Designation: E3131 – 17

Standard Specification for Nucleic Acid-Based Systems for Bacterial Pathogen Screening of Suspicious Visible Powders[1]

This standard is issued under the fixed designation E3131; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

INTRODUCTION

Evaluation of nucleic acid-based detection systems is necessary to ensure that they can achieve required performance metrics for the intended application. These systems should be evaluated in both laboratory and field settings to determine performance, including potential for false positive or negative results, probability of detection (POD), and potential impacts of other substances on system performance such as commonly encountered suspicious powders. Laboratory evaluations establish the best-case performance for a product and also serve as a means to eliminate from consideration those products that have deficiencies or limitations before extensive cost and effort is expended for field testing. Testing should be conducted under conditions recommended by the manufacturer. The statistical derivation used in this specification assumes that conditions during testing remain stable.

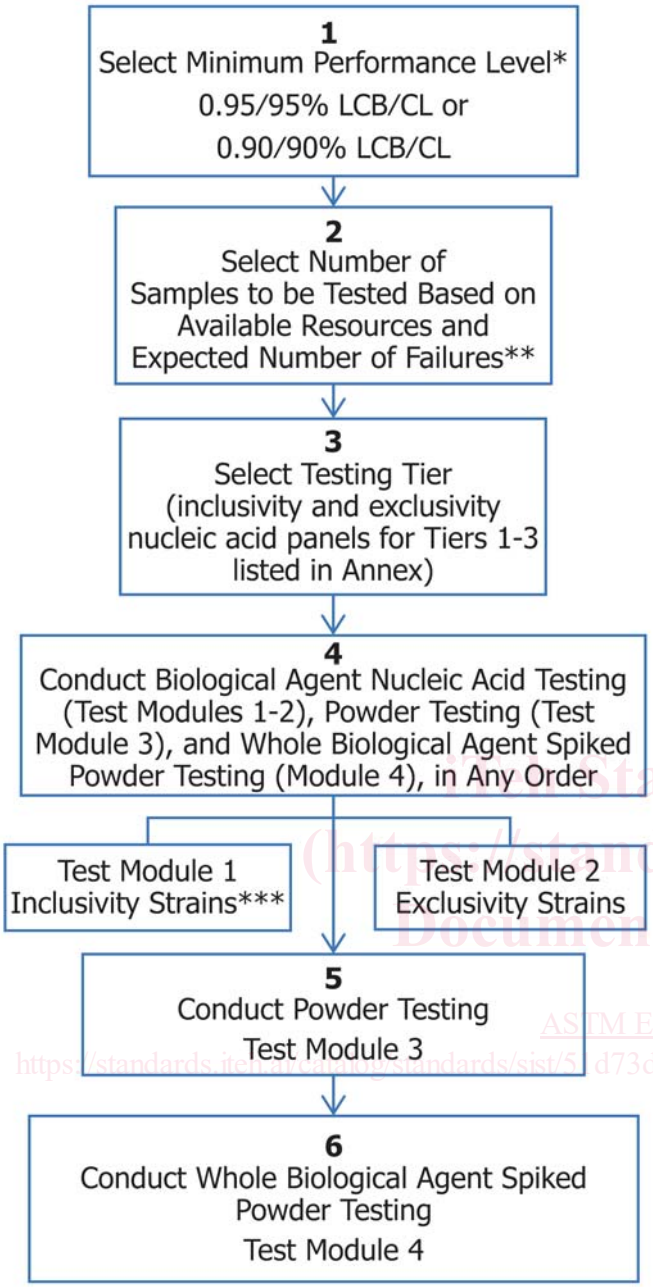
Independent testing of biothreat or biological agent detection systems helps to establish the reliability of results and improves first-responder and supporting agencies' confidence in these tools. It is important that testing requirements balance the need for proven systems with the need for a process that is not cost or time prohibitive and allows the evaluation of new technologies and assays as they are developed. This is particularly true for nucleic acid-based detection systems because new technologies and products continue to emerge on the market and existing assays may be revised which necessitates retesting.

This specification describes a statistically based testing approach for evaluating the performance of nucleic acid-based detection systems. The approach ties performance of the system to a specified lower confidence bound (LCB) on the POD at a known confidence level (CL) (see Fig. 1).

Testing shall be conducted to one of two performance levels (see Figs. 2 and 3): (*1*) ≥95 % POD with 95 % CL, or (*2*) ≥90 % POD with 90 % CL. Four testing modules shall be used to evaluate system performance (see Table 1): (*1*) Test Module 1—Biological agent nucleic acid inclusivity testing; (*2*) Test Module 2—Biological agent nucleic acid exclusivity testing; (*3*) Test Module 3—Suspicious powder testing (commonly encountered hoax powders and environmental material that could interfere with test results, controls, or cause a false positive result); and (*4*) Test Module 4—Whole organism biological agent spiked suspicious powder testing (impact of other material on the ability to detect target biological agents or cause a false negative result). See Table 2 for a listing of suspicious powders and the Annexes for the representative biological agents that shall be tested.

Three different testing tiers are also defined to reduce testing burden by allowing testing of biological agent strain panels with fewer panel members (see Table 1). Inclusivity and exclusivity testing tier panels are provided in Annex A4 and Annex A5. All three testing tiers shall test the full panel of suspicious powders (Table 2) and the whole representative biological agent (see Annex A7) spiked into powders.

While the greatest extensiveness of test panel inclusivity and exclusivity strains and highest POD and CL are always desirable, time and budget constraints often do not permit this extent of testing. While some detection systems may not be able to achieve the highest performance metrics, it is still valuable to know the level to which they can perform.



*LCB/CL—Lower confidence bound/confidence level.

**The number of samples to be tested shall be determined prior to any testing and cannot be modified (for example, if 47 samples are chosen to be analyzed (anticipating no failures) to achieve a 0.95/95 % LCB/CL and a failure occurs, additional samples beyond the original 47 cannot be tested. Testing shall start over).

***Testing Tier 3 does not have any inclusivity strain nucleic acid testing because the representative inclusivity whole biological agent for this tier is tested in Test Module 4.

NOTE 1—Test Modules can be performed in any order.

FIG. 1 Overview of Nucleic Acid-Based Biological Agent Testing

1. Scope

1.1 *General:*

1.1.1 This specification provides system designers, manufacturers, integrators, procurement personnel, end users/practitioners, and responsible authorities a common set of parameters to match the capabilities of biological assessment tools with user needs.

1.1.2 This specification is not meant to provide for all uses. Manufacturers, purchasers, and end users will need to determine specific requirements including, but not limited to, use by hazardous material (HAZMAT) teams and Urban Search and Rescue (US&R) teams, use in explosive or other hazardous environments or atmospheres, use with personal protective equipment (PPE), use by firefighters or law enforcement officers or both, special electromagnetic compatibility needs, extended storage periods, and extended mission time. These specific requirements may or may not be generally applicable to all nucleic acid-based detection systems.

1.2 *Operational Concepts*—Nucleic acid-based detection systems are used to detect, identify, or quantify, or combinations thereof, biological hazards to support short-term tactical decision making to protect responders and the public. The system should provide low false-positive and false-negative rates. Uses of these systems include survey, surveillance, and screening of samples, particularly during a response to a suspected biological agent incident. A field-deployable system should withstand the rigors associated with uses including, but not limited to, high- and low-temperatures and storage conditions, shock and vibration, radio frequency interference, and rapid changes in operating temperature and humidity. Note that this specification does not address testing the potential impact of the rigors associated with use of systems in the field.

1.3 *Nucleic Acid-Based System Detection Capabilities*—Manufacturers or independent third-party testing entities shall document and verify, through testing, the capabilities of the system.

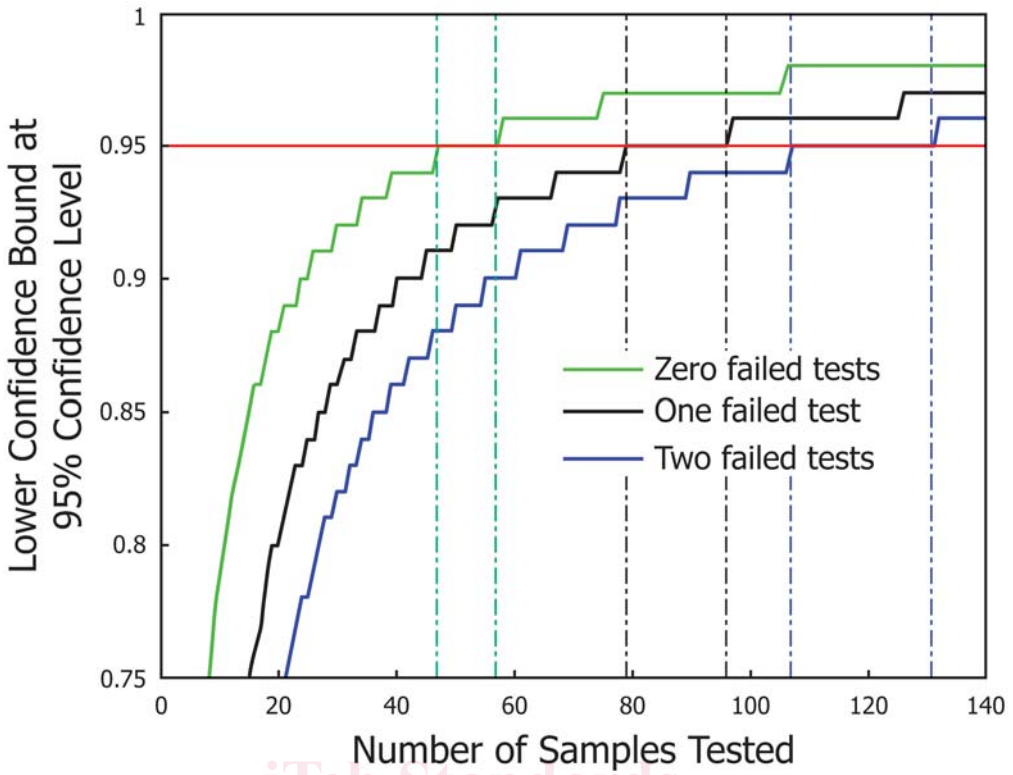
1.4 *Units*—The values stated in SI units are to be regarded as the standard. No other units of measurement are included in this standard. Liquid concentrations of the biohazard materials are presented in number of biological agents or genome equivalents per volume for pathogens such as bacteria and spores (biological agents/mL, genome equivalents/mL (GE/mL), colony forming units/mL (CFU/mL), or spores/mL).

1.5 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety, health, and environmental practices and determine the applicability of regulatory limitations prior to use.*

1.6 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

[1] This specification is under the jurisdiction of ASTM Committee E54 on Homeland Security Applications and is the direct responsibility of Subcommittee E54.01 on CBRNE Sensors and Detectors.

Current edition approved Oct. 15, 2017. Published November 2017. DOI: 10.1520/E3131-17.



NOTE 1—If the number of failed results is 0 out of 47 tested samples, 1 out of 79 samples, or 2 out of 107 samples, then a device meets the POD of 0.95.

FIG. 2 Number of Independent Tests Required to Meet the Performance Criteria of 0.95 LCB (Horizontal Red Line) with 95 % Confidence at the Chosen Test Sample Concentration

2. Referenced Documents

2.1 *ASTM Standards:*[2]

E2677 Test Method for Determining Limits of Detection in Explosive Trace Detectors

2.2 *AOAC Standards:*[3]

SMPR 2010.001 Standard Method Performance Requirements for Polymerase Chain Reaction (PCR) Methods for Detection of *Francisella tularensis* in Aerosol Collection Filters and/or Liquids

SMPR 2010.002 Standard Method Performance Requirements for Polymerase Chain Reaction (PCR) Methods for Detection of *Yersinia pestis* in Aerosol Collection Filters and/or Liquid

SMPR 2010.003 Standard Method Performance Requirements for Polymerase Chain Reaction (PCR) Methods for Detection of *Bacillus anthracis* in Aerosol Collection Filters and/or Liquids

SMPR 2016.006 Standard Method Performance Requirements (SMPRs) for DNA-Based Methods of Detecting *Bacillus anthracis* in Field-Deployable, Department of Defense Aerosol Collection Devices

SMPR 2016.007 Standard Method Performance Requirements (SMPRs) for Detection of *Francisella tularensis* in Aerosol Collection Devices

SMPR 2016.008 Standard Method Performance Requirements (SMPRs) for DNA-Based Methods of Detecting *Yersinia pestis* in Field-Deployable, Department of Defense Aerosol Collection Devices

SMPR 2016.009 Standard Method Performance Requirements (SMPRs) for DNA-Based Methods for Detecting *Brucella suis* in Field-Deployable, Department of Defense Aerosol Collection Devices

SMPR 2016.010 Standard Method Performance Requirements (SMPRs) for DNA-Based Methods of Detecting *Burkholderia pseudomallei* in Field-Deployable, Department of Defense Aerosol Collection Devices

SMPR 2015.011 Standard Method Performance Requirements (SMPRs) for Detection of *Coxiella burnetii*

2.3 *ISO Standard:*[4]

ISO Guide 34 General Requirements for the Competence of Reference Material Producers

2.4 *Federal Standard:*[5]

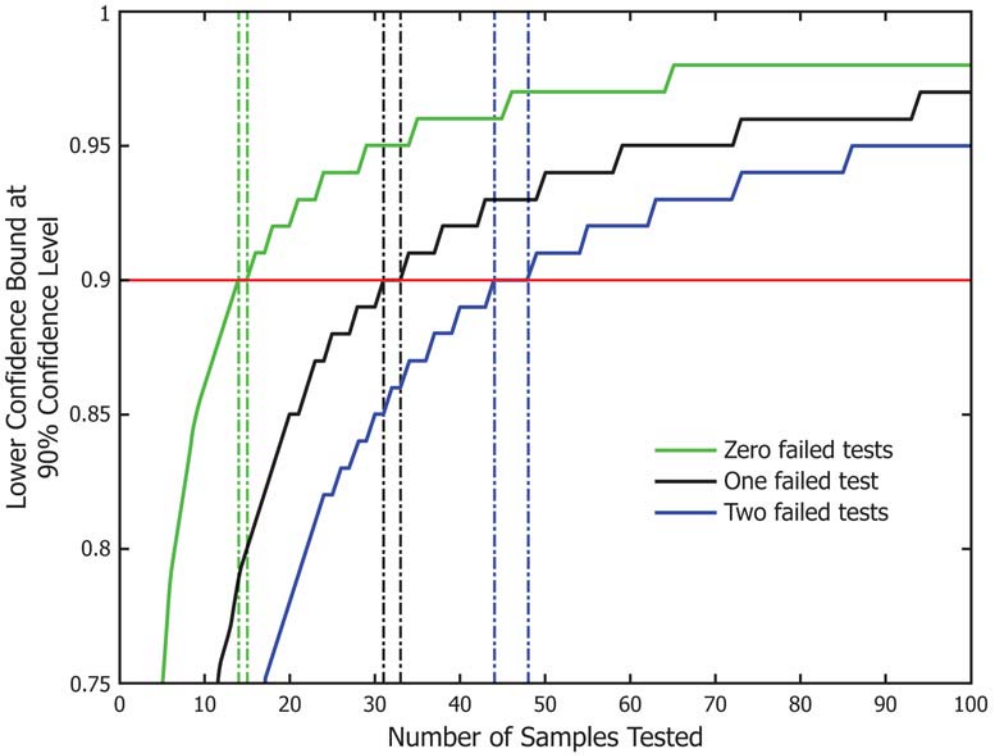
18 USC 178 Definitions

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

[3] Available from AOAC International, 2275 Research Blvd., Suite 300, Rockville, MD 20850-3250, http://www.aoac.org.

[4] Available from International Organization for Standardization (ISO), ISO Central Secretariat, BIBC II, Chemin de Blandonnet 8, CP 401, 1214 Vernier, Geneva, Switzerland, http://www.iso.org.

[5] Available from U.S. Government Printing Office, Superintendent of Documents, 732 N. Capitol St., NW, Washington, DC 20401-0001, http://www.access.gpo.gov.



NOTE 1—If the number of failed results is 0 out of 14 tested samples, 1 out of 31 samples, or 2 out of 44 samples, then a device meets the POD of 0.90.

FIG. 3 Number of Independent Tests Required to Meet the Performance Criteria of 0.90 LCB (Horizontal Red Line) with 90 % Confidence at the Chosen Test Sample Concentration

TABLE 1 Overview of Test Samples for Testing Tiers

	Test Module 1: Inclusivity Nucleic Acid[A]	Test Module 2: Exclusivity Nucleic Acid[B]	Test Module 3: Powders[C]	Test Module 4: Whole Biological Agent Spiked Powders[D]
Testing Tier 1	Full panel	Full panel	22	Single whole agent representative strain
Testing Tier 2	Reduced size panel	Reduced size panel	22	Single whole agent representative strain
Testing Tier 3	None[E]	Reduced size panel	22	Single whole agent representative strain

[A]Inclusivity strain panels for Testing Tiers 1–3 are listed in Annex A4.
[B]Exclusivity strain panels for Testing Tiers 1–3 are listed in Annex A5.
[C]See Table 2 for a list of 22 powders.
[D]Whole biological agent inclusivity strains are listed in Annex A7. Note that whole biological agents are intact spores (for *Bacillus*) or intact cells (all others biological agents).
[E]Testing Tier 3 does not have any inclusivity strain nucleic acid testing because the representative inclusivity strain for this Tier is tested in Test Module 4 as a whole biological agent.

3. Terminology

3.1 *Definitions:*

3.1.1 *accuracy, n*—closeness of agreement between a test result and the accepted reference value.

3.1.2 *assay, n*—quantitative or qualitative test used to determine the presence or absence of a biological material.

3.1.3 *biological agent, n*—any microorganism (including, but not limited to, bacteria, viruses, fungi, rickettsiae, or protozoa); infectious substance; or any naturally occurring, bioengineered, or synthesized component of any such microorganism or infectious substance capable of causing: (*1*) death, disease, or other biological malfunction in a human, an animal, a plant, or other living organism; (*2*) deterioration of food, water, equipment, supplies, or material of any kind; and (*3*) or, deleterious alteration of the environment. **18 USC 178**

3.1.3.1 *Discussion*—Also termed biothreat agent.

3.1.4 *calibration, n*—set of operations that establish, under specified conditions, the relationship between the values of quantities indicated by a measurement instrument or measuring system or values represented by a material measure or a

TABLE 2 Suspicious Powders

Class of Powder	Powder Type
Organic, biological	Brewer's yeast powder
	Dipel dust
Organic, protein-containing	Milk powder
	Infant formula
	White flour
Organic, no protein	Coffee creamer (non-dairy)
	Instant pectin
	Acetaminophen
	Powdered sugar
	Corn starch
	Polyethylene glycol 3350 (for example, MiraLAX, Glycolax)
Inorganic	Toothpaste powder with fluoride
	Baking powder (aluminum free)
	Calcium carbonate (antacid)
	Baking soda
	Epsom salt
	Magnesium carbonate (gym chalk)
	Borax
	Talc
	Road dust (NIST)
	Kaolin clay
	Popcorn salt

reference material and the corresponding values realized by standards. **Eurachem Selection (1)**[6]

3.1.5 *confidence interval, CI, n*—range of values created using a procedure that, when repeated many times, on distinct datasets, generated from the same underlying stochastic process, will bracket the true measure of performance, such as probability of detection, the proportion of times stated.

3.1.6 *confidence level, CL, n*—probability value associated with a confidence interval; the percentage of intervals that can be expected to include the true population parameter in the long run.

3.1.7 *exclusivity panel, n*—collection of near-neighbor biological agents, viruses, or nucleic acids used during testing.

3.1.7.1 *Discussion*—All exclusivity panel members should result in negative detection results.

3.1.8 *genome equivalents, GE, n*—number of genome copies present in a given mass of deoxyribonucleic acid (DNA) that can be calculated by converting the size of a genome in base pairs to micrograms of DNA.

3.1.8.1 *Discussion*—Assuming the average mass of a base pair (bp) is 660 Da (g/mole) and using Avogadro's number (6.022×10^{23} molecules/mole), the genome equivalents can be determined by first multiplying the mass of DNA (ng) × 6.022 × 10^9 ng/g.

$$GE = \frac{(\text{mass of DNA in ng}) \times \left(6.022 \times 10^{23} \frac{\text{molecule}}{\text{mole}}\right)}{(\text{genome length in base pairs}) \times \left(660 \frac{\text{g}}{\text{mole}}\right) \times 10^9 \frac{\text{ng}}{\text{g}}} \quad (1)$$

3.1.9 *inclusivity panel, n*—collection of closely related biological agents, viruses, or nucleic acids used during testing.

3.1.9.1 *Discussion*—All inclusivity panel members should result in positive detection results.

3.1.10 *inhibition, n*—undesirable effect that can result in a false negative result and is typically caused by the presence of compounds that interfere with the assay or detection process.

3.1.11 *limit of detection, LOD, n*—lowest amount of analyte in a sample that can be detected with (stated) probability. **E2677, EP17-A2 (2)**

3.1.12 *lower confidence bound, LCB, n*—lowest value of a one-sided confidence interval; or the lowest value of a half band created using a procedure that when repeated many times, on distinct datasets, generated from the same underlying stochastic process, will include the true measure of performance a proportion of times equal to the stated probability.

3.1.13 *measurement process, n*—process used to detect a material or determine if a system or instrument performs as intended.

3.1.14 *multiplexed assay, n*—assay that is capable of measuring multiple biological agents in a single test sample.

3.1.15 *near neighbor, n*—organism, virus, or nucleic acid that is similar to a desired target biological agent but should not result in a positive detection result.

3.1.15.1 *Discussion*—Exclusivity panel members include near neighbors.

3.1.16 *operator, n*—person operating an on-site biological assessment technology.

3.1.17 *panel, n*—collection of bacteria, spores, viruses, nucleic acids, or suspicious powders used during testing.

3.1.17.1 *Discussion*—Examples of panels include inclusivity panel, exclusivity panel, and suspicious powder panel.

3.1.18 *pooling, v*—act of creating a single test sample (the pooled sample) that contains strains from different biological agents.

3.1.19 *probability of detection, POD, n*—proportion of positive analytical outcomes for a qualitative method for a given matrix at a given biological agent level or concentration. **SMPR 2010.003**

3.1.20 *reference material, n*—material, sufficiently homogenous and stable with respect to one or more specified properties that has been established to be fit for its intended use in the measurement process; properties can be quantitative or qualitative. **ISO Guide 34**

3.1.21 *sensitivity, n*—change in the response of a measuring instrument divided by the corresponding change in the stimulus. **Eurachem Guide (3)**

3.1.22 *specificity/selectivity*—ability of a measurement procedure to determine accurately and specifically the analyte of interest in the presence of other components in the sample matrix under the stated conditions of the test. **Eurachem Guide (3)**

3.1.23 *spiked sample, n*—sample that is created by adding a known quantity of a biological agent to a known quantity of suspicious powder.

3.1.23.1 *Discussion*—Examples of spiked samples include

[6] The boldface numbers in parentheses refer to a list of references at the end of this standard.

pathogens added to suspicious powders that are typically prepared in a liquid (buffer) suspension.

3.1.24 *strain, n*—isolates or variants of the target biological agent(s) that the method can detect (inclusivity strains) or should not detect (exclusivity strains).

3.1.25 *suspicious powder, n*—any material that is used to create a perceived biological threat such as a hoax powder that may visually resemble *Bacillus anthracis (Ba)*, ricin powder, or simply an unknown powder.

3.1.25.1 *Discussion*—Suspicious powders include readily available household items, food products, building materials, and environmental matrices (for example, talcum powder, flour, drywall dust, and road dust).

3.1.26 *testing module, n*—set of samples used to establish the detection technology performance for a particular type of samples.

3.1.26.1 *Discussion*—Testing modules include inclusivity, exclusivity, suspicious powders, and whole biological agent spiked suspicious powders.

4. Statistical Considerations for Testing

4.1 The testing approach described herein uses a score CI (**4**) to define the number of samples that need to be tested to achieve a desired POD estimate as indicated by the value of a given LCB in a one-sided interval with a specified CL. In this specification, two levels have been chosen as providing acceptable degrees of statistical performance: a low level and a high level. Results from selecting and executing one of the testing plans presented in this specification can be used to determine if the chosen level of statistical performance has been met.

4.2 Each of the sample types listed in 4.2.1 – 4.2.4 shall be considered as part of a separate testing module. That is, the number of samples that shall be analyzed to achieve a desired POD as demonstrated by the specified LCB and CL values, as described in 4.3 and 4.4, apply to test modules 4.2.1 – 4.2.4, which includes the following test modules:

4.2.1 Inclusivity nucleic acid testing (Test Module 1),

4.2.2 Exclusivity nucleic acid testing (Test Module 2),

4.2.3 Suspicious powder testing (Test Module 3), and

4.2.4 Whole biological agent spiked suspicious powder testing (Test Module 4).

4.3 Below is listed the minimum number of replicates required for each of the four testing modules listed in 4.2.1 – 4.2.4 for the case of LCB = 0.95 and 95 % CL at the chosen test concentration (see Fig. 2). Testing shall use results from a single round of experiments with a fixed number of tests (that is, if 47 samples are chosen to be tested but a failure occurs, additional samples cannot be tested. Testing shall start over).

4.3.1 Without a single failed result, 47 samples shall be tested;

4.3.2 With no more than a single failed result, 79 samples shall be tested; or

4.3.3 With no more than two failed results, 107 samples shall be tested.

4.4 Below is listed the minimum number of replicates required for each of the four testing modules listed in 4.2.1 – 4.2.4 to achieve an LCB = 0.90 and 90 % CL at the chosen test concentration (see Fig. 3). Testing shall use results from a single round of experiments with a fixed number of tests (that is, if 14 samples are chosen to be tested but a failure occurs, additional samples cannot be tested. Testing shall start over).

4.4.1 Without a single failed result, 14 samples shall be tested;

4.4.2 With no more than a single failed result, 31 samples shall be tested; or

4.4.3 With no more than two failed results, 44 samples shall be tested.

4.4.4 The two testing plans presented in 4.3.1 – 4.3.3 and 4.4.1 – 4.4.3 allow demonstration of the levels of statistical performance chosen for this work: 0.95 LCB/95 % CL or 0.90 LCB/90 % CL. The testing plans shown are designed to be carried out by selecting a total number of tests for the LCB/CL combination desired and then performing the number of tests needed. The total number of failed results observed after testing is completed determines if the performance level selected has been achieved or not.

4.4.5 A more precise characterization of the statistical performance of a detection system (measured as the LCB/CL combination) is always preferable but, as the difference in testing resources needed between the 0.95 LCB/95 % CL level and the 0.90 LCB/90 % level indicates, more extensive characterization comes at a cost. Because the statistical approach in this specification calls for measuring performance using results from a single round of experiments with a fixed number of tests, it is advisable that the user selects the largest number of tests that can be performed within practical limits. For example, performing 79 tests using the 0.95 LCB/95 % CL level and observing 0 failures allows the user to fulfill the requirements for the performance level selected, while simultaneously providing a safeguard in the case a single failure is found.

4.4.6 Similarly, deciding to test 47 samples with the goal of achieving the lower 0.90 LCB/90 % CL performance level allows for up to 3 failures observed during testing while still meeting 0.90 LCB/90 % CL. Selecting the lower 0.90 LCB/90 % performance level allows for up to 4 failures if 79 samples are tested and up to 7 failures if 107 samples are tested while still meeting the low-level performance requirements.

4.4.7 In general, only testing plans with zero, one, or two failed results and 0.95 LCB/95 % CL or 0.90 LCB/90 % CL are explicitly considered in this specification. To find the LCB or CL values or both corresponding to other testing results, as well as determining testing requirements for other LCB/CL combinations, the user is referred to Annex A8.

4.4.8 See Fig. 1 for an overview of how to plan and conduct biological agent testing. Note that the number of samples to be tested shall be determined before any testing and cannot be modified (for example, if you chose to test 47 samples, anticipating no failures, to achieve a 0.95/95 % LCB/CL and a failure occurs, you cannot simply test more samples—you shall start over).

5. Test Criteria

5.1 *Overview:*