UCRL-TR-226839

LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Acceptance Criteria Framework for Autonomous Biological Detectors

J. M. Dzenitis

December 15, 2006

Disclaimer

Acceptance Criteria Framework for Autonomous Biological Detectors

12 December 2006

Point of contact: John M. Dzenitis
Phone: 925-422-6695
e-mail: john.m.dzenitis@llnl.gov

Participants

Chair

Wack, Edward C. (Ed)
    Director of Future Acquisition, Joint Program Executive Office for Chemical and Biological Defense, U.S. Department of Defense.

Panel members

Ackelsberg, Joel, MD, MPH
    Medical Director, Emergency Readiness and Response Unit, New York City Department of Health and Mental Hygiene.
Beaudry, Gary, Ph.D.
    Director, BioThreat Reponse Laboratory, New York City Department of Health and Mental Hygiene.
Burrus, Charles (Chuck), Ph.D.
    Senior Director, Environmental Protection and Industrial Hygiene, Office of System Safety, New York City Transit, New York State Metropolitan Transportation Authority.
Crone, Donald E. (Don)
    Manager, Mail Processing Protection Systems, U.S. Postal Service.
Diaz, Pamela S. (Pam), MD.
    Senior Medical Advisor, Division of Bioterrorism Preparedness and Response, Centers for Disease Control and Prevention.
Dzenitis, John M., Ph.D.
    Principal Investigator and Project Leader, Chemical and Biological Countermeasures Division, Lawrence Livermore National Laboratory.
Golbitz, William C. (Bill), Ph.D.
    Test and Evaluation Director, BioWatch Systems Program Office.
Hazi, Yair, Ph.D.
    BioWatch Jurisdictional Coordinator, Emergency Readiness and Response Unit, New York City Department of Health and Mental Hygiene.
Imbro, Dennis R., Ph.D.
    Deputy Program Manager, BioWatch Systems Program Office.
Kirsch, Anne
    Assistant Chief Safety Officer, Metro-North Railroad, New York State Metropolitan Transportation Authority.
Lesho, Jeffery C. (Jeff), Ph.D.
    Program Manager, Special Applications Branch, Johns Hopkins University Applied Physics Laboratory.
Masciana, Ronald J. (Ron)
    Deputy Chief, Office of Security, MTA, NYC
Malatesta, Mark L., COL.
    Joint Project Manager, Guardian, Joint Program Executive Office for Chemical and Biological Defense, U.S. Department of Defense.
Metz, Michael F. (Mike)
    Manager, Environmental Monitoring & Emergency Response, Office of System Safety, New York City Transit, New York State Metropolitan Transportation Authority.
Meyer, Richard M. (Rich), Ph.D.
    Director, Bioterrorism Rapid Response and Advanced Technology Laboratory, Centers for Disease Control and Prevention.
O'Neill, William J. (Bill).
    Program Manager, Biological Detection, U.S. Postal Service.

Reinhardt, Brian C.
    Program Manager, Defense Threat Reduction Agency, U.S. Department of Defense.
Stiefel, Jeffrey (Jeff), Ph.D.
    Program Manager, BioWatch, U.S. Department of Homeland Security.
Walter, Michael V. (Mike), Ph.D.
    Staff Senior Scientist, Head Science and Special Projects Section, Joint Program Executive Office for
    Chemical and Biological Defense, U.S. Department of Defense.
Zavasky, Dani-Margot, MD
    Medical Director, Counter Terrorism, New York City Police Department.

Table of Contents

Executive Summary

The purpose of this study was to examine a set of user acceptance criteria for autonomous biological detection systems for application in high-traffic, public facilities. The test case for the acceptance criteria was the Autonomous Pathogen Detection System (APDS) operating in high-traffic facilities in New York City (NYC).

The goals of this document are

- To serve as informal guidance for users in considering the benefits and costs of these systems.

- To serve as informal guidance for developers in understanding the needs of users.

Acceptance criteria were assessed as a function of two major categories of parameters: performance criteria and readiness criteria. Performance criteria include technical items such as types of biothreat agents, numbers of agents, speed to reach an actionable result, and instrument sensitivity, as well as operational items such as cost and mean time between failures. Readiness criteria include factors such as system maturity, the supply chain, and funding. A particular biological detection system will not be usable until readiness criteria are met.

The performance measures can be viewed as contributing to cost or benefit for an autonomous biological detection system in the following manner: The benefit of the system is detection of biological threat agents of interest at levels and on timescales that enable effective public-health response to reduce deaths, illness, and exposure. The cost of the system is up-front acquisition and operating cost to the sponsor plus indirect cost to local authorities arising from impacts on operations.

Linking performance parameters closely to cost and benefit categories results in an unusual ordering of the parameters, but all of the standard parameters should be represented here in some form. In some cases where the parameter could fit in more than one place, the best fit was chosen. The document needs to be taken as a whole.

In follow-up work, this framework will be used to systematically document the APDS for appropriateness and readiness for use in NYC.

The review panel had some overall recommendations that go beyond the acceptance criteria themselves:

- The overall prevention-detection-response-restoration system should be laid out so the individual elements (such as a biological detector) can be understood or specified in context. This systems layout could then be used for a systems cost/benefit analysis.

- Threat scenarios should be specified and translated into quantitative technical requirements.

- Current capabilities should be specified and translated into quantitative technical requirements.

- The detailed technical requirements noted above from the threat scenarios and the current capabilities should be compared.

- These criteria and this sort of effort need to be coordinated with BioWatch guidance documents and with Public Health Actionable Assays effort.

1.　Introduction

1.1.　Purpose of the Study

The purpose of this study was to examine a set of user acceptance criteria for autonomous biological detection systems for application in high-traffic, public facilities. The test case for the acceptance criteria was the Autonomous Pathogen Detection System (APDS) operating in high-traffic facilities in New York City (NYC). However, the acceptance criteria were designed to be generally applicable to other biological detection systems in other locations. For such detection systems, "users" will include local authorities (e.g., facility operators, public health officials, and law enforcement personnel) and national authorities [including personnel from the Department of Homeland Security (DHS), the BioWatch Program, the Centers for Disease Control and Prevention (CDC), and the Federal Bureau of Investigation (FBI)]. The panel members brought expertise from a broad range of backgrounds to complete this picture.

The goals of this document are

1. To serve as informal guidance for users in considering the benefits and costs of these systems.

2. To serve as informal guidance for developers in understanding the needs of users.

In follow-up work, this framework will be used to systematically document the APDS for appropriateness and readiness for use in NYC.

1.2.　Review Panel Recommendations

The review panel had some overall recommendations that go beyond the acceptance criteria:

- The overall prevention-detection-response-restoration system should be laid out so the individual elements (such as a biological detector) can be understood or specified in context. This systems layout could then be used for a systems cost/benefit analysis.

- Threat scenarios should be specified and translated into quantitative technical requirements.

- Current capabilities should be specified and translated into quantitative technical requirements.

- The detailed technical requirements noted above from the threat scenarios and the current capabilities should be compared.

- These criteria and this sort of effort need to be coordinated with BioWatch guidance documents and with Public Health Actionable Assays effort.

1.3.　Concept of Operations

These acceptance criteria for autonomous biological detection systems are defined in the context of the following concept of operations (ConOps):

> Autonomous biological detectors are assumed to be unmanned, field-deployed instruments that monitor air in high-traffic, public facilities. The detectors supply biological information to public health officials by text-message alerts and by

detailed data feeds, when necessary. Public health officials decide when the instrument signals truly indicate the presence of a biological threat agent and then alert other stakeholders, including facility operators, national public health officials, and law enforcement officials. Stakeholders respond on the basis of pre-established protocols. Possible responses include closing a facility before manual sampling and follow-up testing are done.

The above ConOps is consistent with that developed and in place for the BioWatch Program and follows from discussions carried out among a broad set of New York City agencies on the use of autonomous biological detectors.

1.4. Organization of User Acceptance Criteria

Acceptance criteria were assessed as a function of two major categories of parameters: performance criteria and readiness criteria. Performance criteria include technical items such as types of biothreat agents, numbers of agents, speed to reach an actionable result, and instrument sensitivity, as well as operational items such as cost and mean time between failures. Readiness criteria include factors such as system maturity, the supply chain, and funding. A particular biological detection system will not be usable until readiness criteria are met.

In theory, the ideal solution for developing acceptance criteria would be to vary all of the performance measures to minimize an all-inclusive cost-benefit function. However, in practical terms, identifying performance measures and acceptable values of each for a mission is a good starting point. The measures can still be viewed as contributing to cost or benefit for an autonomous biological detection system in the following manner: The benefit of the system is detection of biological threat agents of interest at levels and on timescales that enable effective public-health response to reduce deaths, illness, and exposure. The cost of the system is up-front acquisition and operating cost to the sponsor plus indirect cost to local authorities arising from impacts on operations.

The next major section of this report details a set of performance criteria. For each parameter, there is a "Description" section followed by a "Specification and Justification" section. At this stage the specifications are not complete, but the eventual goal is to set quantitative thresholds (minimum requirements) and goals (desired values). The following major section details a set of readiness criteria as a checklist of items that should be accomplished before startup. In outline form, the discussion is organized as follows:

- Performance
  - Detect agents of interest
  - Detect in the required environments
  - Detect at effective levels
  - Enable public-health response
  - Report on an effective timescale
  - Direct acquisition cost is acceptable to the sponsor
  - Direct operating cost is acceptable to the sponsor
  - Indirect operating cost is acceptable to the local authorities
- Readiness
  - Approved by local agencies
  - Approved by local/state/federal regulatory agencies

- Approved by Public Health Actionable Assays process
- SAFETY Act addressed
- Operational Test and Evaluation (OT&E) completed and approved
- Acceptance testing completed and approved
- Supply chain established
- Cost model validated
- Funding secured
- ConOps approved
- Standard Operating Procedures established
- Response protocols established
- QA/QC protocols established
- Training packages developed
- Infrastructures prepared

The high-level criteria here are either cost or benefit, but there are additional low-level factors that will be practical constraints contributing to the criteria.

2.   Performance Criteria

As explained in the previous section, the performance measures are organized by benefit (Detect agents of interest, Detect in the required environments, Detect at effective levels, Enable public-health response, and Report on an effective timescale) and cost (Direct acquisition cost is acceptable to the sponsor, Direct operating cost is acceptable to the sponsor, and Indirect operating cost is acceptable to the local authorities). Each measure has a section for description and a section for specification.

2.1.   Detect Agents of Interest

This section requires that the system detect the agents of interest, in terms of both its capability while functioning and that it remains functional in the field.

2.1.1. Agent Types

Description

Agent types are bacteria, viruses, and toxins (proteins). Some bacteria of concern form spores in addition to vegetative cells, and both forms are assumed to be included.

Specification and Justification

The instrument should be capable of detecting all three agent types for general applicability.

2.1.2. Agent Panel

Description

The agent panel is the set of agents targeted for detection in a deployment.

Specification and Justification

The agent panel will be taken from a DHS/DHHS threat list and must be integrated with the operating BioWatch and LRN system. This is particularly sensitive information and must be carefully controlled.

2.1.3. Number of Signatures

Description

The instrument must be able to correctly and reliably identify multiple agent signatures in a single sample. This can be accomplished by multiplexing (simultaneously identifying multiple signatures) or by serial analysis (running sequential tests for signatures on a sample). Either way, the results for multiple signatures must be available in the specified time (see Section *2.5.2 Time to Results*). The *required* number of signatures will depend on the number of agents (see Section *2.1.2 Agent Panel*) and the signatures per agent needed to get the required breadth and confirmation. Another measure, the *desired* number of signatures, will include room for additional signatures for better speciation and for emerging threats.

Specification and Justification

To be completed later.

2.1.4. Time to Add an Assay

Description

The time to add an assay to the system is measured from when an agent is isolated to when an assay can be operational on the system. This parameter is a quantitative measure of the upgradeability of the instrument. Upgradeability provides better protection because additional speciation or emerging threats can be covered. It also delays system obsolescence.

Specification and Justification

To be completed later.

2.1.5. Probability of Detection

Description

The definition of probability of detection, p_d, is the probability that an input exactly at the limit of detection (LOD) will be detected by the system. In this definition, the LOD is not a fixed value but depends on the desired p_d. Thus, p_d is inseparable from LOD and also quantitatively related to the probability of false positive and often the detection speed (see Section *5.1 ROC curves*). At a given operating condition, increasing input concentrations would normally result in increasing the probabilities of detection. Probability of detection has a value between 0 and 1 and needs to be specified for each agent. The p_d also has an associated confidence level depending on the number of experiments performed.

Specification and Justification

To be completed later.

2.1.6. Operational Availability

Description

Operational availability, A_O, represents the fraction of time that a system performs as desired. For a continuously running system, this is approximately the uptime divided by the total time in the field. Although it is often viewed as a maintenance parameter, A_O is closely related to p_d, in that if the system is not operating, an input will not be detected. Maintenance and cost implications are addressed in Section *2.7.1 Direct Operating Cost*. Operational availability has a value between 0 and 1.

Specification and Justification

To be completed later.

2.2. Detect in the Required Environments

The system must perform to its technical requirements in the environments of interest. It is not possible to predict all of the factors that might be relevant in these environments, but some of the main criteria are given below and some of the additional possible factors are listed.

2.2.1. Environments of Interest

Description

The environments of interest for the ConOps may serve as general guidance for the developer and will be locations for field testing.

Specification and Justification

The current ConOps requires that the system function in indoor environments that are not fully environmentally controlled like subway platforms, commuter rail platforms, passenger concourses, and HVAC equipment rooms. The air of these environments is breathed by 10,000 to 750,000 people per day.

2.2.2. Temperature Range

Description

The temperature range brackets the ambient temperatures that could be experienced during a given mission. Both external effects on equipment and internal effects on an air sample must be considered.

Specification and Justification

The environments of interest are not fully environmentally controlled, so the temperatures are beyond normal room temperatures. MTA/Transit has a standard of -10 to 43 ºC for their platforms. MTA/MNR suggested 0 to 40 ºC.

2.2.3. Humidity Range

Description

The humidity range brackets the ambient humidity that could be experienced during a given mission. Humidity can include water vapor in terms of relative humidity (between 0 and 100 %) and condensed water as mist and rain. Both external effects on equipment and internal effects on the air sample need to be considered.

Specification and Justification

The environments of interest are not fully environmentally controlled, so the humidity is beyond normal room range, but there will not be direct rain. MTA/Transit has a standard of 10 to 95 % relative humidity for equipment on their platforms.

2.2.4. Particle Tolerance

Description

Particle tolerance specifies the upper limit of particles in the air that the system must be able to handle. The air concentration, expressed as particles/liter of air, should be considered for a given size range of dust particles. The composition of particles may also be important. Both external effects on equipment and internal effects on the air sample need to be considered.

Specification and Justification

Since the target particles are 1 to 10 microns, this will be the range of interest when concentrations are specified, later.

2.2.5. Battery Capability

Description

The system must have battery-backup to allow some operation if facility power is lost from either inadvertent or nefarious activities.

Specification and Justification

The system needs to be able to complete one full analysis panel and transmit the results on battery power.

2.2.6. Additional Factors

Description

Additional environmental factors are listed here for consideration, but not specified.

- Vibration, particularly in transport.

- Electromagnetic Compatability (EMC), Electromagnetic Interference (EMI), and Electromagnetic Discharge (ESD).

- Liquid and chemical splashes and spills.

- Human interference.

2.3. Detect at Effective Levels

This section requires that the sensitivity be high enough (limit of detection low enough) that the system covers enough space, release scenarios, or people to be of significant benefit.

2.3.1. Sensitivity as Limit-of-Detection

Description

Sensitivity as limit-of-detection (LOD) is the input signal level that will be detected at a specified probability of detection p_d (see also Sections *2.1.5 Probability of Detection* and *5.1 ROC curves*). The p_d also has an associated confidence level depending on the number of experiments performed. The related parameter of selectivity is also discussed in Section *2.4.3*.

Several points need to be considered:

1. A measure of *concentration* in air should be the basis for sensitivity. Sensitivity in terms of absolute number of particles is not appropriate except to compare techniques using the same collector and sample size. The concentration can be related to a level of threat by considering infectious doses combined with the movements of agents and people in a facility.

2. Strictly speaking, the full time-course of the concentration in air at the analysis point should be specified. It often is not, so there is some arbitrariness. Instantaneous sensors depend mostly on concentration (*e.g.*, c_a, in targets/L air), whereas integrating collectors depend on concentration available for collection over time (*e.g.*, $\overline{c_a t}$, in targets·minutes/L air). We use the time-integrated definition, limiting the release considered to a collection window (see Section *2.5.1 Sampling Period*).

3. The units for targets in the biological materials vary. Possible units for bacteria include particle, colony-forming-unit (cfu), spore, DNA copy, and agent-containing particle. Full specification of the aerosol composition (live bacteria, dead bacteria, extraneous DNA, extraneous protein, particle agglomeration, particle-size distribution) is usually not given.

Similar problems apply to virus and protein units. Because the bacterial source materials are quantified in cfu/mL, we use that unit for liquid and cfu·minutes/L air for air. Similarly, we will use pfu·minutes/L air for viruses and ng·minutes/L air for toxins. For genomic analysis it may be less ambiguous to use copy·minutes/L air for both bacteria and viruses.

4. The sensitivity (LOD) for most systems is variable and depends on one or more threshold settings, which also affects probability of false positive, p_f. Sensitivity also depends on the probability of detection, p_d, that is desired. Thus, p_f can be set independently, but LOD requires p_d and p_f to be specified.

5. Aerosol releases to directly determine sensitivity cannot be done in the field. Conversely, aerosol chambers do not have representative environmental influences to yield accurate p_f, and aerosol release testing is time consuming for determining LOD. It is common to combine field data to determine noise levels and p_f, laboratory spike data to determine p_d response to known liquid concentrations, and aerosol collector testing to relate the liquid and air concentrations.

6. A common medical definition of sensitivity is the number of true positives divided by the number of actual cases (true positives plus false negatives). The result of such a calculation is a type of probability of detection measured for a set of potentially unquantified inputs, not a sensitivity in terms of biothreat agent levels. Such an approach is appropriate for a side-by-side probability comparison of two methods on a fixed set. It is not appropriate for technical specification because results will vary depending on the challenges presented to the system.

Specification and Justification

To be completed later.

2.4. Enable Public Health Response

This section concerns the quality of the information made available to the decision-makers. Some of the parameters that could arguably be in this section (probability of detection, limit of detection, time to results) are addressed in other benefit sections.

2.4.1. Probability of Autonomous False Positives

Description

A false positive is defined as a positive signal that requires a response when there was actually no agent present. Both "rate" and "probability" can be used for false positive. "False positive rate" should have units of (false events)/time. "Probability of false positive" should have units of (false events)/(total events). Probabilities are useful because they are independent of numbers of instruments and sampling rate; however, in most cases a user wants to know how often a false alarm will occur (rate of false positives). The related parameter of false negatives is discussed as probability of detection in Section *2.1.5* and also in *5.1 ROC curves*.

This section specifies the autonomous false positive directly from the instrument, with the ConOps model in mind that biological signals from the instrument will need to be reviewed by a human monitor who makes a decision ("man-in-the-loop", see Figure 1 in Section *5.2 Program Execution*). This, then, becomes the probability that the human monitor will need to review the data when there is actually no agent present.

Specification and Justification

To be completed later.

2.4.2. Probability of Actionable False Positives

Description

The general description of false positive given above applies here as well, but this section specifies the probability that a false positive makes it through human review. This becomes the probability that the Biological Action Committee will need act when there is actually no agent present (see Figure 1 in Section *5.2 Program Execution*).

Specification and Justification

To be completed later.

2.4.3. Selectivity

Description

The general definition of selectivity is a measure of how well a system can distinguish a target in the presence of near-neighbors. Selectivity can be defined quantitatively in terms of the concentrations of target and of near-neighbor that give the same signal level, resulting in values from 0 to 1 for each target/near-neighbor pair. The near-neighbor must be defined, then the selectivity can be specified or measured based on specified or actual interfering concentration.

Selectivity should be distinguished from specificity, which, strictly defined, is perfect selectivity. Further, a common medical definition of specificity is the number of true negatives divided by the number of actual negatives (true negatives plus false positives). The result of such a calculation is actually a type of probability of *not* false positive, or 1 - (probability of false positive), not a selectivity in terms of distinguishing a target from a known near-neighbor.

Specification and Justification

The specification of selectivity will be part of the Public Health Actionable Assays process.

2.4.4. Data Accessibility

Description

Data accessibility states that the data from the instruments are available by remote connection and interpretable by the local public-health employees. They will monitor the data and determine when the biological signals are actionable (see Figure 1 in Section *5.2 Program Execution*). The levels of data accessibility can be defined as "raw data," "summary data," or "alarm status," meaning that those are the finest data resolutions available.

Specification and Justification

A high data accessibility level of "raw data" will be required due to the newness of these systems and the high impact of the monitor's decision.

2.4.5. Remote System Diagnostics

Description

Remote system diagnostics refers to both assay controls and instrument health signals that assure the instrument is functioning properly. One example is positive internal controls for the assays to

indicate that agent negatives are not false negatives. Another example is air flow rate measurement to indicate that sampling flow paths are not blocked. The levels of system diagnostics can be defined as "automated" (meaning the signals are automatically analyzed and warnings are automatically sent to the monitor), "manual" (meaning the signals are available for remote viewing but are not automatically analyzed), or "unavailable" (meaning that there is not remote access to complete system diagnostics).

Specification and Justification

A system diagnostics level of "automated" will be required to minimize the frequency and time of monitoring the instruments.

2.4.6. Data Security

Description

Data security refers to assurance that the data is not monitored or corrupted and that the system operations cannot be influenced through unauthorized access.

Specification and Justification

Signals that travel through any public medium (*e.g.*, wireless or internet) must have at least 128-bit encryption. The nodes must be fire-walled to prevent unauthorized access.

2.4.7. Physical Security

Description

Physical security is assurance that the system cannot be readily disabled by unauthorized personnel.

Specification and Justification

To be completed later.

2.4.8. Biohazard Security

Description

Biohazard security is assurance that if the system collects hazardous material, it is contained in a manner that does not pose a threat to first responders.

Specification and Justification

To be completed later.

2.4.9. Data Time-Stamping Capability

Description

Time-stamping is the identification of the electronic data with common clock time. This time-stamping must be done to enable situational awareness and event reconstruction.

Specification and Justification

The instruments will time-stamp both biological and instrument-health signals and also synchronize their time within one minute of the U.S. Naval Observatory Master Clock in Washington, DC.

2.4.10. Data Archiving Duration

Description

Some amount of electronic data must be archived on the instrument and archived on networked servers or remote back-ups. The instrument data archiving is required in case of communication losses or for verification purposes. Off-instrument archiving is not addressed here because it is determined by the user's operations and does not affect the instrument.

Specification and Justification

The archiving duration on the instrument needs to be at least one month.

2.4.11. Sample Archiving Capability

Description

Archived samples from the system can be used for further lab-based testing of true positives or can be used to determine the cause of false positives. The sample integrity and identity must be maintained for forensics and investigations.

Specification and Justification

Samples identified as potentially positive must be individually archived to avoid dilution and loss of sample temporal identity. Other samples may be pooled but not contaminated by the instrument, for example not combined with assay waste.

2.5. Report on an Effective Timescale

One major goal of autonomous biological detection is reporting information on a shorter timescale than manual systems. This includes the sampling period and the time to results.

2.5.1. Sampling Period

Description

The sampling period is the time from start to end of collection and is the inverse of the sampling rate. Practically, sampling period is also the time resolution, meaning the time period to which an event may be ascribed. Time resolution allows for some identification of when a release occurred, how long a hazardous condition persisted, and who may have been exposed. Due to cost, manual sampling approaches usually operate on 24 hours at low-to-moderate threat conditions. Autonomous systems are expected to greatly reduce this time.

Specification and Justification

The systems should allow a sampling period between 1 and 24 hours. Changes to the sampling period should be possible through automated, semi-automated (remote), and manual control.

2.5.2. Time to Results

Description

Time to results is defined as the time from end of collection to when decision-makers have actionable information. Manual sampling approaches with daily pickup typically take about 8 hours for transportation and analysis. Autonomous systems are expected to greatly reduce this time. The systems will be electronically networked to allow transmission of data nearly

instantaneously, but additional time for manual review and communication of results by public health authorities needs to be considered.

Specification and Justification

To be completed later.

2.6. Direct Acquisition Cost is Acceptable to the Sponsor

Here, costs are divided into direct and indirect costs, and the direct costs are divided into acquisition and operating costs paid by the sponsor. The sponsor could be a local or federal organization.

2.6.1. Direct Acquisition Cost

Description

Direct acquisition cost is paid once by the sponsor to set up the capability. Since the largest component of this cost is likely to be the cost of the autonomous biological detectors themselves, there is reason to express the cost in dollars/instrument, even if this includes other costs and the cost varies with the number of instruments. The other costs should be itemized and include spare parts, site preparation, installation, training, and documentation.

Specification and Justification

Acceptable costs are between 25 and 250 k$/instrument, with the number of units deployed depending on this direct acquisition cost point, as well as operating cost and lifespan.

2.7. Direct Operating Cost is Acceptable to the Sponsor

The second direct cost is a recurring operating cost paid by the sponsor.

2.7.1. Direct Operating Cost

Description

Direct operating cost requires recurring payment for running and maintaining the system. The contributions can include reagents and other consumables, replacement parts, labor of several types, and QA/QC of the operating instruments. The bottom-line effect of reliability and maintainability parameters like Mean Time Between Failures (MTBF) and Mean Time To Repair (MTTR) are in direct operating cost here. MTBF and MTTR may also be specified independently as additional constraints if desired. Similarly, the skill level of maintainers directly affects the cost but may also be specified independently.

Specification and Justification

To be completed later.

2.8. Indirect Operating Cost is Acceptable to Local Authorities

Indirect operating costs are used to represent the impact on the local stakeholders of having the system in their facilities, even if a separate sponsor is paying the direct financial costs. These costs are usually difficult to quantify, but the contributions can be identified and reasonable ranges can be established.

There is an indirect cost of high-consequence actions from false positives, but that parameter is specified in Section *2.4.2 Probability of Actionable False Positives.*

2.8.1. Local Interoperability

Description

Local interoperability refers to the need for the system to fit into the operations of the local authorities in the facility, public-health, and law-enforcement domains.

Specification and Justification

To be completed later.

2.8.2. Facility Labor Requirement

Description

Facility labor requirement specifies the amount of time and the skill level required of facility personnel for the operation.

Specification and Justification

To be completed later.

2.8.3. Biological Monitor Labor Requirement

Description

Biological monitor labor requirement specifies the amount of time and the skill level required of local biological personnel for the operation.

Specification and Justification

To be completed later.

2.8.4. Maintenance Interval

Description

Maintenance interval refers to the time between regularly scheduled visits to the instruments in the field. The labor cost that these visits incur is taken into account in Section *2.7.1 Direct Operating Cost*; this section is meant to document that there may be impacts to the facilities associated with these visits. Manual sampling systems usually operate with daily visits at low-to-moderate threat conditions, but autonomous systems are expected to require fewer visits.

Specification and Justification

One week is an acceptable maintenance interval in the ConOps selected here.

2.8.5. Mean Time Between Failures

Description

The mean time between failures (MTBF) is the time between unscheduled repair visits to the instrument. As was the case for the maintenance-interval metric above, the labor cost that these visits incur is taken into account in Section *2.7.1 Direct Operating Cost*; this section is meant to document impacts to the facilities.

Specification and Justification

If the failures require attention of facility personnel, the MTBF should be greater than 2 weeks.

2.8.6. Power

Description

Electricity costs are likely to be minor, but power needs can impact local facilities significantly if they require new circuits, nonstandard voltages, or backup systems.

Specification and Justification

The instruments should use less than 15 A at 220 V, with a goal of less than 10 A at 110 V. The need for the instrument to survive moderate facility power losses is required in Section *2.2.5 Battery Capability*. Additional backup or redundant systems are not required of the instrument itself.

2.8.7. Size

Description

System size may have an indirect cost if it impacts operations, requires facility modification, or limits the useful locations for a detector.

Specification and Justification

To be completed later.

2.8.8. Visual Impact

Description

Visual impact refers to the shape, configuration, and color of the instrument. In public locations, it is desired to have a visual impact level of "low", meaning that the instrument should not significantly affect employees or customers or violate the historical or landmark consistency requirements.

Specification and Justification

To be completed later.

2.8.9. Noise

Description

Noise has an indirect cost to the facility if affects employees or customers.

Specification and Justification

The instrument should contribute less than 85 dB in the environment.

3. Readiness Criteria

The readiness criteria are milestones showing the capability to use the system in its intended role.

3.1. Approved by Local Agencies

The process for high-level approval of the system by operating agencies (facilities, law enforcement, public health, *et cetera*) needs to be established and executed.

3.2. Approved by Local/State/Federal Regulatory Agencies

The extent of and process for approval of the system by local, state, and federal regulatory agencies need to be established and executed.

3.3. Approved by Public Health Actionable Assays (PHAA) Process

The system should be approved by the Public Health Actionable Assays (PHAA) process established by DHS and DHHS.

3.4. SAFETY Act Addressed

The DHS SAFETY Act and its potential use in the operation need to be considered, and the actions taken need to be documented.

3.5. OT&E Completed and Approved

An operational test and evaluation (OT&E) in the relevant environments should be completed and approved by the local and national stakeholders.

3.6. Acceptance Testing Completed and Approved

Acceptance testing for the particular systems should be completed and approved by the local and national stakeholders.

3.7. Supply Chain Established for Instruments, Parts, Reagents, and Service

A supply chain should be established for instruments, parts, reagents, service, and any other resource that is a continuing requirement. This should include one year's worth of required resources and identified suppliers for all items.

3.8. Cost Model Validated

A cost model for the operation should be developed, validated with measurements during operational testing, and then approved by the sponsor.

3.9. Funding Secured

Funding should be secured for the initial acquisition and the first year of operation, along with an execution plan for three years of continuation.

3.10. Concept of Operations Approved

A Concept of Operations (ConOps) specific to the system and its intended use should be written and approved by local and national stakeholders. The ConOps will guide the creation of detailed procedures and response protocols.

3.11. Standard Operating Procedures Established

Standard operating procedures (SOPs) should be established and documented for all aspects of regular operation of the system, including preventative maintenance, repairs, system monitoring, signal interpretation, facility interactions, and lab support.

3.12. Response Protocols Established

Protocols for response to actionable signals should be established for all participating agencies. This should include emergency operating procedures and investigative protocols.

3.13. QA/QC Protocols Established

Protocols for continuing quality assurance and quality control should be documented and a verification plan established. This will include semi-annual or annual competency assessments and continuing education.

3.14. Training Packages Developed

Training packages should be developed for those who interact directly or indirectly with the new system. These will be specific to the required roles and will incorporate the SOPs and the Response Protocols noted above.

3.15. Infrastructures Prepared on Levels of Facility, Region, and Nation

The physical infrastructure, communications infrastructure, operational and interoperability infrastructures need to be prepared on the levels of the facility, region, and nation.

4. Summary Tables

Table 1 identifies proposed performance criteria and units for each parameter, together with spaces for quantitative thresholds (minimum requirements) and proposed goals (desired values) for each criterion. Table 2 identifies the readiness criteria as a milestone checklist. The two tables are presented as starting points for discussion.

Table 1. Performance criteria.

Parameter	Units	Threshold	Goal
Agent types	Bacteria, viruses, and/or toxins		
Agent panel	list of agents		
Number of signatures	Number of signatures	>	
Time to Add an Assay	Days	<	
Probability of detection, agent #1	Detection events/releases at LOD	>	
Probability of detection, agent #2, *etc*.	Detection events/releases at LOD	>	
Operational availability, A_o	Hours uptime/hours deployed	>	
Environments of Interest	(No units)		
Temperature range	ºC		
Humidity range	% relative humidity		
Particle tolerance	particles of 1 to 10 um/L air	>	
Battery capability	(No units)		
Additional factors	(No units)		
Sensitivity, bacterium #1	cfu·minute/L air	<	
Sensitivity, bacterium #2, *etc*.	cfu·minute/L air	<	
Sensitivity, virus #1	pfu·minute/L air	<	
Sensitivity, virus #2, *etc*.	pfu·minute/L air	<	
Sensitivity, toxin #1	ng·minute/L air	<	
Sensitivity, toxin #2, *etc*.	ng·minute/L air	<	
Probability of autonomous false positive, agent #1	False events/sample	<	
Probability of autonomous false positive, agent #2, *etc*.	False events/sample	<	
Probability of actionable false positive	False events/sample	<	
Selectivity	(Set by PHAA process)		
Data accessibility	raw data, summary data, or alarm status		
Remote system diagnostics	automated, manual, or unavailable		
Data security	encryption type		
Physical security	high, medium, low		
Biohazard security	high, medium, low		
Data time-stamping capability	(No units)		
Data archiving duration	Days	>	

Parameter	Units	Threshold	Goal
Sample archiving capability	Degree of isolation		
Sampling period	Hours from collection start to end		
Time to results	Hours from collection end to actionable result	<	
Direct acquisition cost	$/instrument	<	
Direct operating cost	$/instrument/year	<	
Local interoperability	high, medium, low		
Facility labor requirement	hrs/instrument/week and skill level		
Biological monitor labor requirement	hrs/instrument/week and skill level		
Maintenance interval	Days between service visits	>	
Mean time between failures, MTBF	Days between repairs	>	
Power requirements	Amps and Volts	<	
Size	Cubic meters	<	
Visual impact	high, medium, low		
Noise	dB	<	

Table 2. Readiness criteria.

Milestone	Completed
Approved by local operation agencies	☐
Approved by regulatory agencies	☐
Approved by Public Health Actionable Assays process	☐
SAFETY Act addressed	☐
OT&E completed and approved	☐
Acceptance testing completed and approved	☐
Supply chain established	☐
Cost model validated	☐
Funding secured	☐
Concept of Operations approved	☐
Standard Operating Procedures established	☐
Response protocols established	☐
QA/QC protocols established	☐
Training packages developed	☐
Infrastructures prepared on levels of facility, region, and nation	☐

5.   Special Topics

5.1.   ROC curves

Receiver Operating Characteristic (ROC) curves represent the *relationship between probability of false positive, probability of detection, and limit of detection (p_f, p_d, and LOD) for a particular test*. There are several potential misunderstandings that can be cleared up by considering the full implications of the ROC curves.

1. The particular test must be set. This involves not only the signature itself but also the instrument, the time taken to run the test, the detection algorithm used, and the environment of the samples and instrument.

2. Most tests can be represented by a 3-dimensional surface in probability of false positive, probability of detection, and limit of detection (p_f, p_d, and LOD). These three parameters cannot all be specified independently, and they must be specified in context of each other.

3. The probability of false positive p_f usually depends on one or more threshold settings and is adjustable. It can usually be set based solely on baseline data ("blanks" or noise) without knowledge of p_d or LOD.

4. The LOD and the p_d are interrelated and depend on p_f. At a fixed p_d, the LOD increases (sensitivity decreases) if the threshold is adjusted to decrease p_f. At a fixed p_f, the LOD increases (sensitivity decreases) if larger p_d is specified.

In medical tests with a given data set, ROC curves are often presented as linear p_f *versus* p_d plots with axes from zero to unity. This representation does not work well for analyzing high-quality tests requiring very low p_f, as is required for autonomous biological detection, because the only area of interest is very close to $p_f = 0$. A better display of ROC curve is log(p_f) *versus* LOD for a given p_d.

Practically, one might specify p_f as a hard constraint based on operational considerations and then see what LOD results at reasonable p_d (*e.g.*, 90 %). One could also run an instrument without a threshold setting and continuously feed out p_f for each test result instead. This is not commonly done. However, one should keep in mind that if a p_f is set at 1.0×10^{-6}, there is not much difference between a result at 1.1×10^{-6} that would be considered a negative and a result at 0.9×10^{-6} that would be considered an alarm.

5.2.   Program Execution

A process flow diagram for a proposed program execution for autonomous biological monitoring is shown in Figure 1. This diagram is included to address some of the operational questions that frequently arise.

One important aspect of this process flow is that the biological signals from the instrument do not directly cause a response. Instead, the biological signals are manually interpreted by a biological monitor who decides if the signals are actionable. That information flow is the same as the current manual BioWatch system, where a biologist reviews the laboratory assay results to determine if they are valid or invalid.

Another aspect worth pointing out is that the instrument maintainer, including the field service personnel, does not need access to the reagent details or biological signals. They work based on operational health signals from the instruments.

It is possible that the Integrator, Instrument Manufacturer, and Instrument Maintainer are the same organization, as is the case for the USPS's Biohazard Detection System, but those functions can also be separated.
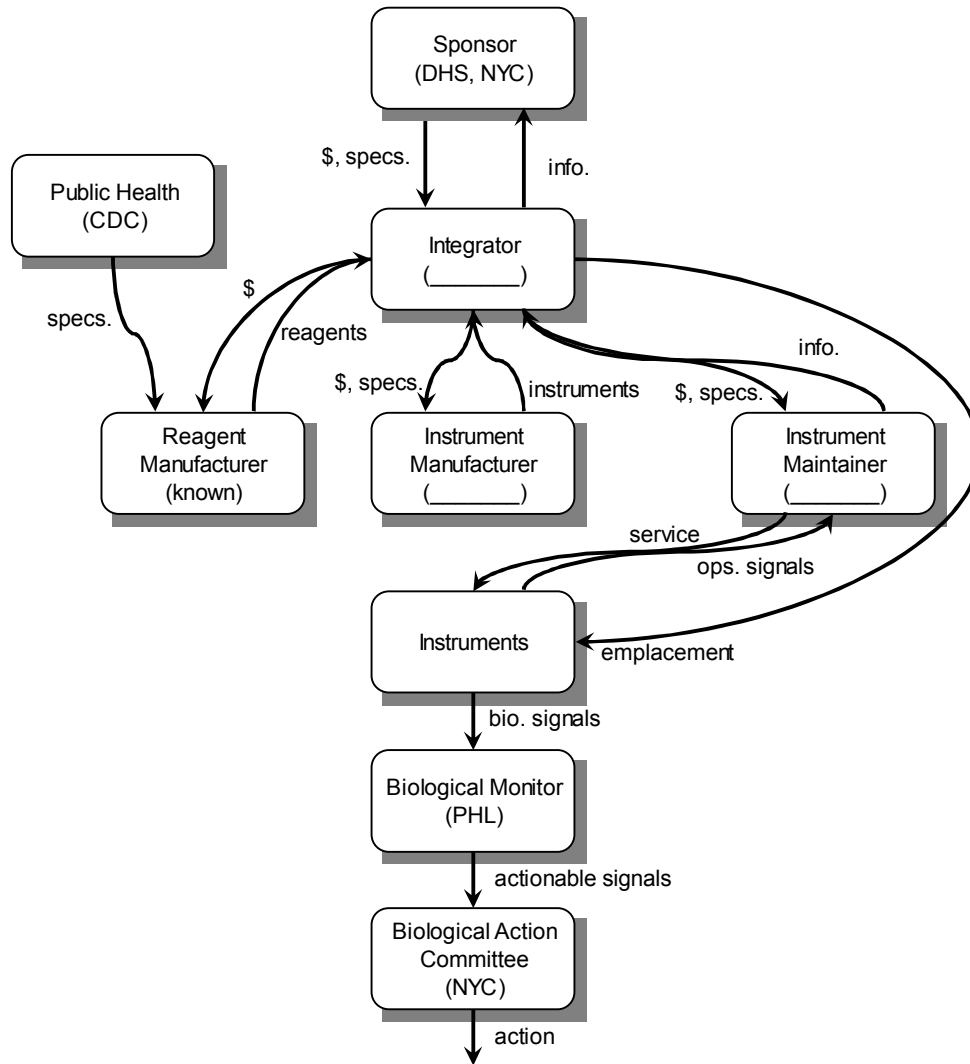

Figure 1. Flow diagram of proposed program execution for autonomous biological monitoring.

6.   Glossary

Actionable Information, Public Health Actionable Information: information of sufficient quality to motivate actions in the field, in particular actions to affect public health such as closing a facility, making public-health announcements, and performing follow-up sampling.

Concept of Operations (ConOps): an overview of how a system is to be used, including organizations and their roles, processes that will be executed, and equipment to be used.

Mean Time Between Failures: the average time of operation before a significant drop in performance occurs; approximately the deployed time divided by the number of failures.

Multiplexing: simultaneously identifying multiple agent signatures.

Selectivity: a measure of how well a system can distinguish a target in the presence of near-neighbors. Selectivity can be defined quantitatively in terms of the concentrations of target and of near-neighbor that give the same signal level, resulting in values from 0 to 1 for each target/near-neighbor pair. Differs from the medical definition of selectivity, which is an estimated 1 - (probability of false positive) for a particular set of samples.

Sensitivity: in terms of limit of detection (LOD), the concentration of agent that will result in a stated probability of false positive p_f and probability of detection p_d. Differs from the medical definition of sensitivity, which is an estimated probability of detection for a particular set of samples.

Signature: an identifying characteristic of a biological threat agent, such as a specific nucleic acid sequence or protein epitope (binding site).