Appendix I: AOAC INTERNATIONAL Methods Committee Guidelines for Validation of Biological Threat Agent Methods and/or Procedures

Table of Contents **Page**

This document provides technical protocol guidelines for the AOAC validation of biological threat agent methods and/or procedures, and covers terms and their definitions associated with the *Performance Tested Methods*SM and *Official Methods of Analysis*SM programs.

The guidelines working group consisted of Sharon Brunelle (AOAC technical consultant), Robert LaBudde (AOAC statistical advisor), Maria Nelson (AOAC technical consultant), and Paul Wehling (AOAC statistical advisor). The guidelines were reviewed by the AOAC Methods Committee on Biological Threat Agents and the Official Methods Board (OMB) and approved by OMB on November 12, 2009.

This work was partially funded by the Department of Defense through DHHS/FDA/OAGS/DCGM, award 223012464, and the Department of Homeland Security Science and Technology Directorate, award HSHQDC-08-C-00012.

1 Scope

The purpose of this document is to provide comprehensive technical guidelines for conducting AOAC INTERNATIONAL (AOAC) validation studies for biological threat agent methods and/or procedures submitted for *Performance Tested Methods*SM (PTM) status and/or for AOAC® *Official Methods of Analysis*SM (OMA) status. These guidelines pertain to bacteria and toxins only at this time. The requirements for method developer validation studies, independent validation studies, and collaborative validation studies for those methods are described. The guidelines describe studies involved in the validation process for methods including acceptable minimum detection level (AMDL), acceptable minimum quantitation level (AMQL), inclusivity/exclusivity testing, matrix studies, robustness, product consistency and stability, instrument variation, and reproducibility.

2 Applicability

These guidelines are intended to be applicable to the validation of candidate biological threat agent methods, whether proprietary or nonproprietary, that are submitted to AOAC for OMA status or PTM certification. Circumstances, unforeseen by AOAC, may necessitate divergence from the guidelines in certain cases. The PTM Program requires a method developer study and an independent laboratory study. The OMA Program requires a single-laboratory validation (SLV) study (also known as the precollaborative study, referred to here as the method developer study) and a collaborative study. A harmonized PTM–OMA Program can be followed in which PTM certification is sought and, if successful, serves as the SLV phase of the OMA Program. This approach provides an interim certification while working toward OMA status. *See* to Table 1 for more detail.

3 Terms and Definitions

Where appropriate, definitions have been taken from international standards and the source is noted. Sources of definitions include the following:

ISO/IEC Guide 99:2007, *International vocabulary of metrology—Basic and general concepts and associated terms* (VIM)

ISO 3534-2:2006, *Statistics—Vocabulary and symbols—Part 2: Applied statistics*

ISO 14971:2007, *Medical devices—Application of risk management to medical devices*

ISO 17511:2003, *In vitro diagnostic medical devices—Measurement of quantities in biological samples—Metrological traceability of values assigned to calibrators and control materials*

ISO 5725-1:1994, *Accuracy (trueness and precision) of measurement methods and results—Part 1: General principles and definitions*

Table 1. AOAC validation study requirements

AOAC program	Study requirements	Relevant guideline sections	
		Qualitative	Quantitative
PTM	Method developer validation study	4.1	5.1
	Independent validation study	4.2	5.2
OMA	Method developer validation study	4.1.2–4.1.5	5.1.2–5.1.5
	Collaborative validation study	4.3	5.3
Harmonized PTM-OMA	Method developer validation study	4.1	5.1
	Independent validation study	4.2	5.2
	Collaborative validation study	4.3	5.3

USP 31:2008, *U.S. Pharmacopeia General Information/<1223> Validation of Alternative Microbiological Methods*

3.1 Acceptable Minimum Detection Level

The predetermined minimum level of a biological threat agent, as specified by the Methods Committee on Biological Threat Agents with input from subject matter experts, which must be detected by the candidate method with an estimated 5% lower confidence limit on the probability of detection (POD) of 0.95 or higher. The AMDL is dependent on the intended use.

3.2 Acceptable Minimum Quantitation Level

The minimum level of a biological threat agent, as predetermined by the Methods Committee on Biological Threat Agents with input from subject matter experts, which must be quantified with predefined precision and accuracy under the experimental conditions by the method under validation. The AMQL is dependent on the intended use.

3.3 Accuracy

Closeness of agreement between a quantity value obtained by measurement and the true value of the measurand (VIM 2007).

3.4 Biological Threat Agent

Biological, virological, or toxic threat agents are identified in the Health and Human Services list of Select Agents (1).

3.5 Bias

The difference between the expectation of the test result or measurement result and the true value (ISO 3534-2).

3.6 Candidate Method

The method submitted for validation.

3.7 Candidate Method Result

The final result of the qualitative or quantitative analysis for the candidate method. For methods with a confirmation phase, only presumptive positive results that confirm positive are considered as positive for the candidate method. All other results are considered as negative for the candidate method.

3.8 Collaborator

An intended end user who participates in the collaborative study.

3.9 Collaborator Data Set

Results from a combination of an analytical instrument, device, or equipment and an operator, technician, or analyst. Results from a single test site may serve as multiple collaborator data sets dependent on nonredundant operators and acceptable separation of effort.

3.10 Exclusivity

The nontarget agents, which are potentially cross-reactive, that are not detected by the method.

3.11 Identification Method

Method of analysis whose purpose is to determine the identity of a biological threat agent.

3.12 Inclusivity

The strains or isolates or variants of the target agent(s) that the method can detect.

3.13 Intended Use

Use for which a product, process, or service is intended according to the specifications, instructions, and information provided by the manufacturer (ISO 14971).

3.14 Matrix

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

3.15 Method

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

3.16 Near Neighbor

Organisms and/or substances selected to be either closely related or potentially cross-reactive with the organism and/or substance under test.

3.17 Precision

Closeness of agreement between quantity values obtained by replicate measurements of a quantity, under specified conditions (VIM 2007). Precision is usually expressed as the standard deviation or relative standard deviation.

3.18 Presumptive versus Confirmatory Analysis

Some candidate methods follow a sequence of steps and only report results for the analyte when a second test method has been carried out on presumptive positives. The intended purpose of the second test is to reduce the frequency of false-positive results of the first test method. In such a case, the first test method result is called a "presumptive" result, and a presumptive positive result which is followed by a second method positive result is called a "confirmed" positive result. Other variations may include confirmation of presumptive negative results, or both presumptive positive and negative results, although confirmation of positive results is by far the more common.

3.19 Probability of Detection

The proportion of positive analytical outcomes for a qualitative method for a given matrix at a given agent level or concentration. POD is concentration dependent. Several POD measures can be calculated, e.g., POD_R (for the reference method), POD_C (for the confirmed candidate method), POD_{CP} (for the candidate presumptive method), and POD_{CC} (for the candidate confirmation method). *See* ref. 2 and *Annex A*.

3.20 Qualitative Method

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

3.21 Quantitative Method

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

3.22 Reference Method or Procedure

Pre-existing recognized analytical method against which the candidate method will be compared. For a validation protocol, the reference method will be designated by the Methods Committee on Biological Threat Agents.

3.23 Repeatability

Precision under repeatability conditions (ISO 5725-1).

3.24 Repeatability Conditions

Conditions where independent test results are obtained with the same method on equivalent test items in the same laboratory by the same operator using the same equipment within short intervals of time.

3.25 Reproducibility

Precision under reproducibility conditions (ISO 5725-1).

3.26 Reproducibility Conditions

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

3.27 Robustness Study

A study which tests the capacity of a method to remain unaffected by small but deliberate variations in method parameters and which provides an indication of its reliability during normal usage (USP 31).

3.28 Sample

The batch of matrix material from which replicate test portions are removed for analysis. The sample (uncontaminated or inoculated) contains agent at one specified level.

3.29 Test Portion

A quantity of subsample or member of a sample set that is taken for analysis by the method.

3.30 Test Sites

Sites that simulate where the method is intended to be used, such as (in order of increasing risk): (*1*) *Traditional laboratory*.—Controlled atmosphere; permanent stationary facility (e.g., public health laboratory). (*2*) *Mobile laboratory.*—Controlled atmosphere; mobile facility (e.g., National Guard mobile laboratory). (*3*) *Nonlaboratory indoor facility.*—Controlled atmosphere; stationary facility (e.g., office building). (*4*) *Outdoor mobile site*.—Uncontrolled atmosphere; mobile facility (e.g., fire truck). (*5*) *Outdoor automated stationary site.*—Uncontrolled atmosphere; stationary site (e.g., environmental monitoring station).

4 Qualitative Methods

Inclusivity panels, exclusivity panels, environmental panels, and standard method performance requirements (SMPRs) have been established for some intended uses (3). If a candidate method intended use is not covered by established SMPRs, the Methods Committee on Biological Threat Agents will determine the appropriate panels and performance requirements.

4.1 Method Developer Validation Study

4.1.1 Scope

The method developer validation study is intended to determine the performance of a biological threat agent method under the controlled conditions of a laboratory. The study is designed to evaluate performance parameters including minimum inclusivity, exclusivity, POD at the AMDL, robustness, between-instrument variation, and product quality (lot-to-lot variability and product stability).

The method developer validation study is normally conducted in a single laboratory, usually the method developer's laboratory. Alternatively, the method developer can contract the work to a contract laboratory.

4.1.2 Inclusivity/Exclusivity Study

4.1.2.1 Analyte Variant Selection

The inclusivity panel should be sufficiently large that the analyte variation is adequately represented. The Methods Committee on Biological Threat Agents, with the advice of subject matter experts, shall determine the specific analyte variants required. Typically, 30 variants will be required, if available.

Exclusivity organisms and/or substances include near neighbors, chosen to adequately cover potentially cross-reactive organisms or substances. The Methods Committee on Biological Threat Agents, with the advice of subject matter experts, shall determine the specific species, strains or substances required for exclusivity evaluation. Typically, 30 will be required, if available.

Species/strains/substances specified for use should be traceable to the source. The source and origin of each species/strain/substance must be documented. Certificate of analysis or other documentation of analyte identity should be on file and available to the Methods Committee on Biological Threat Agents.

4.1.2.2 Study Design

Inclusivity strains or substances are prepared and analyzed as vegetative cells, spores, or components thereof as applicable to the candidate method. The target concentration for testing is the predetermined AMDL. Test one replicate per strain/substance using the candidate method.

Exclusivity species/strains/substances are prepared and analyzed as vegetative cells, spores, or components thereof as applicable to the candidate method. The target concentration for testing is 10 times the AMDL. Test one replicate per strain/substance using the candidate method.

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

If an individual inclusivity panel member tests negative, it may be retested in 96 replicates with no failures allowed to demonstrate a 5% lower confidence limit on the POD of 0.95 or higher for that panel member. If an individual exclusivity panel member tests positive, it may be retested in 96 replicates with no failures allowed to demonstrate a 95% upper confidence limit on the POD of 0.05 or lower for that panel member. All test and retest data must be reported.

4.1.2.3 Data Analysis and Reporting

Analyze the data for a positive or negative response. The data is reported as number of panel members detected relative to the total tested. For example, "Of the 30 specific inclusivity strains tested, 28 were detected and two were not detected. Those strains not detected were the following: …" or "Of the 30 specific exclusivity strains tested, 27 were not detected and three were detected. Those detected were the following: …" The study report should include a table titled "Inclusivity/Exclusivity Panel Results," which lists all analyte variants and near neighbors tested, their source, origin and essential characteristics, and testing outcome.

4.1.3 Laboratory Matrix Study

The purpose of the laboratory matrix study is to measure the POD at the AMDL or lower in a controlled laboratory setting for all matrixes claimed in the intended use statement.

4.1.3.1 Reference Method

If there is a reference method, then the candidate method performance parameters should be compared to those published for the reference method. Designation of an appropriate reference method will be determined by the Methods Committee on Biological Threat Agents.

4.1.3.2 Matrix Categories

AOAC INTERNATIONAL recognizes claims for only the range of matrix categories or specific matrix types included in the method developer study and the collaborative study. The number of different matrixes to be tested depends on the intended use of the method.

Matrixes include, but are not limited to, environmental surfaces (such as stainless steel, plastic, ceramic/glass, nonporous rubber, food-grade painted surfaces, finished wood, sealed concrete), filter samples, aqueous air collection samples, bulk powders, and soils (loam, sandy, clay, subsoil, and silt). Matrixes should be selected to focus on those most probable to be encountered from the intended use statement.

4.1.3.3 Study Design

4.1.3.3.1 Analyte Concentration

Divide each matrix into at least two samples. Inoculate one sample with a target agent at the AMDL or lower. Inoculate the second sample with a near neighbor at 10 times the AMDL. Test portions are to be taken from each sample.

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

4.1.3.3.2 Number of Test Portions

The study design shall include 96 inoculated test portions at the AMDL or a specified lower concentration for each matrix. Ninety-six test portions of matrix inoculated with a near neighbor at 10 times the AMDL are included as negative controls.

4.1.3.3.3 Agent for Matrix Inoculation

The matrix inoculation should be conducted with a pure culture of one strain or a pure solution of one target variant. The inoculum should be in a form that is to be detected. Exceptions may be made at the discretion of the Methods Committee on Biological Threat Agents. The same strain must be used for all matrixes and will be designated by the Methods Committee on Biological Threat Agents. The strains specified for use must be traceable to the source. The source and origin must be documented.

4.1.3.3.4 Preparation of Samples

Samples must be prepared in a consistent manner for all similar studies. Consideration should be given to spore counts versus vegetative cell counts; stage of growth of cells; cell preparation procedures; protein concentration procedures; DNA concentration procedures; relationship between cell/spore count and DNA concentration; and determination of genome equivalents based on DNA concentration. Study Directors and AOAC staff should consult with the Methods Committee on Biological Threat Agents for detailed guidance.

4.1.3.3.4.1 Environmental Surfaces

Dilute target agent in prescribed medium and spread evenly across designated surface at target concentration level. Allow to dry prior to sampling. Dilute near neighbor in prescribed medium and spread evenly across a separate designated surface at 10 times the AMDL to serve as the negative control. Follow sample collection procedures as defined in AOAC *Official Method*^{SM} **2006.04** Standard Practices for Bulk Sample Collection and Swab Sample Collection of Visible Powders Suspected of Being Biological Agents from Nonporous Surfaces (4).

4.1.3.3.4.2 Air Collection Samples, Filters

Dilute agent in H_2O or other suitable diluent and apply to clean filter at the AMDL. Allow to dry prior to testing. Alternatively, dry spores can be inoculated onto filters using an appropriate closed chamber airflow device, if available. Inoculate filters with a near neighbor organism at 10 times the AMDL to serve as negative controls.

4.1.3.3.4.3 Air Collection Samples, Aqueous

Dilute agent in device collection liquid at target concentration level. Inoculate collection liquid with a near neighbor organism at 10 times the AMDL to serve as negative controls.

4.1.3.3.4.4 Soils and Bulk Powders

Dry matrixes such as soils and powders should be inoculated with dry or lyophilized agent. Add dried agent to matrix and mix thoroughly into test sample to achieve target concentration level. Add one dried near neighbor and mix thoroughly in matrix to serve as negative controls.

4.1.3.3.5 Confirmation of Test Portions

If the candidate method includes presumptive and confirmatory steps, then follow the confirmation method as written.

4.1.3.4 Data Analysis and Reporting

Each concentration of analyte in each matrix must be analyzed and reported separately.

4.1.3.4.1 Raw Data Tables

For each matrix and concentration, report each result from each test portion separately in a software-compatible format. *See Annex B* for raw data table format.

4.1.3.4.2 POD

The POD estimate is calculated as the number of positive outcomes divided by the total number of trials.

Estimate the POD with a 95% confidence interval for the candidate method and, if included, the presumptive and confirmed results. *See Annex C* for details.

4.1.3.4.3 Difference of Probability of Detection (dPOD)

If appropriate historical data are available from a previously validated method, then calculate the dPOD.

dPOD is the difference between any two POD values at the same analyte concentration.

Estimate the $dPOD_C$ as the difference between the candidate and reference POD values. Calculate the 95% confidence interval on the $dPOD_C$.

$$dPOD_C = POD_C - POD_R$$

Similar calculations can be carried out for presumptive and confirmed data sets, if needed. *See Annex C* for details. If the confidence interval of a dPOD does not contain zero, then the difference is statistically significant at the 5% level.

4.1.3.4.4 Summary Data Tables

For all matrixes and levels, use the summary table format in *Annex D*.

4.1.3.4.5 Data Analysis and Reporting in the Absence of a Reference Method

If no appropriate reference method is available for the target agent, indicate "Not Applicable" where appropriate in the summary tables.

4.1.4 Intended Use Matrix Study (Required for Intended Uses Outside the Laboratory)

The purpose of the intended use matrix study is to measure the POD at the AMDL or lower under simulated intended use conditions with representative end users for all matrixes claimed in the intended use statement.

4.1.4.1 Intended Use Statement

A statement by the method sponsor of the context, expected matrixes, and expected analytical goals of an analytical method must be provided by the method developer.

4.1.4.2 Matrixes

POD at the AMDL or lower (concentration must be the same as in **4.1.3**) for each matrix claimed in the intended use is estimated under simulated intended use conditions by representative personnel. All matrixes for all claimed intended uses must be tested.

4.1.4.3 Study Design

See **4.1.3.3**.

4.1.5 Environmental Interference Study

Biological threat agent methods are likely to be used in a variety of uncontrolled environments. Biological threat agent methods must be impervious to contaminants likely to be found in the environment. All candidate biological threat agent methods must be investigated for potential interference from a wide variety of contaminants likely to be found in the environment. Interference from environmental contaminants may show up as cross-reactivity or inhibition. Both types of interference must be studied.

The AOAC Methods Committee on Biological Threat Agents in conjunction with input from the stakeholder community shall provide guidance on the contaminants to be used for each class of analytical methodology (i.e., methods based on the polymerase chain reaction techniques would illustrate a class of analytical methodology). The environmental panels established to date are found in the SMPRs (3).

4.1.5.1 Study Design

4.1.5.1.1 Contaminant Concentrations

Environmental panel organisms are tested at 10 times the AMDL in the absence of target agent and in the presence of target agent at the AMDL. Environmental organisms can be tested in pools of no more than 10 organisms per pool. For pooled organisms, the concentration of each is 10 times the AMDL.

Environmental panel substances are tested according to the panel requirements in the applicable SMPRs (3). Test each substance in the presence of analyte at the AMDL and in the presence of a near neighbor at 10 times the AMDL.

4.1.5.1.2 Number of Replicates

Environmental panel organisms and substances are initially tested without replicates in the presence of agent or near neighbor.

If a pool of organisms yields inhibition or cross-reactivity, each organism must be retested individually to determine the cause of the inhibition or cross-reactivity. Any observed inhibition or cross-reactivity may be followed by retesting using a study design with a sufficient number of replicates that provides inhibition POD ≥ 0.95 or cross-reactivity POD ≤ 0.05 with 95% confidence. A study design using a minimum of 96 replicates satisfies these requirements. Results from a study using a minimum of 96 replicates are acceptable if all analyses yield the expected results. All test and retest data must be reported and discrepancies investigated.

4.1.6 Robustness Study

Minor, reasonable variations in a method of a magnitude that might well be expected to occur when the method is used are deliberately introduced and tested. Variations in method parameters that can be influenced by the end user should be tested. Use a screening experimental design (5).

The method developer, in conjunction with the AOAC project manager, is expected to make a good faith effort to determine which parameters are most likely to affect the analytical performance and determine the range of variations that can occur without adversely affecting analytical results.

4.1.6.1 Analyte Concentration

Robustness strains/substances are prepared and analyzed as vegetative cells, spores or components thereof as applicable to the candidate method. One agent strain/substance is tested at the AMDL, and at 10 times the AMDL. If matrix studies were carried out at a level below the AMDL, then conduct robustness studies at that level and at 10 times that level of agent. Use a near neighbor at 10 times the AMDL as a negative control.

4.1.6.2 Study Design

For continuous numerical outputs, test five replicates of each agent concentration and five replicates of near neighbor for each factorial pattern.

For binary sensors, test 96/*P* replicates of agent, where *P* is the number of factorial patterns, at the AMDL and 96/*P* near neighbor replicates at 10 times the AMDL for each pattern.

4.1.6.3 Data Analysis and Reporting

The results are analyzed for variable detection due to parameter variations. Report the appropriate statistical measures of the measured variable(s) (e.g., Ct, absorbance, POD value, etc.) for

each set of replicates for each factorial pattern. This should include at least means, standard deviations, and confidence intervals where appropriate. Use multiple linear regression to determine whether a change in response due to parameter variation is physically important (6, 7).

For binary sensors, only one unexpected result out of 96 replicates is allowed. Calculate the POD value and confidence interval for each factorial variation. Analyze results using a generalized linear model, e.g., logistic regression, in consultation with the statistical advisor (6).

4.1.7 Product Consistency and Stability Study

The product consistency and stability study is a study or set of studies designed to ensure that the performance of the product is consistent from lot-to-lot and over time under normal storage conditions for the shelf life of the product. These studies are required for PTM certification.

4.1.7.1 Analyte Concentration

Strains/substances for product consistency and stability testing are prepared and analyzed as vegetative cells, spores, or components thereof as applicable to the candidate method. One agent strain/substance is tested at the AMDL, and at 10 times the AMDL. If matrix studies were carried out at a level below the AMDL, then conduct product consistency and stability studies at that level and at 10 times that level of agent. Use a near neighbor at 10 times the AMDL as a negative control.

4.1.7.2 Study Design

For the lot-to-lot consistency study, at least six lots of product must be tested and must show consistent results (8).

For product stability, either real-time or accelerated stability (9) data will be accepted for initial PTM certification. If accelerated stability data are submitted, real-time stability data must be submitted within 1 year of product certification to support the shelf life of the kit. For real-time stability, data estimating with confidence intervals the bias and repeatability (for continuous numerical outputs) or POD values (for binary sensors) between time zero and the end of the shelf life must be submitted.

For continuous numerical outputs, test five replicates of each agent and near neighbor concentration, for each lot or time point.

For binary sensors, test 96/L, where L is the number of lots or time points and is no less than six, replicates of agent at the AMDL and 96/L replicates of a near neighbor at 10 times the AMDL for each lot and each time point.

Lot-to-lot and stability testing can be combined into one study by testing at least six production lots whose ages span the shelf life of the product.

4.1.7.3 Data Analysis and Reporting

The results are analyzed for variable detection between lots or time points. Report the appropriate statistical measures of the measured variable(s) (e.g., Ct, absorbance, POD value, etc.) for each set of replicates for each kit/reagent lot and between lots or for each stability time point. This should include at least means, standard deviations, confidence intervals, and time slope, as appropriate.

For binary sensors, only one unexpected result out of 96 replicates is allowed. Calculate the POD value and confidence interval for each lot and each time point. For binary and continuous sensors, determine the standard deviation across lots and the time slope with confidence interval and % performance degradation across time points. For the combined approach, use a method of statistical analysis acceptable to the AOAC statistical advisor, such as a generalized mixed model fitting program to determine the lot to lot variability as a random effect and the time slope as a fixed effect (6, 7).

4.1.8 Between-Instrument Variation Study (if applicable)

This study applies only to methods that require the use of an instrument. The between-instrument variation study is a study designed to ensure that the performance of the method is consistent from instrument to instrument. This study is required for PTM certification.

4.1.8.1 Analyte Concentration

Strains/substances for between-instrument variation testing are prepared and analyzed as vegetative cells, spores, or components thereof as applicable to the candidate method. One agent strain/substance is tested at the AMDL, and at 10 times the AMDL. If matrix studies were carried out at a level below the AMDL, then conduct between instrument variation studies at that level and at 10 times that level of agent. Use a near neighbor at 10 times the AMDL as a negative control.

4.1.8.2 Study Design

At least six instruments must be tested. For continuous numerical outputs, test five replicates of each agent and near neighbor concentration for each instrument.

For binary sensors, test 96/I, where I is the number of instruments, replicates of agent at the AMDL and 96/I replicates of a near neighbor at 10 times the AMDL for each instrument.

4.1.8.3 Data Analysis and Reporting

The results are analyzed for variable detection between instruments. Report the appropriate statistical measures of the measured variable(s) (e.g., Ct, absorbance, POD value, etc.) for each set of replicates for each instrument and between instruments. This should include at least means, standard deviations, and confidence intervals, as appropriate.

For binary sensors, only one unexpected result out of 96 replicates is allowed. Calculate the POD value and confidence interval for each instrument. For binary and continuous sensors, determine the standard deviation across instruments.

4.2 Independent Validation Study

4.2.1 Scope

The independent validation study, preferably conducted under PTM guidance, should verify the analytical results obtained by the method developer. The study traditionally verifies POD in the hands of an independent trained user in a controlled laboratory setting. The independent study also determines the performance of the method under simulated intended use conditions using representative personnel. In addition, inclusivity/exclusivity and environmental interference will be repeated in the independent validation study to verify the method developer's data.

All biological threat agent methods shall be validated under conditions that simulate the intended use. For example, a biological threat agent method intended for use by trained emergency response personnel in personal protective equipment at an outdoor mobile site must be validated under those or similar conditions. Biological threat agent methods intended for a variety of uses shall

be validated for all combinations of the intended use to be claimed as *Official Methods of Analysis*℠.

Only those intended uses successfully validated may be claimed as *Official Methods of Analysis*℠.

4.2.2 Inclusivity/Exclusivity Study

4.2.2.1 Analyte Variant Selection

The inclusivity panel should be sufficiently large that the analyte variation is adequately represented. The Methods Committee on Biological Threat Agents, with the advice of subject matter experts, shall determine the specific analyte variants required. Typically, 30 variants will be required, if available. The panel shall be the same as that used in the method developer study, whenever possible.

Exclusivity organisms and/or substances include near neighbors, chosen to adequately cover potentially cross-reactive organisms or substances. The Methods Committee on Biological Threat Agents, with the advice of subject matter experts, shall determine the specific species, strains, or substances required for exclusivity evaluation. Typically, 30 will be required, if available. The panel shall be the same as that used in the method developer study, whenever possible.

Species/strains/substances specified for use should be traceable to the source. The source and origin of each species/strain/substance must be documented. Certificate of analysis or other documentation of analyte identity should be on file and available to the Methods Committee on Biological Threat Agents.

4.2.2.2 Study Design

Inclusivity strains or substances are prepared and analyzed as vegetative cells, spores, or components thereof as applicable to the candidate method. The target concentration for testing is the predetermined AMDL. Test one replicate per strain/substance using the candidate method.

Exclusivity species/strains/substances are prepared and analyzed as vegetative cells, spores, or components thereof as applicable to the candidate method. Test exclusivity panel members at 10 times the AMDL. Test one replicate per strain/substance using the candidate method.

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

If an individual inclusivity panel member tests negative, it may be retested in 96 replicates with no failures allowed to demonstrate a 5% lower confidence limit on the POD of 0.95 or higher for that panel member. If an individual exclusivity panel member tests positive, it may be retested in 96 replicates with no failures allowed to demonstrate a 95% upper confidence limit on the POD of 0.05 or lower for that panel member. All test and retest data must be reported.

4.2.2.3 Data Analysis and Reporting

Analyze the data for a positive or negative response. The data is reported as number of panel members detected relative to the total tested. For example, "Of the 30 specific inclusivity strains tested, 28 were detected and two were not detected. Those strains not detected were the following: …" or "Of the 30 specific exclusivity strains tested, 27 were not detected and three were detected. Those detected were the following: …" The study report should include a table titled "Inclusivity/Exclusivity Panel Results," which lists all analyte variants and near neighbors tested, their source, origin and essential characteristics, and testing outcome.

4.2.3 Laboratory Matrix Study

The purpose of the laboratory matrix study is to measure the POD at the concentration tested by the method developer for all matrixes claimed in the intended use statement in a controlled laboratory setting.

4.2.3.1 Study Design

The study design for validation of qualitative methods in the independent study follows the method developer validation study design. Analyte concentration, near neighbor concentration, number of test portions, agent for inoculation, preparation of samples, confirmation of test portions, and data analysis and reporting are found in Section **4.1.3.3**.

4.2.3.2 Reference Method

If there is a reference method, then the candidate method performance parameters should be compared to those published for the reference method. Designation of an appropriate reference method will be determined by the Methods Committee on Biological Threat Agents.

4.2.4 Intended Use Matrix Study (Required for Intended Uses Outside the Laboratory)

The purpose of the intended use matrix study is to measure the POD at the concentration tested by the method developer for all matrixes claimed under simulated intended use conditions with a representative end user(s).

4.2.4.1 Study Design

The study design for validation of qualitative methods at an independent test site follows the method developer validation study design. Design characteristics for analyte concentration, near neighbor concentration, number of test portions, agent for inoculation, preparation of samples, confirmation of test portions, and data analysis and reporting are found in Section **4.1.3**.

4.2.5 Environmental Interference Study

Biological threat agent methods are likely to be used in a variety of uncontrolled environments. Biological threat agent methods must be impervious to contaminants likely to be found in the environment. All candidate biological threat agent methods must be investigated for potential interference from a wide variety of contaminants likely to be found in the environment. Interference from environmental contaminants may show up as cross-reactivity or inhibition. Both types of interference must be studied.

The AOAC Methods Committee on Biological Threat Agents, in conjunction with input from the stakeholder community, shall provide guidance on the contaminants to be used for each class of analytical methodology (i.e., methods based on the polymerase chain reaction techniques would illustrate a class of analytical methodology). The environmental panels established to date are found in the SMPRs (3).

4.2.5.1 Study Design

4.2.5.1.1 Contaminant Concentrations

Environmental panel organisms are tested at 10 times the AMDL in the absence of agent and in the presence of agent at the AMDL. Environmental organisms can be tested in pools of no more than 10 organisms per pool. For pooled organisms, the concentration of each is 10 times the AMDL.

Environmental panel substances are tested according to the panel requirements in the SMPRs (3). Test each substance in the presence of agent at the AMDL and in the presence of a near neighbor at 10 times the AMDL.

4.2.5.1.2 Number of Replicates

Environmental panel organisms and substances are initially tested without replicates in the presence of agent or near neighbor.

If a pool of organisms yields inhibition or cross-reactivity, each organism must be retested individually to determine the cause of the inhibition or cross-reactivity. Any observed inhibition or cross-reactivity may be followed by retesting using a study design with a sufficient number of replicates that provides inhibition POD ≥ 0.95 or cross-reactivity POD ≤ 0.05 with 95% confidence. A study design using a minimum of 96 replicates satisfies these requirements. Results from a study using a minimum of 96 replicates are acceptable if all analyses yield the expected results. All test and retest data must be reported and discrepancies investigated.

4.3 Collaborative Study

4.3.1 Scope

The collaborative study is a requirement for *Official Methods*SM and succeeds the PTM independent validation study. The collaborative study characterizes the performance parameters (e.g., POD, repeatability, reproducibility) of the candidate method across collaborators.

All biological threat agent methods shall be validated under simulated conditions of intended use. For example, a biological threat agent method intended for use by trained emergency response personnel in personal protective equipment at an outdoor mobile site must be validated under conditions that simulate the outdoor mobile site and include representative personnel as collaborators. Biological threat agent methods intended for a variety of uses shall be validated for all combinations of the intended use to be claimed as *Official Methods of Analysis*SM.

Collaborators shall receive three types of training in conjunction with the study. They shall have at least 3 h of training on AOAC method processes. If the study follows the alternative collaborative study design, then collaborators shall receive training on specific facility requirements from the laboratories hosting the collaborators. All collaborators shall be trained by the vendor on the test method under study. All training shall be documented.

A trial run should be conducted prior to the initiation of the validation study. The purpose of the trial run is to ensure that logistics, sample handling, and data reporting processes are worked out and understood by all of the collaborators. A small sample set (e.g., 2–4 samples) using an avirulent or nontoxic surrogate, if possible, should be analyzed. The trial run should be conducted under the same conditions as the validation study. A reasonable amount of time should be allotted for troubleshooting after the completion of the trial run, including a discussion with all of the collaborators to address issues and answer questions. The data should not be analyzed or included in the validation report.

4.3.2 Matrix Study

POD is estimated at the level tested in the method developer study for all claimed matrixes in the intended use statement under simulated intended use conditions by representative personnel.

4.3.2.1 Number of Collaborators

The collaborative study must include a minimum of 12 collaborators. After eliminating data sets for assignable cause, the study must yield at least 10 valid data sets.

The study must include a minimum of three test sites with no more than four collaborators at any one test site. Care must be taken to ensure the independence of collaborators at each test site.

4.3.2.2 Levels of Contamination

Divide each matrix into at least two samples. Inoculate one sample with the target agent at the level tested in the method developer study. Inoculate a second sample with a near neighbor at 10 times the AMDL of the target to serve as the negative control. Analyze test portions from each sample.

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

4.3.2.3 Number of Test Portions

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

4.3.2.4 Agent for Matrix Inoculation

See Section **4.1.3.3.3**.

4.3.2.5 Preparation of Samples

See Section **4.1.3.3**.

4.3.2.6 Confirmation of Test Portions

If the candidate method includes presumptive and confirmatory phases, then follow the method as written.

4.3.2.7 Data Analysis and Reporting

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

4.3.2.7.1 Raw Data Tables

For each matrix and concentration level, report each result from each test portion separately. *See Annex B* for raw data table in a software-compatible format.

4.3.2.7.2 Between-Collaborator Standard Deviation

For each matrix, level, and collaborator, calculate the POD. For each matrix and level, calculate the standard deviation of collaborator PODs (sPOD).

4.3.2.7.3 Between-Collaborator Probability of Detection (CPOD)

Report the CPOD estimates with 95% confidence intervals for the candidate method and, if included, the presumptive and confirmed results. *See Annex E* for details.

4.3.2.7.4 Summary Data Tables

For all matrixes and levels, use the summary table from *Annex F*.

4.3.2.7.5 Collaborator Comments

Comments on the candidate method should be collected from all collaborators and reported in the collaborative study report.

5 Quantitative Methods

Inclusivity panels, exclusivity panels, environmental panels, and SMPRs have been established for some intended uses (3). If a candidate method intended use is not covered by established SMPRs, the Methods Committee on Biological Threat Agents will determine the appropriate panels and performance requirements.

5.1 Method Developer Validation Study

5.1.1 Scope

The method developer validation study is intended to determine the performance of a biological threat agent method under the controlled conditions of a laboratory. The study is designed to evaluate performance parameters including minimum inclusivity, exclusivity, repeatability, bias, robustness, and product quality (lot-to-lot variability and product stability).

The method developer study is normally conducted in a single laboratory, usually the method developer's laboratory. Alternatively, the method developer can contract the work to an independent site.

5.1.2 Inclusivity/Exclusivity Study

5.1.2.1 Analyte Variant Selection

The inclusivity panel should be sufficiently large that the analyte variation is adequately represented. The Methods Committee on Biological Threat Agents, with the advice of subject matter experts, shall determine the specific analyte variants required. Typically, 30 variants will be required, if available.

Exclusivity organisms and/or substances include near neighbors, chosen to adequately cover potentially cross-reactive organisms or substances. The Methods Committee on Biological Threat Agents, with the advice of subject matter experts, shall determine the specific species, strains or substances required for exclusivity evaluation. Typically, 30 will be required, if available.

Species/strains/substances specified for use should be traceable to the source. The source and origin of each species/strain/substance must be documented. Certificate of analysis or other documentation of analyte identity should be on file and available to the Methods Committee on Biological Threat Agents.

5.1.2.2 Study Design

Inclusivity strains or substances are prepared and analyzed as vegetative cells, spores, or components thereof as applicable to the candidate method. The target concentration for testing is the AMQL of the method. Test one replicate per strain or substance using the candidate method.

Exclusivity species/strains/substances are prepared and analyzed as vegetative cells, spores, or components thereof as applicable to the candidate method. The target concentration for testing is 10 times the AMQL for the target of the method. Test one replicate per strain or substance using the candidate method.

Inclusivity and exclusivity evaluations shall be performed together as one study. Inclusivity and exclusivity test samples must be blind-coded and intermingled so the analysts do not know the identity or concentration of the test samples.

5.1.2.3 Data Analysis and Reporting

Analyze the data for a positive or negative response. The data is reported as number of species/strains/substances detected. For example, "Of the 30 specific inclusivity strains tested, 28 were detected and two were not detected. Those strains not detected were the following: …" or "Of the 30 specific exclusivity strains tested, 27 were not detected and three were detected. Those detected were the following: …"

The study report should include a table titled "Inclusivity/Exclusivity Panel Results," which lists all species/strains/substances tested, their source, origin and essential characteristics, and testing outcome.

5.1.3 Laboratory Matrix Study

The purpose of the laboratory matrix study is to measure the linearity, repeatability, and bias of the candidate method in a controlled laboratory setting for all matrixes claimed in the intended use statement.

5.1.3.1 Reference Method

If there is a reference method, then the candidate method performance parameters should be compared to those published for the reference method. Designation of an appropriate reference method will be determined by the Methods Committee on Biological Threat Agents.

5.1.3.2 Matrix Categories

AOAC recognizes claims for only the range of matrix categories or specific matrix types included in the method developer study and the collaborative study. The number of different matrixes to be tested depends on the intended use of the method.

Matrixes include, but are not limited to, environmental surfaces (such as stainless steel, plastic, ceramic/glass, nonporous rubber, food-grade painted surfaces, finished wood, sealed concrete), filter samples, liquid air collection samples, bulk powders, and soils (loam, sandy, clay, subsoil, and silt).

5.1.3.3 Study Design

5.1.3.3.1 Contamination Levels

Three levels (high, medium, and low) inoculated with a target agent and one level inoculated with a near neighbor are required. The low target level should be the AMQL or lower, and the medium and high levels should be chosen to span the quantitative range of the method. Inoculate the nontarget sample with a near neighbor at 10 times the AMQL.

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

5.1.3.3.2 Number of Test Portions

Five test portions per level (inoculated and uninoculated) are analyzed per method.

5.1.3.3.3 Agent for Matrix Inoculation

The matrix inoculation should be conducted with a pure culture of one strain or pure analyte. The inoculum should be in a form that is to be detected. The same strain/substance must be used for all matrixes and will be designated by the Methods Committee on Biological Threat Agents. The strains/substances specified for use must be traceable to the source. The source and origin should be documented.

5.1.3.3.4 Preparation of Samples

Samples must be prepared in a consistent manner for all similar studies. Consideration should be given to spore counts versus vegetative cell counts; stage of growth of cells; cell preparation procedures; protein concentration procedures; DNA concentration procedures; relationship between cell/spore count and DNA concentration; and determination of genome equivalents based on DNA concentration. Consult with the Methods Committee on Biological Threat Agents for detailed procedures.

5.1.3.3.4.1 Environmental Surfaces

Dilute target agent in prescribed medium and spread evenly across designated surface at target concentration levels. Allow to dry prior to sampling. Dilute near neighbor in prescribed medium and spread evenly across a separate designated surface at 10 times AMQL. Follow sample collection procedures as defined in AOAC *Official Method*^{SM} **2006.04** (4).

5.1.3.3.4.2 Air Collection Samples, Filters

Dilute agent in H_2O and apply to clean filter at target concentration levels. Allow to dry prior to testing. Inoculate filters with a near neighbor organism at 10 times the AMQL to serve as negative controls.

5.1.3.3.4.3 Air Collection Samples, Aqueous

Dilute agent in device collection liquid at target concentration levels. Inoculate collection liquid with a near neighbor organism at 10 times the AMQL to serve as negative controls.

5.1.3.3.4.4 Soils and Bulk Powders

Dry matrixes such as soils and powders should be inoculated with dry or lyophilized agent. Add dried agent to matrix and mix thoroughly into test sample to achieve target concentration levels. Add one dried near neighbor and mix thoroughly to matrix to serve as negative controls.

5.1.3.3.5 Confirmation of Test Portions

If the candidate method includes presumptive and confirmatory phases, then follow the method as written.

5.1.3.4 Data Analysis and Reporting

5.1.3.4.1 General Considerations

Data often do not show a statistically normal distribution. In order to normalize the data, perform a logarithmic transformation on the reported CFU/unit or spores/unit (including any zero results) as follows:

$$\text{Log}_{10}\,[\text{CFU/unit} + (0.1)f]$$

where f is the reported CFU/unit corresponding to the smallest reportable result, and unit is the reported unit of measure (e.g., g, mL, filter). For details, *see Annex G*.

5.1.3.4.2 Initial Review of Data: Linearity

Plot the candidate method result versus the agent concentration. The vertical *y*-axis (dependent variable) is used for the candidate method and the horizontal *x*-axis (independent variable) for the agent concentration. This independent variable *x* is considered to be accurate and have known values. Usually major discrepancies will be apparent.

5.1.3.4.3 Outliers

It is often difficult to make reliable estimations (average, standard deviation, etc.) with a small bias and in presence of outliers. Data should be examined to determine whether there exists an occasional result that differs from the rest of the data by a greater amount than could be reasonably expected or found by chance alone. Perform outlier tests (Cochran and Grubbs) in order to discard significantly outlying values (10). There must be an explanation for every excluded result; no results can be excluded on a statistical basis only. To view the data adequately, construct a stem-leaf display, a letter-value display, and a box plot (11).

Results excluded as outliers should not be included in the statistical analysis.

5.1.3.4.4 Repeatability (s_r)

The repeatability is within test site precision, designated s_r, or the standard deviation from independent replicate results obtained by the candidate method on identical test material, under the same conditions (e.g., apparatus, operator, test site, and incubation time).

5.1.3.4.5 Bias

The bias is the difference between the mean of transformed values from replicate analyses by the candidate method at one concentration and the known inoculation level. Report the bias with a 95% confidence interval.

5.1.4 Intended Use Matrix Study (Required for Intended Uses Outside the Laboratory)

The purpose of the intended use matrix study is to measure the linearity, repeatability, and bias of the candidate method under simulated intended use conditions using representative personnel for all matrixes claimed in the intended use statement.

5.1.4.1 Matrixes

Linearity, bias, and repeatability for each matrix claimed by the method developer are estimated under simulated intended use conditions using representative personnel. All matrixes for all claimed intended uses must be tested.

5.1.4.2 Study Design

The study design for the intended use matrix study is the same as for the laboratory matrix study. Contamination levels, number of test portions, agent for inoculation, preparation of samples, confirmation of test portions, and data analysis and reporting are found in Section **5.1.3**.

5.1.5 Environmental Interference Study

Biological threat agent methods are likely to be used in a variety of uncontrolled environments. Biological threat agent methods must be impervious to contaminants likely to be found in the environment. All candidate biological threat agent methods must be investigated for potential interference from a wide variety of contaminants likely to be found in the environment. Interference from environmental contaminants may show up as cross-reactivity or inhibition. Both types of interference must be studied.

The AOAC Methods Committee on Biological Threat Agent s shall determine the contaminants to be used for each class of analytical methodology (i.e., methods based on the polymerase chain reaction techniques would illustrate a class of analytical methodology) or intended use. Established environmental panels are found in the SMPRs (3).

5.1.5.1 Study Design

5.1.5.1.1 Contamination Levels

Environmental panel organisms are tested at 10 times the AMQL in the absence of target agent and in the presence of target agent at two times the AMQL. Environmental panel organisms can be tested in pools of no more than 10 organisms per pool, with each organism represented at 10 times the AMQL

Environmental panel substances are tested according to the panel requirements in the SMPRs (3). Test each substance in the presence of agent at two times the AMQL and in the presence of a near neighbor at 10 times the AMQL.

5.1.5.1.2 Number of Replicates

Environmental panel organisms and substances are initially tested without replicates in the presence of agent or near neighbor.

Any observed interference may be followed by retesting using five replicates of agent at two times the AMQL with and without interferent added to determine if the interferent affects bias or precision or both. If a pool of organisms yields unexpected results, each organism must be retested individually to determine the cause of the unexpected results. This retest must follow the retest study design.

5.1.6 Robustness Study

Minor, reasonable variations in a method of a magnitude that might well be expected to occur when the method is used in the "field" are deliberately introduced and tested. Variations in method parameters that can be influenced by the end user should be tested.

The method developer is expected to make a good faith effort to determine which parameters are most likely to affect the analytical performance and determine the range of variations that can occur without adversely affecting analytical results.

5.1.6.1 Analyte Concentration

Robustness strains/substances are prepared and analyzed as vegetative cells, spores, or components thereof as applicable to the candidate method. One agent strain/substance is tested at two times the AMQL and at 10 times the AMQL. Use a near neighbor at 10 times the AMQL as a negative control.

5.1.6.2 Study Design

Five replicates of each agent concentration and near neighbor are tested for each variation.

5.1.6.3 Data Analysis and Reporting

The results are analyzed for effects on bias and repeatability. Standard deviations (s_r) at each concentration are compared to determine if any robustness parameter value causes more than a three-fold increase in s_r.

5.1.7 Product Consistency and Stability Study

The product consistency and stability study is a study or set of studies designed to ensure that the performance of the product is consistent from lot-to-lot and over time under normal storage conditions for the shelf life of the product. These studies are required for PTM certification.

5.1.7.1 Analyte Concentration

Strains/substances for product consistency and stability testing are prepared and analyzed as vegetative cells, spores, or components thereof as applicable to the candidate method. One agent strain is tested at two times the AMQL and at 10 times the AMQL. Use a near neighbor at 10 times the AMQL as a negative control.

5.1.7.2 Study Design

For the lot-to-lot consistency study, at least six lots of product must be tested and must show consistent results.

For product stability, either real-time or accelerated stability (9) data will be accepted for initial PTM certification. If accelerated stability data are submitted, real-time stability data must be submitted within 1 year to support the shelf life of the kit. For real-time stability, data estimating with confidence intervals the bias and repeatability between time zero and the end of the shelf life must be submitted.

Five replicates of each agent concentration and near neighbor are tested for each lot or time point.

Lot-to-lot and stability testing can be combined into one study by testing at least six production lots whose ages span the shelf life of the product.

5.1.7.3 Data Analysis and Reporting

The results are analyzed for variable bias and repeatability between lots or time points. Report the mean and standard deviation for each set of replicates at each concentration for each product lot or stability time point and between lots. Standard deviations (s_r) at each concentration are compared to determine if the s_r varies by more than three-fold.

For the combined approach, use a generalized mixed model fitting program to determine the lot to lot variability as a random effect and the time slope as a fixed effect.

5.1.8 Between-Instrument Variation Study (if applicable)

This study applies only to methods that require the use of an instrument. The between-instrument variation study is a study designed to ensure that the performance of the method is consistent from instrument to instrument. This study is required for PTM certification.

5.1.8.1 Analyte Concentration

Strains/substances for between instrument variability testing are prepared and analyzed as vegetative cells, spores, or components thereof as applicable to the candidate method. One agent strain/substance is tested at two times the AMQL and at 10 times the AMQL. Use a near neighbor at 10 times the AMQL as a negative control.

5.1.8.2 Study Design

At least six instruments must be tested and show consistent results. Five replicates of each agent and near neighbor concentration are tested for each instrument.

5.1.8.3 Data Analysis and Reporting

The results are analyzed for variable bias and repeatability between instruments. Report the mean and standard deviation for each set of replicates at each concentration for each instrument. Standard deviations (s_r) at each concentration are compared to determine if the sr varies by more than three-fold.

5.2 Independent Validation Study

5.2.1 Scope

The independent validation study, preferably conducted under PTM guidance, should verify the analytical results obtained by the method developer. The study traditionally verifies repeatability in the hands of an independent trained user in a controlled laboratory setting. The independent study also determines the performance of the method under simulated intended use conditions using representative personnel. In addition, inclusivity/exclusivity and environmental interference will be repeated in the independent validation study to verify the method developer's data.

All biological threat agent methods shall be validated under conditions that simulate the intended use. For example, a biological threat agent method intended for use by trained emergency response personnel in personal protective equipment at an outdoor mobile site must be validated under those or similar conditions. Biological threat agent methods intended for a variety of uses shall be validated for all combinations of the intended use to be claimed as *Official Methods of Analysis*℠.

Only those intended uses successfully validated may be claimed as *Official Methods of Analysis*℠.

5.2.2 Inclusivity/Exclusivity Study

5.2.2.1 Analyte Variant Selection

The inclusivity panel should be sufficiently large that the analyte variation is adequately represented. The Methods Committee on Biological Threat Agents, with the advice of subject matter experts, shall determine the specific analyte variants required. Typically, 30 variants will be required, if available. The panel shall be the same as that used in the method developer study, whenever possible.

Exclusivity organisms and/or substances include near neighbors, chosen to adequately cover potentially cross-reactive organisms or substances. The Methods Committee on Biological Threat Agents, with the advice of subject matter experts, shall determine the specific species, strains, or substances required for exclusivity evaluation. Typically, 30 will be required, if available. The panel shall be the same as that used in the method developer study, whenever possible.

Species/strains/substances specified for use should be traceable to the source. The source and origin of each species/strain/substance must be documented. Certificate of analysis or other documentation of analyte identity should be on file and available to the Methods Committee on Biological Threat Agents.

5.2.2.2 Study Design

Inclusivity strains/substances are prepared and analyzed as vegetative cells, spores, or components thereof as applicable to the candidate method. The target concentration for testing is the AMQL. Test one replicate per strain/substance using the candidate method.

Exclusivity species/strains/substances are prepared and analyzed as vegetative cells, spores, or components thereof as applicable to the candidate method. The target concentration for testing is 10 times the AMQL. Test one replicate per strain/substance using the candidate method.

Inclusivity and exclusivity evaluations shall be performed together as one study. Inclusivity and exclusivity test samples must be blind-coded and intermingled so the analysts do not know the identity or concentration of the test samples.

If a pool of exclusivity organisms yields positive results, each organism must be retested individually. All test and retest data must be reported.

5.2.2.3 Data Analysis and Reporting

Analyze the data for a positive or negative response. The data is reported as number of species/strains/substances detected. For example, "Of the 30 specific inclusivity strains tested, 28 were detected and two were not detected. Those strains not detected were the following: …" or "Of the 50 specific exclusivity strains tested, 47 were not detected and three were detected. Those detected were the following: …"

The study report should include a table titled "Inclusivity/Exclusivity Panel Results," which lists all species/strains/substances tested, their source, origin and essential characteristics, and testing outcome.

5.2.3 Laboratory Matrix Study

The purpose of the laboratory matrix study is to test all matrixes claimed by the method developer in a controlled laboratory setting to confirm the bias and repeatability of the candidate method as determined by the method developer.

5.2.3.1 Study Design

The study design for validation of quantitative methods at an independent test site follows the method developer validation study design. Contamination levels, number of test portions, source of contamination, preparation of samples, confirmation of test portions, and data analysis and reporting are found in Section **5.1.3**.

5.2.3.2 Reference Method

If there is a reference method, then the candidate method performance parameters should be compared to those published for the reference method. Designation of an appropriate reference method will be determined by the Methods Committee on Biological Threat Agents.

5.2.4 Intended Use Matrix Study (Required for Intended Uses Outside of Laboratory)

The purpose of the intended use matrix study is to measure the bias and repeatability for all matrixes claimed under simulated intended use conditions with a representative end user(s).

5.2.4.1 Study Design

The study design for validation of quantitative methods at an independent test site follows the method developer validation

study design. Contamination levels, number of test portions, source of contamination, preparation of samples, confirmation of test portions, and data analysis and reporting are found in Section **5.1.3**.

5.2.5 Environmental Interference Study

Biological threat agent methods are likely to be used in a variety of uncontrolled environments. Biological threat agent methods must be impervious to contaminants likely to be found in the environment. All candidate biological threat agent methods must be investigated for potential interference from a wide variety of contaminants likely to be found in the environment. Interference from environmental contaminants may show up as cross-reactivity or inhibition. Both types of interference must be studied.

The AOAC Methods Committee on Biological Threat Agents shall determine the contaminants to be used for each class of analytical methodology (i.e., methods based on the polymerase chain reaction techniques would illustrate a class of analytical methodology) or intended use. Established environmental panels are found in the SMPRs (3).

5.2.5.1 Study Design

5.2.5.1.1 Contaminant Concentrations

Environmental panel organisms are tested at 10 times the AMQL in the absence of target agent and in the presence of target agent at two times the AMQL. Environmental panel organisms can be tested in pools of no more than 10 organisms per pool, with each organism represented at 10 times the AMQL.

Environmental panel substances are tested according to the panel requirements in the SMPRs (3). Test each substance in the presence of agent at two times the AMQL and in the presence of a near neighbor at 10 times the AMQL.

5.2.5.1.2 Number of Replicates

Environmental panel organisms and substances are initially tested without replicates in the presence of agent and near neighbor.

Any observed interference may be followed by retesting using five replicates of agent at two times the AMQL with and without interferent added to determine if the interferent affects bias or precision or both. If a pool of organisms yields unexpected results, each organism must be retested individually to determine the cause of the unexpected results. This retest must follow the retest study design.

5.3 Collaborative Study

5.3.1 Scope

The collaborative study is a requirement for *Official Methods*℠ and succeeds the PTM independent validation study. The collaborative study characterizes the performance parameters (e.g., bias, repeatability, reproducibility) of the candidate method across collaborators.

All biological threat agent methods shall be validated under conditions that simulate the intended use using representative personnel. For example, a biological threat agent method intended for use by trained emergency response personnel in personal protective equipment at an outdoor mobile site must be validated under conditions that simulate the outdoor mobile site and include representative personnel as collaborators. Biological threat agent methods intended for a variety of uses shall be validated for all combinations of the intended use to be claimed as *Official Methods of Analysis*℠.

Collaborators shall receive three types of training in conjunction with the study. They shall have at least 3 h of training on AOAC method processes. If the study follows the alternative collaborative study design, then collaborators shall receive training on specific facility requirements from the laboratories hosting the collaborators. All collaborators shall be trained by the vendor on the test method under study. All training shall be documented.

A trial run should be conducted prior to the initiation of the validation study. The purpose of the trial run is to ensure that logistics, sample handling and data reporting processes are worked out and understood by all of the collaborators. A small sample set (e.g., 2–4 samples) using an avirulent or nontoxic surrogate, if possible, should be analyzed. The trial run should be conducted under the same conditions as the validation study. A reasonable amount of time should be allotted for troubleshooting after the completion of the trial run, including a discussion with all of the collaborators to address issues and answer questions. The data should not be analyzed or included in the validation report.

5.3.2 Matrix Study

Bias, repeatability, and reproducibility are estimated for all claimed matrixes in the intended use statement under simulated intended use conditions by representative personnel.

5.3.2.1 Number of Collaborators

The collaborative study must include a minimum of 12 collaborators. After eliminating data sets for assignable cause, the study must yield at least eight valid data sets.

The study must include a minimum of three test sites with no more than four collaborators at any one test site. Care must be taken to ensure the independence of collaborators at each test site.

5.3.2.2 Contamination Levels

Divide each matrix into four samples. Three levels (high, medium, and low) inoculated with a target agent and one level inoculated with a near neighbor are required. The low target level should be the AMQL or lower, and the medium and high levels should be chosen to span the quantitative range of the method. Inoculate the nontarget sample with a near neighbor at 10 times the AMQL.

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

5.3.2.3 Number of Test Portions

The number of test portions per sample per matrix per test site is two. Test portions are to be randomized and blind-coded when sent to participating test sites for analysis.

5.3.2.4 Agent for Matrix Inoculation

See Section **5.1.3.3.3**.

5.3.2.5 Preparation of Samples

See Section **5.1.3.3.4**.

5.3.2.6 Confirmation of Test Portions

If the candidate method includes presumptive and confirmatory phases, then follow the method as written.

5.3.2.7 Data Analysis and Reporting

5.3.2.7.1 General Considerations

Data often do not show a statistically normal distribution. In order to normalize the data, perform a logarithmic transformation on the reported CFU/unit or spores/unit (including any zero results) as follows:

$$\text{Log}_{10}\,[\text{CFU/unit} + (0.1)f]$$

where *f* is the reported CFU/unit or spores/unit corresponding to the smallest reportable result, and unit is the reported unit of measure (e.g., g, mL, or filter). For details, *see Annex G*.

5.3.2.7.2 Initial Review of Data

Construct a Youden plot. For a given matrix–level–laboratory combination, plot replicate pairs as first replicate versus second replicate. Usually major discrepancies will be apparent: displaced means, unduly spread replicates, outlying values, differences between methods, consistently high or low test site rankings, etc.

Plot the candidate method result versus the agent concentration. The vertical *y*-axis (dependent variable) is used for the candidate method and the horizontal *x*-axis (independent variable) for the agent concentration. This independent variable *x* is considered to be accurate and have known values. Usually major discrepancies will be apparent.

5.3.2.7.3 Outliers

It is often difficult to make reliable estimations (average, standard deviation, etc.) with a small bias and in presence of outliers. Data should be examined to determine whether any test site shows consistently high or low values or an occasional result that differs from the rest of the data by a greater amount than could be reasonably expected or found by chance alone. Perform outlier tests (Cochran and Grubbs) in order to discard significantly outlying values (10). There must be an explanation for every excluded test site or result; no results can be excluded on a statistical basis only. To view the data adequately, construct a stem-leaf display, a letter-value display, and a box plot (11).

Results excluded as outliers should not be included in the statistical analysis.

5.3.2.7.4 Performance Indicators

Performance indicators for quantitative methods include bias, repeatability, and reproducibility of the transformed data.

5.3.2.7.5 Bias

The bias is the difference between the mean of transformed values from replicate analyses by the candidate method at one concentration and the known inoculation level. Report the bias with a 95% confidence interval.

5.3.2.7.6 Repeatability (s_r)

The repeatability is designated s_r, or the standard deviation from independent replicate results obtained by the candidate method on identical test material, under the same conditions (e.g., apparatus, operator, test site, and incubation time).

5.3.2.7.7 Reproducibility (s_R)

The reproducibility is across collaborator precision, designated s_R, or the standard deviation of the mean test results on identical test material using the candidate method and obtained by different collaborators using separate equipment.

5.3.2.7.8 Calculations

For details, refer to OMA Program Manual, Part 9 (10).

6 Biological Threat Agent Procedures

Biological threat agent procedures may include the collection, handling, transport, and/or manipulation of biological threat agents. The evaluation of a procedure differs markedly from the evaluation of a method validation. There are no "analytical" parameters such as "POD" to evaluate for a procedure. The performance characteristics of a procedure are typically: recovery; reproducibility; and stability of the target agent.

Such procedures are often not performed in a laboratory even though the resulting samples may (or may not) go to a laboratory. Therefore the concept of interlaboratory reproducibility does not necessarily apply, whereas the variability contributed by intended users at the intended use sites is paramount and must be determined.

In some cases, collection, handling, transport, or manipulation procedures may be tied to one or more analytical methods. In these cases, the performance characteristics of the procedure can be evaluated: (*1*) separately as a stand-alone procedure, or (*2*) together with an analytical method. Validation of a stand-alone procedure can be of more value because the resulting validation can be matched to other analytical methods when the sample resulting from the collection, handling procedure is equivalent to the starting sample for an analytical method. However, there is no obligation to validate (collection, etc.) procedures separately from methods. The procedures and methods may be validated as one single method.

6.1 Performance Tested Methods℠ Evaluation

There is no PTM Program for evaluation of biological threat agent procedures. All biological threat agent procedures are evaluated under the OMA Program.

6.2 Official Methods of Analysis℠ Evaluation of a Procedure

The purpose of OMA evaluations of a procedure is to confirm that a variety of trained users can perform the procedure equivalently under claimed end-use or simulated end-use conditions.

OMA evaluations also investigate: (*1*) the effectiveness of training; (*2*) the clarity of the method instructions; as well as (*3*) the robustness of the method itself. If significant variability is observed across sites or there appears to be one or two "outlier" sites, the sites should be interviewed to determine if the adverse results are due to deviation from protocol, training failure, unclear protocol instructions, or adverse conditions. It is important to capture comments on the ease of use and other aspects of the method from the multiple sites as these comments may either support the utility of the method or provide recommendations for improvement of the method or method instructions.

6.3 Evaluation Parameters

The range of procedures includes collection, handling, transport, or manipulation procedures. Recovery; reproducibility; and stability of the target agent are the main evaluation parameters for evaluation of a procedure.

7 Nomenclature

Biological threat agent methods awarded PTM status shall be designated as "AOAC *Performance Tested Methods*℠ for Biological Threat Agents."

Biological threat agent methods and/or procedures awarded *Official Methods*℠ status shall be designated as "AOAC *Official Methods of Analysis*℠ for Biological Threat Agents."

8 Publications

AOAC is incorporated and resides as a corporate entity in the United States of America. Therefore, AOAC shall comply with the requirements of U.S. national security. Information generated in the validation of biological threat agent methods and/or procedures are confidential until and unless the appropriate governing agency or agencies have granted permission to distribute information. Certain sections of validation reports, such as the identification of inclusivity strains and exclusivity species, may be confined to annexes and retained for limited distributed subject to the discretion of appropriate governing U.S. agencies.

9 Safety

All manipulations of biological threat agents are required to be performed in accordance with biosafety and biosurety practices stipulated by each institution. Use the equipment and facilities indicated for the test agent. For recommendations on safe handling, refer to the CDC/NIH *Biosafety in Microbiological and Biomedical Laboratories* (BMBL) manual (12). Exact adherence to the method, good laboratory practices, and quality control are required for proficiency and validity of the results.

10 References

(1) http://www.cdc.gov/od/sap/docs/salist.pdf

(2) Wehling, P., LaBudde, R., Brunelle, S., & Nelson, M. (2011) *J. AOAC Int*. **94**, 335–347

(3) Stakeholder Panel on Agent Detection Assays (SPADA) (2011) *J. AOAC Int.* **94**, 1338, 1342, 1347, 1352, 1356

(4) *Official Methods of Analysis*℠ (2011) 18th Ed., AOAC INTERNATIONAL, Gaithersburg, MD, USA, Method **2006.04** (The method was approved and published by ASTM as ASTM E2458-06, Standard Practices for Bulk Sample Collection and Swab Sample Collection of Visible Powders Suspected of Being Biological Agents from Nonporous Surfaces. ASTM E2458-06 was revised in November 2010 to better reflect the needs of the responder and public health communities. The revised sampling standard, published by ASTM as E2458-10, is supported by operational guidance E2770-10, which was approved by ASTM International in October 2010. The ASTM standards are available at http://www.astm.org/standards/e2770.htm and http://www.astm.org/standards/e2458.htm)

(5) Box, G.E.P., Hunter, J.S., & Hunter W.G. (2005) *Statistics for Experimenters: Design, Innovation, and Discovery,* 2nd Ed., Wiley-Interscience

(6) Jiang, J. (2007) *Linear and Generalized Linear Mixed Models and Their Applications*, Springer

(7) Pinheiro, J.C., & Bates, D.M. (2002) *Mixed Effects Models in S and S-Plus,* Springer

(8) LaBudde, R.A. (2009) "Statistical analysis of interlaboratory studies, XXIII. Lot-lot variation and stability testing in a single laboratory study," TR299, Least Cost Formulations, Ltd., Virginia Beach, VA, USA, http://www.lcfltd.com/Documents/tr299%20stat%20anal%20inter%20XXIII%20lot-lot%20variation%20and%20.pdf

(9) Porterfield, R.I., & Capone, J.J. (1984) *Application of Kinetic Models and Arrhenius Methods to Product Stability Evaluations*, Medical Device & Diagnostic Industry, pp 45–50

(10) *Official Methods of Analysis*℠ Program Manual (2010) Part 9, AOAC INTERNATIONAL, Gaithersburg, MD, USA

(11) Velleman, P.F., & Hoaglin, D.C. (1981) *Applications, Basics and Computing of Exploratory Data Analysis*, Duxbury Press, Boston, MA, USA

(12) *Biosafety in Microbiological and Biomedical Laboratories* (BMBL) (2007) 5th Ed., U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention and National Institutes of Health, Washington, DC, USA

(13) LaBudde, R.A. (2009) "Coverage accuracy for binomial proportion 95% confidence intervals for 12 to 100 replicates," TR297, Least Cost Formulations, Ltd., Virginia Beach, VA, USA, http://www.lcfltd.com/Documents/tr297%20coverage%20accuracy%20binomial%20proportions.pdf

(14) LaBudde, R.A. (2009) "Statistical analysis of interlaboratory studies, XXII. Statistical analysis of a qualitative multicollaborator study as a quantitative study under the large sample approximation," TR296, Least Cost Formulations, Ltd., Virginia Beach, VA, USA, http://www.lcfltd.com/Documents/tr296%20stat%20anal%20inter%20XXII%20qual%20study%20char%20v3.pdf

(15) LaBudde, R.A. (2008) "Statistical analysis of interlaboratory studies, XX. Measuring the performance of a qualitative test method," TR290, Least Cost Formulations, Ltd., Virginia Beach, VA, USA, http://www.lcfltd.com/Documents/tr290%20stat%20anal%20inter%20XX%20meas%20perf%20qual%20test%20metho.pdf

ANNEX A
Understanding the POD Model

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

The single parameter of interest is the 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, a high probability of obtaining a positive response

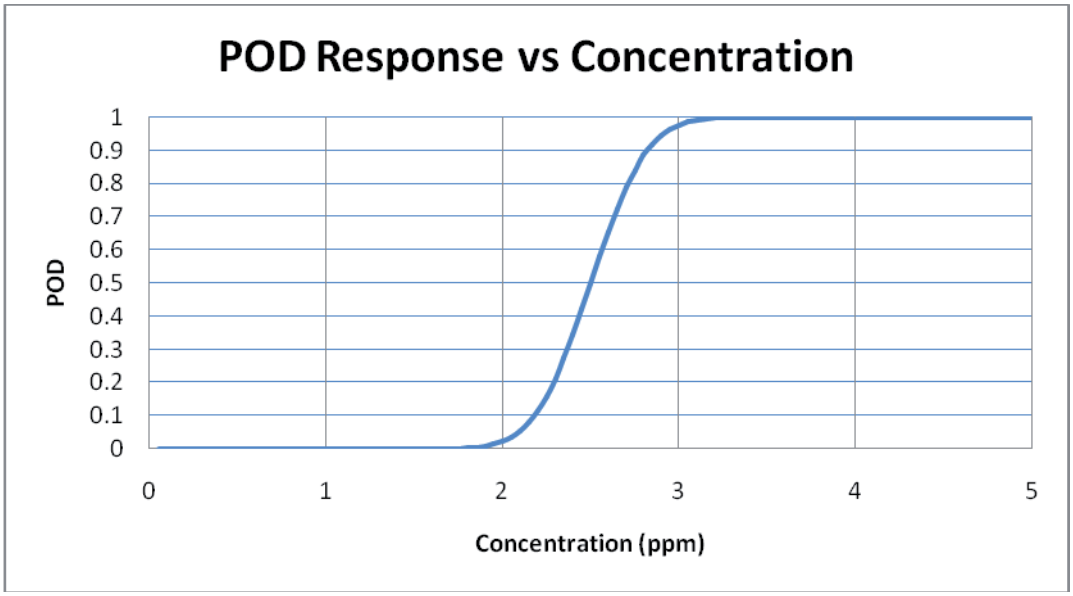

Figure A1. Theoretical POD curve for a qualitative detection method.

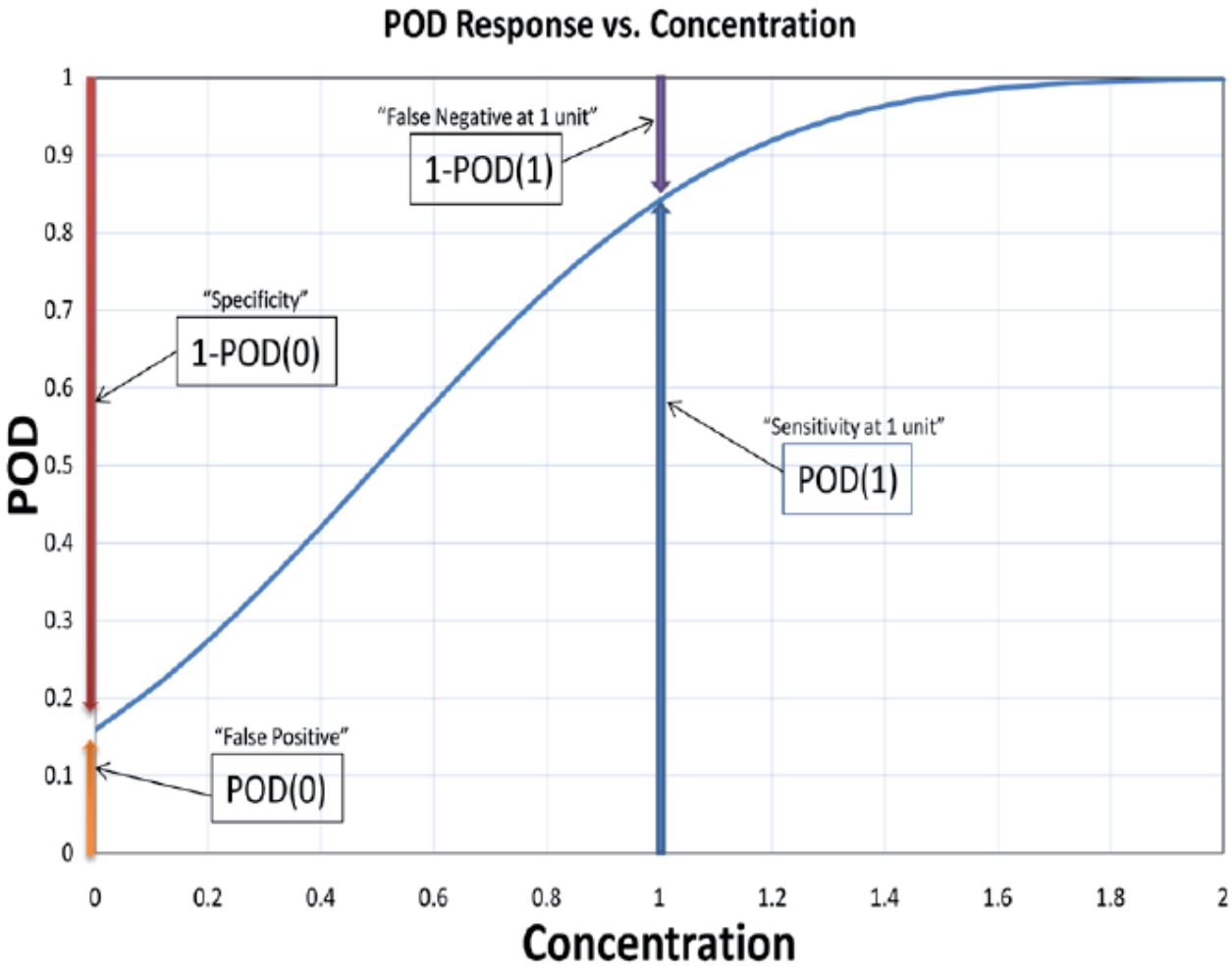

Figure A2. Comparison of the POD model terminology to other obsolete terms. Use of the terms "sensitivity," "specificity," "false positive," and "false negative" is denigrated under the POD model.

Table A1. Terminology denigrated under the POD model

Denigrated 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 concn = 0	The POD curve value at concn = 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

is desired. 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).

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

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

Older validation models have used the terms "sensitivity," "specificity," "false positive," and "false negative" to describe method performance. The POD model has incorporated all of the performance concepts of these systems into a single parameter, POD (*see* Figure A2).

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 concn = 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 (*see* Table A1), which is always assumed to vary by concentration. In other models, the terms "false positive," "false negative," "sensitivity," and "specificity" have been defined in a variety of ways, usually not conditional on concentration. For these reasons, their use is denigrated under this model.

ANNEX B
Raw Format Data Table Template and Example for Qualitative Method: Method Developer Studies, Independent Studies, and Collaborative Studies

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

Each row (record) in the raw format data table should contain the following columns (fields):

(*1*) *Matrix type*.—An identifier indicating the matrix involved, such as "FILTERS." The same exact identifier must be used for the same matrix.

(*2*) *Concentration level*.—The concentration/test portion for the level.

(*3*) *Test site*.—An identifier uniquely indicating the test site involved, such as "S1."

(*4*) *Collaborator team*.—An identifier uniquely specifying the collaborator team across test sites, e.g., "C01."

(*5*) *Instrument*.—An identifier uniquely specifying the apparatus used in testing, across test sites and collaborator teams, e.g., "I01."

(*6*) *Method*.—An identifier indicating the test method used, such as "R" for the reference method, "CP" for the candidate presumptive method, or "CC" for the candidate confirmation method.

(*7*) *Replicate*.—A unique identifier for the test portion involved. If this identifier is common to two rows in the table, this implies

the results are matched by test portion. Example identifiers might be "01," "001," or "A1."

(*8*) *Result*.—"0" for absence or "1" for presence (detection).

In computer format, the raw format data table should be given either as: (*1*) a "fixed-format" file with fixed column widths and blanks or tabs as separators and a file extension of ".txt" or ".xls"; or (*2*) a "comma-separated value" file with commas as separators between columns and identifiers within quotes, and a file extension of ".csv".

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

An example file named "banthracis.csv" might be:

"matrix", "level", "site", "collab", "instrument", "method", "replicate", "result"

"filter", "2.20", "S1", "C01", "I01", "CP", "001", 0

"filter", "2.20", "S1", "C01", "I01", "CC", "002", 1

"filter", "2.20", "S1", "C01", "I01", "R", "003", 1

"filter", "2.20", "S1", "C01", "I01", "CP", "004", 1

"filter", "2.20", "S1", "C01", "I01", "CC", "005", 1

"filter", "2.20", "S1", "C01", "I01", "R", "006", 1

etc.

ANNEX C
Calculation of POD and dPOD Values from Qualitative Method: Method Developer and Independent Validation Data

In general, four different probabilities of detection (PODs) are to be calculated: POD_R (for the reference method), POD_C (for the confirmed candidate method), POD_{CP} (for the candidate presumptive method), and POD_{CC} (for the candidate confirmation method).

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

$$POD = x/N$$

where POD is POD_C, POD_R, etc.

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

(*1*) For the case where $x = 0$,

$$POD = 0$$

$$LCL = 0$$

$$UCL = \frac{3.8415}{N + 3.8415}$$

(*2*) For the case where $x = N$,

$$POD = 1$$

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

$$UCL = 1$$

(*3*) For the case where $0 < x < N$,

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

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

The confidence interval corresponds to the uncorrected Wilson-score method, modified for $x = 1$ and $x = N$–1 to improve coverage accuracy on the boundary (13).

The differences in proportions detected are estimated by (14):

$$dPOD_C = POD_C - POD_R$$

$$dPOD_{CP} = POD_{CP} - POD_{CC}$$

The associated 95% confidence interval (LCL, UCL) for the expected value of $dPOD = POD_1 - POD_2$ is estimated by:

$$LCL = d\,POD - \sqrt{(POD_1 - LCL_1)^2 + (POD_2 - UCL_2)^2}$$

$$UCL = d\,POD + \sqrt{(POD_1 - UCL_1)^2 + (POD_2 - LCL_2)^2}$$

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

Example:

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

$$POD = 1.00$$

$$LCL = N/(N + 3.8415)$$
$$= 96/99.8415$$
$$= 0.962$$

$$UCL = 1.000$$

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

$$POD = 95/96$$
$$= 0.990$$

$$LCL = \{x + 1.9207 - 1.9600 \sqrt{[x - x^2/N + 0.9604]}\}/[N + 3.8415]$$
$$= \{96.9207 - 1.9600 \sqrt{1.9500}]\}/99.8415$$
$$= 0.943$$

$$UCL = 1.000 \text{ (because } x = N - 1)$$

Suppose finally that $x = 94 = N - 2$. Then

$$POD = 94/96$$
$$= 0.979$$

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

$$UCL = \{x + 1.9207 + 1.9600 \sqrt{[x - x^2/N + 0.9604]}\}/[N + 3.8415]$$
$$= 0.994$$

Table D1. Example data summary table: Detection of *Bacillus anthracis* spores on nitrocellulose aerosol collection filters by PCR[a]

Organism	Concentration	*N*	*x*	POD_C	95% LCL	95% UCL
B. anthracis Ames	20,000 spores/filter	96	95	0.9896	0.9433	1.0000
B. cereus E33L	200,000 spores/filter	96	1	0.0104	0.000	0.0567

[a] *N* = number of replicates tested; *x* = number of positive outcomes; and POD_C = POD for the confirmed candidate method.

ANNEX D
Summary Data Table for Qualitative Method: Method Developer and Independent Validation Studies

Prepare one table per matrix and indicate laboratory or intended use conditions (*see* Table D1).

ANNEX E
Calculation of CPOD and dCPOD Values from Qualitative Method Collaborative Study Data

Calculations are done in four steps.

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

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

Step 3.—Both s_C and s_r are used to estimate a 95% confidence interval for the expected value of CPOD. Also estimated is the reproducibility standard deviation s_R, which is the standard deviation of measurement of a single *x* from a single collaborator, including both the collaborator–collaborator and repeatability error sources (15). A 95% confidence interval is estimated for the expected value of s_R.

Step 4.—A *t* statistic based on the χ^2 distribution or Fisher randomization test of homogeneity of POD across laboratories is performed to determine if the observed collaborator effect is detectably greater than zero.

CPOD

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

$$\text{CPOD} = \frac{\sum x_i}{\sum n_i} = \frac{x}{N} \qquad i = 1, 2, 3, ..., C$$

where $N = \Sigma\, n_i$ is the total number of data, $x = \Sigma\, x_i$ is the total number of positive detections, and *C* is the number of collaborators.

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

Repeatability Standard Deviation (s_r)

Estimate the repeatability standard deviation:

$$s_r^2 = \frac{\sum_{i=1}^{C}(n_i - 1)s_i^2}{\sum_{i=1}^{C}(n_i - 1)} = \frac{\sum\left[x_i - \left(x_i^2 / n_i\right)\right]}{N - C}$$

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

where s_i^2 is the variance of the results from collaborator *i*, x_i is the number of positive detections from collaborator *i*, n_i is the number of observations from collaborator *i*, *N* is the total number of data, and *C* is the number of collaborators.

Collaborator Effect Standard Deviation (s_C)

Estimate the collaborator effect standard deviation:

$$s_C^2 = \max\left\{0,\ s_{\text{POD}}^2 - \frac{s_r^2}{n}\right\}$$

where

$$n = \left(N - \frac{\sum n_i^2}{N}\right)\left(\frac{1}{C-1}\right)$$

is the weighted average number of replicates per collaborator and

$$s_{\text{POD}}^2 = \frac{\sum(\text{POD}_i - \text{CPOD})^2}{C-1}$$

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

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

Reproducibility Standard Deviation (s_R)

Calculate the reproducibility standard deviation:

$$s_R^2 = s_r^2 + s_C^2$$

If the collaborator effect s_C is near zero, then

$$s_r \approx s_R \approx \sqrt{\text{CPOD}(1 - \text{CPOD})}$$

Confidence Intervals for CPOD, s_r, s_C, and s_R

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

For CPOD, calculate 95% confidence limits as follows:

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

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

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

If CPOD <0.15 or CPOD > 0.85:

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

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

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

If CPOD = 0:

$$\text{LCL} = 0$$

$$\text{UCL} = 3.8415/(N + 3.8415)$$

If CPOD = 1:

$$\text{LCL} = N/(N + 3.8415)$$

$$\text{UCL} = 1$$

dCPOD Estimates and Confidence Intervals

The dCPOD statistics are differences between the CPOD estimates:

$$\text{dCPOD}_C = \text{CPOD}_C - \text{CPOD}_R$$

$$\text{dCPOD}_{CP} = \text{CPOD}_{CP} - \text{CPOD}_{CC}$$

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

$$s_{d\text{POD}} = \sqrt{\frac{\Sigma(d\text{POD}_i - d\text{CPOD})^2}{C - 1}}$$

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

Test of Intercollaborator Variability

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

$$T = \sum\left[\frac{\left(x_i - n_i\, x/N\right)^2}{n_i\, x/N} + \frac{\left(n_i - x_i - \frac{n_i(N-x)}{N}\right)^2}{\frac{n_i(N-x)}{N}}\right]$$

$$= \sum\frac{\left(x_i - n_i \cdot \text{CPOD}\right)^2}{n_i \cdot \text{CPOD}\left(1 - \text{CPOD}\right)}$$

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

Example:

Suppose the reference method in an interlaboratory study gave the results in Table E1 when 12 replicate test portions were tested by each of 10 collaborators.

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

Table E1. Example: Reference method results in an interlaboratory study

Lab	Positive	Negative	Total	POD
1	7	5	12	0.5833
2	9	3	12	0.7500
3	6	6	12	0.5000
4	10	2	12	0.8333
5	5	7	12	0.4167
6	7	5	12	0.5833
7	5	7	12	0.4167
8	7	5	12	0.5833
9	11	1	12	0.9167
10	9	3	12	0.7500
All	76	44	120	

Table E2. Example: Summary of statistical results for an interlaboratory study

Parameter	Value
CPOD	0.6333
s_r	0.4735
s_C	0.1046
s_R	0.4850
p-Value for *t*-test	0.1703

The repeatability standard deviation

$$s_r^2 = \frac{\sum\left[x_i - \left(x_i^2 / n_i\right)\right]}{N - C} = \frac{\left[\left(7 - 49/120\right) + \left(9 - 81/120\right) + \ldots + \left(9 - 81/120\right)\right]}{120 - 10}$$
$$= 0.2242$$
$$s_r = \sqrt{s_r^2} = \sqrt{0.2242} = 0.4735$$

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

The among-collaborator standard deviation is

$$s_C^2 = \max\left\{0,\ \frac{\sum(\text{POD}_i - \text{CPOD})^2}{C-1} - \frac{s_r^2}{n}\right\}$$
$$= \max\left\{0,\ \frac{\left[(0.5833 - 0.6333)^2 + \ldots + (0.75 - 0.6333)^2\right]}{10-1} - \frac{0.2242}{12}\right\}$$
$$= \max\{0, 0.02963 - 0.0187\}$$
$$= 0.01093$$

and $s_C = \sqrt{0.01093} = 0.1045$, which is noticeably less than s_r, as expected.

The reproducibility standard deviation is

$$s_R^2 = s_r^2 + s_C^2$$
$$= 0.01093 + 0.2242$$
$$= 0.2351$$

so $s_R = \sqrt{0.2351} = 0.4849 \approx s_r$.

The results are summarized in Table E2.

The "homogeneity test" reported above is the *t* statistic based on the χ^2 distribution, so the *p*-value of 0.1703 should be compared to 0.10 to observe if an intercollaborator effect was detectable. The test indicates the observed value of $s_C = 0.1046$ is not statistically significant, so the study was not large enough to reliably detect an intercollaborator effect of this size.

ANNEX F
Data Summary Table Template and Example for Qualitative Method Collaborative Studies

Prepare one table per matrix (*see* Table F1) and plot CPOD by concentration for each method (*see* Figure F1) and dCPOD by concentration (*see* Figure F2) with 95% confidence interval error bars.

Table F1. Comparative results for the detection of *B. anthracis* Ames on aerosol collection filters by the candidate and reference methods in a collaborative study

	Concentration					Candidate result (C)			Reference method (R)			C vs R
Statistic	CFU/mL	CFU/filter	Site	Collaborator	Instrument	*N*	*x*	POD_C, %	*N*	*x*	POD_R, %	$dCPOD_{C,R}$, %
			01	01	01	12	0	0	12	0	0	0
			01	02	01	12	0	0	12	0	0	0
			01	03	01	12	0	0	12	0	0	0
			—	—	—			—			—	—
			03	10	03	12	0	0	12	0	0	0
Estimate	0.00	0.00		All		120	0	0	120	0	0	0
LCL	0.00	0.00						0			0	–3
UCL	0.00	0.01						3			3	3
			01	01	01	12	11	92	12	12	100	–8
			01	02	01	—	—	—	—	—	—	—
			01	03	01	—	—	—	—	—	—	—
			—	—	—	—	—	—	—	—	—	—
			03	10	03	—	—	—	—	—	—	—
Estimate	1830	18300		All		120	119	99	120	120	100	–1
LCL	400	4000						95			97	–5
UCL	9150	91500						100			100	2
			01	01	01	12	12	100	12	12	100	0
			01	02	01	—	—	—	—	—	—	—
			01	03	01	—	—	—	—	—	—	—
			—	—	—	—	—	—	—	—	—	—
			03	10	03	—	—	—	—	—	—	—
Estimate	21200	212000		All		120	120	100	120	120	100	0
LCL	4240	42400						97			97	–3
UCL	106000	1060000						100			100	3

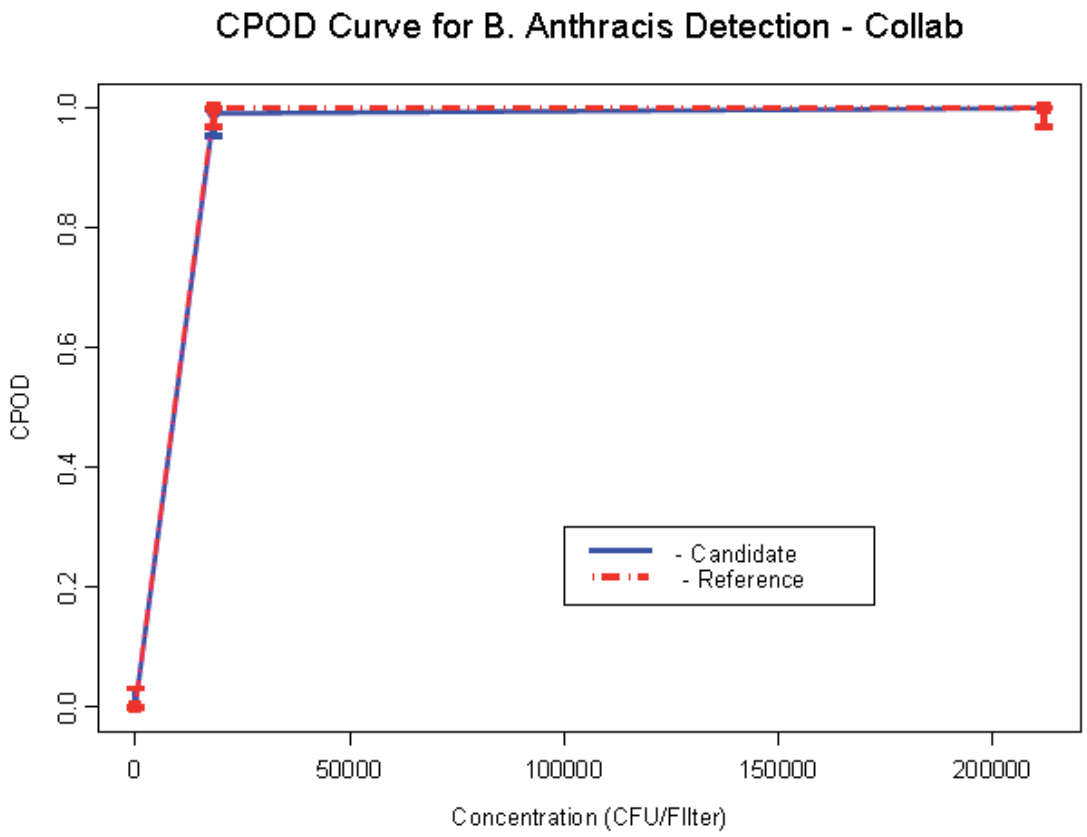

Figure F1. Plot of example CPOD data from Table F1.

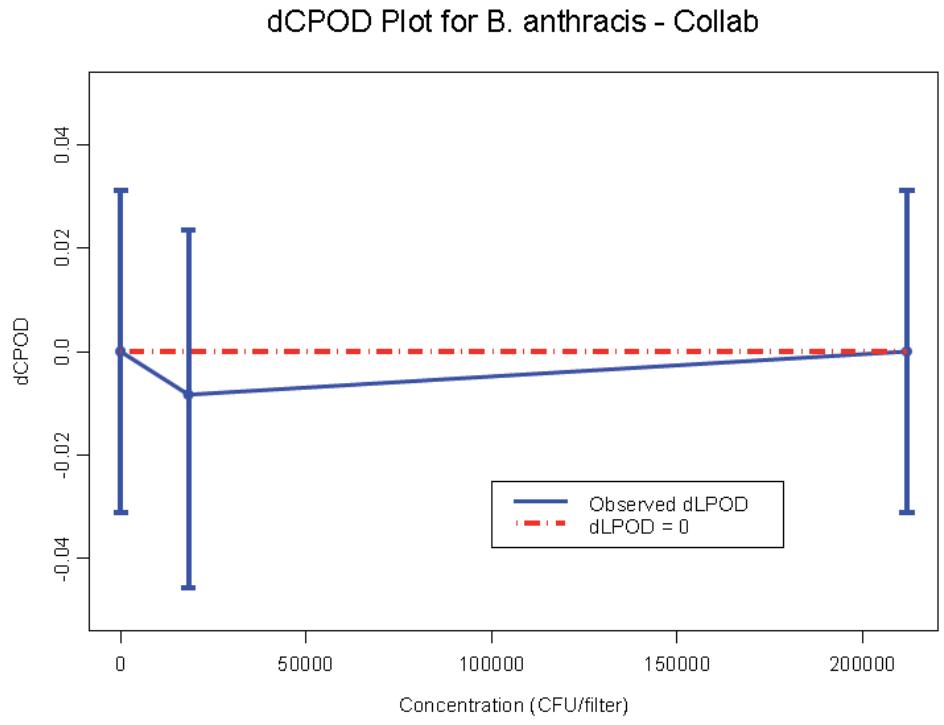

Figure F2. Plot of example dCPOD data from Table F1.

ANNEX G
Logarithmic Transformation of Data from Quantitative Method Developer, Independent, and Collaborative Data

Quantitative microbacterial count data from experiments spanning multiple dilutions often do not show a Poisson nor a Gaussian statistical distribution. When the underlying physical mechanism allows for "clustering," typically a logarithmic transformation will normalize the data.

Perform a logarithmic transformation on the reported CFU/unit or spores/unit (including any zero results) as follows:

$$Y = \log_{10} [\text{CFU/unit} + (0.1)f]$$

where f is the reported CFU/unit or spores/unit corresponding to the smallest reportable result, and "unit" is the reported unit of measure (e.g., g, mL, filter).

Example:

(*1*) For the control concentration, the CFU/g is reported as "<0.003." So CFU/unit = 0.0, and $Y = \log_{10} [0.0 + (0.1)(0.003)] = -3.52$.

(*2*) For the low concentration, the CFU/g is 0.042. So $Y = \log_{10} [0.042 + (0.1)(0.003)] = -1.37$.

(*3*) For the high concentration, the CFU/g is 0.231. So $Y = \log_{10} [0.231 + (0.1)(0.003)] = -0.64$.

ANNEX H
Systems Suitability Requirements for Biological Threat Agent Methods

The controls listed in Table H1 shall be embedded in assays as appropriate. Manufacturer must provide written justification if controls are not appropriate to an assay.

Table H1. Controls

Positive control	Designed to demonstrate an appropriate test response. Positive control should be included at a low, but easily, detectable concentration, and should monitor the performance of the entire assay. The purpose of using a low concentration of positive control is to avoid contamination of the test sample and/or instrument.	Single use per sample (or sample set) run	**Success**: Control detected at expected levels **Failure**: Control not detected or at levels outside of the expected range
Negative control	Designed to demonstrate that the assay itself does not produce a detection in the absence of the target organism. The purpose of this control is to rule out contamination in the assay or test.	Single use per sample (or sample set) run	**Success**: No detections made **Failure:** Detections made
Inhibition control	Designed to specifically address the impact of a sample or sample matrix on the assay's ability to detect the target organism.	Single use per sample run	**Success:** Control detected at expected levels **Failure:** Control not detected or at levels outside of the expected range