U.S. DEPARTMENT OF **ENERGY**

Calculating Confidence, Uncertainty, and Numbers of Samples When Using Statistical Sampling Approaches to Characterize and Clear Contaminated Areas

GF Piepel
BD Matzke
LH Sego
BG Amidan

April 2013

Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by **Battelle** *Since 1965*

Calculating Confidence, Uncertainty, and Numbers of Samples When Using Statistical Sampling Approaches to Characterize and Clear Contaminated Areas

GF Piepel
BD Matzke
LH Sego
BG Amidan

April 2013

Executive Summary

This report discusses selected methods, formulas, and inputs needed to make characterization and clearance decisions for *Bacillus anthracis*-contaminated and uncontaminated (or decontaminated) areas using a statistical sampling approach. Specifically, the report includes the methods and formulas for calculating the

- number of samples required to achieve a specified confidence in characterization and clearance decisions

- confidence in making characterization and clearance decisions for a specified number of samples

for two common, statistically based, environmental sampling approaches. In particular, the report addresses an issue raised by the Government Accountability Office by providing methods and formulas to calculate the confidence that a decision area is uncontaminated (or successfully decontaminated) if all samples collected according to a statistical sampling approach have negative results. Key to addressing this topic is the probability that an individual sample result is a false negative, which is commonly referred to as the *false negative rate (FNR)*.

The two statistical sampling approaches currently discussed in this report are 1) *hotspot sampling* to detect small isolated contaminated locations during the characterization phase, and 2) *combined judgment and random (CJR) sampling* during the characterization or clearance phase. Typically, if contamination is widely distributed in a decision area, it will be detectable via judgment sampling during the characterization phrase. Hotspot sampling is appropriate for characterization situations where contamination is not widely distributed and may not be detected by judgment sampling. CJR sampling is appropriate during the characterization or clearance phase when it is desired to augment judgment samples with statistical (random) samples. Only these two statistical sampling approaches are addressed in this report because they are the only two approaches mentioned in a sampling reference document being developed by the Validated Sampling Plan Working Group (VSPWG). It was by agreement with the sponsor of this work (the Department of Homeland Security) that this report be limited to those two statistical sampling approaches for that reason. Other approaches which could reduce the number of samples required are being studied and may be included in future evaluations.

The hotspot and CJR statistical sampling approaches are discussed in the report for four situations:

1. qualitative data (detect and non-detect) when the FNR = 0 or when using statistical sampling methods that account for FNR > 0

2. qualitative data when the FNR > 0 but statistical sampling methods are used that assume the FNR = 0

3. quantitative data (e.g., contaminant concentrations expressed as CFU/cm^2) when the FNR = 0 or when using statistical sampling methods that account for FNR > 0

4. quantitative data when the FNR > 0 but statistical sampling methods are used that assume the FNR = 0.

For Situation 2, the hotspot sampling approach provides for stating with Z% confidence that a hotspot of specified shape and size with detectable contamination will be found. Also for Situation 2, the CJR

approach provides for stating with X% confidence that at least Y% of the decision area does not contain detectable contamination. The forms of the confidence statements for the other three situations are discussed in Section 2.2.

Statistical methods that account for FNR > 0 currently only exist for the hotspot sampling approach with qualitative data (or quantitative data converted to qualitative data). This report documents the current status of methods and formulas for the hotspot and CJR sampling approaches. Limitations of these methods are identified. Extensions of the methods that are applicable when FNR = 0 to account for FNR > 0, or to address other limitations, will be documented in future revisions of this report if future funding supports the development of such extensions.

For quantitative data, this report also presents statistical methods and formulas for

1. quantifying the uncertainty in measured sample results

2. estimating the true surface concentration corresponding to a surface sample

3. quantifying the uncertainty of the estimate of the true surface concentration.

All of the methods and formulas discussed in the report were applied to example situations to illustrate application of the methods and interpretation of the results.

Acknowledgments

Acronyms and Notation

ACL	acceptable contamination level
AOZ-ACS	accept-on-zero attribute compliance sampling
BA	*Bacillus anthracis*
CDC	Centers for Disease Control and Prevention
CFU	colony forming unit
$C_{CFU,S}^{M}$	measured concentration (CFU/cm^2) for a single surface sample obtained with a specific set S of SAP methods and sampling conditions (e.g., surface material, surface concentration, humidity)
$C_{CFU,S}^{T}$	true, unknown concentration (CFU/cm^2) for a single surface sample obtained with a specific set S of SAP methods and sampling conditions (e.g., surface material, surface concentration, humidity)
$\hat{C}_{CFU,S}^{T}$	estimate of the true, unknown concentration (CFU/cm^2) for a single surface sample obtained with a specific set S of SAP methods and sampling conditions (e.g., surface material, surface concentration, humidity)
$C_{Spores,S}^{T}$	true, unknown concentration (CFU/cm^2) of spores for a single surface sample obtained with a specific set S of SAP methods and sampling conditions (e.g., surface material, surface concentration, humidity)
CJR	combined judgment and random
DHS	U.S. Department of Homeland Security
DOE	U.S. Department of Energy
$F_{CFU:Spore,S}$	fraction of CFUs per spore achieved with a specific analysis method (including culturing and counting) and a specific set S of sampling conditions (e.g., surface material, surface concentration, humidity), considered as a random variable with a true, unknown mean [denoted Mean($F_{CFU:Spore,S}$)] and a true, unknown %RSD [denoted %RSD($F_{CFU:Spore,S}$)]
FN	false negative
FNR	false negative rate = PFN
FPR	false positive rate
GAO	Government Accountability Office
INL	Idaho National Laboratory
LOD	limit of detection
Mean ($F_{CFU:Spore,S}$)	true, unknown mean of the random variable $F_{CFU:Spore,S}$, which is the fraction of CFUs per spore achieved with a specific analysis method (including culturing and counting) and a specific set S of sampling conditions (e.g., surface material, surface concentration, humidity)

$\overline{\text{Mean}}\,(F_{\text{CFU:Spore},S})$	estimate of the true, unknown mean of the random variable $F_{\text{CFU:Spore},S}$, which is the fraction of CFUs per spore achieved with a specific analysis method (including culturing and counting) and a specific set S of sampling conditions (e.g., surface material, surface concentration, humidity)
$\text{Mean}\,(R_{A,S})$	true, unknown mean of the random variable $R_{A,S}$ which is the recovery efficiency for a specific sample analysis method (including culturing and counting) and a specific set of conditions (all denoted by S)
$\overline{\text{Mean}}\,(R_{A,S})$	estimate of the true, unknown mean of the random variable $R_{A,S}$ which is the recovery efficiency for a specific sample analysis method (including culturing and counting) and a specific set of conditions (all denoted by S)
$\text{Mean}\,(R_{HST,S})$	true, unknown mean of the random variable $R_{HST,S}$, which is the recovery efficiency for a specific sample handling, storage, and transport method and a specific set S of conditions (all denoted by S)
$\overline{\text{Mean}}\,(R_{HST,S})$	estimate of the true, unknown mean of the random variable $R_{HST,S}$, which is the recovery efficiency for a specific sample handling, storage, and transport method and a specific set S of conditions (all denoted by S)
$\text{Mean}\,(R_{P,S})$	true, unknown mean of the random variable $R_{P,S}$, which is the recovery efficiency for a specific sample processing/extraction method and a specific set S of conditions (all denoted by S)
$\overline{\text{Mean}}\,(R_{P,S})$	estimate of the true, unknown mean of the random variable $R_{P,S}$, which is the recovery efficiency for a specific sample processing/extraction method and a specific set of conditions (all denoted by S)
$\text{Mean}\,(R_{SC,S})$	true, unknown mean of the random variable $R_{SC,S}$, which is the recovery efficiency for a specific sample collection method and specific S set of conditions (all denoted by S)
$\overline{\text{Mean}}\,(R_{SC,S})$	estimate of the true, unknown mean of the random variable $R_{SC,S}$, which is the recovery efficiency for a specific sample collection method and specific S set of conditions (all denoted by S)
$\text{Mean}\,(R_{SAP,S})$	true, unknown mean of the random variable $R_{SAP,S}$, which is the recovery efficiency for a specific set S of SAP methods and sampling conditions (e.g., surface material, surface concentration, humidity)
$\overline{\text{Mean}}\,(R_{SAP,S})$	estimate of the true, unknown mean of the random variable $R_{SAP,S}$, which is the recovery efficiency for a specific set S of SAP methods and sampling conditions (e.g., surface material, surface concentration, humidity)
NIST	National Institute of Standards and Technology
PFN	probability of false negative = FNR
PFP	probability of false positive

$\%\text{RSD}(F_{CFU:Spore,S})$	true, unknown percent relative standard deviation of the random variable $F_{CFU:Spore,S}$, which is the fraction of CFUs per spore achieved with a specific analysis method (including culturing and counting) and a specific set S of sampling conditions (e.g., surface material, surface concentration, humidity)
$\overline{\%\text{RSD}}(F_{CFU:Spore,S})$	estimate of the true, unknown percent relative standard deviation of the random variable $F_{CFU:Spore,S}$, which is the fraction of CFUs per spore achieved with a specific analysis method (including culturing and counting) and a specific set S of sampling conditions (e.g., surface material, surface concentration, humidity)
$\%\text{RSD}(R_{A,S})$	true, unknown percent relative standard deviation of the random variable $R_{A,S}$, which is the recovery efficiency for a specific sample analysis method (including culturing and counting) and a specific set of conditions (all denoted by S)
$\overline{\%\text{RSD}}(R_{A,S})$	estimate of the true, unknown percent relative standard deviation of the random variable $R_{A,S}$, which is the recovery efficiency for a specific sample analysis method (including culturing and counting) and a specific set of conditions (all denoted by S)
$\%\text{RSD}(R_{HST,S})$	true, unknown percent relative standard deviation of the random variable $R_{HST,S}$, which is the recovery efficiency for a specific sample handling, storage, and transport method and a specific set S of conditions (all denoted by S)
$\overline{\%\text{RSD}}(R_{HST,S})$	estimate of the true, unknown percent relative standard deviation of the random variable $R_{HST,S}$, which is the recovery efficiency for a specific sample handling, storage, and transport method and a specific set S of conditions (all denoted by S)
$\%\text{RSD}(R_{P,S})$	true, unknown percent relative standard deviation of the random variable $R_{P,S}$, which is the recovery efficiency for a specific sample processing/extraction method and a specific set S of conditions (all denoted by S)
$\overline{\%\text{RSD}}(R_{P,S})$	estimate of the true, unknown percent relative standard deviation of the random variable $R_{P,S}$, which is the recovery efficiency for a specific sample processing/extraction method and a specific set S of conditions (all denoted by S)
$\%\text{RSD}(R_{SC,S})$	true, unknown percent relative standard deviation of the random variable $R_{SC,S}$, which is the recovery efficiency for a specific sample collection method and specific S set of conditions (all denoted by S)
$\overline{\%\text{RSD}}(R_{SC,S})$	estimate of the true, unknown percent relative standard deviation of the random variable $R_{SC,S}$, which is the recovery efficiency for a specific sample collection method and specific S set of conditions (all denoted by S)
$\%\text{RSD}(R_{SAP,S})$	true, unknown percent relative standard deviation of the random variable $R_{SAP,S}$, which is the recovery efficiency for a specific set S of SAP methods and sampling conditions (e.g., surface material, surface concentration, humidity)
$\overline{\%\text{RSD}}(R_{SAP,S})$	estimate of the true, unknown percent relative standard deviation of the

	random variable $R_{SAP,S}$, which is the recovery efficiency for a specific set S of SAP methods and sampling conditions (e.g., surface material, surface concentration, humidity)
PNNL	Pacific Northwest National Laboratory
$R_{A,S}$	recovery efficiency for a specific sample analysis method (including culturing and counting) and a specific set of conditions (all denoted by S), considered as a random variable with a true, unknown mean [denoted $\text{Mean}(R_A)$] and a true, unknown percent relative standard deviation [denoted $\%\text{RSD}(R_A)$]
RE	recovery efficiency
$R_{HST,S}$	recovery efficiency for a specific sample handling, storage, and transport method and a specific set S of conditions (all denoted by S), considered as a random variable with a true, unknown mean [denoted $\text{Mean}(R_{HST,S})$] and a true, unknown percent relative standard deviation [denoted $\%\text{RSD}(R_{HST,S})$]
$R_{P,S}$	recovery efficiency for a specific sample processing/extraction method and a specific set S of conditions (all denoted by S), considered as a random variable with a true, unknown mean [denoted $\text{Mean}(R_{P,S})$] and a true, unknown percent relative standard deviation [denoted $\%\text{RSD}(R_{P,S})$]
$R_{SC,S}$	recovery efficiency for a specific sample collection method and specific S set of conditions (all denoted by S), considered as a random variable with a true, unknown mean [denoted $\text{Mean}(R_{SC,S})$] and a true, unknown percent relative standard deviation [denoted $\%\text{RSD}(R_{SC,S})$]
$R_{SAP,S}$	recovery efficiency for a specific set S of SAP methods and sampling conditions (e.g., surface material, surface concentration, humidity). This is considered as a random variable with a true, unknown mean [denoted $\text{Mean}(R_{SAP,S})$] and a true, unknown percent relative standard deviation [denoted] $\%\text{RSD}(R_{SAP,S})$
SAP	sampling and analysis process
SD	standard deviation
UTL	upper tolerance limit
VSP	Visual Sample Plan. Freely-available software for developing environmental sampling plans and related data analysis (VSP Development Team, 2012).
VSPWG	Validated Sampling Plan Working Group

Contents

Figures

Tables

1.0 Introduction

This report discusses selected methods, formulas, and inputs needed to make characterization and clearance decisions for *Bacillus anthracis*-contaminated and uncontaminated (or decontaminated) areas using a statistical sampling approach. Specifically, the report includes the methods and formulas for calculating the

- confidence in making characterization and clearance decisions for a specified number of samples, and

- number of samples required to achieve a specified confidence in characterization and clearance decisions

for two common, statistically based, environmental sampling approaches. In particular, the report addresses issues raised by the Government Accountability Office (GAO) by providing methods and formulas to calculate the confidence that a decision area is uncontaminated (or successfully decontaminated) if all samples using a statistical sampling approach have negative results (see Section 1.1).

The two statistical sampling approaches currently discussed in this report are 1) *hotspot sampling* to detect small isolated contaminated locations during the characterization phase, and 2) *combined judgment and random (CJR) sampling* during the characterization or clearance phase. Typically, if contamination is widely distributed in a decision area, it will be detectable via judgment sampling during the characterization phrase. Hotspot sampling is appropriate for characterization situations where contamination is not widely distributed and may not be detected by judgment sampling. CJR sampling is appropriate during the characterization or clearance phase when it is desired to augment judgment samples with statistical (random) samples. Only these two statistical sampling approaches are addressed in this report because they are the only two approaches mentioned in a sampling reference document being developed by the Validated Sampling Plan Working Group (VSPWG). It was by agreement with the sponsor of this work (the Department of Homeland Security) that this report be limited to those two statistical sampling approaches for that reason.

Some fundamental terms are now defined. A *decision area* is the geographical or spatial area about which investigators wish to draw a conclusion on the basis of sampling and other relevant information. Decision areas typically include all or part of a building, but they may be any contaminated (or potentially contaminated) area for which characterization or clearance is required. In this report, a *statistical sampling* approach refers to any sampling approach based on a mathematical or statistical model that provides confidence or probability statements about the decision area. Statistical sampling approaches typically involve taking samples at random locations, or at the nodes of a randomly located grid or lattice that is superimposed on the decision area. The term *probabilistic sampling* is often used to describe statistical sampling designs that exclusively use randomly located samples. However, in this report the term *random sampling* is used instead of probabilistic sampling, because it is more descriptive of what is done. Also, *judgment samples* play a vital role in the characterization and clearance process because they include the informed prior belief and expert knowledge of personnel involved in the sampling process. Consequently, statistical sampling techniques have recently been expanded to include both judgment and random samples. This methodology is referred to as the combined judgmental and random (CJR) approach (Sego et al. 2007, 2010). The types of statistical sampling approaches are explained in Section 1.1 and discussed further in Section 3.

The work presented in the report addresses the main categories of uncertainty in the sampling and analysis process (SAP) to make a characterization or clearance decision. These include uncertainties associated with 1) statistical sampling, 2) sample collection, 3) sample handling, storage, and transportation, 4) sample processing/extraction, and 5) sample analysis. The work also addresses factors that affect the results, such as the: contaminant concentration; statistical sampling approach; surface material sampled; and sample collection, processing/extraction, and analytical methods.

The original plans for this work envisioned this report summarizing the current status of statistical sampling and uncertainty methods, with subsequent extensions of the methodology and updates of this report. However, subsequent work (now underway) may appear in stand-alone technical reports or journal articles. The plans are contingent on continued funding for this work and work by others to supply the data needed to complete the methods, formulas, and data required to calculate the numbers of samples and confidence for characterization and clearance decisions.

Section 1.1 presents the background and justification for the work in this report. Section 1.2 discusses the performing organization and funding for the work. Section 1.3 outlines the organization of the remainder of the report.

1.1 Background and Justification

In 2001, letters containing *Bacillus anthracis* (BA) contaminated the Hart Senate office building in Washington, DC and postal facilities that processed the letters. This experience illuminated areas of concern in the procedures and methods used to characterize and clear buildings contaminated by BA. A congressional inquiry and the GAO identified two main concerns (GAO 2005a, 2005b). The first was the reliance on sampling specific areas where it was thought BA would be found. This type of sampling approach is referred to as *judgment sampling*. The GAO reports identified the need to use *statistical sampling* so that when all results are negative, a building (or area within a building) can be cleared with a known level of statistical confidence. The second main concern was that the sample collection, handling/storage/transportation, processing/extraction, and analytical methods used were not validated. The lack of validated methods raised questions about the reliability of the negative sampling results from sampling the facilities in 2001.

The Validated Sampling Plan Working Group (VSPWG) was formed in July 2006 in response to the congressional inquiry and GAO reports. The VSPWG is headed by Department of Homeland Security (DHS) Science and Technology Directorate and includes experts from the Environmental Protection Agency, the Centers for Disease Control and Prevention (CDC), National Institute of Standards and Technology (NIST), the Department of Defense, the Federal Bureau of Investigation, and U.S. national laboratories including Pacific Northwest National Laboratory (PNNL) and Sandia National Laboratories. The VSPWG is developing a sampling reference document and working toward validating methods for steps in the SAP. These steps include: 1) sampling approach (e.g., appropriate uses of judgment, random, and CJR sampling techniques), 2) sample collection methods, 3) handling, storage, and transportation of samples, 4) sample processing and extraction methods (i.e., extracting the contaminant from samples), and 5) sample analysis (i.e., analytical methods). These steps and the uncertainties associated with them are discussed further in Section 2.3.

This report addresses the first GAO concern by documenting the methods, formulas, and data to calculate the numbers of samples, and confidence in characterization and clearance decisions, based on statistical sampling approaches. These methods and formulas provide statistical approaches for 1) detecting contamination with specified confidence, and 2) clearing an uncontaminated or decontaminated area with specified confidence when all samples give negative results.

1.2 Performing Organization and Funding

The work in this report was performed by staff in the Applied Statistics and Computational Modeling group at PNNL. The specific contributors are listed as authors of this report. This effort was influenced by discussions and other work by PNNL coworkers and members of the VSPWG. Specific individuals who provided inputs or guidance are listed in the Acknowledgments.

The PNNL work summarized in this report was funded by the Biological Research and Development Branch of the Chemical and Biological Division, in the Science and Technology Directorate of the U.S. Department of Homeland Security (DHS). The interest and support of Lance Brooks and Chris Russell (Program Managers, Chemical and Biological Division) for this work in support of the VSPWG is gratefully acknowledged.

The work was funded under the prime contract (Contract DE-AC05-76RL01830) between the U.S. Department of Energy (DOE) and Battelle (the operator of PNNL for research, testing, evaluation, and/or development activities) pursuant to Section 309(a)(1)(c) of the *Homeland Security Act of 2002* (Public Law 107-296), which authorizes DHS to task the DOE national laboratories on a "work for others" basis.

1.3 Report Organization

The remainder of this report describes the concepts, methods, and formulas for calculating the uncertainty and confidence in characterization and clearance decisions based on a statistical sampling approach. The report is organized as follows. The steps of the SAP (from sample collection through analysis) and uncertainties for characterization and clearance goals of sampling are discussed in Section 2. Section 3 introduces and presents an overview of the statistical sampling approaches for characterization and clearance mentioned in the sampling reference document being developed by the VSPWG. Section 4 discusses methods for estimating the false negative rate (FNR) of a single sample and available data that estimate the FNR. The next two sections present methods and formulas for determining the number of samples and confidence for characterization decisions (Section 5) and clearance decisions (Section 6) using statistical sampling approaches. Section 7 presents methods for quantifying the uncertainty of a measured contaminant concentration and for estimating and quantifying the uncertainty in the estimate of a true contaminant concentration when the SAP yields quantitative data. Section 8 summarizes the work.

2.0 The Sampling and Analysis Process and Uncertainties for Characterization and Clearance

The goal of the work documented in this report is to develop methods and formulas for calculating the number of samples and overall confidence in making characterization and clearance decisions when using a statistical sampling approach. To achieve this goal, the type of data (qualitative or quantitative) and the uncertainties in each step of the SAP must be taken into account. Existing methods and formulas typically account only for uncertainty resulting from sampling a fraction of the decision area, and do not account for uncertainties in the steps of the SAP. Hence, the goal of the work in this document is to extend existing methods and formulas to account for the uncertainties in the SAP.

Quantitative and qualitative data are discussed in Section 2.1. Statistical statements for characterization and clearance situations are discussed in Section 2.2. The steps of the SAP and the uncertainties associated with each step are discussed in Section 2.3.

2.1 Correct and Incorrect Sample Results and Uncertainties for Quantitative and Qualitative Data

Data resulting from a SAP can be quantitative or qualitative in nature. In this report, quantitative data are assumed to be surface concentrations expressed as numbers of colony forming units (CFU) per unit surface area (i.e., CFU/cm^2). However, the statistical sampling methods discussed herein can accommodate virtually any quantitative measure of a contaminant. Qualitative data represent a binary outcome, such as 1) a quantitative sample result being classified as acceptable or unacceptable as defined by an action level threshold, 2) contamination being detected or not detected, or 3) the presence or absence of a particular quality. The criteria that define the qualitative, binary outcome must be clearly stated and understood by all interested parties. Clearly, these criteria depend on the context of the situation, and they may involve the type of contaminant, the event scenario, and/or the SAP. Thus, for the sake of convenience and generality, we will refer to the binary outcomes of qualitative data using the labels *acceptable* and *unacceptable*.

An acceptable/unacceptable data point (i.e., a qualitative sample result) can be classified in one of four ways as shown in Figure 2.1. A correct classification occurs when a sample is taken from an acceptable (or unacceptable) location and is classified as acceptable (or unacceptable). A *false negative* occurs when a sample is taken from an unacceptable location, but the sample is incorrectly classified as acceptable. A *false positive* occurs when a sample is taken from an acceptable location, but the sample is incorrectly classified as unacceptable.

Truth

		Acceptable	Unacceptable
Outcome from an Observed Sample	Acceptable	Correct	False Negative
	Unacceptable	False Positive	Correct

Figure 2.1. Four Possible Outcomes from an Observed Sample Based on Qualitative (Acceptable/Unacceptable) Data

For each sample in a given situation, there is a probability that each of these four events occurs. Uncertainties for qualitative data are expressed in terms of "probability units" of the four possible outcomes in. Figure 2.1 We are principally concerned with incorrect outcomes: the probability of a false negative (PFN) and the probability of a false positive (PFP). In the literature, the PFN is commonly referred to as the false negative rate (FNR) and the PFP is referred to as the false positive rate (FPR). Hence, in this report we subsequently use FNR and FPR. The values of probabilities FNR and FPR may depend on the surface material being sampled and the concentration of surface contamination, as well as on methods of sample collection, handling/storage/transportation, processing/extraction, analysis, and other factors. Hence, for qualitative data, uncertainties are expressed in terms of FNR and FPR.

For quantitative data 1) systematic/bias uncertainties are expressed as differences or ratios relative to the true value, and 2) random uncertainties are expressed as standard deviations (SD) or percent relative standard deviations (%RSD). However, uncertainties for quantitative data can also be expressed in terms of FNR and FPR, provided acceptable and unacceptable outcomes are appropriately defined (i.e., in terms of a quantitative measurement, the true value of the measurand, and an acceptable contamination level (ACL)). For example, a false negative (FN) could be defined as a measured contamination level being at or below the ACL—and therefore classified as acceptable—when the true contamination level is unacceptable (i.e. above the ACL). However, the FNR will depend on how far the true contamination level is above the ACL. Using this definition of a FN for quantitative data, the FNR will be the largest when the true contamination level is just larger than the ACL, and will decrease as the true contamination level increases. If there is no ACL (i.e., ACL = 0), then a FN occurs when a zero contamination level is measured when in fact there is contamination. In that case, the FNR will increase to 1.0 as the true contamination level decreases sufficiently. On the other hand, the FNR approaches zero as the true contamination level increases, regardless of whether the data are quantitative or qualitative.

If the true contamination level is large enough, the FNR can be zero for quantitative data even though they are subject to uncertainties (systematic/bias and random). For example, a surface sampling and analysis procedure may have an overall efficiency (for all steps of the SAP) of only 20%, but if the contamination level is high enough, the 20% efficiency will not result in false negatives, and the FNR will be zero.

2.2 Statistical Statements for Characterization and Clearance

Two key phases of the response to a biological contamination incident are *characterization* and *clearance*. The goal of characterization is to detect contamination if it is present, and usually to determine the extent of contamination. The goal of clearance is to determine whether or not a decision area is 1) uncontaminated by a biological agent, or 2) has had the contamination successfully removed or inactivated to the clearance level with a desired level of confidence.

The nature of characterization and clearance based on statistical sampling depends on whether sample results are quantitative or qualitative (acceptable/unacceptable). For characterization and clearance with acceptable/unacceptable data, only the presence and extent of contamination can be determined (by positive sample results), not the amount of contamination. For quantitative data, it is assumed there is an ACL corresponding to the characterization or clearance goals. In this report, we discuss statistical sampling approaches that provide for two types of statistical confidence statements, depending on whether the objective is characterization or clearance:

ChS: <u>Characterization Statement</u>: There is Z% confidence[a] of identifying within the decision area an unacceptable region of a specified shape and size (or larger), provided such a region exists.

ClS: <u>Clearance Statement</u>: There is X% confidence that at least Y% of the decision area is acceptable, provided all samples are acceptable.

The definition of *acceptable* depends on three factors: 1) whether false negatives are possible (FNR > 0), 2) whether the statistical method accounts for FNR > 0, and 3) the type of data. In Table 2.1, we consider four possible definitions of *acceptable* under these various conditions.

Table 2.1. Definition of *Acceptable* Decision Areas and Samples for Various Situations and Data Types

Situation	Type of Data	Definition of an *Acceptable Region* or *Decision Area*[a]	Definition of an *Acceptable Sample*
1. FNR = 0, or 2. FNR > 0 with methods that account for FNR > 0	Qualitative	Contamination is not present	Contamination is not detected
	Quantitative	Contamination is ≤ ACL	Measured contamination is ≤ ACL
3. FNR > 0 but using methods that assume FNR = 0	Qualitative	Contamination is not detected	Contamination is not detected
	Quantitative	Measured contamination is ≤ ACL	Measured contamination is ≤ ACL

(a) The definition of "acceptable" may apply to either a decision area, or to a region within a decision area.

---

[a] Statisticians sometimes prefer to use the term "probability" in this context rather than "confidence." However, for consistency in terminology for both characterization and clearance approaches and statements, the term "confidence" is used. That way, both characterization and clearance statements can be referred to as "confidence statements."

Applying the definitions of Table 2.1 to the general ChS and ClS statements above, we can formulate eight closely related, yet distinct, statistical statements that reflect the nuances of characterization and clearance, qualitative and quantitative data, and the capability of the statistical methods to account for $FNR > 0$. These statements are listed below.

FNR = 0, or FNR > 0 with Methods that Account for FNR > 0

ChS1a   Characterization, Qualitative Data: There is Z% confidence of identifying within the decision area a contaminated region of a specified shape and size (or larger), provided such a region exists.

ChS2a   Characterization, Quantitative Data: There is Z% confidence of identifying within the decision area a region of a specified shape and size (or larger) that is contaminated above the ACL, provided such a region exists.

ClS3a   Clearance, Qualitative Data: There is X% confidence that at least Y% of the decision area is not contaminated, provided no samples detect contamination.

ClS4a   Clearance, Quantitative Data: There is X% confidence that at least Y% of the decision area has contamination at or below the ACL, provided no samples registered contamination above the ACL.

FNR > 0 but Using Methods that Assume FNR = 0

ChS1b   Characterization, Qualitative Data: There is Z% confidence of identifying within the decision area a region of detectable contamination of a specified shape and size (or larger), provided such a region exists.

ChS2b   Characterization, Quantitative Data: There is Z% confidence of identifying within the decision area a region of specified shape and size (or larger) with contamination that would be measured above the ACL, provided such a region exists.

ClS3b   Clearance, Qualitative Data: There is X% confidence that at least Y% of the decision area does not contain detectable contamination, provided no samples detect contamination.

ClS4b   Clearance, Quantitative Data: There is X% confidence that at least Y% of the decision area does not contain contamination that would be measured above the ACL, provided that none of the sample measurements exceeded the ACL.

In the rest of this report, the statements (ChS, ChS1a, ChS1b, ChS2a, ChS2b) for characterization are referred to as *Z% characterization statements*, and the statements (ClS, ClS3a, ClS3b, ClS4a, ClS4b) for clearance are referred to as *X%/Y% clearance statements*. There is a slight difference in wording of statements above ending in "a" versus "b". In situations where the FNR > 0 but methods are used that assume FNR = 0, the statements ending in "b" refer to "detectable contamination" or "contamination that would be measured above the ACL". These are in contrast to "contamination" or "contamination exceeding the ACL" for the statements ending in "a" when FNR = 0, or FNR > 0 with methods that account for FNR > 0. It is preferable to make statistical statements about actual contamination rather than detectable or measured contamination, but to do so, either the FNR must be zero, or statistical methods and formulas must account for FNR > 0. The consequence of using methods that assume FNR = 0 when

in fact FNR > 0 is that one or more of the Z%, X%, and Y% values are overestimated. Equivalently, the number of samples that would be calculated to achieve specified Z% or X%/Y% would be too small.

The number of samples required to state with a high confidence that no (i.e., zero) contamination exists is almost always prohibitive, unless strong assumptions are made *a priori* that contamination does not exist in the decision area. With any sampling approach (random, judgment, or combined), there is some chance of not sampling any contaminated locations if they exist. Another concern is whether contamination is correctly detected (for qualitative data) or is measured above the specified ACL (for quantitative data) when a contaminated location is sampled.[a] Both of these aspects must be accounted for in formulas to calculate Z% characterization statements and X%/Y% clearance statements, as well as the number samples required to achieve desired values of Z% and X%/Y%.

2.3 Sampling and Analysis Process Steps and Uncertainties

The steps of the SAP for making characterization and clearance decisions include 1) statistical sampling plan (approach and sampling design), 2) sample collection, 3) sample handling, storage, and transportation, 4) sample processing/extraction, and 5) sample analysis. There are uncertainties associated with each of these steps, where uncertainties may be systematic and/or random. Systematic uncertainty is also referred to as *bias*, and random uncertainty is also referred to as *imprecision*.

How uncertainties in the results from the SAP are expressed depends on whether the measurements of contamination are quantitative or qualitative. As noted in Section 2.1, quantitative data in this report are assumed to be surface concentrations expressed in terms of CFU/cm^2. Hence, uncertainties of quantitative data are also assumed to be surface concentrations expressed in terms of CFU/cm^2. For qualitative data, uncertainties are expressed in terms of "probability units" for probabilities of the incorrect outcomes (false negative, false positive) discussed in Section 2.1.

The following subsections briefly describe each step of the SAP and discuss the relevant uncertainties. Section 2.3.1 discusses statistical sampling and the relevant uncertainty. Sections 2.3.2 to 2.3.5 discuss bias and random uncertainties expressed in terms of recovery efficiencies. Simply stated, a recovery efficiency is the percentage of contaminant on a surface that is recovered by a given step, or all steps, of the SAP. Quantitative data are required to estimate recovery efficiencies associated with the steps of the SAP. If the data from samples are qualitative (e.g. detection or non-detection of contamination), it is not possible to directly quantify recovery efficiencies. However, recovery efficiencies, if low enough, may affect the probabilities of the possible outcomes discussed in Section 2.1 (e.g., FNR). So, even if recovery efficiency itself cannot be quantified because of qualitative data, the probabilities of possible outcomes can be quantified using tests where it is known whether samples are contaminated or not, and if so, at what concentration. In summary, the concepts of recovery efficiencies for the steps of the SAP are applicable regardless of whether the sampling results are quantitative or qualitative.

---

[a] This is an issue because the sampling and analysis methods can have combined recovery efficiency significantly below 100%.

2.3.1   Statistical Sampling and Uncertainty

Statistical sampling accounts for the fact that samples are collected from only a fraction (i.e., less than 100%) of the decision area to be characterized or cleared so that decisions can be made with high confidence. For any given sample, there is a probability of that sample not being assigned to a contaminated location. Various statistical sampling approaches can be used to determine the number and locations of samples so that at least one sample is taken from a contaminated area with high confidence. These approaches account in different ways for the probability that a given sample is not from a contaminated location, based on beliefs about how the contaminant is distributed over the decision area. Selected statistical sampling approaches are discussed in Section 3.

For characterization, the fundamental uncertainty associated with statistical sampling is the probability of not sampling at least one contaminated location in the decision area. The number of samples for a given sampling approach must be sufficient to make this probability small. In other words, sufficient samples must be collected using a specific statistical sampling approach so that there is high probability of sampling at least one contaminated location in the decision area. Because there are uncertainties in the remaining steps of the SAP, sufficient samples must be collected using a specific statistical sampling approach so that there is high probability of detecting contamination in at least one sample from the decision area.

For clearance, the fundamental uncertainty regards the probability that unsampled portions of the decision area satisfy the clearance criteria, given that all clearance samples (be they judgment or random) are acceptable. As with characterization, uncertainties associated with the SAP influence the uncertainty for clearance decisions.

2.3.2   Sample Collection and Uncertainties

Sample collection refers to the method and procedure for collecting a sample. In this report we are concerned with collecting surface samples of BA contamination using swab, wipe, and vacuum methods. Sample collection methods are subject to bias uncertainties because typically only a fraction of the spores on a surface is removed by a sample collection method. For a specific sample collection method and specific set S of conditions, we refer to the fraction of spores recovered as the recovery efficiency ($R_{SC,S}$), where $R_{SC,S}$ is dimensionless and $0 \leq R_{SC,S} \leq 1$. Sample collection typically will also be subject to imprecision, because of random differences in sample collection materials (e.g., a pre-moistened wipe) and in applying the sample collection method according to the procedure. Hence, we consider $R_{SC,S}$ as a random variable with a true, unknown mean [denoted Mean($R_{SC,S}$)] and a true, unknown percent relative standard deviation (%RSD) [denoted %RSD($R_{SC,S}$)]. Estimates of these parameters are denoted by $\overline{\text{Mean}}(R_{SC,S})$ and $\overline{\%\text{RSD}}(R_{SC,S})$. The true values and their estimates may depend on the contaminant concentration, surface material, and other factors.

2.3.3   Sample Handling, Storage, and Transportation Uncertainties

Sample handling, storage, and transportation include any steps of the SAP between 1) sample collection and 2) sample processing/extraction at the analytical laboratory. The handling, storage, and transportation of samples may cause bias uncertainty because of loss or "death" of spores, or more difficulty in culturing and counting spores. For example, longer storage time or low storage temperatures

2.6

could have such effects. For a specific sample handling, storage, and transport method and a specific set S of conditions (all denoted by S), we refer to the fraction of spores remaining after the effects of sample handling, storage, and transportation as the recovery efficiency related to these factors ($R_{HST,S}$), where $R_{HST,S}$ is dimensionless and $0 \leq R_{HST,S} \leq 1$. Sample handling, storage, and transportation effects on samples could also cause random uncertainties. Hence, we consider $R_{HST,S}$ as a random variable with a true, unknown mean [denoted Mean($R_{HST,S}$)] and a true, unknown %RSD [denoted %RSD($R_{HST,S}$)]. Estimates of these quantities are denoted by $\overline{\text{Mean}}(R_{HST,S})$ and $\overline{\%RSD}(R_{HST,S})$. The true values and their estimates may depend on the amount of contamination on the sample media, the specifics of the handling, storage, and transportation methods, and other factors.

2.3.4   Sample Processing and Uncertainties

Sample processing includes all preparations of the sample for culturing and counting, including extraction of the contaminant from the sample collection media, and preparation of the extractant to put on culture plates. Because not all of the spores on the sample media may be extracted, or the quantity may otherwise be reduced during processing steps, sample processing is subject to bias uncertainty. For a specific sample processing/extraction method and a specific set S of conditions (all denoted by S), we refer to this bias uncertainty as the recovery efficiency related to sample processing ($R_{P,S}$), where $R_{P,S}$ is dimensionless and $0 \leq R_{P,S} \leq 1$. Sample processing may also be subject to random uncertainty. Hence, we consider $R_{P,S}$ as a random variable with a true, unknown mean [denoted Mean($R_{P,S}$)] and a true, unknown %RSD [denoted %RSD($R_{P,S}$)]. Estimates of these quantities are denoted by $\overline{\text{Mean}}(R_{P,S})$ and $\overline{\%RSD}(R_{P,S})$. The true values and their estimates may depend on the amount of contamination on the sample media, the specifics of the processing methods, and other factors.

2.3.5 Sample Analysis and Uncertainties

Sample analysis by culturing includes preparation of the culture plates, putting the prepared sample extractant on the culture plates, the growth period for the culture plates, any preparations for counting, and counting the CFUs from the culture plates. Some of the contaminant spores in the extractant may not be transferred to culture plates, some spores on culture plates may not produce CFUs, and the counting process may undercount the number of CFUs. There is also the possibility of cross-contamination, where uncontaminated samples are analyzed using contaminated equipment.[a] Hence, sample analysis is subject to bias uncertainty. For a specific sample analysis method (including culturing and counting) and a specific set of conditions (all denoted by S), we refer to this bias uncertainty as the recovery efficiency related to sample analysis (denoted $R_{A,S}$), where $R_{A,S}$ is dimensionless and $0 \leq R_{A,S} \leq 1$. Sample analysis may also be subject to random uncertainty. Hence, we consider $R_{A,S}$ as a random variable with a true, unknown mean [denoted Mean($R_{A,S}$)] and a true, unknown %RSD [denoted %RSD($R_{A,S}$)]. Estimates of these quantities are denoted by $\overline{\text{Mean}}(R_{A,S})$ and $\overline{\%RSD}(R_{A,S})$. The true values and their estimates may depend on the amount of contamination in the extractant, the specifics of the analysis method, and other factors.

---

[a] Other analysis methods (e.g., polymerase chain reaction, PCR) would have different steps of the process and different possible sources of bias and uncertainty.

We also denote by $F_{CFU:Spore,S}$ the fraction of CFUs per spore obtained with a specific analysis process (including culturing and counting) and a specific set S of sampling conditions (e.g., surface material, surface concentration, humidity). If there is one CFU for each spore, then $F_{CFU:Spore,S} = 1$. If there are fewer CFUs than spores, then $F_{CFU:Spore,S} < 1$. We consider $F_{CFU:Spore,S}$ as a random variable with a true, unknown mean [denoted Mean($F_{CFU:Spore,S}$)] and a true, unknown %RSD [denoted %RSD($F_{CFU:Spore,S}$)]. Estimates of these quantities are denoted by $\overline{\text{Mean}}$ ($F_{CFU:Spore,S}$) and $\overline{\text{%RSD}}$ ($F_{CFU:Spore,S}$).

3.0   Overview of Statistical Sampling Approaches for Characterization and Clearance

There are many statistical sampling approaches (Gilbert 1987) that may be appropriate for characterization and clearance situations. Several of these approaches are implemented in the Visual Sample Plan (VSP) software (VSP Development Team 2012; Matzke et al. 2010). However, there are too many possible sampling approaches to address in this work. Hence, it was decided to focus the work for this report on the sampling approaches mentioned in the sampling reference document being developed by the VSPWG. These and other environmental sampling approaches are implemented in VSP. However, those implementations do not account for the uncertainties in the SAP, with one exception discussed subsequently. The methods and formulas for the sampling approaches implemented in VSP are included in the software documentation (Matzke et al. 2010 and help files in VSP), related technical reports, and the refereed literature (references cited in subsequent sections). In Appendix A we provide instruction how to locate the various sampling approaches in VSP that are discussed in this report.

This section presents an overview of statistical sampling approaches for characterization and clearance that are mentioned in the sampling reference document being developed by the VSPWG. Future versions of this PNNL report will include the results of any new work to extend the methods and formulas associated with these sampling approaches to account for uncertainties (systematic/bias and random) in the SAP, either directly or via the FNR.

Section 3.1 identifies the sampling approaches discussed in this document. Sections 3.2 and 3.3 describe the methods and formulas for these approaches in the characterization and clearance situations, respectively.

3.1   Possible Sampling Approaches by Zones

The sampling reference document being developed by the VSPWG discusses possible sampling approaches for four *zones* within a building or other decision area. These zones are:

Zone 1: Definitely Contaminated or Assumed to be Contaminated (i.e., extremely high likelihood of being contaminated)

Zone 2: High Likelihood of Being Contaminated

Zone 3: Low Likelihood of Being Contaminated

Zone 4: Extremely Low Likelihood of Being Contaminated.

The four types of zones and the possible sampling approaches for each are described in the following paragraphs.

Zone 1: Because this zone is known or assumed to be contaminated, detailed characterization sampling is probably not required. It is assumed that this zone will automatically undergo decontamination. However, if data are needed to confirm the assumed contamination, or additional data are needed to plan decontamination, judgment samples are generally used.

Zone 2: By definition, the likelihood of contamination in this zone is high, but before characterization there is no obvious evidence of contamination. If feasible, some judgment samples should be taken first in locations that are the most likely to be contaminated and or where the contaminant is most likely to persist. If any of those judgment samples are unacceptable (refer to Table 2.1 for the definition of what constitutes an acceptable sample under various sampling conditions), this zone can be reclassified as a Zone 1 with little or no further characterization sampling required. The exception would be if it were necessary to identify the extent of contamination within the zone, which would require additional sampling. If none of the initial judgment samples are unacceptable, then more judgment samples or a more extensive statistical sampling approach may be necessary. The statistical sampling approaches for Zone 2 mentioned in the sampling reference document being developed by the VSPWG[a] are the *hotspot sampling approach* and the *CJR sampling approach*. These approaches are described in more detail in Sections 3.2 and 3.3.

Zone 3: In a Zone 3 decision area, there is no prior evidence of contamination and consequently there is a low likelihood that the area is contaminated. These are zones the response team does not believe are contaminated, but they do not have sufficient confidence to support that conclusion. The sampling approach mentioned in the sampling reference document being developed by the VSPWG[a] is judgment sampling, which may detect contamination if it exists. If judgment sampling does not detect contamination and it is desired to clear the decision area without decontamination, one option is the CJR sampling approach, which is discussed in more detail in Section 3.3.

Zone 4: These zones have an extremely low likelihood of being contaminated. If there is sufficient evidence that there is no known plausible pathway for the contaminant to have entered this zone, then the project team may decide not to obtain any samples. This designation of Zone 4 is only included for completeness, so an entire building or other decision area can be represented, including areas where no samples will be required. If there is some non-negligible chance that contamination is present (albeit a very low chance), then the team should classify the area as a Zone 3 and use a relevant sampling approach.

3.2 Statistical Sampling Approaches for Characterization

The only statistical sampling approach for characterization mentioned in the sampling reference document being developed by the VSPWG[a] is the *hotspot sampling approach* (Gilbert 1987, Chapter 10). However, other approaches, including the CJR approach with samples distributed in a systematic grid, may also support a hotspot sampling objective during characterization while simultaneously supporting a clearance decision objective if no unacceptable samples are obtained. Although the CJR approach may be applicable during characterization, for the remainder of this report it will be discussed under the clearance phase objectives. The hotspot sampling approach provides for collecting a sufficient number of samples taken on a regular grid to have a sufficiently high confidence (Z%) of collecting at least one unacceptable sample from an unacceptable region of a specified shape (usually circular or elliptical) and size. Refer to Table 2.1 for the definition of what constitutes an *acceptable region* and an *acceptable sample* under various sampling conditions. It is assumed that the surface area sampled with

---

[a] This document is not issued yet, and hence cannot be referenced, but is tentatively titled "Reference Guide for Developing and Executing *Bacillus anthracis* Sampling Plans in Indoor Settings".

one sample is much smaller than the size of a hotspot. The grid typically consists of a square, rectangular, or triangular pattern.

The term "hotspot sampling" should not be thought of as only applicable for contamination appearing in small areas (hotspots). Rather, it should be thought of as a way to take sufficient samples to have high confidence of detecting contamination as long as it occurs in one or more areas larger than the specified shape and size within the decision area. The size of the smallest-size contaminated area (hotspot) that is desired to be detected can be increased when contamination is expected to occur over larger areas, which would lead to a relatively small numbers of samples. However, when contamination is expected to be sufficiently widespread if it occurs, judgment samples will typically be sufficient to detect the contamination. Hence, the hotspot sampling approach is probably most appropriate when contamination is sparsely distributed or appears in isolated hotspots, such that judgment sampling may not detect the contamination. Hotspot sampling approaches (or another statistical sampling approach) may also be needed if contamination exists at lower concentrations and the FNR is sufficiently high. In such cases, a few judgment samples may all result in false negatives, whereas more samples in a statistical sampling design would increase the likelihood of detecting the contamination despite a higher FNR.

The VSP software implements the hotspot sampling approach for two situations, when: 1) FNR = 0, and 2) FNR > 0. In both situations, the methods implemented in VSP assume that the FPR = 0. The FNR and FPR concepts are defined and discussed for both quantitative and qualitative data in Section 2.1. For quantitative data, a false negative is defined more specifically in VSP as a measured result being as large as the ACL when the true contamination is greater than the ACL. Hence, the FNR is defined in VSP as the probability that a sample measurement indicates contamination is less than or equal to the ACL, when in fact it is just larger than the ACL. As noted in Section 2.1, it is when the true contamination level is equal to the ACL that the FNR is the largest. Hence, using this definition of FNR with quantitative data is conservative for implementing a statistical sampling approach.

The hotspot sampling approach implemented in VSP when FNR = 0 depends on probability theory for having at least one sample on a given grid pattern falling inside a hotspot (Gilbert 1987, Singer and Wickman 1969) of a specified shape and size. The approach for FNR > 0 modifies the FNR = 0 theory and formulas to account for one assumed known value of FNR (Sego and Wilson 2007). Neither the methodology for FNR > 0 nor VSP currently address the situation where the FNR may vary over a decision area because of different 1) sample collection methods, 2) surface materials, and/or 3) contaminant surface concentrations. Further, the methodology for FNR > 0 currently only applies for hotspots of circular shape.

The hotspot sampling approach provides for making a Z% characterization statement using one of the following forms from Section 2.2: ChS1a, ChS2a, ChS1b, and ChS2b. These statements can be made more specific by defining the shape and size of the contaminated region, where the shape may be circular or elliptical (when FNR = 0) or circular only (when FNR > 0).

In summary, methodology exists for the hotspot sampling approach when FNR = 0 and FNR > 0, and these methods are applicable for qualitative (acceptable/unacceptable) and quantitative data. These methods are discussed in more detail in Section 5.

3.3 Statistical Sampling Approaches for Clearance

An *upper tolerance limit* (*UTL*) *sampling approach* provides for collecting a sufficient number of samples to make X%/Y% clearance statements as discussed in Section 2.2. An UTL sampling approach is also sometimes referred to as an *upper percentile sampling approach* because tolerance statements are confidence statements about percentiles of a distribution. Three UTL sampling approaches are discussed in this report. The first approach, sometimes called *compliance sampling,* is for quantitative data (Schilling and Neubauer 2009, Hahn and Meeker 1991). The second approach, sometimes called *accept-on-zero attribute compliance sampling* (AOZ-ACS) is for qualitative data (or for quantitative data dichotomized with respect to the ACL) and is based on the hypergeometric distribution (Squeglia 1994; Bowen and Bennett 1988). The third is the CJR sampling approach (Sego et al. 2007, 2010) which may be applied to qualitative data (or to quantitative data dichotomized with respect to the ACL). These three options are discussed in Sections 3.3.1, 3.3.2.1, and 3.3.2.2, respectively. Note that some of the text in the following subsections was adapted from the documentation accessible in VSP.

For all UTL approaches, it is recommended that X% be calculated for various values of Y%, and then choosing the largest value of Y% that yields an acceptably high value of X%, such that Y% > X%. That is, it is recommended that an X%/Y% clearance statement be made about as large a percentage of the decision area (i.e., Y%) being acceptable as possible, provided the confidence level (i.e., X%) is suitably high. As an example, a 95%/99% clearance statement would be preferred to the much weaker, 99%/95% clearance statement.

The specifics of statistical clearance sampling depend on whether the sample results are quantitative or qualitative (acceptable/unacceptable), as discussed in the balance of Section 3.3.

3.3.1 Quantitative Measurement Results

If the sample measurement results are quantitative, then either a parametric or nonparametric UTL approach for statistical sampling may be used. Both approaches are implemented in VSP.

For the nonparametric approach, no assumption is made regarding the statistical distribution of the underlying population of all possible measurement values. VSP determines the number of samples required so that the maximum sample result will be the UTL. Thus, if the maximum sample result is above the ACL, then the X%/Y% clearance statement cannot be made and the decision area would have to be decontaminated and re-sampled for clearance. If the maximum sample result is below the ACL, then the X%/Y% clearance statement can be made and the decision area would be considered acceptable, i.e., sufficiently uncontaminated (or decontaminated if the sampling occurs after decontamination).

If the underlying population of measurement values over the area being sampled can reasonably be assumed to have a normal distribution, then the parametric UTL approach can be used. This results in fewer required samples than the nonparametric UTL approach. However, the validity of the normal (Gaussian) distribution assumption must be 1) defended prior to sampling, and 2) tested after the samples have been collected and analyzed. After the normal distribution is confirmed as appropriate, the UTL is calculated via a formula of the form

$$UTL = \bar{x} + ks,$$

(3.1)

where $\bar{x}$ is the mean of the sample results, s is the standard deviation of the sample results, and k is a UTL multiplier. More details about k are given by Natrella (1966, Section 2.5), Odeh and Owen (1980), and Hahn and Meeker (1991). After the UTL is calculated, it is compared to the ACL. If the UTL exceeds the ACL, then the X%/Y% clearance statement cannot be made and the decision area would likely be subject to additional investigation, decontamination, and re-sampling for clearance. If the UTL is less than the ACL, then the X%/Y% clearance statement can be made and the decision area would be considered sufficiently uncontaminated (or sufficiently decontaminated if the sampling occurs after decontamination).

3.3.2 Qualitative (Acceptable/Unacceptable) Measurement Results

As discussed in Section 2.1, a qualitative measurement consists of a binary outcome (acceptable/unacceptable) which may arise because the measurements themselves are binary (presence/absence or detect/non-detect) or because quantitative measurements are dichotomized (i.e., greater than the ACL, less than or equal to the ACL).

Qualitative measurement results require that all surfaces in the decision area to be sampled be divided into non-overlapping, equal-size grid cells of specified size (e.g., 10 cm × 10 cm). The size of the grid cell should correspond to the surface area for a given sampling method (i.e., the area sampled by a swab, wipe or vacuum). If more than one sampling method will be used in a decision area, the size of the grid cell could be chosen to match the sampling methodology with the smallest sampling area. However, this practice could result in a larger number of samples than would otherwise be needed if the vast majority of samples are collected using methods with larger surface areas (e.g., wipes and vacuums) and only a few samples are collected using a method with small surface area (e.g., swabs).

Both of the clearance sampling methodologies for qualitative data discussed below in Sections 3.3.2.1 and 3.3.2.2 can provide an X%/Y% clearance statement (as defined in Section 2.2), so long as *all* sample outcomes are acceptable. Thus, only one unacceptable sample reduces the confidence to less than X%, which could trigger a potentially expensive response action. If all samples must be collected and processed at once, these methods may not be practical unless it is likely that none (or very few) of the grid units in the decision area are unacceptable. This would be a reasonable assumption for a decontaminated decision area, or a decision area that is believed to have a very small chance of containing contamination (e.g., a Zone 3 area).

3.3.2.1 Accept-on-Zero Attribute Compliance Sampling

When sample outcomes are qualitative and only random sampling is used, accept-on-zero attribute compliance sampling (AOZ-ACS) (Bowen and Bennett 1988; Squeglia 1994) is appropriate for clearance sampling. The AOZ-ACS method is implemented in VSP. The total number of grid cells is denoted by N. After specifying values for N, X and Y, VSP calculates the number, n, of the N grid cells ($n < N$) that must be sampled and measured. If the measurement of one or more of the n grid cells is unacceptable (refer to Table 2.1 for the definition of *acceptable*), the X%/Y% clearance statement cannot be made.

The AOZ-ACS approach assumes that the sample collection and analysis process always obtains the correct result for each sample (i.e., FNR = 0 and FPR = 0). If the FNR > 0, the clearance statement must

use the appropriate definition of an *acceptable decision area*, as indicated in Table 2.1 and clearance statements ClS3b and ClS4b (in Section 2.2).

3.3.2.2   Combined Judgment and Random Sampling Approach for Clearance

The CJR sampling approach uses Bayesian acceptance sampling methodology to make an X%/Y% clearance statement (see Section 2.2).   It augments the results of judgment samples with samples whose locations are either randomly chosen or positioned on nodes of a grid with a random start location. The CJR approach is discussed in detail by Sego et al. (2007, 2010), but is briefly summarized here and in Section 6.

The grid cells in the decision area are divided into two groups, high-risk and low-risk, corresponding to the risk of being contaminated. The CJR sampling approach presumes that *all* the high-risk grid cells are sampled judgmentally and that a random sample is taken from the low-risk grid cells. This approach permits judgment samples to be obtained from locations perceived as most likely to be contaminated (by knowledgeable experts) while also collecting randomly located samples in case contamination occurs in locations not investigated by judgment samples. *It would be a misuse of the CJR approach to arbitrarily pick a larger number of judgment samples in order to reduce the number of random samples.* That is, samples that are selected judgmentally should include all locations genuinely believed to be of higher risk than the remainder of the decision area.

In addition to X% and Y%, there are other important parameters involved in the CJR method and formulas. These include:

- the number of judgment samples (n_J)

- the number of random samples (n_R)

- the *a priori* probability of that a judgmental sample, taken from a high-risk location, will be unacceptable. This probability (which is assumed to be the same for all judgment sample locations) is denoted P_J, where $0 < P_J < 1$. Equivalently, P_J can also be thought of as the *a priori* rate of unacceptable high-risk grid cells.

- how much more likely it is that a judgment sample location is unacceptable compared to the possible random sample locations. This ratio is denoted $r_{J:R}$, where $r_{J:R} \geq 1$.

The last parameter ($r_{J:R}$) facilitates combining the results from judgment and random samples. For specified values of X%, Y%, P_J, and $r_{J:P}$, the CJR methodology in VSP can be used to determine the required number and placement of random samples (n_R) given the planned number of judgment samples (n_J). For predetermined values of P_J , $r_{J:P}$, , n_J, n_R, and Y%, the CJR methodology can be used to determine the achieved level of confidence, X%, corresponding to an X%/Y% clearance statement. Additional details of the CJR sampling approach for these two applications are discussed in Section 6.

Wright (1992), Grieve (1994), and Axelrod (2005) present a Bayesian model which addresses a very similar problem to the one addressed by the CJR approach, except the population is homogeneous (i.e., the risk is the same for all members of the population). The Wright, Grieve, and Axelrod models are exactly equivalent to a special case of the CJR approach when $n_J = 0$ and $r_{J:P} = 1$ (Sego et al. 2010). Traditional frequentist (i.e., non-Bayesian) approaches to acceptance sampling (Schilling and Neubauer 2009; Bowen and Bennett 1988) also do not consider subpopulations with different levels of risk. It is

also interesting to note that the AOZ-ACS method gives results that are virtually identical to the special case of the CJR method when $P_J = 0.5$, $n_J = 0$ and $r_{J:P} = 1$ (Sego et al. 2010).

4.0   Estimating the False Negative Rate for a Single Sample

As discussed in Section 2.2, there are two aspects of the formulas for calculating Z% characterization statements and X%/Y% clearance statements, as well as the number of samples to achieve these statements with desired values of Z% or X%/Y%.

The first aspect concerns the variability or error that arises from statistical sampling. For characterization, this error arises from the probability of not sampling at least one contaminated location in the decision area. The number of samples collected according to the sampling approach must be sufficiently large so that this probability is small. Stated another way, a sufficient number of samples must be collected according to an appropriate statistical sampling approach so that there is a high probability of sampling at least one contaminated location. For clearance, the statistical sampling error arises from predicting the probability that unsampled portions of the decision area also satisfy the clearance criteria.

The second aspect is related to the probability of failing to detect contamination with a single sample of a contaminated location. This is referred to as the FNR (or PFN), as discussed in Section 2.1.

Sampling methodology and formulas for characterization (Section 5) and clearance (Section 6) that account for the FNR > 0 require quantitative estimates of the FNR. Typically, quantitative estimates of the FNR must be obtained from experimental studies for a given SAP (i.e., sample collection, handling/storage/transportation, processing/extraction, and analysis). Because the FNR may depend on the surface material being sampled and the concentration of surface contamination (see Section 2.1), the effects of these factors must be investigated in the experimental studies for each given set of SAP methods. This is accomplished by determining the experimental values of the FNR for a range of contaminant concentrations on each of several surface materials spanning the region of surface characteristics. Then, a regression equation can be developed to relate the experimentally determined FNR values to the contaminant concentrations and surface material characteristics. The regression equation makes it possible to calculate the FNR for any contaminant concentration and surface material within the ranges of concentration and surface material characteristics tested experimentally.

Krauter et al. (2012) implemented most of the above approach for the set of SAP methods specified by the CDC for sponge-wipe sampling (Rose et al. 2011). They experimentally determined FNR values for eight contaminant concentrations at each of six nonporous surface materials. Using these data, they developed regression equations for the FNR as a function of contaminant concentration separately for each of the six surface materials, and showed that FNR increases with increasing surface roughness index. It was beyond the scope of their work to develop FNR equations as functions of contaminant concentration and surface roughness index, although this certainly could be done.

In a review of 20 laboratory studies on the performance of various combinations of SAP methods with *Bacillus anthracis* (or surrogate) contaminants, Piepel et al. (2011a, 2011b) noted that only the Krauter et al. (2012) study experimentally quantified FNR values. They identified this as a significant gap that needs to be addressed by new studies and evaluations of previous studies to focus on quantifying FNR as a function of contaminant concentration and surface material characteristics. Any such future work will be documented in this section in future revisions of this report.

5.0 Methods and Formulas for Determining the Number of Samples and Confidence for Characterization Decisions Using Statistical Sampling Approaches

This section summarizes the sampling methodology and formulas for determining sampling plans and confidence statements for characterization situations with and without accounting for the uncertainties in the SAP. Accounting for systematic/bias and random uncertainties in the SAP is accomplished by including the FNR in the statistical sampling design. Situations with qualitative (acceptable/ unacceptable) and quantitative data are addressed. These approaches provide for collecting a sufficient number of samples taken on a regular grid to have a sufficiently high confidence (Z%) of collecting at least one unacceptable sample from an unacceptable region of a specified shape and size. Refer to Table 2.1 for the definition of what constitutes an *acceptable region* and an *acceptable sample* under various sampling conditions.

Hotspot sampling is often described as having the objective of determining whether "hotspots" (small contaminated areas) are present. It is from such situations that hotspot sampling gets its name. However, the hotspot sampling approach is applicable for detecting contamination regardless of whether it occurs in small hotspots or is more widely distributed within a decision area. The hotspot sampling approach typically uses samples laid out over the decision area in a square, rectangular, or triangular (equilateral) grid pattern to achieve a specified confidence that at least one sample will detect a hotspot of a specific shape and size. If preliminary information indicates that contamination may be more widely distributed rather than in a single, small hotspot, then the responder can specify a larger hotspot size to be detected, thus reducing the number of samples required. However, if contamination is widespread enough and occurs in expected locations, then judgment sampling will likely be sufficient to detect it. Hence, hotspot sampling may only be needed in practice when it is suspected that contamination exists in only one (or a few) small, isolated locations (hotspots) within a decision area that may not be detected by the judgment sampling approach.

The methods and formulas for the following characterization sampling cases are presented in the subsections shown.

- Hotspot sampling approach with qualitative (acceptable/unacceptable) data
  - FNR = 0, Section 5.1.1
  - FNR > 0, Section 5.1.2
  - Example, Section 5.1.3
- Hotspot sampling approach with quantitative data
  - FNR = 0, Section 5.2.1
  - FNR > 0, Section 5.2.2
  - Example, Section 5.2.3

Only methods and formulas for the hotspot sampling approach (explained more in Section 5.1) are presented in Section 5 for this initial version of the report. The reason for this is that under the most

likely deposition scenarios for contamination incidents, the contaminant will tend to be widely distributed, or it may be distributed along paths that can be identified based on knowledge of the facility (e.g., mail paths in a postal facility or office building) and information about the point of contaminant release. In such situations, judgment sampling is likely to be sufficient to detect the contamination during the discovery and/or characterization phases in responding to an incident. Hence, hotspot sampling would be appropriate when the contamination occurs in only one or a few small, isolated or unexpected locations within a decision area, such that the contamination may not be detected by judgment sampling.

Section 5.3 addresses the situation where a hotspot design is implemented and all samples are found to be acceptable. Section 5.4 briefly mentions the improvements needed to address limiting assumptions in the hotspot sampling methodology discussed and illustrated in Sections 5.1 and 5.2.

5.1 Methods and Formulas for the Hotspot Sampling Approach with Qualitative (Acceptable/Unacceptable) Data

This section addresses qualitative (acceptable/unacceptable) data and summarizes methods for determining the grid spacing and number of samples needed to detect with Z% confidence an elliptical (or circular) hotspot of a specific shape and size in a decision area. Section 5.1.1 summarizes the methods when the FNR = 0, while Section 5.1.2 summarizes the methods when the FNR > 0. Section 5.1.3 illustrates results of the methods and formulas summarized in Sections 5.1.1 and 5.1.2.

5.1.1 Methods and Formulas for the Hotspot Sampling Approach with Qualitative Data when the False Negative Rate is Zero

This section gives an overview discussion and cites references for the method and formula to determine locations and numbers of samples needed to achieve a specified Z% confidence in detecting a hotspot of a given shape and size when using the hotspot sampling approach with qualitative (acceptable/unacceptable) data when the FNR = 0. Specifically, the method and formula for sampling to detect an elliptical (or circular) hotspot shape is presented. While contamination is unlikely to occur in nice regular shapes such as an ellipse (or circle), the corresponding methodology is still useful in providing guidance on the locations and number of samples to collect when the hotspot sampling approach is appropriate. The hotspot methodology assumes that the size of the sample footprint is negligible with respect to the size of the decision area, in which case the samples are called "point samples.

An algorithm used to calculate the probability of at least one sample being placed in a <u>single</u> elliptical (or circular) hotspot (assuming the hotspot exists), was developed and refined by Singer and Wickman (1969), Singer (1972), and Singer (1975). The methodology is summarized in Gilbert (1987) and additional refinements were made by Davidson (1995). The following steps of the procedure for determining the grid spacing and number of samples are based on the Gilbert (1987) explanation, with Step 4 updated to include Davidson's (1995) revisions.

1. Specify L, the half-length of the major (i.e., long) axis of the smallest elliptical hotspot important to detect.

2. Specify the shape parameter (S) of the elliptical hotspot, where S is the ratio of the half-length of the minor (i.e., short) axis of the ellipse to L. Note that $0 < S \leq 1$, and that $S = 1$ for a circle. If S

is not known in advance, a conservative approach is to assume a rather thin elliptical shape, perhaps $S = 0.5$, to give a smaller grid spacing than if a more circular or wider ellipse is assumed. Alternatively, designing for a small circular hotspot would also be conservative. Sampling on a finer grid is used to compensate for lack of knowledge about the target shape.

3. Specify the required Z% percent confidence (where $Z = 100[1 - \beta]$) that the hotspot will be found, where β is the acceptable probability of not finding the hotspot, given that it exists.

4. Select the grid pattern (square, rectangular, or triangular) for grid sampling.

5. Using Davidson's (1995) refined methodology implemented in a computer program (and available in VSP), compute s (the spacing between the grid lines) using L, S, β, and the selected grid pattern (square, rectangle, equilateral triangle). Specifically, s represents the length of a side of a square or triangular grid cell. When the grid pattern is rectangular, s represents the length of the short side of a rectangular grid cell, and an additional parameter ρ represents the ratio of the lengths of the long and short sides of a rectangular grid cell.

6. Calculate the approximate number of samples required ($\tilde{n}$) using the formula

$$\tilde{n} = A_{DA} / A_S \tag{5.1a}$$

where A_{DA} is the area of the decision area to be sampled and A_S is the area of a grid cell given by

$$A_S = \begin{cases} s^2 & \text{square grid pattern} \\ \rho\, s^2 & \text{rectangular grid pattern} \\ \dfrac{\sqrt{3}}{4} s^2 & \text{equilateral triangle grid pattern} \end{cases} \tag{5.1b}$$

The exact number of samples (n) can be determined by placing the grid pattern onto the decision area to be sampled (with a randomly determined starting location) and counting the number of samples inside the decision area.

Gilbert (1987) specifies five assumptions for the above hotspot sampling methodology. The VSP help files and documentation expand the assumptions as follows:

a. At least one circular or elliptical hotspot of the designated size is hypothesized to exist in the decision area.

b. The level of contamination that defines a hotspot is well defined.

c. The location of the hotspot is unknown, and if a hotspot is present, all locations within the sampling area are equally likely to contain the hotspot.

d. With a randomly determined starting location, samples are taken on a square, rectangular, or triangular (equilateral) grid pattern that covers the decision area.

e. Each sample is collected, handled, measured, or inspected using approved methods that yield sufficiently precise measurements.

f. A very small proportion of the surface of the decision area will be sampled. The area sampled by a single sample is much smaller than the hotspot of interest.

g. The sample methodology and sample analysis process is the same for all sample locations.

h. There are no classification errors. If a hotspot is sampled, then contamination is detected (i.e., no false negatives). If an uncontaminated area is sampled, it is not mistakenly identified as a hotspot (i.e., no false positives).

i. A hotspot is detected if any part of the hotspot lies within the area covered by one or more samples, and contaminant is detected on at least one of these samples. For example, if only a portion of the corner of a single sample footprint intersects the edge of a hotspot, the hotspot will be detected.

If there is more than one hotspot in the decision area (Assumption "a"), the methodology is conservative, i.e., the probability of detecting contamination is greater than specified. Note that Assumption "h" corresponds to assuming FNR = 0 and FPR = 0. See Section 2.1 for discussion of the FNR and FPR concepts and terminology.

Sego and Wilson (2007) gave a heuristic explanation of the Singer-Wickman algorithm. To simplify the explanation, they focused on the case of using a square grid with point samples to detect a circular hotspot with radius R. In Figure 5.1, (from Sego and Wilson (2007), used with permission) the black square represents the area between four adjacent sample points, and the four red circles of radius R are centered on these sample points. It is assumed that the hotspot will be centered inside the square (or inside an equivalent square formed by four adjacent sample points in the grid). If the center of the hotspot lies inside one of the red circles as shown with the green circle in Figure 5.1, it will overlap a sample point and will be detected. If the center of the hotspot does not lie inside one of the red circles as shown with the blue circle in Figure 5.1, it will not be detected.

The strategy of the Singer-Wickman algorithm is to calculate the percentage of the area of the square that is occupied by the intersection of the square and the union of the four red circles. In order to obtain relatively high probabilities of detecting a hotspot, the grid size must be small enough for the four circles that are centered on the sample points to cover most, or all, of the square.
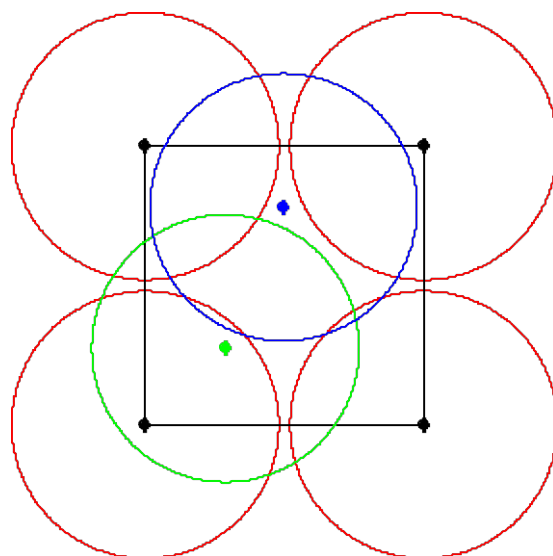
Figure 5.1. Illustration of the Singer-Wickman Algorithm. All circles have a radius R = 0.48 units. The length of the side of the square is 1 unit. For simplicity in this illustration, the black points represent point samples. Hotspots with radius 0.48 units whose centers lie inside one of the red circles will be detected (e.g., green circle). Hotspots whose centers do not lie inside a red circle will not be detected (e.g., blue circle).

VSP implements the methodology described in this subsection to calculate the grid spacing and number of samples given the grid pattern and the desired Z% confidence of detecting a circular hotspot of specified radius. This is illustrated with an example in Section 5.1.3.

5.1.2 Methods and Formulas for the Hotspot Sampling Approach with Qualitative Data when the False Negative Rate is Greater Than Zero

This section summarizes the method and formula for determining the locations and number of samples needed to achieve a specified Z% confidence in detecting a circular hotspot of a given size when using the hotspot sampling approach with qualitative (acceptable/unacceptable) data when the FNR > 0. The methods discussed in this section assume the same FNR > 0 applies to all samples within a decision area. This may be an unrealistic assumption, given that different sample collection methods (e.g., swab, wipe, and vacuum) with different SAP methods may be used to sample locations that have different surface materials (e.g., vinyl, carpet, glass). The FNR is likely to vary with the sample collection method and surface material, and may vary depending on other steps of the SAP and environmental conditions (e.g., humidity) as noted by Krauter et al. (2012). A conservative approach in applying the existing hotspot sampling methodology is to assume the largest FNR that may apply for any of the samples to be collected given the surface materials to be sampled and the specific sampling and analysis methods to be used. It is an open need to extend the hotspot sampling approach to accommodate more realistic situations where different combinations of SAP methods and surface materials have different FNRs. This topic is discussed in further in Section 5.4.

This section gives an overview discussion and cites references for the method and formula to determine locations and numbers of samples needed to achieve a specified Z% confidence in detecting a circular hotspot of a given size when using the hotspot sampling approach with qualitative (acceptable/unacceptable) data when the FNR > 0. While contamination is unlikely to occur in nice circular shapes, the methodology is still useful in providing guidance on the locations and number of samples to collect when the hotspot sampling approach is appropriate.

An algorithm to calculate the probability of detecting a circular hotspot when the FNR > 0 was developed by Sego and Wilson (2007). The algorithm builds on the method discussed in Section 5.1.1 and makes possible calculating the grid size, locations, and number of samples. The algorithm has only been developed for "point samples" (i.e., samples that cover a negligible area). If the FNR remains the same, regardless of how much (or how little) of the sample footprint intersects with the hotspot, treating samples with a non-negligible footprint as if they were point samples is conservative. This means that the actual probability of detecting a hotspot using sample collection methods may be slightly larger than anticipated. However, if the FNR depends on the fraction of the sample footprint that intersects with the hotspot, it is possible that the actual probability may be slightly less than the desired Z% confidence.

The steps of the method to apply the algorithm are:

1. Specify R, the radius of the smallest circular hotspot important to detect.

2. Specify the required Z% confidence (where $Z = 100[1 – \beta]$) that the hotspot will be found, where β is the acceptable probability of not finding the hotspot, given that it exists.

3. Specify the false negative rate (η).

4. Select the grid pattern (square, rectangular, or triangular) for grid sampling. Denote by s the length of a side of a square or triangular (equilateral) grid cell, or the length of the short side of a rectangular grid cell. When the grid pattern is rectangular, an additional parameter ρ represents the ratio of the lengths of the long and short sides of the rectangular grid cell.

5. Using the Sego and Wilson (2007) methodology, compute s (the spacing between grid lines) using R, Z, η, the selected grid pattern (square, rectangle, equilateral triangle), and, for rectangular grids, ρ.

6. The approximate ($\tilde{n}$) and exact (n) number of samples required to achieve Z% confidence are calculated using Equations (5.1a) and (5.1b) as described in Step 6 of Section 5.1.1.

An overview of the Sego and Wilson (2007) methodology and associated formulas for circular hotspots are presented in the balance of this section. A methodology for detecting elliptical hotspots is not currently available. However, the theory could be developed for the elliptical case, or at the very least, the probabilities of detection could be simulated.

Recall from Section 5.1.1 and Figure 5.1 that to obtain relatively high probabilities of detecting a hotspot, the grid size must be small enough for the four circles that are centered on the sample points to cover most, or all, of the square. For FNR > 0, high probabilities of detection will require substantial overlap of the circles shown in Figure 5.1. A hotspot that is centered in a region of overlap will be sampled more than once, which will increase the probability of detecting the hotspot. The strategy for extending the algorithm in Section 5.1.1 to account for false negatives is to identify areas within the square where there is overlap with one, two, three, or four of the red circles. Conditional probabilities of

detection can be calculated for each of these areas, and then a total probability of detection is calculated by summing the products of the areas' conditional probabilities and proportions of the total area. The methods and formulas for rectangular (including square) and triangular grids are now briefly summarized.

Method for Rectangular (or Square) Grid Sampling and a Circular Hotspot: In Figure 5.2 (from Sego and Wilson (2007), used with permission), an example similar to Figure 5.1 is shown with one rectangle formed by sample points from a rectangular grid pattern. This example is simplified for the case where circles in one rectangle do not overlap circles in another rectangle, resulting in a maximum of four circles overlapping. The distribution of the number of samples within the hotspot is determined by calculating the proportion of the total area within the rectangle that is overlapped by different numbers of circles. The proportion of the area that is numbered "1" is the probability that one sample is taken in the circular hotspot, the proportion numbered "2" is the probability that two samples are taken in the circular hotspot, and so on. The overall probability of detecting the hotspot (accounting for the FNR) is given by:

$$P(r,\eta,\rho) = \frac{1}{\rho} \sum_{k=1}^{4} |A_k(r)|(1-\eta^k) \qquad (5.2)$$

where

$P(r, \eta, \rho)$ = probability of detecting a hotspot with standardized radius r and false negative rate η when using grid sampling for a hotspot of radius R with grid parameters s and ρ.

R = radius of the circular hotspot

s = length of the short side of a rectangular grid cell

r = R/s, the standardized radius of the hotspot, where it is assumed that $0 < r \leq 1$

ρ = ratio of the length of the long side to the length of short side of a rectangular grid cell. Hence, ρs is the length of the long side of a rectangular grid cell. Note that $\rho = 1$ gives a square grid pattern.

$|A_k(r)|$ = area (calculated using Euclidean geometry) of the region $A_k(r)$ inside a standardized rectangular grid cell where a hotspot centered in that region would be sampled k times

η = the FNR, assumed to be the same for all samples.

Note that (5.2) can be solved numerically for R or s to determine the hotspot radius R that can be detected with Z% confidence and grid size s, or to determine the grid size s required to detect a hotspot of size R with a specified Z% confidence. As the FNR increases, smaller grid spacing (e.g., smaller rectangles in Figure 5.2) is required to compensate for the reduction in probability of detection. For smaller grid spacing, circles in the grid of interest will begin to overlap with circles from adjacent rectangular grid cells when $r > 1$, resulting in more than four circles overlapping. The algorithm is more complex when $r > 1$, so Sego and Wilson (2007) discuss and recommend an approximation algorithm for calculating the probability of detection when $1 < r \leq 2$.
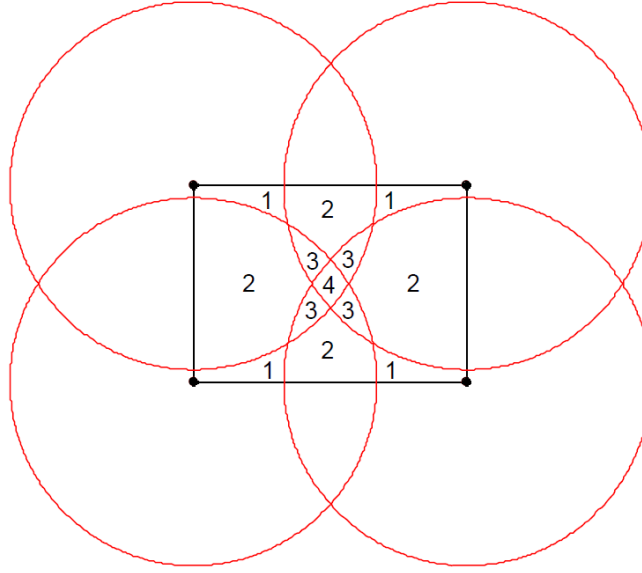
Figure 5.2. Regions of Overlap within a Rectangular Sample Grid when Using the Hotspot Sampling Approach with a Circular Hotspot. The numbers in the figure represent the number of times a hotspot would be sampled if it were centered in that region.

Approach for Triangular (Equilateral) Grid Sampling and a Circular Hotspot: For triangular grid sampling, the approach is the same except the grid cell is an equilateral triangle defined by three points instead of a rectangle defined by four points (Sego and Wilson 2007). The only difference in the formula for the probability of detection with a triangular grid when $0 < r \leq 1$ is that the $|A_k(r)|$ must be scaled by

$\sqrt{3}/4$, the area of a standardized equilateral triangle (with sides of unit length). Hence, for triangular grids, the overall probability of detection can be stated as

$$P(r,\eta) = \frac{4}{\sqrt{3}} \sum_{k=1}^{4} |A_k(r)|(1-\eta^k), \quad 0 < r \leq 1 \tag{5.3}$$

where all notation is as defined following Equation (5.2). Equation (5.3) can be solved numerically for R or s to determine the hotspot radius R that can be detected with $Z\%$ confidence and grid size s, or to determine the grid size s required to detect a hotspot of size R with a specified $Z\%$ confidence. As with the rectangular problem, the calculations become increasingly complex when the FNR is large enough that $r > 1$, so Sego and Wilson (2007) discuss and recommend an approximation algorithm for calculating the probability of detection when $1 < r \leq 2$.

For rectangular, square, or triangular grid patterns, after the grid spacing (s) is determined using the preceding methods and equations, the approximate number of samples required ($\tilde{n}$) can be calculated using Equation (5.1a) with the appropriate portion of Equation (5.1b) for rectangular, square, or triangular grid patterns. The exact number of samples (n) can be determined by overlaying the grid pattern onto the decision area to be sampled (with a randomly determined starting location) and counting the number of samples that lie in the decision area.

The assumptions for the methodology presented in this subsection, based on the assumptions provided in the VSP documentation, are as follows:

    a. At least one circular hotspot of the designated size is hypothesized to exist in the decision area.

    b. The level of contamination that defines a hotspot is well defined.

    c. The location of the hotspot is unknown, and if a hotspot is present, all locations within the sampling area are equally likely to contain the hotspot.

    d. With a randomly determined starting location, samples are taken on a square, rectangular or triangular (equilateral) grid pattern to cover the decision area.

    e. Each sample is collected, handled, measured or inspected using approved methods that yield sufficiently precise measurements.

    f. A very small proportion of the surface of the decision area will be sampled. The area sampled by a single sample is much smaller than the hotspot of interest.

    g. The sample methodology and sample analysis process is the same for all sample locations.

    h. There are no false positives (the sample of an uncontaminated location is not mistakenly identified as contaminated).

    i. The false negative rate is known and is the same for all sample determinations.

    j. The false negative rate is the same regardless of whether the footprint of a sample location lies completely inside or only partially inside a hotspot.

If there is more than one hotspot in the decision area (Assumption "a"), the methodology is conservative (i.e., the probability of detecting contamination is greater than specified). Assumption "i" may not be realistic if there are different surface materials in the decision area that will be sampled and analyzed using different sample collection, processing, and analysis methods. This assumption could be satisfied by using the largest FNR that applies for any of the sampling and analysis methods to be used, which would be conservative. Ideally, this methodology would be extended in the future to a stratified approach that allows different FNR values for different strata consisting of combinations of sampling and analysis methods and surface materials. This is discussed in Section 5.4.

VSP implements the methodology in this subsection to calculate the grid spacing and number of samples given the grid pattern, false negative rate, and the desired Z% confidence of detecting a circular hotspot of specified radius. This is illustrated for an example in Section 5.1.3.

5.1.3 Examples of Applying the Formulas for the Hotspot Sampling Approach with Qualitative Data

This section presents examples of applying the methods and formulas for hotspot sampling described in Sections 5.1.1 and 5.1.2. The example is based on a remote and unoccupied two-story office building (denoted PBF-632) on the Idaho National Laboratory (INL) site. The INL PBF-632 building was utilized in field tests denoted INL-1 (Amidan et al. 2007, 2009) and INL-2 (Piepel et al. 2009). The first floor has 11 offices, a reception area (lobby), men's and women's restrooms, and a mechanical room. The second floor has 15 offices, two storage rooms, men's and women's restrooms, and a mechanical room. Example calculations were performed for 1) a typical office (approximately 9.5 ft by 21 ft), and 2) one floor

(~ 4025 sq ft) using VSP to implement the methods and formulas summarized in Sections 5.1.1 and 5.1.2. All example calculations were for a circular hotspot, a triangular grid pattern, samples covering an area of 1 ft × 1 ft when the FNR = 0, and point samples when the FNR > 0. As noted previously in Section 5.1.2, the methodology has only been developed for point samples when the FNR > 0.

To develop a characterization sampling design using the hotspot sampling approach for a contamination incident, the response team should first identify the range of contaminant concentrations that may be present in the decision area to be sampled during the characterization phase. If the minimum concentration expected is higher than where false negatives begin to appear, given the SAP and the surface materials to be sampled, then calculations can be performed using the methods and formulas for FNR = 0 (Section 5.1.1). The response team then needs to decide on 1) the shape (ellipse or circle when FNR = 0, and circle when FNR > 0) and size (e.g., radius of circle) of hotspot to be detected, 2) the desired confidence with which it should be detected (e.g., 95%), and the grid pattern (square, rectangular, triangular) to use. Then the grid spacing required to meet those detection goals using the hotspot sampling approach can be determined and the approximate and exact number of samples to be collected over the decision area can be determined using the methods in Section 5.1.1 (FNR = 0) or in Section 5.1.2 (FNR > 0).

Figure 5.3 summarizes the number of statistical samples required to sample a typical office of INL PBF-632 and detect, with 95% confidence, circular hotspots with radii ranging from 1.5 ft to 4 ft and FNRs ranging from 0 to 0.50. The calculations were performed using VSP that implements methods from Section 5.1.1 for FNR = 0 and Section 5.1.2 for FNR > 0. When the FNR = 0, the number of samples varies from 3 to detect a 4 ft radius hotspot to 22 to detect a 1.5 ft diameter hotspot. The number of samples increases as the FNR increases, such that for FNR = 0.50 the number of samples varies from 18 to detect a 4 ft radius hotspot to 128 (not shown in Figure 5.3) to detect a 1.5 ft radius hotspot.
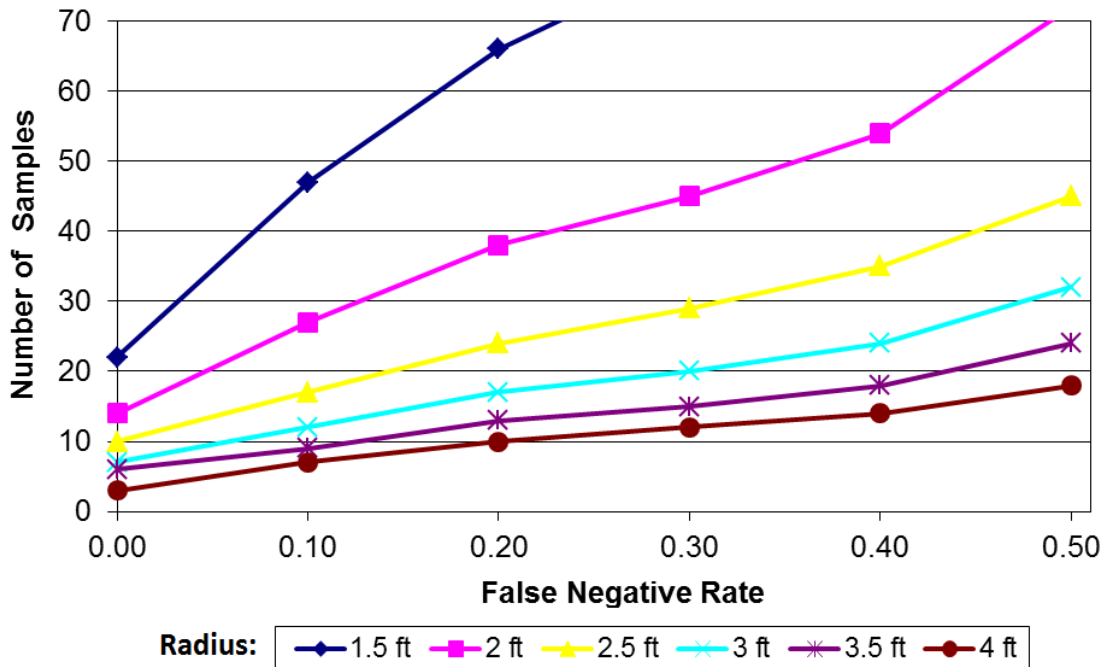
Figure 5.3. Number of Samples Required to Detect with 95% Confidence a Circular Hotspot of a Given Radius (represented by the colored lines) within a Typical Room of the INL PBF-632 Building

Figure 5.4 summarizes the results of calculating the number of samples required to sample a single floor of INL PBF-632 and detect a circular hotspot with 95% confidence when the radius of a circular hotspot was varied from 3 ft to 10 ft and the FNR was varied from 0 to 0.50. When the FNR = 0, the number of samples varies from 9 to detect a 10 ft radius hotspot to 79 to detect a 3 ft radius hotspot. The number of samples increases as the FNR increases, such that for FNR = 0.50 the number of samples varies from 43 to detect a 10 ft radius hotspot to 455 to detect a 3 ft radius hotspot.

Figure 5.3 and Figure 5.4 show how the numbers of statistical samples substantially increase as the FNR increases and the size of hotspot to be detected decreases, both when sampling a single office or a single floor of a small office building. However, it should be kept in mind that the number of statistical samples in Figure 5.3 and Figure 5.4 are for samples covering a 1 ft × 1 ft area when FNR = 0, but are "point samples" when FNR > 0. Consequently, the number of samples may be slightly overestimated for the cases when FNR > 0. As explained in Section 5.1.3, this is because the hotspot sampling methodology has not been developed to account for sample areas when the FNR > 0.
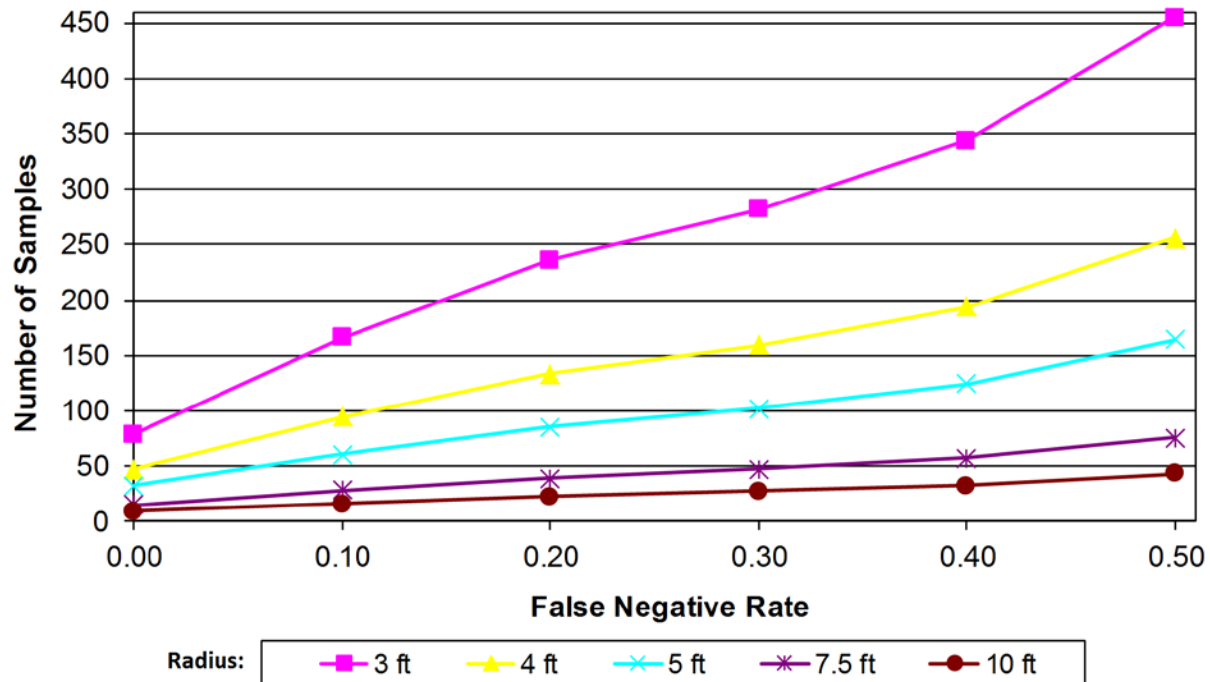
Figure 5.4. Number of Samples Required to Detect with 95% Confidence a Circular Hotspot of a Given Radius (represented by the colored lines) on a Single Floor of the INL PBF-632 Building

5.2   Methods and Formulas for the Hotspot Sampling Approach with Quantitative Data

Consider characterization sampling situations where quantitative data (e.g., contaminant surface concentrations, CFU/cm^2) greater than zero are obtained for some or all samples from a hotspot sampling plan. Hence, there may be some quantitative data and some qualitative (non-detect) data. Although quantitative values less than a limit of detection (LOD) are often reported as "non-detect" values due to uncertainty in the measurement process, it is better statistically to use quantitative measured values whenever they are available, despite the uncertainty (Gilbert 1987, Section 14.1). However, even if this is possible for some data points that would otherwise be reported as non-detects (or less than LOD), there may still be samples for which there are no quantitative values and only non-detects (less than LOD) can be reported.

When the hotspot sampling approach is used and some or all sample results are quantitative, and the measured value or the LOD is less than or equal to an ACL, the current practice is to dichotomize the quantitative results with respect to the ACL (which may be zero if no contamination is acceptable). Specifically, quantitative results are mapped into two qualitative outcomes: "above ACL," or "unacceptable" (for measured values > ACL) and "below ACL," or "acceptable" (for measured values and/or non-detects ≤ ACL). Hence, the hotspot sampling methods and formulas for qualitative data (as discussed in Section 5.1) can be applied. This current practice is discussed in Sections 5.2.1 and 5.2.2 for when the FNR = 0 and FNR > 0, respectively.

5.2.1 Methods and Formulas for the Hotspot Sampling Approach with Quantitative Data when the False Negative Rate is Zero

Consider contamination incidents in which the contamination to be detected by a hotspot sampling approach has concentrations much larger than the ACL (so that FNR = 0). For such situations, the current practice of converting quantitative data to qualitative data and applying the qualitative hotspot sampling approach (see Section 5.2) could not be improved upon by using statistical methods that make use of quantitative data (which do not currently exist). The preceding sentence presumes the sampling objective is to detect the presence of hotspots defined by the ACL, and not measure the intensity of those hotspots. Hence, the approach of converting any quantitative data to qualitative data and applying the qualitative hotspot sampling approach is appropriate for situations where, if contamination is present, it is expected to occur at concentrations well above the ACL such that FNR = 0. For this situation, the methods discussed in Section 5.1.1 are applicable. The following subsection addresses the remaining situations.

5.2.2 Methods and Formulas for the Hotspot Sampling Approach with Quantitative Data when the False Negative Rate is Greater Than Zero

Consider contamination incidents in which the contaminant concentrations to be detected using the hotspot sampling approach are not significantly greater than the ACL. In such situations, it is possible that measured contaminant concentrations may be less than the ACL (because of inefficiencies and/or random uncertainties in the steps of the SAP, as discussed in Section 4.0), even though the actual contaminant concentrations are greater than the ACL. These would lead to false negatives (i.e., FNR > 0). For such situations, any quantitative data could be converted to qualitative data and the qualitative hotspot sampling approach applicable when FNR > 0 (see Section 5.1.2) could be applied. This approach compensates for the FNR > 0 by collecting larger numbers of samples than would be the case when FNR = 0 (and using the methodology in Section 5.1.1).

5.2.3 Examples of Applying the Formulas for the Hotspot Sampling Approach with Quantitative Data for a Range of False Negative Rates

Because the only existing method for applying the hotspot sampling approach with quantitative data is to convert the quantitative data to qualitative data (i.e., "< ACL" is classified as "acceptable", "≥ ACL" is classified as "unacceptable") the examples in Section 5.1.3 for qualitative data suffice. If hotspot sampling methodology for quantitative data is developed as part of future work, it may be documented in a revision of Section 5.2.2, and illustrated in a revision of this section.

5.3 When All Samples Results from Hotspot Sampling Are Acceptable

After a hotspot sampling methodology is implemented, if none of the samples are found to be unacceptable, investigators will likely want to know how likely is it that un-sampled portions of the decision are also acceptable, possibly in anticipation of making an X%/Y% clearance statement regarding the decision area. In this situation, we recommend using the accept-on-zero compliance sampling (Section 3.3.2.1) or the CJR method (Sections 3.3.2.2 and Section 6) to 1) calculate the confidence

achieved by the hotspot design, and/or 2) determine the number of additional samples required to make the desired confidence statement. This situation is discussed in more detail in Section 6.3.

5.4   Possible Improvements in the Hotspot Sampling Methodology

Possible improvements in the hotspot sampling methodologies discussed in Sections 5.1 and 5.2 are now briefly mentioned. These improvements are related to assumptions made in developing the statistical sampling theory and formulas for the hotspot sampling approaches with qualitative and quantitative data when FNR = 0 and when FNR > 0. Extensions of the hotspot sampling methodology to address these items are possible with sufficient resources to develop the more complicated algorithms and formulas that will be needed.

False Negative Rate = 0 with Qualitative Data (or Quantitative Data Converted to Qualitative Data)

- Extend methodology to allow samples obtained with different collection methods (e.g., swab, wipe, and vacuum) and/or different surface materials within a decision area to have different sample footprints (areas).

False Negative Rate > 0 with Qualitative Data (or Quantitative Data Converted to Qualitative Data)

- Extend methodology to allow samples covering a specified footprint (area), instead of the current limitation of "point samples" (samples covering negligible area).

- Extend methodology to allow samples obtained with different collection methods (e.g., swab, wipe, and vacuum) and/or different surface materials within a decision area to have different sample footprints (areas) and different FNRs.

- Extend methodology to allow hotspots to have elliptical shapes, rather than only circular shapes.

The extensions in the first and second bullets would be very complicated, in that the hotspot would have to be well-defined to account for partial overlap of the hotspot with the sample footprint. A more complicated FNR model would also be required to account for the case when a sample footprint partially intersects with the edge of a hotspot. An important question is whether these complicated extensions would be worth the effort, especially when the concept of a well-defined hotspot isn't compatible with most contaminant-dispersion models.

In the third bullet, there may be little value-added to extending the methodology to handle elliptical as well as circular hotspots, because ultimately the concept of a hotspot is a heuristic to develop a sampling design if smaller, isolated areas may be unacceptable. The shapes of such unacceptable areas would likely not be either circular or elliptical, but elliptical may at least be more realistic than circular. Hence, this methodology extension may add value for applications with smaller, isolated unacceptable areas.

6.0 Methods and Formulas for Calculating the Number of Samples and Confidence for Clearance Decisions Using Statistical Sampling Approaches

This section summarizes the methods and formulas for determining sampling plans and statistical statements for clearance situations with and without accounting for the uncertainties (systematic/bias and random) in the SAP process (via accounting for the FNR). Situations with qualitative (acceptable/unacceptable) and quantitative data are addressed. This section specifically summarizes the methods and formulas for determining the

A. number and locations of samples required when using a specific statistical sampling approach to support making an X%/%Y clearance statement (see Section 2.2) with specified X and Y, presuming that all samples will yield acceptable results (qualitative data) or results less than the ACL (quantitative data)

B. values of X and Y in an X%/Y% clearance statement (see Section 2.2) when *n* samples, collected according to a specific statistical sampling approach, all yield acceptable results (qualitative data) or results less than the ACL (quantitative data)

for clearance situations. The methods and formulas for the following clearance sampling cases are presented in the subsections shown.

- CJR sampling approach with qualitative (acceptable/unacceptable) data
  - FNR = 0, Section 6.1.1
  - FNR > 0, Section 6.1.2
  - Example, Section 6.1.3
- CJR sampling approach with quantitative data
  - FNR = 0, Section 6.2.1
  - FNR > 0, Section 6.2.2
  - Example, Section 6.2.3

Section 3.3 discusses UTL sampling approaches for meeting clearance objectives when the data are qualitative and quantitative. The approaches include methods that use random samples (as defined in Section 1.0) only, as well as the CJR approach that uses judgment and random samples. However, the only statistically-based sampling approach for clearance sampling mentioned in the sampling reference document being developed by the VSPWG is the CJR sampling approach (Section 3.3.2). This is reasonable, because judgment samples will generally always be taken, and it may be necessary to augment them with random samples (e.g., the CJR approach) to make an X%/Y% clearance statement (see Section 2.2). Furthermore, the CJR method can be implemented without judgment samples, if desired. In addition, the qualitative UTL methods discussed in Section 3.3.1 are essentially special cases of the CJR methodology (Sego et al. 2010). Hence, only methods and formulas for the CJR sampling approach to address Items A and B above are presented in this section for this initial version of the report.

Section 6.3 discusses how the CJR method may be used for clearance by analyzing and augmenting characterization sampling designs when all the samples results are acceptable. Section 6.4 briefly mentions the improvements needed to address limiting assumptions in the CJR sampling methodology discussed and illustrated in Sections 6.1, 6.2, and 6.3.

6.1 Methods and Formulas for the CJR Sampling Approach with Qualitative (Acceptable/Unacceptable) Data

The CJR sampling approach involves taking judgment samples in portions of a decision area subject to high risk of contamination and supplementing with random samples in low-risk portions of the decision area. The goal is to make an X%/Y% clearance statement (see Section 2.2) if all samples results are "acceptable," as defined by the investigators. For qualitative samples, an acceptable sample result may be a non-detect. For quantitative samples, an acceptable result may be defined as the sample measurement being $\leq$ ACL. See Table 2.1 for a definition of "acceptable" under a variety of sampling situations.

This section summarizes the CJR sampling methodology and formulas for addressing Items A and B in Section 6.0 for clearance situations where the data are qualitative (acceptable/unacceptable). Section 6.1.1 details the methods used when the FNR = 0. Section 6.1.2 discusses methods for the situation when the FNR > 0. Finally, Section 6.1.3 illustrates results of the methods and formulas for determining Items A and B listed at the start of Section 6.0.

In this section, the form of X%/Y% clearance statement supported by the CJR sampling methodology for qualitative data is either ClS3a or ClS3b (see Section 2.2), depending on whether 1) FNR = 0 or FNR > 0 with methods that account for FNR > 0, or 2) FNR > 0 but using methods that assume FNR = 0.

6.1.1 Methods and Formulas for the CJR Sampling Approach with Qualitative Data when the False Negative Rate is Zero

Section 6.1.1.1 summarizes the method and formula for determining the locations and number of random samples needed to augment a specified number of judgment samples to achieve an X%/Y% clearance statement with specified X% and Y% values when using the CJR sampling approach with qualitative (presence-absence) data when the FNR = 0. Section 6.1.1.2 discusses the "reverse use" of these methods with qualitative data when the FNR = 0 to determine the X% and Y% values achieved after a contamination incident in which the CJR sampling approach was applied with specified numbers of judgment and random samples and locations, and all sample results were acceptable.

6.1.1.1 Method and Formula for Determining the Number and Locations of Random Samples to Augment Judgment Samples when Using the CJR Sampling Approach with Qualitative Data when the False Negative Rate is Zero

This section gives an overview discussion and cites references for the method and formula to determine numbers and locations of samples needed to achieve a specified X%/Y% clearance statement when using the CJR sampling approach with qualitative (acceptable/unacceptable) data when the FNR = 0. The details of the methodology are discussed by Sego (2007, 2010), with summaries of the steps and assumptions of the methodology discussed in the following text.

Applying the CJR method when FNR = 0 to determine the number and locations of random samples to augment a specified number of judgment samples involves the following steps:

1. Divide all surfaces in the decision area to be sampled into non-overlapping grid cells with a size and shape corresponding to the footprint of the sampling methodology. See the discussion in Section 3.3.2 if more than one sample collection method with different sampling footprints will be used to sample the decision area.

2. Separate the decision area to be sampled into two groups of grid cells that are areas of high-risk and low-risk of being unacceptable. Expert opinion or judgment will be required to determine which grid cells are assigned to the high- and low-risk groups. All high-risk grid cells in the decision area are sampled using judgment sampling, while the remaining low-risk grid cells in the decision area are sampled randomly.

3. Specify the a priori expected probability that a judgment sample will be unacceptable (denoted P_J, where $0 < P_J < 1$). This can also be thought of as the a priori expected rate of unacceptable high-risk grid cells[a].

4. Specify the likelihood that a judgmentally sampled location is unacceptable relative to a randomly sampled location (denoted $r_{J:R}$, where $r_{J:R} \geq 1$). For example, if $r_{J:R} = 2$, then high-risk grid cells are assumed to be twice as likely, on average, to be unacceptable than low-risk grid cells.

5. Using the preceding information and the number of high-risk cells (i.e., the number of judgment samples, n_J), determine an initial X%/Y% statement that can be made assuming the judgment samples all yield unacceptable results. If the Y% does not meet the required percentage (e.g., 99%) with the desired X% confidence (e.g., 95%), then determine the number of random samples (n_R) required (when combined with the judgment samples) to achieve the desired values of X% and Y%.

6. Determine the locations of the calculated number of random samples using *random sampling* of the low-risk grid cells. Simple random, adaptive fill, or systematic random (i.e. grid) sampling may be used. If maximizing the coverage of the decision area is important, an adaptive fill or a grid method with a randomized starting point could be applied to the low-risk grid cells. The simple random and adaptive fill methods are implemented for the CJR approach in VSP.

The assumptions for the above CJR sampling methodology, based on documentation in VSP, are as follows:

a. The total number of grid cells in the decision area is known and each grid cell is the same size.

b. The size of the grid cell is appropriate for the chosen sampling methodology. If more than one sampling methodology is employed in a decision area, the size of the grid cell is chosen to match the sampling methodology with the smallest footprint.

c. The outcome from each sample will be qualitative (i.e., acceptable or unacceptable).

---

[a] In VSP, the parameter P_J is input as 1- P_J, i.e., the a priori expected probability that a judgment sample will be acceptable.

d. Each sample is collected, handled, measured, or inspected using approved methods that yield sufficiently precise measurements.

e. The measurement (inspection) method correctly classifies each sample as being acceptable or unacceptable. That is, an acceptable grid cell is not classified as being unacceptable (a false positive) and an unacceptable grid cell is not classified as being acceptable (a false negative).

f. The potential measured outcomes of from all grid cells are independent. If spatial correlation is present, the CJR method is conservative, (i.e., more samples are required than would otherwise be needed).

g. In the decision area, there are n_J high-risk grid cells which are more likely to be unacceptable than the remaining grid cells which are low-risk grid cells.

h. All n_J high-risk grid cells are sampled with judgment samples.

i. A high-risk cell is, on average, $r_{J:R}$ times more likely to be unacceptable than a low-risk cell.

j. Acceptable outcomes from judgment samples increase the confidence that low-risk cells are also acceptable.

k. Before sampling takes place, the expected probability of a judgment sample being unacceptable is P_J.

l. The probability that a high-risk cell is unacceptable has a Beta prior distribution with the first shape parameter equal to 1 and the second shape parameter equal to $(1 - P_J) / P_J$. This ensures that the expected value of the Beta prior is P_J.

m. A random sample of size n_R is taken from the low-risk grid cells. The sample locations may be selected using simple random, systematic random, or adaptive fill sampling.

Note that assumption "d" corresponds to assuming FNR = 0 and FPR = 0. See Section 2.1 for discussion of the FNR and FPR concepts and terminology. The CJR method has two special cases. For $r_{J:R} = 1$ and $n_J = 0$, all portions of the decision area are assumed to have the same likelihood of contamination. For $r_{J:R} = 1$, $n_J = 0$, and $P_J = 0.50$, the CJR method is virtually equivalent[a] to the AOZ-ACS acceptance sampling methods for qualitative data discussed in Section 3.3.2.1

The number of random samples is calculated by numerically solving the posterior predictive confidence function for n_R. This confidence function can be explained in words as follows: There is X% confidence that the number of acceptable low-risk grid cells constitutes at least Y% of the total decision area, given that all judgment and all random samples are acceptable. The mathematical formulas to calculate the posterior confidence function are complex, and cannot be easily summarized in this report. Interested readers should refer to VSP help documentation for a succinct overview of the CJR method, or to Sego et al. (2010) for in-depth information. VSP implements the CJR methodology to calculate the number of random samples given: 1) the desired X and Y values for an X%/Y% clearance statement, 2) the area and dimensions of the decision area to be sampled with judgment and random samples, 3) the

---

[a] The precise nature of this near equivalence is discussed in detail in Section 3.4 of Sego et al. 2010.

number of judgment samples, 4) the a priori expected probability P_J, and 5) the relative likelihood $r_{J:R}$. This is illustrated for an example in Section 6.1.3.1.

6.1.1.2 Method and Formula for Calculating the X%/Y% Clearance Statement Given the Number and Locations of Judgment and Random Samples when Using the CJR Sampling Approach with Qualitative Data and when the False Negative Rate is Zero

This subsection discusses the "reverse use" of the methods described in Section 6.1.1.1 for qualitative data to calculate the X% confidence that at least Y% of the decision area is acceptable after applying the CJR sampling approach with specified numbers and locations of judgment and random samples (with random samples collected using simple random sampling, adaptive fill, or systematic grid sampling), and all sample results are acceptable. If the CJR sampling approach was designed, and applied as designed, for the contamination incident, then the X% and Y% values are the same as those specified during the design of the CJR sampling plan. However, if any of the samples are unacceptable, the X%/Y% clearance statement does not apply.

Provided all sample results are acceptable, the X%/Y% values can also be calculated if a CJR sampling plan was performed that was not designed to achieve specific X and Y values, or if the CJR sampling plan actually performed differed from the one that was designed (e.g., some samples lost, or extra samples collected). However, the original values of P_J and $r_{J:R}$ that were chosen before sampling should be preserved. For this situation, there is not a unique combination of X% and Y% values, but rather a set of X%/Y% combinations that correspond to the CJR sampling plan that was conducted. Within the set of X%/Y% values, as X% increases Y% decreases, and vice versa. It is recommended that Y% > X%, since it is preferable to make a clearance statement about the highest possible Y% of the decision area not having contamination, as long as the X% is large enough to be defensible. Hence, it is recommended that 1) X% values be calculated for a series of desired Y% values such that Y% > X%, and 2) the combination most desirable for that clearance situation be chosen.

VSP calculates the X% confidence, given the desired Y%, the size of samples (assumed square), the number of square grid cells in the decision area, the numbers and locations of judgment and random samples, and the parameters P_J and $r_{J:R}$ (discussed in Section 6.1.1.1). Hence, the recommendation above can easily be implemented in VSP by calculating the X% confidence for several choices of Y%. This is illustrated for an example in Section 6.1.3.2.

6.1.2 Methods and Formulas for the CJR Sampling Approach with Qualitative Data when the False Negative Rate is Greater Than Zero

Currently there are no statistical methods and formulas for the CJR sampling approach with qualitative data that specifically account for FNR > 0. However, this work is currently funded and in progress. Sego et al. (2010) point out that the CJR method could be adapted to account for a single FNR > 0, perhaps using the approach discussed by Axelrod (2005). It would also be desirable to accommodate environmental samples obtained via different SAP methodologies (e.g., swab, wipe, vacuum) in the same decision area. Different SAP methodologies can have different sample areas and the FNR can vary with the 1) concentration level of the contaminant, 2) surface being sampled, 3) sample collection method, 4) specifics of the handling/storage/transportation, processing/extraction, and analytical steps of the SAP, 5)

environmental conditions, and 6) other factors. This section of the report may be completed if this report is updated in the future.

In the interim, the method and formulas discussed in Section 6.1.1 for FNR = 0 may be applied when FNR > 0, although doing so will only support making an X%/Y% clearance statement of the form ClS3b listed and discussed in Section 2.2. This is because qualitative data only indicate whether contamination was detected or not detected given the set of SAP methods used. If the SAP methods used have FNR > 0, this weaker clearance statement allows for the undesirable situation of contamination existing but not being detected (i.e., false negatives). Ideally, statistical X%/Y% clearance statements should be about whether or not contamination is present, not whether contamination is detectable by SAP methods that generally have substantial inefficiencies. To do that, methods that account for FNR > 0 must be developed. However, as the FNR increases, the number of samples required to make an X%/Y% clearance statement of the form ClS3a (see Section 2.2) also increases. Conversely, the X% and Y% values that can be achieved for a given number of samples get smaller as the FNR increases.

6.1.3 Examples of Applying the Formulas for the CJR Sampling Approach with Qualitative Data when the False Negative Rate is Zero

This section currently only illustrates the use of the methods and formulas discussed in Section 6.1.1 for the CJR sampling approach with qualitative data when the FNR = 0. The example may be extended in a future revision of this report when the methods and formulas in Section 6.1.2 are developed for the CJR sampling approach with qualitative data when the FNR > 0.

The methods described in Sections 6.1.1.1 and 6.1.1.2 are illustrated assuming one floor of the INL PBF-632 building (as discussed in Section 5.1.3) is to be cleared. For the examples in this section, it is assumed that the decision area to be cleared has been decontaminated or was not expected to be contaminated, and that initial judgment samples all yielded negative results. VSP was used to perform the example calculations using a square grid with sample area assumed to be 1 ft × 1 ft.

6.1.3.1 Examples of Applying the Methods and Formulas for Determining the Number and Locations of Random Samples to Augment Judgment Samples when Using the CJR Sampling Approach with Qualitative Data when the False Negative Rate is Zero

Figure 6.1 summarizes the number of random samples (as a function of the number of judgment samples) required to make a 95%/97% clearance statement of the form ClS3a (see Section 2.2) about a decision area consisting of a single floor in the INL PBF-632 building when using the CJR sampling approach. Three parameters were allowed to vary: 1) the number of judgment samples, 2) the a priori probability (P_J, expressed as a percentage) that a high-risk cell is unacceptable (10% or 30%), and 3) the likelihood that a high-risk, judgment sample location is unacceptable as compared to a low-risk, random sample location (e.g., 1× indicates a high risk location just as likely as a low risk location, while 3× indicates a high-risk location is three times as likely as a low-risk location). VSP was then used to calculate the number of random samples for various combinations of parameter values. Figure 6.1 shows that the number of random samples decreases roughly linearly with the increase in judgment samples. Also, the smallest numbers of random samples occurs for the 10%/3× case, whereas the largest numbers of random samples occurs for the 30%/1× case. These results make sense, in that 1) the smaller the a

priori probability that a judgment sample is unacceptable, and 2) the higher the relative likelihood of judgment samples being unacceptable relative to random samples, the fewer the number of random samples required to clear the decision area.

Consider a specific example when (i) there is a 30% a priori probability that a high-risk judgment sample location is unacceptable and (ii) a judgment sample location is just as likely (1×) to be unacceptable as a random sample location. Figure 6.1 shows that with 20 acceptable judgment samples, 74 acceptable random samples would be required to have 95% confidence that 97% of the floor is acceptable.
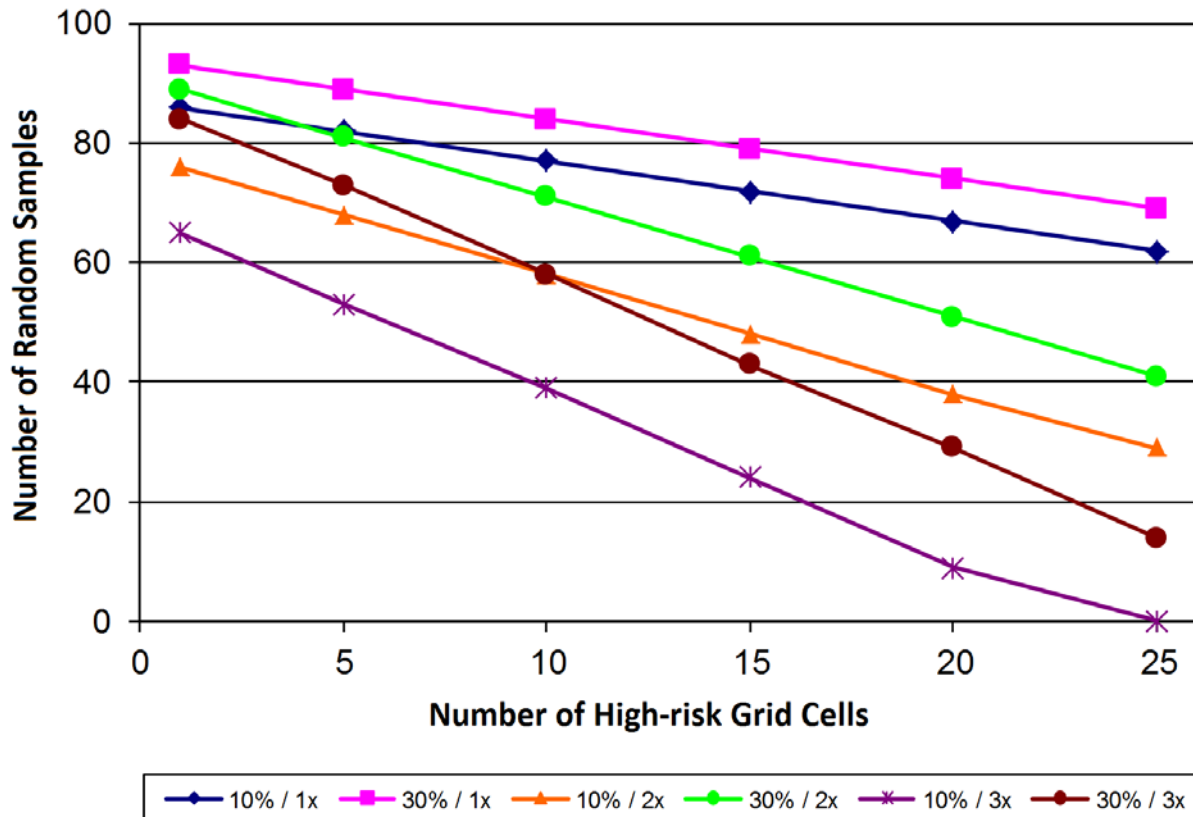
Figure 6.1. Number of Acceptable Random Samples Required to be 95% Confident that at least 97% of a Typical Floor in the INL PBF-632 Building is Acceptable Given Various Numbers of High-risk Grid Cells that are all Sampled Judgmentally and Found to be Acceptable. Colored lines represent selected combinations of the a priori probability that a judgment sample will unacceptable (P_J) and the relative likelihood of a judgment-sample location being unacceptable relative to a random-sample location.

Caution is needed in interpreting and using Figure 6.1. As discussed in Section 6.1.1.1, the number of judgment samples should correspond only to locations that have a "high risk" of contamination, and all such high-risk locations should be sampled with judgment samples. It would be a misuse of the CJR approach to arbitrarily pick a larger number of judgment samples in order to reduce the number of random samples.

6.1.3.2 Examples of Applying the Method and Formula for Calculating the X%/Y% Clearance Statement Given the Number and Locations of Judgment and Random Samples when Using the CJR Sampling Approach with Qualitative Data and when the False Negative Rate is Zero

This subsection presents an example applying the CJR sampling methodology in Section 6.1.1.2 to a situation when all sample results (judgment and random) yield acceptable results. Specifically, the methodology is used to calculate the X% confidence that at least Y% of the decision area is acceptable given the numbers of judgment and random samples, the size of samples/grid cells, the a priori probability that a judgment sample will be unacceptable, and the relative likelihood of a judgment-sample location versus a random-sample location being unacceptable.

Consider an example in which it is desired to determine the X%/Y% clearance statement of the form ClS3a in Section 2.2 that can be made when a CJR sampling design (with 20 judgment samples and 40 random samples) was performed on the first floor of the INL PBF-632 building after decontamination and all sample results were acceptable. Figure 6.2 shows the possible values of Y% (such that $Y \geq X = 95\%$) for the same six combinations of "a priori probability" and "relative likelihood" used in Figure 6.1. All calculations were performed in VSP assuming the FNR = 0. Figure 6.2 shows that the value of Y% increases with increasing "relative likelihood" (from 1× to 2× to 3×) and when the a priori probability P_J decreases from 30% to 10%.
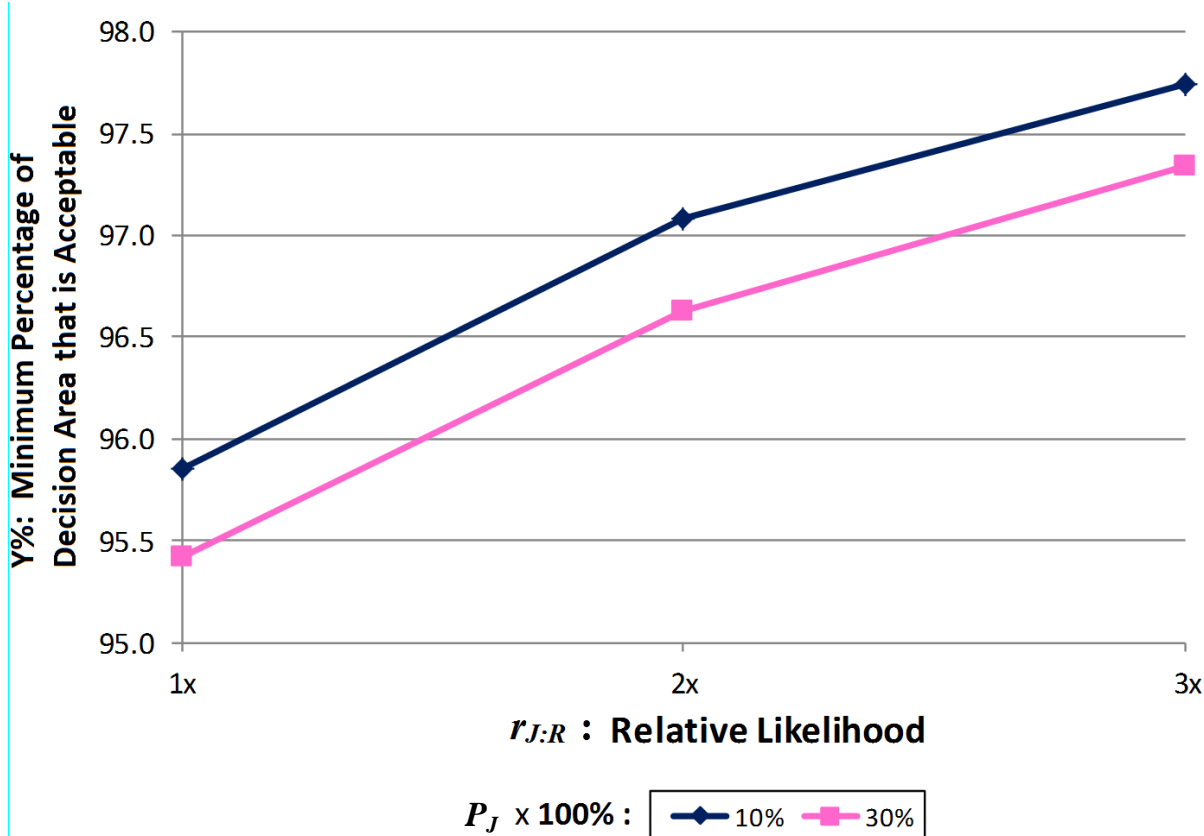
Figure 6.2. Plot for 95% Confidence that at Least Y% of the First Floor of the INL PBF-632 Building is Acceptable Based on a CJR Sampling Design with 20 Judgment Samples and 40 Random Samples, All of Which Yielded Acceptable Results. Colored lines represent selected combinations of the a priori probability that a high-risk grid cell will be unacceptable and the relative likelihood of a high-risk grid cell being unacceptable relative to a low-risk grid cell.

6.2 Methods and Formulas for the CJR Sampling Approach with Quantitative Data

Consider clearance sampling situations where quantitative data (e.g., contaminant surface concentrations, CFU/cm^2) greater than zero are obtained for some or all samples from a CJR sampling plan. Hence, there may be some quantitative data and some qualitative (non-detect) data. Although quantitative values less than an LOD are often reported as "non-detect" values due to uncertainty in the measurement process, it is better statistically to use quantitative measured values whenever they are available, despite that uncertainty (Gilbert 1987, Section 14.1). However, even if this is possible for some data points that would otherwise be reported as non-detects (or less than LOD), there may still be samples for which there are no quantitative values and only non-detects (less than LOD) can be reported.

When the CJR sampling approach is used and some or all sample results are quantitative, the current practice is to compare the quantitative results to an ACL (which may be zero if no contamination is acceptable) and convert the results into "Above ACL" (measured value > ACL) and "Below ACL" (measured value ≤ ACL) qualitative data. Then the CJR sampling methods and formulas for qualitative

data (as discussed in Section 6.1) can be applied. This current practice is discussed in Sections 6.2.1 and 6.2.2 when the FNR = 0 and FNR > 0, respectively.

In this section, the form of X%/Y% clearance statement supported by the CJR sampling methodology for quantitative data is either ClS4a or ClS4b (see Section 2.2), depending on whether 1) FNR = 0 or FNR > 0 with methods that account for FNR > 0, or 2) FNR > 0 but using methods that assume FNR = 0.

6.2.1 Methods and Formulas for the CJR Sampling Approach with Quantitative Data when the False Negative Rate is Zero

Consider clearance situations in which any contamination present in the decision area has concentrations much larger than the ACL (so that FNR = 0). Presumably this would only occur for a decision area that was mistakenly judged unlikely to be contaminated and not decontaminated prior to clearance sampling. For such situations, the current practice of converting quantitative data to qualitative data and applying the qualitative CJR sampling approach (see Section 6.1) is appropriate and could not be improved upon by developing and using a quantitative CJR sampling methodology. Such methods do not currently exist, and there are substantial conceptual and theoretical challenges to developing such methods.

The methods discussed in Section 6.1 are applicable and provide for making X%Y% clearance statements of the form ClS4a (see Section 2.2). However, if the ACL is a concentration at which false negatives may occur, then applying the methods discussed in Section 6.1 would only support making an X%/Y% clearance statement of the form ClS4b (see Section 2.2). The following subsection addresses these latter situations.

6.2.2 Methods and Formulas for the CJR Sampling Approach with Quantitative Data when the False Negative Rate is Greater Than Zero

Currently there are no statistical methods and formulas for the CJR sampling approach with quantitative data (which may also contain non-detect results) that specifically account for FNR > 0. It was noted in Section 6.1.2 that work is underway to extend the CJR methods for qualitative data with FNR = 0 to the situation of qualitative data with FNR > 0. When the extended method is developed, it could be applied to quantitative data for a single FNR > 0 by first converting the quantitative data to qualitative data (i.e., "Below ACL" and "Above ACL"). However, since the CJR sampling approach for qualitative data with FNR = 0 has not yet been extended to FNR > 0, it cannot currently be applied in this way for quantitative data with FNR > 0. If such methodology is developed in the future, it could be discussed in a future revision of this section.

The only alternative that could be applied now is to convert the quantitative data to qualitative data and apply the CJR sampling methods for qualitative data with FNR = 0, as discussed in Section 6.1.1. Doing this would only support making X%/Y% clearance statements of the form in ClS3b and ClS4b, as discussed in Section 2.2. That is, statements could only be made about detectable contamination rather than actual contamination. This alternative implementation of the CJR approach would ignore the potentially large inefficiencies (biases) and random uncertainties in the steps of the SAP process (see Section 4.0). This could lead to the measured surface concentrations significantly underestimating their actual values. This in turn could lead to measured surface concentrations of the contaminant being less

than the ACL, but actual concentrations being greater than the ACL (i.e., false negatives with FNR > 0). Hence, the quantitative-to-qualitative CJR sampling approach assuming FNR = 0 when in fact FNR > 0 could significantly overestimate the X% and/or Y% values of an X%/Y% clearance statement. This in turn could mean that the number of acceptable judgment and random samples needed to achieve a given X%/Y% clearance statement would be underestimated.

This section may be revised in a future version of this report when the modified methods and formulas discussed in the first paragraph are developed for the CJR sampling approach with quantitative data when the FNR > 0.

6.2.3 Examples of Applying the Formulas for the CJR Sampling Approach with Quantitative Data for a Range of False Negative Rates

Because the only existing method for applying the CJR sampling approach with quantitative data is to convert the quantitative data to qualitative data (i.e., "< ACL" is classified as "non-detect", "≥ ACL" is classified as "detect") and assume FNR = 0, the examples in Section 6.1.3 for qualitative data suffice. If CJR sampling methods for quantitative data and FNR = 0 or FNR > 0 are developed as part of future work, they will be documented in a revision of Section 6.2 and illustrated in a revision of this section.

6.3 Methods for Making Clearance Statements When All Characterization Samples are Acceptable

If the sample results from a characterization design (e.g., a hotspot design) are acceptable, investigators will likely want to know how probable is it that un-sampled portions of the decision are also acceptable, possibly in anticipation of making an X%/Y% clearance statement regarding the decision area. In this situation, we recommend using the accept-on-zero, attribute compliance sampling (AOZ-ACS) approach, discussed in Section 3.3.2.1) or the CJR approach (Sections 3.3.2.2 and 6.1). These approaches can be used to calculate 1) the clearance confidence achieved by the characterization design, and/or 2) determine the number of additional samples required to make a desired X%/Y% clearance statement. These two cases are discussed in Sections 6.3.1 and 6.3.2, respectively.

6.3.1 Calculating the Clearance Confidence of an Implemented Characterization Design

If all samples from a characterization design are acceptable (refer to Table 2.1 for definitions of "acceptable" under various sampling scenarios), an X%/Y% confidence statement can be obtained using the AOZ-ACS method (as discussed in Section 3.3.2.1) or the CJR method (as discussed in Section 6.1.1.2). These methods are implemented in VSP. If the CJR model is used, investigators should carefully assess each of the required inputs of the CJR method. This is especially true of the a priori expected probability that a high-risk, judgment sample location will be unacceptable (P_J). Unless definitive evidence exists from a previous study (i.e., prior to the characterization sampling campaign), P_J should be set to 0.5, which corresponds to a non-informative prior distribution. After all, P_J is the *a priori belief* that judgment samples are unacceptable. After observing potentially many acceptable characterization samples, retrospective "estimates" of P_J could be overly optimistic. This would result in underestimating P_J, which, in turn, would overestimate the achieved clearance confidence of the

characterization sampling design. If no information about P_J is available (presumably from previously collected data) and no judgment samples were collected during characterization, the number of high-risk cells in the CJR method should be set to 0, $r_{J:R}$ should be set to 1, and P_J should be set to 0.5. This special case of the CJR method will produce confidence estimates that are virtually identical to those given by the AOZ-ACS method (Sego et al. 2010). If the desired confidence level is not achieved, additional clearance samples may be taken, as discussed in Section 6.3.2.

6.3.2 Calculating the Number of Samples Required to Augment an Existing Characterization Design to Make a Clearance Statement

In the event that the number of acceptable characterization samples that were collected is not sufficient to make the desired clearance statement, the CJR method (implemented in VSP) can be used to calculate the additional number of required judgment and/or random samples. Because the characterization sampling campaign and the collection of additional samples during the clearance campaign will likely be separated in time, this combined sampling strategy requires assuming that state of the decision area remains static (i.e., unchanged) between characterization and clearance sampling. This would imply, for example, that no cross-contamination occurs and that contamination does not increase anywhere in the decision area due to re-suspension of contaminant particles, air flow, attachment/detachment of spores to/from electrostatic surfaces, etc. Before beginning a clearance sampling campaign, it may be advisable to take a few additional samples at (or very near) the locations of the original characterization samples to examine the assumption that the decision area is static.

Provided investigators feel comfortable assuming a static decision area, the characterization design may be augmented to make a clearance statement using the following procedure:

1. Determine which of the characterization samples were taken at high-risk locations and which were taken at low-risk locations. Those taken at high-risk locations would become part of the collection of judgment samples in the new CJR design. Those taken at low-risk locations would constitute some (or perhaps all) of the random samples of the new CJR design (*even if they were not originally determined in a random fashion*). Denote the number of high-risk and low-risk characterization samples as n_{1J} and n_{1R}, respectively.

2. Determine whether there are high-risk areas that were not sampled during characterization. Grid cells in these high-risk areas should be sampled judgmentally as part of the CJR design. Denote the number of these as yet unsampled, high-risk locations as n_{2J}.

3. Estimate $r_{J:R}$. Special care should be taken not to overestimate this parameter.

4. Chose the value of P_J. As discussed in Section 6.3.1, unless definitive results from a study prior to the characterization sampling are available, P_J should be set to 0.5, thereby avoiding an overly optimistic estimate of P_J, influenced by the n_{1J} acceptable characterization results already observed.

5. Specify the parameters X% and Y% that give the desired confidence statement.

6. Use the CJR model to calculate the confidence achieved using $n_{1J} + n_{2J}$ judgmental samples, n_{1R} random samples, and the values of $r_{J:R}, P_J$, and Y% determined in Steps 1-5 above. If the

confidence meets or exceeds the desired X% confidence specified in Step 5, then all that remains to make the desired confidence statement is to collect *all*[a] of the n_{2J} judgmental samples and find them all to be acceptable.

7. If the achieved confidence calculated in Step 6 is less than the desired X% confidence specified in Step 5, the number of additional random samples, n_{2R}, required to achieve the desired confidence level can be calculated. Then, if the additional n_{2J} judgmental samples and n_{2R} random samples are collected and found to be acceptable, the desired confidence statement can be made.

6.4   Possible Improvements in the CJR Sampling Methodology

Possible improvements in the CJR sampling methodologies discussed in Sections 6.1 and 6.2 are now briefly mentioned. These improvements are related to assumptions made in developing the statistical sampling theory and formulas for the CJR sampling approaches with qualitative and quantitative data when FNR = 0 and when FNR > 0. Extensions of the CJR sampling methodology to address these items are possible with sufficient resources to develop the more complicated algorithms and formulas that will be needed.

Less Restrictive Assumptions

- One of the important assumptions for the CJR approach is that the number of judgment samples is equal to the number of high-risk grid cells in the decision area that are determined a priori. While this assumption simplifies the mathematics considerably, it would be more realistic to not require that all of the high-risk area be sampled.

- A generalization of the CJR method would allow there to be more than two risk groups in the decision area. In particular, this would accommodate different surface types, their propensity to attract and/or retain contamination, and different sampling methods (swab, wipe, vacuum). Some progress has been made on this stratified clearance sampling method, and a preliminary, non-public version exists in VSP. However, more work will be required to fully validate this methodology and implement it in VSP.

False Negative Rate = 0 with Qualitative Data (or Quantitative Data Converted to Qualitative Data)

- Extend methodology to allow samples obtained with different collection methods (e.g., swab, wipe, and vacuum) and/or different surface materials within a decision area to have different sample footprints (areas).

False Negative Rate > 0 with Qualitative Data (or Quantitative Data Converted to Qualitative Data)

- Extend methodology for FNR = 0 with qualitative data (or quantitative data converted to qualitative data) to address FNR > 0.

---

[a] It is possible that observing n_{2J} acceptable samples during the clearance sampling phase will result in a considerably higher confidence level than desired. Nonetheless, it is important that all n_{2J} samples be collected in order to satisfy the CJR assumptions. In this situation, the value of Y% may be increased (instead of, or in addition to, increasing X%), yielding a stronger clearance statement.

- Extend methodology to allow 1) samples obtained with different collection methods (e.g., swab, wipe, and vacuum) and/or different surface materials within a decision area to have different sample areas, and 2) different strata with different values of FNR > 0. Different strata could be defined by the collection method and other methods in the SAP, as well as the surface being sampled.

7.0 Measured and Estimated True Contaminant Concentrations and Their Uncertainties in the Sampling and Analysis Process

Section 7.1 presents methods and formulas for quantifying the uncertainty (bias and random) in measured contaminant concentrations when the SAP yields quantitative data. Section 7.2 presents methods and formulas for estimating the true contaminant concentration of a sample and the random uncertainty in the estimate.

7.1 Statistical Models and Random Uncertainties for a Measured Concentration from a Single Sample

In this section we assume quantitative data and present a *statistical model* for the measured contaminant concentration (CFU/cm^2) of a single sample subject to uncertainties associated with 1) sample collection, 2) handling, storage, and transportation, 3) processing/extraction, and 4) analysis. By "statistical model" we mean one that accounts for the uncertainties (systematic and random) of the terms in the model that correspond to sources of uncertainty in the steps of the SAP. The statistical model for a measured contaminant concentration from a single sample is used in Section 7.2 to develop formulas for estimating the true contaminant concentration and its uncertainty for a single sample.

7.1.1 Statistical Model and Random Uncertainty for a Measured Concentration from a Single Sample when Information is Available Only on the Combined Recovery Efficiency

Section 7.1.1 presents the statistical model for the measured contaminant concentration and the formula for random uncertainty of a single sample when systematic and random uncertainty information is available only on the combined (over all steps of the SAP) recovery efficiency.

In this case, the statistical model for the measured contaminant concentration of a single surface sample is given by

$$C_{CFU,S}^{M} = R_{SAP,S} C_{CFU,S}^{T} \tag{7.1}$$

where $C_{CFU,S}^{M}$ = measured concentration (CFU/cm^2) for a single surface sample obtained with a specific set S of SAP methods and sampling conditions (e.g., surface material, surface concentration, humidity)

$R_{SAP,S}$ = recovery efficiency for a specific set S of SAP methods and sampling conditions (e.g., surface material, surface concentration, humidity). This is considered a random variable with a true, unknown mean [denoted $Mean(R_{SAP,S})$] and a true, unknown percent relative standard deviation [denoted] $\%RSD(R_{SAP,S})$

$$C_{CFU,S}^T \quad = \quad \text{true, unknown concentration (CFU/cm}^2\text{) for a single surface sample obtained with}$$

a specific set S of SAP methods and sampling conditions (e.g., surface material, surface concentration, humidity).

The values of the systematic and random uncertainties in $R_{SAP,S}$ may depend on other factors, such as the 1) concentration of the contaminant, 2) surface material being sampled, 3) sample collection method, 4) specifics of handling, storage, and transportation, 5) specifics of sample processing and extraction, and 6) specifics of the analytical process. This dependence is denoted by the subscript "S" in the notation.

When only information about systematic and random uncertainties in $R_{SAP,S}$ are available, the random uncertainty (represented by the standard deviation) in $C_{CFU,S}^M$ is approximated by

$$\% RSD(C_{CFU,S}^M) \approx \% RSD(R_{SAP,S}) \tag{7.2}$$

as developed in Section B.1 of the Appendix B. Equation (7.2) is applicable when estimates of $\% RSD(R_{SAP,S})$ are available. Piepel et al. (2011a, 2011b) summarized $\% RSD(R_{SAP,S})$ values for different sets S of SAP methods and conditions investigated in 20 laboratory sampling studies from the literature. Section 7.3.1 presents values of the preceding and other sampling statistics for CDC-recommended swab and wipe methods, as well as a selected vacuum method, which are representative uncertainties of those sampling methods.

7.1.2 Statistical Model and Random Uncertainty for a Measured Concentration from a Single Sample when Information is Available on the Recovery Efficiencies for All Steps of the Sampling and Analysis Process

Section 7.1.2 presents the statistical model for the measured contaminant concentration and the formula for random uncertainty of a single sample when systematic and random uncertainty information is available for all steps of the SAP.

In this case, the statistical model for the measured contaminant concentration of a single surface sample is given by

$$
\begin{aligned}
C_{CFU,S}^M &= R_{SAP,S}\, C_{CFU,S}^T \\
&= R_{SAP,S}\, F_{CFU:Spore,S}\, C_{Spores,S}^T \\
&= (R_{SC,S} R_{HST,S} R_{,SP} R_{A,S})\, F_{CFU:Spore,S}\, C_{Spores,S}^T
\end{aligned}
\tag{7.3}
$$

where

$F_{CFU:Spore,S}$ = fraction of CFUs per spore achieved with a specific analysis method (including culturing and counting) and a specific set S of sampling conditions (e.g., surface material, surface concentration, humidity), considered as a random variable with a true, unknown mean [denoted Mean($F_{CFU:Spore,S}$)] and a true, unknown %RSD [denoted %RSD($F_{CFU:Spore,S}$)]

7.2

$C_{Spores,S}^T$ = true, unknown concentration (CFU/cm^2) of spores for a single surface sample obtained with a specific set S of SAP methods and sampling conditions (e.g., surface material, surface concentration, humidity)

$R_{SC,S}$ = recovery efficiency for a specific sample collection method and specific set S of conditions (all denoted by S), considered as a random variable with a true, unknown mean [denoted $Mean(R_{SC,S})$] and a true, unknown percent relative standard deviation [denoted $\%RSD(R_{SC,S})$]

$R_{HST,S}$ = recovery efficiency for a specific sample handling, storage, and transport method and a specific set S of conditions (all denoted by S), considered as a random variable with a true, unknown mean [denoted $Mean(R_{HST,S})$] and a true, unknown percent relative standard deviation [denoted $\%RSD(R_{HST,S})$]

$R_{P,S}$ = recovery efficiency for a specific sample processing/extraction method and a specific set S of conditions (all denoted by S), considered as a random variable with a true, unknown mean [denoted $Mean(R_{P,S})$] and a true, unknown percent relative standard deviation [denoted $\%RSD(R_{P,S})$]

$R_{A,S}$ = recovery efficiency for a specific sample analysis method (including culturing and counting) and a specific set of conditions (all denoted by S), considered as a random variable with a true, unknown mean [denoted $Mean(R_{A,S})$] and a true, unknown percent relative standard deviation [denoted $\%RSD(R_{A,S})$]

and other notation is as defined in Section 7.1.1. The values of the various systematic and random uncertainties in $R_{SC,S}$, $R_{HST,S}$, $R_{P,S}$, $R_{A,S}$, and $F_{CFU:Spore,S}$ may depend on other factors, such as the 1) concentration of the contaminant, 2) surface material being sampled, 3) sample collection method, 4) specifics of handling, storage, and transportation, 5) specifics of sample processing and extraction, and 6) specifics of the analytical process.

In this case, the random uncertainty (represented by the standard deviation) in $C_{CFU,S}^M$ is approximated by

$$\%RSD(C_{CFU,S}^M) \approx \left[\begin{array}{l} \{\%RSD(R_{SC,S})\}^2 + \{\%RSD(R_{HST,S})\}^2 + \{\%RSD(R_{P,S})\}^2 \\ + \{\%RSD(R_{A,S})\}^2 + \{\%RSD(F_{CFU:Spore,S})\}^2 \end{array} \right]^{0.5} \tag{7.4}$$

as developed in Section B.2 of Appendix B. When estimates of the five %RSDs are available, an estimate of the total %RSD is given by

$$\overline{\%RSD}(C_{CFU,S}^M) \approx \left[\begin{array}{l} \{\overline{\%RSD}(R_{SC,S})\}^2 + \{\overline{\%RSD}(R_{HST,S})\}^2 + \{\overline{\%RSD}(R_{P,S})\}^2 \\ + \{\overline{\%RSD}(R_{A,S})\}^2 + \{\overline{\%RSD}(F_{CFU:Spore,S})\}^2 \end{array} \right]^{0.5}. \tag{7.5}$$

Note that Piepel et al. (2011a, 2011b) did not summarize uncertainties at this level of detail because there was limited information in the literature on uncertainties of the individual steps of the SAP. However, if estimates of the five %RSD values in Equation (7.5) are available, they can be inserted into the equation to calculate the total random uncertainty (%RSD) in the measured contaminant concentration value from a single surface sample.

Another alternative to quantify the uncertainty in measured contaminant concentration values $C_{CFU,S}^M$ is to assume a statistical distribution (e.g., uniform, normal, or log-normal) for each random variable in Equation (7.3), and then perform a Monte Carlo simulation (Krose et al. 2011). That would generate a distribution of simulated $C_{CFU,S}^M$ values that would graphically show the uncertainty in a measured contaminant concentration for a single surface sample. However, it is not clear that there are sufficient data on the uncertainties for each of the variables in Equation (7.3) for different sets of SAPs to provide guidance on the inputs required by a Monte Carlo approach.

Examples of the formulas in this section are presented in Section 7.3.2.

7.2 Estimating the True Contaminant Concentration for a Single Sample and Quantifying the Uncertainty of the Estimate

In this section the results of Section 7.1 are used to obtain formulas for an estimate of the true contaminant concentration for a single sample (Section 7.2.1) and the random uncertainty of the estimate (Section 7.2.2).

7.2.1 Estimating the True Contaminant Concentration for a Single Sample

Solving the statistical model in Equation (7.1) for the true contaminant concentration and substituting an estimate of the mean of $R_{SAP,S}$ (for a given set S of sampling conditions) yields a formula for an estimate of the true concentration from a single sample. The formula is

$$\hat{C}_{CFU,S}^T = \frac{C_{CFU,S}^M}{\overline{Mean}(R_{SAP,S})} \tag{7.6}$$

where $\hat{C}_{CFU:S}^T$ = estimate of the true, unknown concentration (CFU/cm^2) for a single surface sample obtained with a specific set S of SAP methods and sampling conditions (e.g., surface material, surface concentration, humidity)

$\overline{Mean}(R_{SAP,S})$ = estimate of the true, unknown mean of the random variable $R_{SAP,S}$, which is the recovery efficiency for a specific set S of SAP methods and sampling conditions (e.g., surface material, surface concentration, humidity).

and $C_{CFU,S}^M$ is the same as previously defined. The mean of $R_{SAP,S}$ may depend on the surface concentration, surface material, humidity, etc., (referred to as sampling conditions), as well as different methods for sample collection, handling/storage/transportation, processing/extraction, and analysis. Hence, Equation (7.6) is a general equation that may have different values for $\overline{Mean}(R_{SAP,S})$ depending on the specific SAP methods and sampling conditions.

Equation (7.6) is applicable when estimates of $\overline{Mean}(R_{SAP,S})$ are available. Piepel et al. (2011a, 2011b) summarized combined recovery efficiency data for many specific sets of SAP methods and sampling conditions investigated in 20 laboratory studies documented in the literature. If data on the means of $R_{SC,S}$, $R_{HST,S}$, $R_{P,S}$ and $R_{A,S}$ are available, then Equation (7.6) can be expanded as

$$\hat{C}_{CFU,S}^T = \frac{C_{CFU,S}^M}{\overline{Mean}(R_{SC,S})\,\overline{Mean}(R_{HST,S})\,\overline{Mean}(R_{P,S})\,\overline{Mean}(R_{A,S})\,\overline{Mean}(F_{CFU:Spore,S})} \qquad (7.7)$$

where all notation is as previously defined.

The means and %RSDs of $R_{SC,S}$, $R_{HST,S}$, $R_{P,S}$ and $R_{A,S}$ may depend on the surface concentration, surface material, and other sampling conditions, as well as different methods for sample collection, handling/storage/transportation, processing/extraction, and analysis. Hence, Equation (7.7) is a general equation that may have different inputs for the means and %RSDs depending on the specific SAP methods and sampling conditions.

7.2.2   The Random Uncertainty in an Estimate of the True Contaminant Concentration for a Single Sample

Estimates of the true contaminant concentration for a single sample yielded by Equations (7.6) and (7.7) are subject to random uncertainty because the measured values ($C_{CFU:S}^M$) and the means of $R_{SAP,S}$ [Equation (7.6)] and $R_{SC,S}$, $R_{HST,S}$, $R_{P,S}$ and $R_{A,S}$ [Equation (7.7)] estimated from experimental data are subject to random uncertainties. Hence, it is important to quantify the random uncertainty in estimates of $\hat{C}_{CFU:S}^T$ calculated using Equations (7.6) and (7.7). There are two ways to do this, using 1) the error propagation method (Hahn and Shapiro 1967) to calculate the standard deviation and/or %RSD of $\hat{C}_{CFU:S}^T$, and 2) the Monte Carlo simulation method (Krose et al. 2011) to generate the statistical probability distribution of the estimate $\hat{C}_{CFU:S}^T$. The former requires only the means and standard deviations of the random variables in Equation (7.6) or Equation (7.7), while the latter also requires an assumed statistical distribution of each random variable (uniform, normal, and log-normal distributions are commonly used).

Applying the first-order error propagation methodology in Equation (7-4) of Hahn and Shapiro (1967) to Equation (7.6) yields

$$SD(\hat{C}_{CFU,S}^{T}) = SD\left[\frac{C_{CFU,S}^{M}}{\overline{Mean}(R_{SAP,S})}\right]$$

$$= \left\{\left[\frac{SD(C_{CFU,S}^{M})}{\overline{Mean}(R_{SAP,S})}\right]^{2} + \left[\frac{C_{CFU,S}^{M}}{[\overline{Mean}(R_{SAP,S})]^{2}}\right]^{2}[\overline{SD}(\overline{Mean}(R_{SAP,S}))]^{2}\right\}^{0.5}$$

$$= \left\{\left[\frac{SD(C_{CFU,S}^{M})}{\overline{Mean}(R_{SAP,S})}\right]^{2} + \left[\frac{C_{CFU,S}^{M}}{\overline{Mean}(R_{SAP,S})}\right]^{2}\left[\frac{\overline{SD}[\overline{Mean}(R_{SAP,S})]}{\overline{Mean}(R_{SAP,S})}\right]^{2}\right\}^{0.5}$$

$$= \left\{\left[\frac{SD(C_{CFU,S}^{M})}{\overline{Mean}(R_{SAP,S})}\right]^{2} + \left[\frac{C_{CFU,S}^{M}}{\overline{Mean}(R_{SAP,S})}\right]^{2}\left[\frac{\overline{\%RSD}[\overline{Mean}(R_{SAP,S})]}{100}\right]^{2}\right\}^{0.5}$$

$$= \left\{\left[\frac{SD(C_{CFU,S}^{M})}{\overline{Mean}(R_{SAP,S})}\right]^{2} + \left[\frac{C_{CFU,S}^{M}}{\overline{Mean}(R_{SAP,S})}\right]^{2}\left[\frac{\overline{\%RSD}(R_{SAP,S})}{100\sqrt{n}}\right]^{2}\right\}^{0.5} \tag{7.8}$$

where all notation is as previously defined. The modification of Equation (7.8) from the next-to-last line to the last line above is possible if the estimated mean $\overline{Mean}(R_{SAP,S})$ for a given set of SAP methods and sampling conditions is obtained by averaging "n" replicate observations of $R_{SAP,S}$.

The first-order error propagation method could be applied to Equation (7.7) in a similar fashion, but the resulting formula would be quite a bit longer than Equation (7.8), which corresponds to Equation (7.6). Because there are limited data on the mean and %RSD values for the sources of uncertainty corresponding to the terms in Equation (7.7), we do not present the results of applying first-order error propagation to Equation (7.7).

7.3   Illustrative Examples

The subsections in this section provide illustrative examples of using the formulas in Sections 7.1 and 7.2 for the measured and estimated true contaminant concentrations and their uncertainties in the SAP.

7.3.1   Examples of Estimated Random Uncertainty for a Measured Concentration from a Single Sample when Information is Available Only on the Combined Recovery Efficiency

Equation (7.2) in Section 7.1.1 presents the formula for the %RSD of a measured contaminant concentration of a single sample when systematic and random uncertainty information is available only on the combined (over all steps of the SAP) recovery efficiency. Piepel et al. (2011a, 2011b) summarized

$C_{CFU:S}^M$, $\overline{Mean}(R_{SAP,S})$, and $\%RSD(R_{SAP,S})$ values for different sets S of SAP methods and sampling conditions investigated in 20 laboratory sampling studies from the literature. Specifically, $C_{CFU:S}^M$, $\overline{Mean}(R_{SAP,S})$, and $\%RSD(R_{SAP,S})$ values were summarized for swab, wipe, and vacuum methods of sample collection and the corresponding handling/storage/transportation, processing/extraction, and analytical steps of the SAP.

Table 7.1 lists values of $C_{CFU:S}^M$, $\overline{Mean}(R_{SAP,S})$, and $\%RSD(R_{SAP,S})$ for CDC's currently recommended methods for swab and wipe samples, as well as selected results for vacuum samples. Values are listed for stainless steel surfaces, since that is the most commonly tested surface. Values for some other surface materials are also provided. The $\overline{Mean}(R_{SAP,S})$ and $\%RSD(R_{SAP,S})$ values can vary significantly with the specific methods of the SAP, surface concentration, surface material, and other environmental factors. The values in Table 7.1 give representative ranges of the values for recommended (swab, wipe) and recent (vacuum) studies, which can be used as an indicator of recovery efficiencies, their uncertainties, and uncertainties of measured surface concentrations.

Table 7.1. Ranges of Measured Surface Concentrations [$C_{CFU:S}^M$], Mean Recovery Efficiencies [$\overline{Mean}(R_{SAP,S})$], and %RSDs of Recovery Efficiencies [%RSD($R_{SAP,S}$)] for Single Swab, Wipe, and Vacuum Samples of Selected Sampling and Analysis Methods as Summarized by Piepel et al. (2011a, 2011b)

Sampling Method	Study	Surface Material	$C_{CFU:S}^M$ (CFU/cm^2)	$\overline{Mean}(R_{SAP,S})$ (%)	%RSD($R_{SAP,S}$) (%)
Macrofoam swab	Estill et al. (2011)	Stainless steel	0.0018–0.18	3.4–5.1	81–560
Macrofoam swab	Estill et al. (2011)	Carpet	0.0012–0.15	12.0–14.0	110–260
Macrofoam swab	Hodges et al. (2010)	Stainless steel	0.48–1064	15.8–83.1	15–59
Sponge-stick wipe	Rose et al. (2011)	Stainless steel	0.013–15.48	24.4–32.4	28–76
Sponge-stick wipe	Krauter et al. (2012)	Stainless steel	0.0015–0.974	35.8–63.6	8.9–106[b]
Sponge-stick wipe	Krauter et al. (2012)	Five materials[a]	0.0–0.994	4.0–76.9	0–316[b]
Vacuum, HEPA[c] sock	Estill et al. (2011)	Stainless steel	0.0043–0.10	3.7–5.5	79–82
Vacuum, HEPA sock	Estill et al. (2011)	Carpet	0.0011–0.062	3.7–6.3	100–130

(a) Ceramic tile, vinyl tile, faux leather, painted wallboard, plastic light cover panel.
(b) Includes only sample-within-run uncertainty.
(c) HEPA = high efficiency particulate air (filter).

7.3.2 Examples of Estimating the Random Uncertainty for a Measured Concentration from a Single Sample when Information is Available on the Recovery Efficiencies for All Steps of the Sampling and Analysis Process

Equation (7.5) in Section 7.1.2 presents the formula for the estimated %RSD of a measured contaminant concentration of a single sample when systematic and random uncertainty information is

available only on separate steps of SAP process. As noted previously, Piepel et al. (2011a, 2011b) did not summarize such information at that level of detail. However, we illustrate the use of Equation (7.5) using some hypothetical values of the five input %RSD values on the right side of Equation (7.5).

Suppose that $\overline{\%RSD}(R_{SC,S}) = 28.7$, $\overline{\%RSD}(R_{HST,S}) = 10.2$, $\overline{\%RSD}(R_{P,S}) = 15.3$, $\overline{\%RSD}(R_{A,S}) = 19.1$, and $\overline{\%RSD}(F_{CFU:Spore,S}) = 12.5$. Then, substituting these into Equation (7.5) yields

$$\overline{\%RSD}(C_{CFU,S}^{M}) \approx \left[\begin{array}{l} \{\overline{\%RSD}(R_{SC,S})\}^2 + \{\overline{\%RSD}(R_{HST,S})\}^2 + \{\overline{\%RSD}(R_{P,S})\}^2 \\ + \{\overline{\%RSD}(R_{A,S})\}^2 + \{\overline{\%RSD}(F_{CFU:Spore,S})\}^2 \end{array} \right]^{0.5}$$

$$= [(28.7)^2 + (10.2)^2 + (15.3)^2 + (19.1)^2 + (12.5)^2]^{0.5} = 41.0 .$$

Hence, for this example the total random uncertainty [over all of the sources of uncertainty represented in Equation (7.5)] in the measured surface concentration of a single sample obtained using a set S of SAP methods and sampling conditions is given by $\overline{\%RSD}(C_{CFU,S}^{M}) = 41.0$.

7.3.3 Examples of Random Uncertainty in an Estimate of the True Contaminant Concentration for a Single Sample when Information is Available Only on the Combined Recovery Efficiency

Equation (7.6) gives the formula for estimating the true contaminant concentration from a single sample when systematic and random uncertainty information is available only on the combined (over all steps of the SAP) recovery efficiency. Equation (7.8) gives the formula for the random uncertainty of this estimate of the true concentration from a single sample. An example illustrating the use of these equations is now presented.

Consider $C_{CFU,S}^{M} = 1.5$ CFU/cm^2, the measured surface concentration of a single sample obtained from a contaminated (with *Bacillus anthracis*) vinyl tile surface using the sponge-stick wipe method discussed by Rose et al. (2011) and Krauter et al. (2012). Krauter et al. (2012) reported that the recovery efficiency of the sponge-stick wipe sampling and analysis methods did not depend on surface concentration over the range they investigated (0.0025–1.86 CFU/cm^2). Rose et al. (2011) also reported that the recovery efficiency did not depend on surface concentration for the range they tested (0.013–17.13 CFU/cm^2). For a vinyl tile surface, Krauter et al. (2012, Table 2) reported a mean recovery efficiency of 25.6% [i.e., $\overline{Mean}(R_{SAP,S}) = 0.256$]. Then, applying Equation (7.6) yields

$$\hat{C}_{CFU,S}^{T} = \frac{C_{CFU,S}^{M}}{\overline{Mean}(R_{SAP,S})} = \frac{1.5}{0.256} = 5.86 \, CFU / cm^2 \tag{7.9}$$

as the estimate of the true surface concentration for the location sampled. Dividing the measured concentration by the average recovery efficiency of the method in essence bias-corrects the measured concentration to yield an estimate of the true concentration.

Both the measured concentration $(C_{CFU,S}^M)$ and mean estimate of the recovery efficiency $[\overline{Mean}(R_{SAP,S})]$ in Equation (7.9) are subject to random uncertainty. Considering the results in Table 1 of Krauter et al. (2012) for the tests on a vinyl surface with highest surface concentration, we chose $\overline{\%RSD}(C_{CFU,S}^M) = 10.6$ and $\overline{\%RSD}(R_{SAP,S}^M) = 10.6$ (it is a coincidence that both values have the same value when rounded to one decimal place). From $\overline{\%RSD}(C_{CFU,S}^M) = 10.6$ we can calculate $\overline{SD}(C_{CFU,S}^M) = (10.6/100)1.5 = 0.159$ CFU/cm². Inserting these and previously discussed values into Equation (7.8) provides for estimating the random uncertainty (SD) in the estimated true surface concentration

$$
SD(\hat{C}_{CFU,S}^T) = \left\{ \left[\frac{SD(C_{CFU,S}^M)}{\overline{Mean}(R_{SAP,S})} \right]^2 + \left[\frac{C_{CFU,S}^M}{\overline{Mean}(R_{SAP,S})} \right]^2 \left[\frac{\overline{\%RSD}(R_{SAP,S})}{100\sqrt{n}} \right]^2 \right\}^{0.5}
$$

(7.10)

$$
= \left\{ \left[\frac{0.159}{0.256} \right]^2 + \left[\frac{1.5}{0.256} \right]^2 \left[\frac{10.6}{100\sqrt{10}} \right]^2 \right\}^{0.5} = 0.65 \text{ CFU/cm}^2
$$

where the $\sqrt{10}$ is a result of Krauter et al. (2012) calculating the mean and %RSD of recovery efficiency for the sponge-stick wipe method applied to vinyl surfaces based on 10 replicate coupons.

In summary, for this example of using the sponge-stick wipe method applied to a vinyl surface and obtaining a measured concentration from a single sample of 1.5 CFU/cm², the estimate of the true surface concentration is 5.86 CFU/cm² from Equation (7.9) with a standard deviation of 0.65 CFU/cm² from Equation (7.10).

8.0   Summary and Future Needs

Section 8.1 summarizes the work that has been performed so far and documented in this first version of the report. Section 8.2 discusses the extensions of the statistical sampling methods and formulas, as well as data inputs needed, to finalize the methods and formulas to address situations that may occur in contamination incidents.

8.1   Summary

This report documents the methodology, formulas, and inputs needed to make characterization and clearance decisions for *Bacillus anthracis*-contaminated and uncontaminated (or decontaminated) areas using a statistical sampling approach. Specifically, the report includes the methods and formulas for calculating the

- number of samples required to achieve a specified confidence in characterization and clearance decisions

- confidence in making characterization and clearance decisions for a specified number of samples

for two common statistically based environmental sampling approaches, discussed subsequently.

The purpose of the work in this report is to address a key issue raised by the Government Accountability Office (GAO) in reports (2005a, 2005b) following the 2001 anthrax letter attacks. That issue is the importance of having high confidence that a decision area is uncontaminated (or successfully decontaminated) if all samples collected have negative results. Statistical sampling approaches provide the structure for making clearance statements with high confidence. To be able to state with high confidence that there is virtually no contamination based on all samples having negative results, it is important to have knowledge of the probability that an individual sample result is a false negative. This probability is commonly referred to as the *false negative rate (FNR).*

The two statistical sampling approaches currently discussed in this report are 1) *hotspot sampling* to detect small isolated contaminated locations during the characterization phase, and 2) *combined judgment and random (CJR) sampling* to meet clearance objectives. If contamination is widely distributed in a decision area, it will typically be detectable via judgment sampling during the characterization phrase. Hotspot sampling provides for sampling in characterization situations where contamination is not widely distributed and may not be detected by judgment sampling. The CJR sampling approach provides for making a clearance statement about a decision area with high confidence. These two statistical sampling approaches are discussed in this report because they are the two mentioned in a forthcoming sampling reference document being developed by the VSPWG. Other statistical sampling approaches could be included in future revisions of this report (not the VSPWG document) if needed.

The hotspot and CJR statistical sampling approaches are discussed in the report for four situations:

1. qualitative data (detect and non-detect) when the FNR = 0 or when using statistical sampling methods that account for the FNR > 0

2. qualitative data when the FNR > 0 but statistical sampling methods are used that assume the FNR = 0

3. quantitative data (e.g., contaminant concentrations expressed as CFU/cm^2) when the FNR = 0 or when using statistical sampling methods that account for the FNR > 0

4. quantitative data when the FNR > 0 but statistical sampling methods are used that assume the FNR = 0.

For these situations, the hotspot sampling approach provides Z% confidence of identifying within the decision area an unacceptable region of a specified shape and size (or larger), provided such a region exists. The CJR approach provides X% confidence that at least Y% of the decision area is acceptable, provided all samples are acceptable. Various forms of these "characterization statements" and "clearance statements," as well as a definition of what constitutes an "acceptable" sample and decision area under these four circumstances, are provided in Section 2.2.

Statistical methods that account for FNR > 0 currently exist only for the hotspot sampling approach with qualitative data (or quantitative data converted to qualitative data). This report documents the current status of methods and formulas for the hotspot and CJR sampling approaches. Section 8.2 identifies limitations of these methods and possible extensions of the methods applicable to account for FNR > 0.

For quantitative data, this report also presents statistical methods and formulas for 1) quantifying the uncertainty in measured sample results, 2) estimating the true surface concentration corresponding to a surface sample, and 3) quantifying the uncertainty of the estimate of the true surface concentration. These and other methods and formulas discussed in the report were applied to example situations to illustrate application of the methods and interpretation of the results.

8.2 Future Needs

This section identifies extensions and data inputs that are needed to complete methods, formulas, and data for calculating numbers of samples and confidence in characterization and clearance decisions using two statistical sampling approaches. Section 8.2.1 discusses future data needs, while Sections 8.2.2 and 8.2.3 list, respectively, possible future extensions of the hotspot and CJR sampling methods and formulas. Relevant portions of this document may be updated and completed if future funding supports the work required to develop method extensions and additional data needed as inputs.

8.2.1   Data Needs

The primary data needed as inputs to methods and formulas for calculating the 1) numbers of samples required in statistically based sampling plans, and 2) confidence in characterization and clearance decisions, are quantitative performance measures that capture the principal sources of uncertainty in the main steps of the SAP. The principal sources of uncertainty are sample collection, sample handling/transportation/storage, sample processing/extraction, and sample analysis (e.g., culturing and counting).

The main quantitative performance measures are recovery efficiency (RE) [i.e., systematic/bias uncertainty], random uncertainty in recovery efficiency [e.g., %RSD(RE)], limit of detection (LOD), false negative rate (FNR), and false positive rate (FPR). Piepel et al. (2011a, 2011b) summarized these

performance measures for 20 laboratory studies from the literature that investigated various sampling and analysis methods for swab, wipe, and vacuum sampling. They focused on combined estimates of RE and its random uncertainty, where "combined" refers to including the four main sources of uncertainty identified at the end of the previous paragraph. Having information on the performance of each main step of a sampling and analysis method can be useful in improving individual steps and hence the overall method performance. However, it is time consuming and costly to collect sufficient data to estimate well the performance measures for the four main steps of a sampling and analysis process. Collecting sufficient data to quantify combined performance measures is sufficiently time consuming and costly that Piepel et al. (2011a, 2011b) found limited or incomplete results for many SAP methods whose performance has been studied.

Hence, we recommend that future efforts focus on adequately estimating performance measures (RE, %RSD(RE), LOD, FNR, FPR) over the four main steps of the SAP (sample collection, sample handling/storage/transportation, sample processing/extraction, sample analysis) for the one or two recommended (or most widely used) set(s) of methods for each of swab, wipe, and vacuum sampling. FNR is a key performance measure that directly determines the confidence that a decision area is not contaminated when all sample results are negative. Piepel et al. (2011a, 2011b) noted that only one of the 20 studies they summarized reported FNR results. Because the performance measures can depend on the concentration of contaminant and/or the surface material being sampled, performance measures must be estimated over the range of concentrations and surface materials that may be sampled in *Bacillus anthracis* contamination incidents. This includes low concentrations below which false negatives may occur for a given set of sampling and analysis methods when sampling a range of surface materials. See Piepel et al. (2011a, 2011b) for more detailed discussions of data gaps and needs for the performance of swab, wipe, and vacuum sampling methods.

The sponge-stick wipe method recommended by CDC has the best accumulation of performance data, given the recent results of Krauter et al. (2012) for the FNR performance of the method over a range of surface concentrations and surface materials, along with the previous results of Rose et al. (2011). The results of Hodges et al. (2010) for the macrofoam swab set of methods recommended by CDC need to be augmented with results summarizing the FNR performance for a range of surface concentrations and surface materials, similar to the study by Krauter et al. (2012) for the sponge-stick wipe set of methods.

8.2.2   Needs for Hotspot Sampling Methodology

The possible improvements or extensions of the hotspot sampling approach are discussed in Section 5.4, and are listed in the following bullets.

False Negative Rate = 0 with Qualitative Data (or Quantitative Data Converted to Qualitative Data)

- Extend methodology to allow samples obtained with different collection methods (e.g., swab, wipe, and vacuum) and/or different surface materials within a decision area to have different sample footprints (areas).

False Negative Rate > 0 with Qualitative Data (or Quantitative Data Converted to Qualitative Data)

- Extend methodology to allow samples covering a specified footprint (area), instead of the current limitation of "point samples" (samples covering negligible area).

- Extend methodology to allow samples obtained with different collection methods (e.g., swab, wipe, and vacuum) and/or different surface materials within a decision area to have different sample footprints (areas) and different FNRs.

- Extend methodology to allow hotspots to have elliptical shapes, rather than only circular shapes.

The extensions in the first and second bullets would be very complicated, in that the hotspot would have to be well-defined to account for partial overlap of the hotspot with the sample footprint. A more complicated FNR model would also be required to account for the case when a sample footprint partially intersects with the edge of a hotspot. An important question is whether these complicated extensions would be worth the effort, especially when the concept of a well-defined hotspot isn't compatible with most contaminant-dispersion models.

In the third bullet, there may be little value-added to extending the methodology to handle elliptical as well as circular hotspots, because ultimately the concept of a hotspot is a heuristic to develop a sampling design if smaller, isolated areas may be unacceptable. The shapes of such unacceptable areas would likely not be either circular or elliptical, but elliptical may at least be more realistic than circular. Hence, this methodology extension may add value for applications with smaller, isolated unacceptable areas.

8.2.3   Needs for Combined Judgment and Random Sampling Methodology

The possible improvements or extensions of the CJR sampling approach are discussed in Section 6.4, and are listed in the following bullets.

Less Restrictive Assumptions

- One of the important assumptions for the CJR approach is that the number of judgment samples is equal to the number of high-risk grid-cells in the decision area that are determined a priori. While this assumption simplifies the mathematics considerably, it would be more realistic to not require that all of the high-risk area be sampled.

- A generalization of the CJR method would allow there to be more than two risk groups in the decision area. In particular, this would accommodate different surface types, their propensity to attract and/or retain contamination, and different sampling methods (swab, wipe, vacuum). Some progress has been made on this stratified clearance sampling method, and a preliminary, non-public version exists in VSP. However, more work will be required to fully develop this methodology and implement it in VSP.

False Negative Rate = 0 with Qualitative Data (or Quantitative Data Converted to Qualitative Data)

- Extend methodology to allow samples obtained with different collection methods (e.g., swab, wipe, and vacuum) and/or different surface materials within a decision area to have different sample footprints (areas).

False Negative Rate > 0 with Qualitative Data (or Quantitative Data Converted to Qualitative Data)

- Extend methodology for FNR = 0 with qualitative data (or quantitative data converted to qualitative data) to address FNR > 0.

- Extend methodology to allow 1) samples obtained with different collection methods (e.g., swab, wipe, and vacuum) and/or different surface materials within a decision area to have different sample areas, and 2) different strata with different values of FNR > 0. Different strata could be defined by the collection method and other methods in the SAP, as well as the surface being sampled.

9.0   References

Amidan BG, GF Piepel, BD Matzke, JJ Filliben, and B Jones. 2007. *Experimental Design for the INL Sample Collection Operational Test*. PNNL-17129, Pacific Northwest National Laboratory, Richland, Washington. Accessed March 29, 2013 at http://www.pnl.gov/main/publications/external/technical_reports/PNNL-17129.pdf.

Amidan BG, BA Pulsipher, and BD Matzke. 2009. *Statistical Analyses of Second Indoor Bio-Release Field Evaluation Study at Idaho National Laboratory*. PNNL-18932, Pacific Northwest National Laboratory, Richland, Washington. Accessed March 29, 2013 at http://www.pnl.gov/main/publications/external/technical_reports/PNNL-18932.pdf.

Axelrod M. 2005. *Using Ancillary Information to Reduce Sample Size in Discovery Sampling and the Effects of Measurement Error*. UCRL-TR-216206, Lawrence Livermore National Laboratory, Livermore, California. Accessed March 29, 2013 at https://e-reports-ext.llnl.gov/pdf/324013.pdf.

Bowen MW and CA Bennett. 1988. *Statistical Methods for Nuclear Material Management*. NUREG/CR-4604. U.S. Nuclear Regulatory Commission, Washington, D.C.

Davidson JR. 1995. *ELIPGRID-PC: Upgraded Version.* ORNL/TM-13103, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Estill, CF, PA Baron, JK Beard, MJ Hein, LD Larsen, L Rose, FW Schaefer III, J Noble-Wang, L Hodges, HAD Lindquist, GJ Deye, and MJ Arduino. 2009. "Recovery Efficiency and Limit of Detection of Aerosolized *Bacillus anthracis* Sterne from Environmental Surface Samples." *Applied and Environmental Microbiology* 75:4297-4306.

Government Accountability Office (GAO). 2005a. *Anthrax Detection: Agencies Need to Validate Sampling Activities in Order to Increase Confidence in Negative Results* (Report to the Chairman, Subcommittee on National Security, Emerging Threats, and International Relations, House Committee on Government Reform, House of Representatives). GAO-05-251, U.S. Government Accountability Office, Washington, D.C., March 2005. Accessed March 29, 2013 at http://www.gao.gov/products/GAO-05-251.

Government Accountability Office (GAO). 2005b. *Anthrax Detection: Agencies Need to Validate Sampling Activities in Order to Increase Confidence in Negative Results* (Testimony before the Chairman, Subcommittee on National Security, Emerging Threats, and International Relations, House Committee on Government Reform, House of Representatives). GAO-05-493T, U.S. Government Accountability Office, Washington, D.C., April 2005. Accessed March 29, 2013 at http://www.gao.gov/products/GAO-05-493T.

Gilbert RO. 1987. *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold, New York.

Grieve AP. 1994. "A Further Note on Sampling to Locate Rare Defectives with Strong Prior Evidence." *Biometrika* 81:787–789. Accessed March 29, 2013 at http://biomet.oxfordjournals.org/content/81/4/787.abstract.

Hahn GJ and WQ Meeker. 1991. *Statistical Intervals, A Guide to Practitioners*. John Wiley & Sons, Inc., New York.

Hahn GJ and S Shapiro. 1967. *Statistical Models in Engineering*. John Wiley and Sons, Inc., New York.

Hodges, LR, LJ Rose, H O'Connell, and MJ Arduino. 2010. "National Validation Study of a Swab Protocol for the Recovery of Bacillus anthracis Spores from Surfaces." *Journal of Microbiological Methods* 81:141–146. Accessed March 29, 2013 at http://www.sciencedirect.com/science/article/pii/S0167701210000692.

Krauter PA, GF Piepel, R Boucher, M Tezak, BG Amidan, and W Einfeld. 2012. "False Negative Rate and Other Performance Measures of a Sponge-Wipe Surface Sampling Method for Low Contaminant Concentrations." *Journal of Applied and Environmental Microbiology* 78:846–854.

Krose, DP, T Taimre, and ZI Botev. 2011. *Handbook of Monte Carlo Methods*. John Wiley and Sons, New York.

Matzke BD, JE Wilson, LL Nuffer, ST Dowson, JE Hathaway, NL Hassig, LH Sego, CJ Murray, BA Pulsipher, B Roberts, and S McKenna. 2010. *Visual Sample Plan Version 6.0 User's Guide*. PNNL-19915, Pacific Northwest National Laboratory, Richland, Washington. Accessed March 29, 2013 at: http://vsp.pnl.gov/docs/pnnl19915.pdf.

Natrella MG. 1966. *Experimental Statistics*. National Bureau of Standards Handbook 91. John Wiley and Sons, New York.

Odeh RE and DB Owen. 1980. *Tables for Normal Tolerance Limits, Sampling Plans and Screening*. Marcel Dekker, New York.

Piepel GF, BG Amidan, and BD Matzke. 2009. *Experimental and Sampling Design for the INL-2 Sample Collection Operational Test*. PNNL-18187, Pacific Northwest National Laboratory, Richland, Washington. Accessed March 29, 2013 at http://www.pnl.gov/main/publications/external/technical_reports/PNNL-18187.pdf.

Piepel GF, BG Amidan, and R Hu. 2011a. *Laboratory Studies on Surface Sampling of Bacillus anthracis Contamination: Summary, Gaps, and Recommendations*. PNNL-20910, Pacific Northwest National Laboratory, Richland, Washington. Accessed March 29, 2013 at http://www.pnnl.gov/main/publications/external/technical_reports/PNNL-20910.pdf.

Piepel, G.F., B.G. Amidan, and R. Hu. 2012. "Laboratory Studies on Surface Sampling of *Bacillus anthracis* Contamination: Summary, Gaps, and Recommendations," *Journal of Applied Microbiology*, doi: 10.1111/j.1365-2672.2012.05381.x.

Rose LJ, L Hodges, H O'Connell, and J Noble-Wang. 2011. "National Validation Study of a Cellulose Sponge-Wipe Processing Method for Use after Sampling *Bacillus anthracis* Spores from Surfaces." *Applied Environmental Microbiology* 77:8355–8359. Accessed March 29, 2013 at http://aem.asm.org/content/77/23/8355.full.

Schilling EG and DV Neubauer. 2009. *Acceptance Sampling in Quality Control*, Second Edition. Chapman & Hall/CRC Press, Boca Raton, Florida.

Sego LH and JE Wilson. 2007. *Accounting for False Negatives in Hotspot Detection*. PNNL-16812, Pacific Northwest National Laboratory, Richland, Washington. Accessed March 29, 2013 at http://www.pnl.gov/main/publications/external/technical_reports/PNNL-16812.pdf.

Sego LH, KK Anderson, BD Matzke, WK Sieber, S Shulman, J Bennett, M Gillen, JE Wilson, and BA Pulsipher. 2007. *An Environmental Sampling Model for Combining Judgment and Randomly Placed Samples*. PNNL-16636, Pacific Northwest National Laboratory, Richland, Washington. Accessed March 29, 2013 at http://www.pnl.gov/main/publications/external/technical_reports/PNNL-16636.pdf.

Sego LH, S Shulman, KK Anderson, JE Wilson, BA Pulsipher, and WK Sieber. 2010. *Acceptance Sampling Using Judgmental and Randomly Selected Samples*. PNNL-19315, Pacific Northwest National Laboratory, Richland, Washington. Accessed March 29, 2013 at http://www.pnl.gov/main/publications/external/technical_reports/PNNL-19315.pdf.

Singer DA. 1972. "ELIPGRID: A Fortran IV Program for Calculating the Probability of Success in Locating Elliptical Targets with Square, Rectangular and Hexagonal Grids." *Geocom Bulletin/Programs* 4:1–16.

Singer DA. 1975. "Relative Efficiencies of Square and Triangular Grids in the Search for Elliptically Shaped Resource Targets." *Journal of Research of the U.S. Geological Survey* 3(2):163–167.

Singer DA and FE Wickman. 1969. *Probability Tables for Locating Elliptical Targets with Square, Rectangular, and Hexagonal Point Nets*. Special Publication 1-69, Pennsylvania State University, University Park, Pennsylvania. Accessed March 29, 2013 at http://www.libraries.psu.edu/content/dam/psul/up/emsl/documents/specialpub/specialpub69-1.pdf.

Squeglia NL. 1994. *Zero Acceptance Number Sampling Plans*. ASQ Quality Press, Milwaukee, Wisconsin.

VSP Development Team. 2012. *Visual Sample Plan: A Tool for Design and Analysis of Environmental Sampling*, Version 6.0. Pacific Northwest National Laboratory, Richland, Washington. Available at http://vsp.pnnl.gov.

Wright T. 1992. "A Note on Sampling to Locate Rare Defectives with Strong Prior Evidence." *Biometrika* 79:685–691.

Appendix A

Instructions for Locating Sampling Approaches in VSP

Appendix A

Instructions for Locating Sampling Approaches in VSP

In this Appendix, we illustrate how to access the various sampling designs discussed in this report in Visual Sample Plan Version 6.4.

A.1 Hotspot Designs

From the pull-down menu:

Sampling Goals → Locate Hot Spots → Assume no false negative errors

Sampling Goals → Locate Hot Spots → Account for false negative errors

A.2 The AOZ-ACS Design

From the pull-down menu:

Sampling Goals → Show that at least some high % of the sampling area is acceptable → Using presence/absence measurements
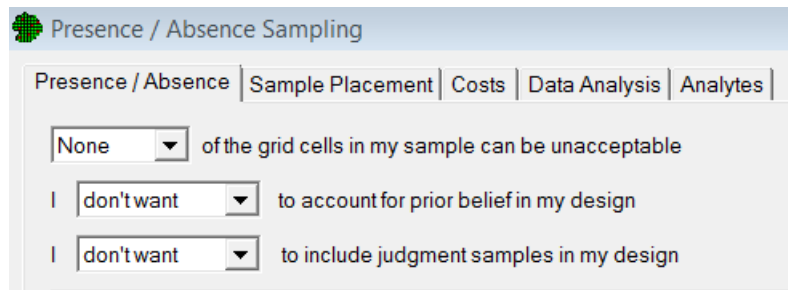
Then choose the options shown in Figure A.1

Figure A.1. Options to Select the AOZ-ACS Design

A.3 The CJR Design

From the pull-down menu:

Sampling Goals → Show that at least some high % of the sampling area is acceptable → Using presence/absence measurements

then choose the options shown in Figure A.2.

Figure A.2. Options to Select the CJR Design

The two special cases of the CJR design discussed in Section 6.1.1.1 can also be obtained as illustrated in Figure A.3 and Figure A.4.

Figure A.3. Options to Select the CJR with a Non-informative Prior (effectively no prior belief)

Figure A.4. Options to Select a Bayesian Design that Accounts for Prior Belief with No Judgment Samples

Appendix B

Development of Equations in Section 7

Appendix B

Development of Equations in Section 7

This appendix presents the development of equations in Sections 7.1 and 7.2 of Section 7.

B.1 Development of Uncertainty Equation in Section 7.1.1

This section presents the development of Equation (7.2) in Section 7.1.1. Equation (7.1) is a multiplicative model, which can be written in additive form by taking the natural logarithms of both sides

$$\ln(C_{CFU,S}^{M}) = \ln(R_{SAP,S}) + \ln(C_{CFU,S}^{T}) \tag{B.1}$$

where all of the notation is as defined in Section 7.1.1 and ln(x) denotes the natural logarithm of x. A list of acronyms and notation, with their definitions, is located in the front of the report.

The random uncertainty (represented by the standard deviation) in $\ln(C_{CFU,S}^{M})$ is given by

$$SD[\ln(C_{CFU,S}^{M})] = SD[\ln(R_{SAP,S})] \tag{B.2}$$

because $C_{CFU,S}^{T}$ is a constant. Applying the approximate relationship $SD[\ln(X)] \approx \%RSD(X)/100$ to Equation (B.2) yields

$$\%RSD(C_{CFU,S}^{M}) \approx \%RSD(R_{SAP,S}) \tag{B.3}$$

This equation is listed as Equation (7.2) in Section 7.1.1.

B.2 Development of Uncertainty Equation in Section 7.1.2

This section presents the development of Equation (7.4) in Section 7.1.2. Equation (7.3) is a multiplicative model, which can be written in additive form by taking the natural logarithms of both sides

$$\begin{aligned}
\ln(C_{CFU,S}^{M}) &= \ln(R_{SAP,S}) + \ln(C_{CFU,S}^{T}) \\
&= \ln(R_{SAP,S}) + \ln(F_{CFU:Spore,S}) + \ln(C_{Spores,S}^{T}) \\
&= \ln(R_{SC,S}) + \ln(R_{HST,S}) + \ln(R_{P,S}) + \ln(R_{A,S}) + \ln(F_{CFU:Spore,S}) + \ln(C_{Spores,S}^{T})
\end{aligned} \tag{B.4}$$

where all of the notation is as previously defined and ln(x) denotes the natural logarithm of x. A list of acronyms and notation, with their definitions, is located in the front of the report.

The random uncertainty (represented by the standard deviation) in $\ln(C_{CFU,S}^M)$ is given by

$$SD[\ln(C_{CFU,S}^M)] = SD[\ln(R_{SC,S}) + \ln(R_{HST,S}) + \ln(R_{P,S}) + \ln(R_{A,S})$$

$$+ \ln(F_{CFU:Spore,S}) + \ln(C_{Spores,S}^T)] \ .$$

(B.5)

Under the assumption that the random variables $R_{SC,S}$, $R_{HST,S}$, $R_{P,S}$, $R_{A,S}$, and $F_{CFU:Spore,S}$ are statistically independent (which is reasonable because they are associated with different steps of the sampling and analysis process), Equation (B.5) becomes

$$SD[\ln(C_{CFU,S}^M)] = [\{SD[\ln(R_{SC,S})]\}^2 + \{SD[\ln(R_{HST,S})]\}^2 + \{SD[\ln(R_{P,S})]\}^2 + \{SD[\ln(R_{A,S})]\}^2$$

$$+ \{SD[\ln(F_{CFU:Spore,S})]\}^2]^{0.5}$$

(B.6)

Then, using the approximate relationship $SD[\ln(X)] \approx \%RSD(X)/100$, Equation (B.6) becomes

$$\frac{\%RSD(C_{CFU,S}^M)}{100} \approx \left[\begin{array}{l} \left\{ \dfrac{\%RSD(R_{SC,S})}{100} \right\}^2 + \left\{ \dfrac{\%RSD(R_{HST,S})}{100} \right\}^2 + \left\{ \dfrac{\%RSD(R_{P,S})}{100} \right\}^2 \\[2mm] + \left\{ \dfrac{\%RSD(R_{A,S})}{100} \right\}^2 \end{array} \right]^{0.5}$$

(B.7)

which reduces to

$$\%RSD(C_{CFU,S}^M) \approx \left[\begin{array}{l} \{\%RSD(R_{SC,S})\}^2 + \{\%RSD(R_{HST,S})\}^2 + \{\%RSD(R_{P,S})\}^2 \\[2mm] + \{\%RSD(R_{A,S})\}^2 + \{\%RSD(F_{CFU:Spore,S})\}^2 \end{array} \right]^{0.5} .$$

(B.8)

This is listed as Equation (7.4) in Section 7.1.2.

Distribution

EXTERNAL DISTRIBUTION

3   **Department of Homeland Security**
     Randy Long (randolph.long@dhs.gov)
     Tina Sanders
       (christina.a.sanders@associates.hq.dhs.gov)
     Segaran Pillai (segaran.pillai@dhs.gov)

8   **Environmental Protection Agency**
     Erica Canzler (canzler.erica@epa.gov)
     Schatzi Fitz-James (fitz-james.schatzi@epa.gov)
     Romy Campisano
       (campisano.romy@epamail.epa.gov)
     Dino Mattorano (mattorano.dino@epa.gov)
     Marissa Mullins (mullins.marissa@epa.gov)
     Tonya Nichols (nichols.tonya@epamail.epa.gov)
     Shawn Ryan (ryan.shawn@epa.gov)
     Sanjiv Shah (shah.sanjiv@epa.gov)

1   **Sandia National Laboratories**
     Bob Knowlton (rgknowl@sandia.gov)

1   **National Institute of Standards & Technology**
     Jayne Morrow (jayne.morrow@nist.gov)

3   **Johns Hopkins University Applied Physics Laboratory**
     Elizabeth Corson (Elizabeth.corson@jhuapl.edu)
     Rachel Quizon (rachel.quizon@jhuapl.edu)
     Eric Van Gieson (eric.van.gieson@jhuapl.edu)

1   **American College of Forensic Examiners**
     John Bridges (JHBridges@vzw.blackberry.net)

2   **Federal Bureau of Investigation**
     Doug Anders (douglas.anders@ic.fbi.gov)
     Doug Beecher (douglas.beecher@ic.fbi.gov)

8   **Centers for Disease Control and Protection**
     Matthew Arduino (marduino@cdc.gov)
     Lisa Delaney (ldelaney1@cdc.gov)
     Cherie Estill (clf4@cdc.gov)
     Richard Kellogg (rbk1@cdc.gov)
     Stephen Morse (sam1@cdc.gov)
     Laura Rose (lmr8@cdc.gov)
     Angela Weber (aweber@cdc.gov)
     Betsy Weirich (eweirich@cdc.gov)

2   **Institute for Defense Analyses**
     Jeff Grotte (jgrotte@ida.org)

2   **Homeland Security Institute**
     Ed Hildebrand (carl.hildebrand@hsi.dhs.gov)
     Eric Sylwester (eric.sylwester@hsi.dhs.gov)

1   **Department of Defense – Joint Program Executive Office (JPEO)**
     Emma Wilson (emma.wilson@jpeocbd.osd.mil)

1   **Department of Defense –ECBC**
     Vipin Rastogi (vipin.rastogi@us.army.mil)

INTERNAL DISTRIBUTION

6   **Pacific Northwest National Laboratory**
     Brett Amidan (brett.amidan@pnnl.gov)
     Brett Matzke (brett.matzke@pnnl.gov)
     Greg Piepel (greg.piepel@pnnl.gov)
     Brent Pulsipher (brent.pulsipher@pnnl.gov)
     Landon Sego (landon.sego@pnnl.gov)
     Information Release