laboratory validation	Reference Method Comparison ³ Applicable Range Bias ⁴ Precision Recovery	Reference Method Comparison ³ Applicable Range Bias ⁴ Precision Recovery Limit of Detection (LOD) Limit of Quantitation (LOQ)	Reference Method Comparison ³ Inclusivity/Selectivity Exclusivity/Specificity Environmental Interference Laboratory Variance Bias ⁴ Probability of Detection ⁶	Reference Method Comparison ³ Inclusivity/Selectivity Exclusivity/Specificity Environmental Interference Laboratory Variance Bias ⁴ Probability of Detection (POD) at the AMDL ⁸	Reference Method Comparison ³ Inclusivity /Selectivity Exclusivity/Specificity Precision Environmental Interference Bias ⁴
	Independent	TBD ⁵	TBD ⁵	TBD ⁵	Probability of Detection (POD) at the AMDL ⁸	Bias ⁴
	Collaborative Study	Reproducibility	Reproducibility	POD (0) POD (c) Laboratory Probability of Detection ⁸	POD (0) POD (c) Laboratory Probability of Detection ⁸	POD (0) POD (c) Laboratory Probability of Detection ⁷

Notes:

1. ≥100 g/kg
2. <100 g/kg
3. If a reference method is available
4. If a reference material is available.
5. To be determined by the Topic Advisor/ General Referee
6. At a critical level.
7. If a reference method is available. LPOD = CPOD.
8. Acceptable Minimum Detection Level (AMD L)
9. See Appendix B for additional information on classification of methods.

DRAFT

Table 2: Recommended Definitions

Bias	The 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	The overall fractional response (mean POD = CPOD) for the method calculated from the pooled POD_j responses of the individual laboratories ($j = 1, 2, \dots, L$). ¹ See Appendix A.
Limit of Detection (LOD)	The minimum concentration or mass of analyte that can be detected in a given matrix with no greater than 5% false positive risk and 5% false negative risk.
Limit of Quantitation (LOQ)	The minimum concentration or mass of analyte in a given matrix that can be reported as a quantitative result
Maximum Level (ML)	Maximum or minimum or normative level. The expected level or concentration of the analyte.
POD(0)	The probability of the method giving a (+) response when the sample is truly without analyte
POD (c)	The probability of the method giving a (-) response when the sample is truly without analyte
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. Consult Appendix A for a definition.

¹ AOAC INTERNATIONAL Methods Committee Guidelines for Validation of Biological Threat Agent Methods and/or Procedures. Appendix E. Calculation of CPOD and dCPOD Values from Qualitative Method Collaborative Study Data.

DRAFT

Precision (repeatability)	The 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. ²
Recovery	Fraction or percentage of the analyte that is recovered when the test sample is analyzed using the entire method is the recovery. 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). ³
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.

² ISO 5725-1-1994

³ Appendix D: *Guidelines for Collaborative Study Procedures To Validate Characteristics of a Method of Analysis*, Official Methods of Analysis Manual; AOAC ; 2002.

DRAFT

Table 3: Recommendations for Evaluation

Accuracy	The closeness of agreement between a test result and the accepted reference value. The term accuracy, when applied to a set of test results, involves a combination of random components and a common systematic error or bias component.
Bias (if a reference material is available)	The 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	Analyze test portions containing a specified concentration of one environmental materials panel member. Materials may be pooled. Consult with AOAC statistician.
Exclusivity	Analyze one test portion containing a specified concentration of one exclusivity panel member. More replicates can be used. Consult with AOAC statistician.
Inclusivity	Analyze one test portion containing a specified concentration of one inclusivity panel member. More replicates can be used. Consult with AOAC statistician.
Limit of Detection (LOD)	Measure 10 blank samples. Calculate the mean average and standard deviation of the results $LOD^4 = \text{average (blank)} + z s_0 \text{ (blank)}$; where $s_0 = \text{standard deviation}$ $z = 2 \times \text{Gaussian critical value} = 2 \times 1.645 = 3.3$
Limit of Quantitation (LOQ)	Estimate the LOQ = average (blank) + 10 * s0 (blank); Measure blank samples with analyte at the estimated LOQ. Calculate the mean average and standard deviation of the results Guidance ⁵ : For ML ≥ 100 ppm (0.1 mg/kg): LOD = ML * 1/5 For ML < 100 ppm (0.1 mg/kg): LOD = ML * 2/5
POD(0)	Use data from collaborative study
POD (c)	
Repeatability	Prepare and homogenize 3 unknown samples at different concentrations to represent the full, claimed range of the method. Analyze each unknown sample by the candidate method 7 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

⁴ ISO 16140:2003

⁵ Codex Alimentarius Codex Procedure Manual

DRAFT

	method). All of the analyses for one unknown sample should be performed within as short 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 Probability of Detection at a critical concentration. Consult with table 7 to determine the number of test portions required to demonstrate the desired Probability of Detection.	
Recovery	<p>Determined from spiked blanks or samples with at least 7 independent analyses per concentration level at a minimum of 3 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 7 independent replicates.</p> <p>Marginal % recovery = $(C_f - C_u) \times 100 / C_A$ Total % recovery = $100(C_f)/(C_u + C_A)$</p> <p>Where C_f = concentration of fortified samples, C_u = concentration of unfortified samples; and C_A = concentration of analyte added to the test sample.⁶</p> <p>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.⁷</p>	
Relative standard deviation (RSD)	$RSD = s_i \times 100 / \bar{x}$	
Reproducibility (Collaborative study)	Quantitative methods	Recruit 10-12 collaborators Must have 8 valid data sets 2 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 6 replicates at five concentrations for each analyte/matrix combination to each collaborator
Standard deviation (s_i)	$s_i = [\sum(x_i - \bar{x})^2 / n]^{0.5}$	

⁶ Appendix D: Guidelines for Collaborative Study Procedures To Validate Characteristics of a Method of Analysis, Official Methods of Analysis Manual; AOAC ; 2002.

⁷ AOAC Guidelines for Single Laboratory Validation of Chemical Methods for Dietary Supplements and Botanicals

DRAFT

Table 4: Expected Precision (repeatability) as a Function of Analyte Concentration

Analyte %	Analyte Ratio	Unit	RSD%
100	1	100%	1.3
10	10 ⁻¹	10%	1.9
1	10 ⁻²	1%	2.7
0.01	10 ⁻³	0.1%	3.7
0.001	10 ⁻⁴	100 ppm (mg/kg)	5.3
0.0001	10 ⁻⁵	10 ppm (mg/kg)	7.3
0.00001	10 ⁻⁶	1 ppm (mg/kg)	11
0.000001	10 ⁻⁷	100 ppb (µg/kg)	15
0.0000001	10 ⁻⁸	10 ppb (µg/kg)	21
0.00000001	10 ⁻⁹	1 ppb (µg/kg)	30

Table excerpted from: AOAC Peer-Verified Methods Program, Manual on policies and procedures, Arlington, Va., USA (1998).

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 5: Expected Recovery as a Function of Analyte Concentration

Analyte %	Analyte Ratio	Unit	Mean Recovery (%)
100	1	100%	98-102
10	10 ⁻¹	10%	98-102
1	10 ⁻²	1%	97-103
0.01	10 ⁻³	0.1%	95-105
0.001	10 ⁻⁴	100 ppm	90-107
0.0001	10 ⁻⁵	10 ppm	80-110
0.00001	10 ⁻⁶	1 ppm	80-110
0.000001	10 ⁻⁷	100 ppb	80-110
0.0000001	10 ⁻⁸	10 ppb	60-115
0.00000001	10 ⁻⁹	1 ppb	40-120

Table excerpted from: AOAC Peer-Verified Methods Program, Manual on policies and procedures, Arlington, Va., USA (1998).

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.

DRAFT

Table 6: Predicted Relative Standard Deviation of Reproducibility (PRSD_R)

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	0.000000001	45

Table excerpted from: Definitions And Calculations Of Horrat Values From Intralaboratory Data, AOAC INTERNATIONAL, Horrat for SLV.doc, 2004-01-18.

Predicted relative standard deviation = PRSD(R) or PRSD_R. The reproducibility relative standard deviation calculated from the Horwitz formula:

$$PRSD(R) = 2C^{-0.15}; \text{ where C is expressed as a mass fraction.}$$

This table provides the calculated PRSD(R) for a range of concentrations. See Appendix C for additional information.

Table 7: POD and and Number of Test Portions

6/4/10 Copyright 2010 by Least Cost Formulations, Ltd., All rights reserved.

Sample Size Required for Proportion HIGH Version: 0.0

- 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.

- NOTES: 1. Based on modified Wilson score 1-sided confidence interval.
 2. AOQL = Average Outgoing Quality Level

Minimum Probability rho	Sample Size N	Minimum Number Events x	Maximum Number Non-events y	1-sided Lower Confidence Limit on rho	Expected Lower Confidence Limit on rho	Expected Upper Confidence Limit on rho	Effective AOQL 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%	77.8%	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	58	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%	83.8%
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%	90.8%
90%	40	40	0	93.7%	91.2%	100.0%	95.8%
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	98.0%	98.6%	100.0%	99.3%
99%	480	479	1	99.1%	98.8%	100.0%	99.4%

Table excerpted from: Technical Report TR308; Sampling plans to verify the proportion of an event exceeds of falls below a specified value; LaBudde, Robert; June 4, 2010. Not published. This 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 Institute for Health.

Figure 1: Relationship between Precision versus Trueness (Bias)

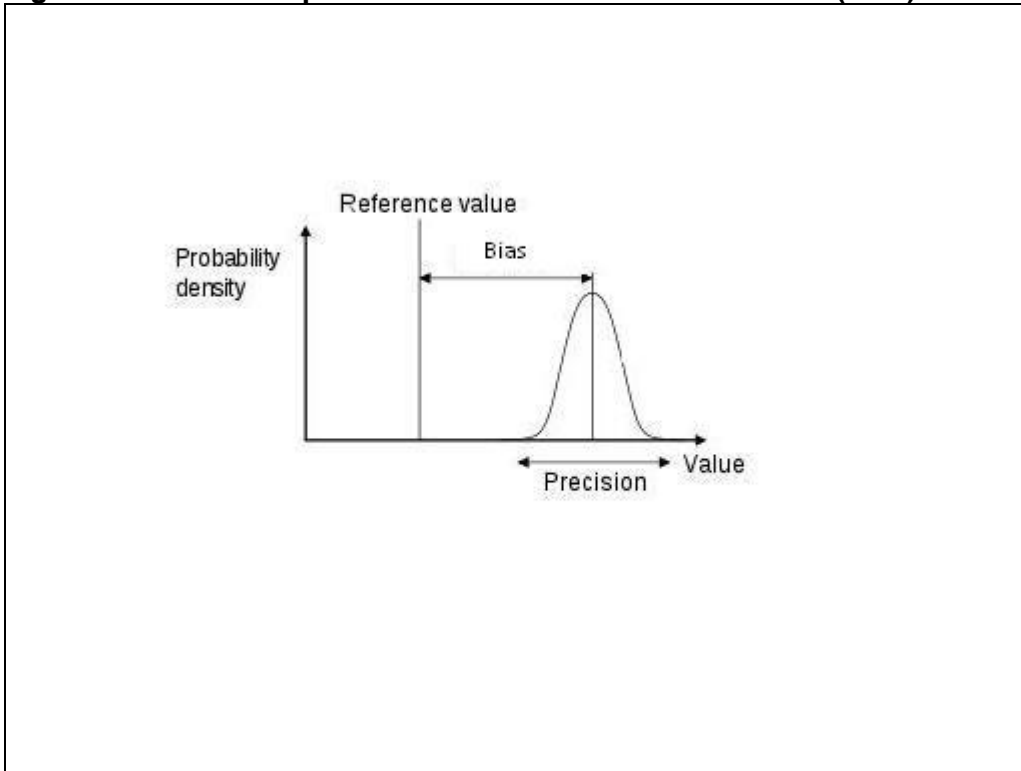


Figure 1 illustrates the relationship between trueness (bias) and precision. Trueness is reported as bias. Bias is defined as the difference between the test results and an accepted reference value.

Figure 2: Relationship between LOD and LOQ

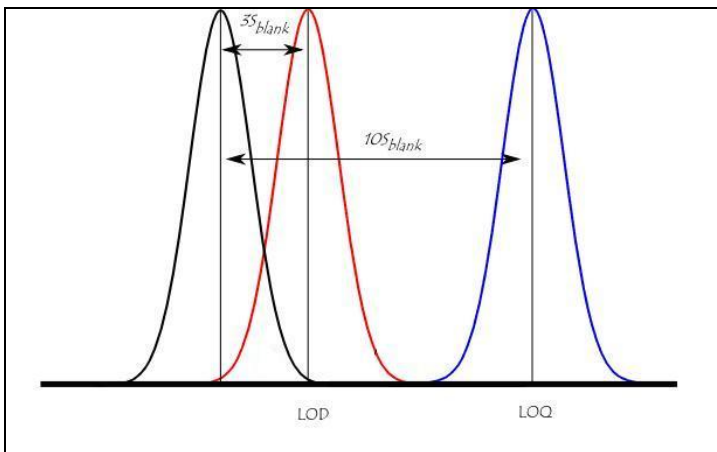
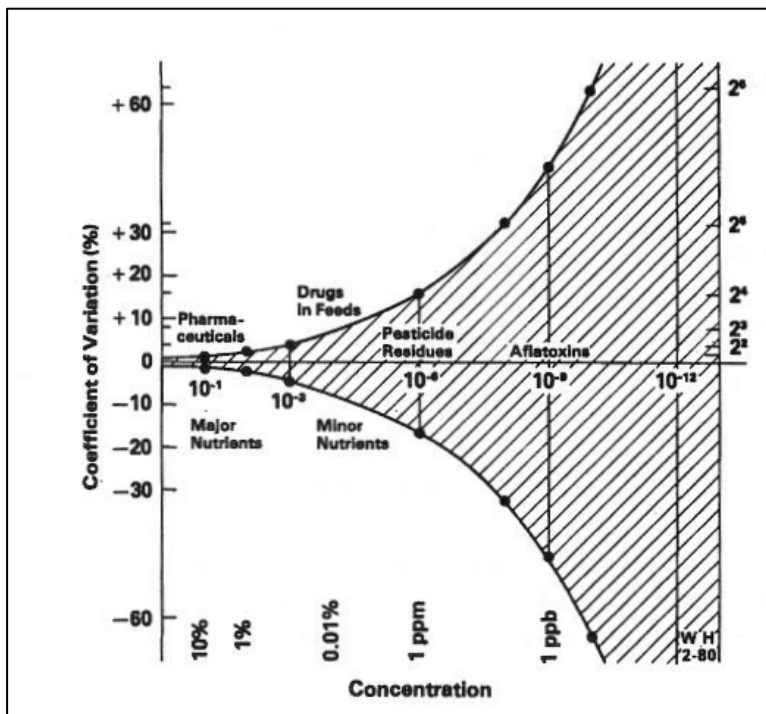


Figure 2 illustrates the relationship between the Limit of Detection (LOD) and Limit of Quantitation (LOQ). LOD is defined as the lowest quantity of a substance that can be distinguished from the absence of that substance (a *blank value*) within a stated confidence limit. LOQ is the level above which quantitative results may be obtained with a stated degree of confidence.

Figure 3: Horwitz Curve



J. AOAC Int. **89**, 1095 (2006)

Figure 3 illustrates the exponential increase in the coefficient of variation as the concentration of the analyte decreases.

Appendix A. Understanding the POD Model

Excerpted from: AOAC INTERNATIONAL Methods Committee Guidelines for Validation of Biological Threat Agent Methods and/or Procedures, 2010, AOAC INTERNATIONAL.

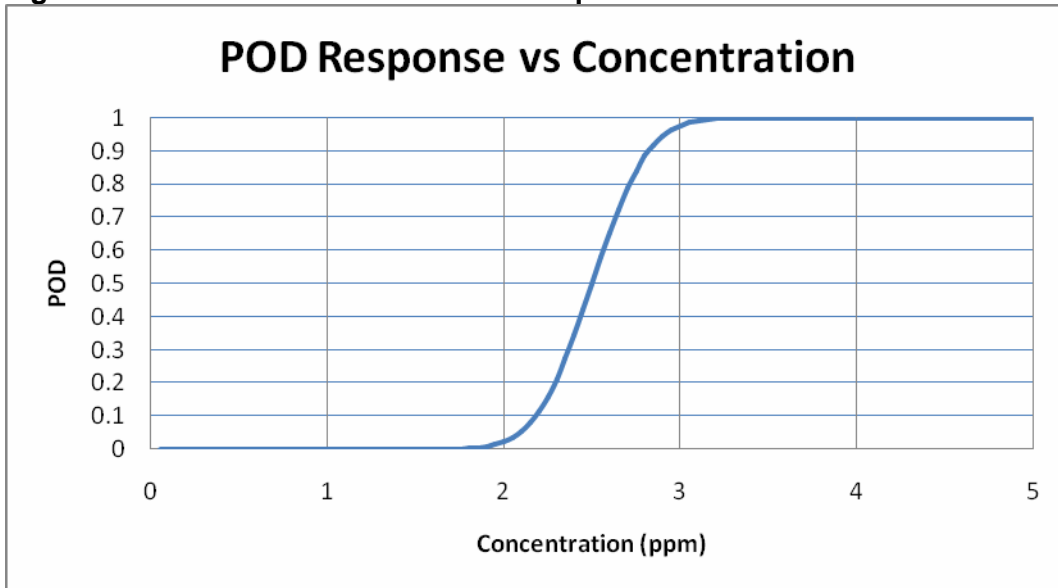
The Probability of Detection (POD) model is a way of characterizing the performance of a qualitative (binary) method. A binary qualitative method is one that gives a result as one of two possible outcomes, either positive or negative, or presence/absence or +/-.

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 A1).

Figure A1: 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.

DRAFT

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.

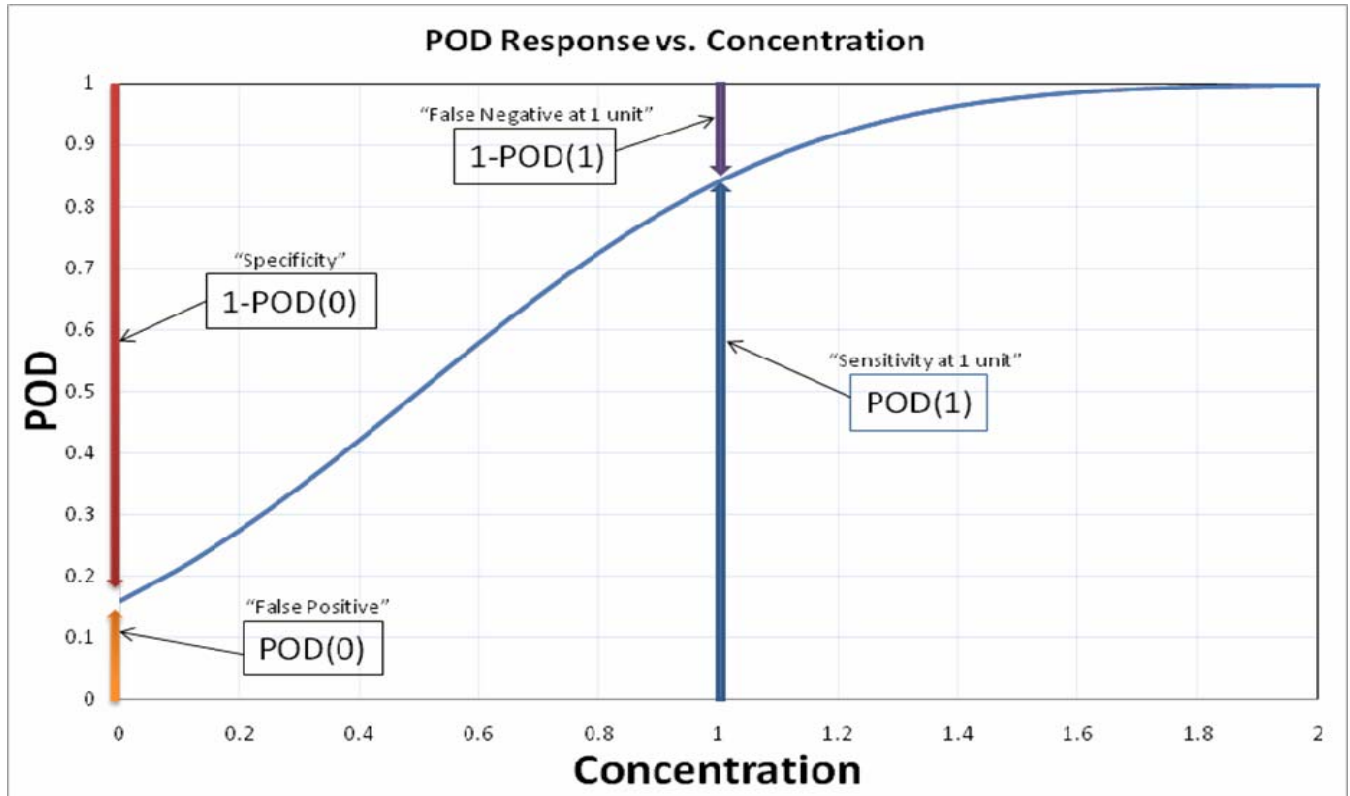
Similarly, False Negative has sometimes been defined as the probability of a negative response when the sample is truly positive (concentration >0). In the POD curve, this would always be specific to a given sample concentration, but would be represented as the distance from the POD curve to the POD = 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, these terms are obsolete under this model.

Table A1: Terminology

Traditional Terminology	Concept	POD Equivalent	Comment
False Positive	The probability of the method giving a (+) response when the sample is truly without analyte	POD(0) POD at Conc = 0	The POD curve value at Conc = 0 – The “Y-intercept” of the POD curve
Specificity	The probability of the method giving a (-) response when the sample is truly without analyte	1-POD(0)	The distance along the POD axis from POD=1 to the POD curve value.
False Negative (at a given concentration)	The probability of a (-) response at a given concentration	1-POD(c)	The distance from the POD curve to the POD = 1 “top axis” in the vertical direction
Sensitivity (at a given concentration)	The probability of a (+) response at a given concentration	POD(c)	The 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 A2: Comparison of POD Model Terminology to Other Obsolete Terms



The terms "Sensitivity", "Specificity", "False Positive" and "False Negative" are obsolete under the POD model.

Appendix B: Classification of Methods

The following guidance may be used to determine which performance parameters in Table 1 apply to different classifications of methods.

AOAC INTERNATIONAL does not recognize the term “semi-quantitative” as a method classification. Methods that have been self-identified as “semi-quantitative” 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 4 and 5.

In some cases and for some purposes, methods with less accuracy and precision than recommended in tables 4 and 5 may be acceptable. Method developers should consult with the appropriate method committee to determine if the recommendations in tables 4 and 5 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 Cut Off 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.

Appendix C: 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 are 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 are 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.

1.6 Mass fraction. Concentration, C, expressed as a decimal fraction. For calculating and reporting statistical parameters, the 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, the 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, **C1** 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, **C1**.

DRAFT

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:

$$HORRAT = RSD(R) / PRSD(R)$$

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 PRSDR calculated from reproducibility data because this parameter is more predictable than the parameter calculated from repeatability data:

$$HORRAT(R) = RSD_R / PRSD(R)$$

$$HORRAT(r) = RSD_r / PRSD(R)$$

Some expected, predicted relative standard deviations are given in the following summary table:

Table C1: 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	0.000000001	45

2.0 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 one half to two thirds the among-laboratory relative standard deviations.

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 the following table at 1/2 of PRSDR, which represents the best case.

Table C2: 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

2.2.2 HORRAT(r)

Based on experience and for the purpose of exploring the extrapolation of HORRAT values to single-laboratory validation (SLV) studies, take as the minimum acceptability one half of the lower limit ($0.5 \times 0.5 \approx 0.3$) and as the maximum acceptability two thirds 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);

DRAFT

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., $10\ ng/g$; $10\ ppb$).

Excerpts from Development and Use of In-House Reference Materials



Appendix D: *Development and Use of In-House Reference Materials*



Development and Use of In-House Reference Materials, rev. 2, 2009
Copyright 2005 by the Technical Division on Reference Materials (TDRM), a
division of AOAC International, 481 North Frederick Avenue, Suite 500,
Gaithersburg, MD 20877-2417 USA

All rights reserved. No part of this booklet may be reproduced, in part or in
whole, by any form or by any means, without the written permission of the
TDRM.

Appendix D: *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 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 **certified reference material (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?

Certified reference materials 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.

Appendix D: *Development and Use of In-House Reference Materials*

What Certified Reference Materials are Currently Available?

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

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

A number of websites provide general overviews and catalogues 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/programmes/aqcs>

<http://www.aaccnet.org/checksample>

<http://www.irmm.irc.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?

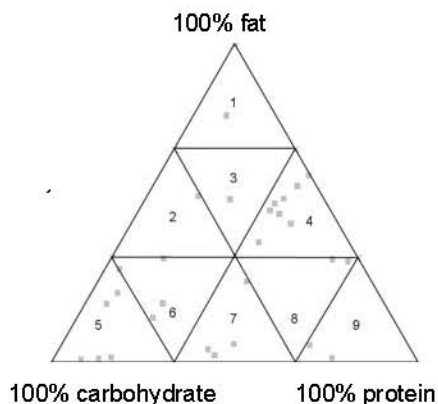
There are many benefits to the use of a CRM. CRMs have been prepared to be homogeneous and, if stored under the proper conditions, stable. You are provided with a certified value as well as the statistical data for the concentration of your analyte; this is about as close as you can come to knowing the 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 below, but the same processes apply to the

Appendix D: Development and Use of In-House Reference Materials

development of in-house RMs 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 this triangle, below, and are provided in the list that follows.



Sector	RM Number	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 I
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

These food-matrix reference materials are spread through all sectors of the triangle, thereby making it likely that

Appendix D: *Development and Use of In-House Reference Materials*

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 (3 jars, each containing 170 g of 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 CRM⁴. 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 TDRM as well as in ISO Guides 34⁵ and 35⁶.

For more information about the AOAC Technical Division on Reference Materials, visit <http://aoac.org/divisions/tdrm>

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. *Fresenius' J. Anal. Chem.* **1995**, *352*, 73-76.
3. Wolf, W. R. *Methods of Analysis for Nutrition Labeling*, Sullivan, D. R.; Carpenter, D. E., Eds.; AOAC International: Arlington, VA, 1993.
4. European Reference Materials. Comparison of a Measurement Result with the Certified Value. Application Note 1. 2005. European Reference Materials
5. ISO. Guide 34 General Requirements for the Competence of Reference Material Producers. 2nd. 2009. Geneva, Switzerland, ISO.
6. Guide 35 Certification of Reference Materials -- General and Statistical Principles. 2006. Geneva, Switzerland, ISO.