ISSUE PAPER:

ESTIMATES OF SAMPLE SIZES REQUIRED FOR A GENERATOR TO DEMONSTRATE A WASTE QUALIFIES FOR EXEMPTION UNDER HWIR

Prepared For:

U.S. Environmental Protection Agency
Office of Solid Waste
Economics, Methods, and Risk Analysis Division
401 M Street, SW
Washington, DC 20460

Prepared By:

Science Applications International Corporation
1710 Goodridge Drive
McLean, VA 22102

May 21, 1999

EPA Contract No. 68-W6-0068, WA 1-9, QRT 1
SAIC Project No. 06-5140-08-6663-601

Contents

1. Introduction 1
   - 1.1 Background 1
   - 1.2 Preferred Options 2

2. Data Quality Objectives 3

3. Description and Analysis of Options 5
   - 3.1 Option No. 1: Using a “No Exceedance” Decision Rule to Qualify a Waste 5
     - 3.1.1 Description 5
     - 3.1.2 Advantages 5
     - 3.1.3 Limitations 5
     - 3.1.4 Approach 6
     - 3.1.5 Number of Samples Required 7
     - 3.1.6 Can Composite Sampling Be Used with the “No Exceedance” Rule? 10
     - 3.1.7 How is Subsequent Testing Performed After the Initial Exemption is Demonstrated? 11
   - 3.2 Option No. 2: Using an Upper Confidence Limit on the Mean to Qualify a Waste for the HWIR Exemption 12
     - 3.2.1 Description 12
     - 3.2.2 Advantages 12
     - 3.2.3 Limitations 12
     - 3.2.4 Approach 13
     - 3.2.5 Number of Samples Required 13
     - 3.2.6 Can Composite Sampling Be Used to Calculate the UCL on the Mean? 15
     - 3.2.7 How is Subsequent Testing Performed After the Initial Exemption is Demonstrated? 15
   - 3.3 Option No. 3: Using a Sample Mean that is Less than the Exemption level with the Maximum Sample Value Less than a Multiple of the Exemption level to Qualify a Waste for Exemption 16
     - 3.3.1 Description 16
     - 3.3.2 Advantages 17
     - 3.3.3 Limitations 18
     - 3.3.4 Approach 18
     - 3.3.5 Number of Samples Required 18
     - 3.3.6 Can Composite Sampling Be Used? 20
     - 3.3.7 How is Subsequent Testing Performed After the Initial Exemption is Demonstrated? 20

4. Summary and Recommendations 20

Estimates of Sample Sizes Required for a Generator to Demonstrate a Waste Qualifies for Exemption Under HWIR

1. Introduction

EPA's Hazardous Waste Identification Rule (HWIR) sampling subteam requires estimates of the number of samples that a waste generator will need to take to demonstrate that a waste qualifies for exemption under the rule. In addition, the subteam is interested in the number of samples that will be required in subsequent testing to demonstrate maintenance of the exemption. Estimates of the number and frequency of sampling will be used by EPA for two purposes: (1) as inputs to cost and economic analyses in support of the proposed rule (e.g., for use in the supporting statement for the ICR and in the Regulatory Impact Analysis (RIA)), and (2) to develop regulatory and preamble language for inclusion in the proposed rule.

1.1 Background

In 1995, EPA published the proposed HWIR to amend its regulations under the Resource Conservation and Recovery Act (RCRA) by establishing constituent-specific exit levels for low- risk solid wastes that are designated as hazardous because they are listed, or have been mixed with, derived from, or contain listed hazardous wastes. Under this proposal, generators of listed hazardous wastes that meet the self-implementing exemption levels would no longer be subject to the hazardous waste management system under Subtitle C of RCRA as listed hazardous wastes (see draft implementation language from 1995 in Appendix A).

Currently, EPA is working to re-propose the HWIR. EPA plans to publish a proposed rule by October 31, 1999, and a final rule by April 30, 2001. As part of the work to develop the new rule, in 1997 EPA assembled an HWIR sampling subteam composed of representatives from various EPA program offices including OSW, ORD, and NEIC. An extensive and extended dialogue took place to explore options for sampling and testing to demonstrate that a waste qualifies for exclusion under HWIR. Several options were identified, and EPA explored the advantages and limitations of each.

In this issue paper we summarize the preferred options (as identified by EPA management and the sampling subteam), identify key sampling objectives, state the advantages and limitations of each of the options, and present an estimate of the number of samples required by a waste generator to determine if a waste qualifies for the exemption under HWIR. <u>One of the key underlying assumptions of this analysis is that samples will be obtained using a random sampling design from a defined population (mass) of waste. Furthermore, it is assumed that minimal sampling and analytical error are introduced in the data collection process.</u> Other assumptions and supporting analyses are presented in the text and appendices.

This paper *does not* address the following:

(1) how best to define the appropriate volumes of waste to be sampled (i.e., the size of the DQO decision unit[1]): EPA is considering proposing an upper bound on the mass of waste subject to

[1] A "decision unit" is simply a volume or mass of material (such as waste or soil) about which a decision will be made.

testing, however, the generator will have the option of defining decision units that are smaller than the upper bound specified by EPA. For highly variable (heterogeneous) wastes, relatively small decision unit -- such as individual drums or "batches" of waste -- may be appropriate to minimize the chance of mixing hazardous waste with non-hazardous waste. For relatively homogeneous wastes, larger decision units could be used to reduce sampling and analysis costs while maintaining a high level of confidence in the decision. The revised version of Chapter Nine, SW-846 recommends the generator "define the smallest, most appropriate subsets of the population (sub-populations), waste, or media to be characterized based on spatial or temporal boundaries. The boundaries will define the unit of waste or media about which a decision will be made."

(2) sampling frequency. EPA is considering proposing a minimum sampling frequency in the rule, however, it will remain the generator's responsibility to ensure that a waste always meets the exemption requirements for all HWIR exemption chemicals, regardless of which chemical the generator is required to test and how often such testing is performed.

(3) estimates of the number of samples a regulatory enforcement agency may take to test generator compliance with the exemption levels.

1.2 Preferred Options

In developing the preferred options for testing compliance with the HWIR exemption levels, EPA sought to identify approaches that:

- balance the need for more data (to reduce uncertainty) with the need to limit data collection costs.
- are relatively simple and easy to implement.
- are scientifically and statistically defensible.
-
- are protective of human health and the environment.

While EPA's sampling subteam identified a number of options, three primary options emerged for consideration in the proposed rule. The three options include:

Option 1 **"Not To Exceed" rule** in which the maximum sample value must be less than or equal to the HWIR exemption level for the waste to qualify for an exemption (This option was selected by EPA management as the preferred option);

Option 2 **Confidence Limit on the Mean** in which the upper confidence limit (UCL) on the mean must be less than or equal to the HWIR exemption level to qualify for the exemption; and

Option 3 **Sample mean less than the Exemption level with the Maximum capped by a multiple of the Exemption level.** The mean of the samples must remain below the exemption level and the maximum sample value must be less than some multiple of the HWIR exemption level (e.g., 2.8X) to qualify for an exemption.

2. Data Quality Objectives

To further clarify and describe each of the above options and provide a basis for estimating the number of samples required to obtain the initial exemption, it is first necessary to address several key questions that normally are addressed in Step Five of EPA's Data Quality Objective (DQO) Process (USEPA 1994): "Develop a Decision Rule."

Table 1 presents a summary of outputs of this step for each of the three options under consideration.

Table 1. Selected DQO Outputs for Each Decision Rule Option

DQO Step	DQO Output	Decision Rule Options		
		Option No. 1	**Option No. 2**	**Option No. 3**
		"No Exceedance" of HWIR Exemption level	**Upper Confidence Limit on the mean must be less than or equal to HWIR Exemption level**	**Mean must be less than HWIR exemption level, with no single sample value greater than [x] times the Exemption level**
Develop a Decision Rule (DQO Step Five)	Parameter of interest	Maximum (since the maximum is unknown, use the UCL on an upper percentile as an estimate of the maximum)	Mean (since the true mean is unknown, use the UCL on the sample mean)	Sample mean (a point estimate of the true mean) and the maximum (since the maximum is unknown, use the UCL on an upper percentile as an estimate of the maximum)
	Action level.	HWIR Exemption level	HWIR Exemption level	The sample mean is compared to the HWIR exemption level. The maximum is compared to some multiple of the HWIR Exemption level [e.g., as set by policy or as set by facility using a prediction limit.].
	An "if ...then..." statement that defines the conditions that would cause the decision maker to choose among alternative actions.	"If the maximum sample value is less than or equal to the HWIR exemption level for all constituents reasonably expected to be present in the waste, then the waste is not a hazardous waste. Otherwise, the waste is a hazardous waste."	"If the mean concentration (as measured by the UCL on the mean) is less than or equal to the HWIR exemption level for all constituents reasonably expected to be present in the waste, then the waste is not a hazardous waste. Otherwise, the waste is a hazardous waste."	"If the sample mean is less than the HWIR exemption level AND the maximum sample value is less than or equal to [a multiple of the HWIR exemption level] for all constituents reasonably expected to be present in the waste, then the waste is not a hazardous waste. Otherwise, the waste is a hazardous waste."

3. Description and Analysis of Options

3.1 Option No. 1: Using a "No Exceedance" Decision Rule to Qualify a Waste

3.1.1 Description

To apply a "no exceedance" decision rule, the maximum sample value must be less than or equal to the HWIR exemption level for the waste to qualify for an exemption. It is not possible to identify the maximum sample value in a waste.[2] However, we can use an upper percentile as a reasonable approximation of the maximum.

3.1.2 Advantages

- Simple.
- Easy to apply.
- Easy to enforce.
- Provides consistency with the manner in which LDR numeric treatment standards and Delisting levels are enforced.

3.1.3 Limitations

- Less waste will be eligible for exclusion under the "no exceedance" rule compared to the other options.

- Even if all samples are below the exemption level, the samples do not demonstrate compliance. They only show a lack of non-compliance.

- The more samples collected, the more likely that one sample will exceed the exemption level. That is, with large numbers of samples a facility is more likely to be disqualified by an "unlucky" rare but high concentration sample. This creates an incentive for generators to obtain as few samples as possible.

- Comparing the maximum concentration in a waste is not consistent with what the exemption levels are intended to represent. HWIR exemption levels are long-term (average) chronic risk-based standards. When chronic risks are of concern, EPA usually uses the mean as the parameter of interest (e.g., in Superfund risk assessments and cleanups).

- Use of the maximum was proposed as the preferred option in 1995. Commenters on the 1995 rule objected to the maximum test. Examples of comments include the following:

    - "The exemption levels were modeled based on long-term average concentrations, not

---

[2] It is important to note that the size, shape, and orientation of each physical sample of waste, as well as the manner in which the sample is handled and subsampled in the laboratory, will have a substantial effect on the analytical result and the distribution of the sample analysis data. For the purpose of this discussion, it is assumed that samples will be obtained, handled, and analyzed so that bias and random variability are kept within tolerable limits. Guidance for doing so is provided in the revised Chapter Nine of SW-846 (1999).

maximum exposures.”

S "Minor deviations [above the exemption level] on an isolated basis would not pose a significant threat to human health or the environment.”

S “The universe of hazardous waste qualifying for an exemption and/or implementing an exemption will be extremely tiny.”

S Waste management practices and units cannot quickly be switched between Subtitle C and D based on an exceedance of a single sample.

S The probability of failure increases with the number of samples because each sample increases the chance of the maximum exceeding a specific value.

S Even when the true level of each constituent does not change, the wastestream will test "out" (i.e., return to Subtitle C) a given percent of the time strictly because of random sample and test variation.

S The maximum approach would reduce the incentive for generators to seek an exemption under HWIR and “undermine EPA's stated goal of reducing the problem of "over-regulation" that exists under the current Subtitle C program.”

S “For the HWIR exemption to have any real utility, it must be predicated on a compliance system that does not impose such harsh punitive results for short-term compliance excursions.”

S The maximum approach “would make an already conservative exemption become totally unworkable.”

S The maximum approach “combined with the excessive waste characterization requirements, present such high risks and costs to a business, especially a small business, that they will preclude participation in the HWIR program.”

- The size, shape, and orientation of each physical sample will have a substantial effect on the analytical result. As such, the “maximum” obtained from a set of samples can not be estimated as reliably as the mean due to random variability introduced in the sampling and analytical processes.

3.1.4 Approach

It is not possible to demonstrate using statistical methods that 100 percent of any waste will be below an HWIR exemption level 100 percent of the time; however, we can use statistical tests to check if “most of the waste is below the exemption level most of the time.”

Under the “no exceedance” decision rule, it will be necessary for the generator to ensure that almost the entire distribution lies below the HWIR exemption level and that the mean is well below the exemption level (see Figure 1). There are two statistical approaches available to check if an upper proportion of the waste is less than the exemption level:

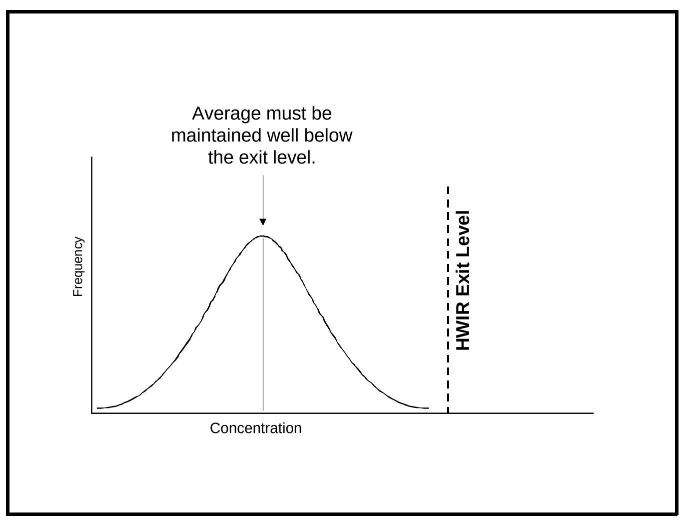


Figure 1. Under the “not to exceed” decision rule, the entire distribution must lie below the exemption level.

(1) exact and large sample non-parametric tests of proportions based on the binomial distribution, and

(2) a parametric test for percentiles based on an upper confidence limit on a percentile (equivalent to a tolerance limit), which assumes the data have a normal distribution.

The choice of the appropriate test is driven by the distributional characteristics of the data. If the data exhibit a normal distribution or can be transformed to a normal distribution (for example, by using a log transformation), then the parametric test can be used. Otherwise, the non-parametric approach should be used (such as when more than 50 percent of the data are reported as "non detect"). In fact, the non-parametric approach would be preferred where there are many constituents of interest each with a different variance and mean (such as will be encountered when making the HWIR claimant make the initial exit demonstration). Moreover, the assumption of normality may be somewhat optimistic with many data reported as non-detect. The revised version of Chapter Nine SW-846 contains extensive guidance and flowcharts regarding the choice of statistical tests when a "no exceedance" decision rule is used.

3.1.5 Number of Samples Required

This section describes two methods used to estimate the number of samples an HWIR claimant might require for an HWIR exit demonstration: a parametric approach that assumes the underlying population exhibits a normal distribution, and a non-parametric approach which does not make an assumption about the distribution of the population.

Parametric approach: If the assumption of normality is reasonable, the number of samples required to estimate an upper confidence limit on an upper percentile can be determined by calculating the k ("kappa") factor that corresponds to a number of samples listed in tables of k factors for estimating upper 100($1-a$)% confidence limit on the 99th percentile (see Table B-1 in Appendix B). The k factor associated with a particular sample size is determined by rearranging the equation for the upper confidence limit on an upper percentile (equivalent to a tolerance limit):

$$\frac{UL_{1-a}(x_p) - \bar{x}}{s} = k_{1-a,\,p}$$

where:

$k_{1-a,p}$	is the upper $1-a$ factor for the *P*th percentile with n sample measurements.
$UL_{1-a}(x_p)$	is the upper percentile of interest,
$\bar{x}$	estimated mean, and
s	estimated standard deviation.

By inputing assumed values for the mean ($\bar{x}$) and standard deviations (s), a k factor can be estimated. The k factor can then matched with a sample size (n) for the desired $1-a$. Note that the approach does not specify of a Type II error rate (b).

If we set the HWIR exemption level equal to unity (one), make assumptions about the relationship between the true mean and the exemption level and the relative standard deviation (or coefficient of variation, s/m), sample sizes can be estimated. An example of a sample size calculation calculate an upper confidence limit on an upper percentile is provided below.

Example Calculation of Number of Samples Required to Estimate an Upper Confidence Limit on an Upper Percentile

Assume the HWIR exit level for a given constituent is 1, the true (but unknown) mean is 0.25, and the standard deviation is 0.25 (i.e., the CV=1.0). What is the appropriate number of samples required to determine if a 90 percent upper confidence limit on the 99th percentile is less than or equal to the HWIR exemption level?

We can set EL equal to the upper limit (UL) on the 99th percentile and calculate the k factor as:

$$k = \frac{EL - m}{s} = \frac{1.0 - 0.25}{0.25} = 3.0$$

The k factor falls between two values in Table B-1 (appendix B) for $1 - a = 0.90$ (2.987 for n =23 and 3.007 for n =22). We find that about 22 samples are needed to estimate a 90% UCL on the 99th percentile that is less than or equal to the Exemption Level.

The results of this analysis are summarized in Tables 2, 3, and 4 below. The sample size estimates in the table are based on the assumption that certain properties of the underlying population (such as the mean and variance) are *known* in advance (or at least can be estimated from a pilot study, for example). In practice, these properties will not be known in advance and the actual number of samples required may differ from those given in the tables. Appendix C provides graphical representations of population distributions with the characteristics given in the tables.

Note that one of the limitations of the approach is that as the CV increases to about 1 or greater, the proportion of the distribution that extends below zero increases effectively producing distributions that are right-skewed. For these right-skewed distributions, the assumption of normality may no longer hold. If the assumption of normality cannot be made, sample sizes can be estimated using the non-parametric approach.

Table 2. Estimated Number of Samples Needed to Demonstrate That the Upper 99th Percentile Is Less Than or Equal to the HWIR Exemption level When the True Mean is 0.1 of the Exemption level

$(1-a)$	Coefficient of Variation				
	0.1 (see Figure C.1)	0.5 (see Figure C.2)	1.0 (see Figure C.3)	2.0 (see Figure C.4)	3.0 (see Figure C.5)
.90	4*	4*	4*	6	23
.95	4*	4*	4	8	35
.99	4*	4	5	13	70

*Indicates the calculated sample size is less than four.

Table 3. Estimated Number of Samples Needed to Demonstrate That the Upper 99th Percentile Is Less Than or Equal to the HWIR Exemption level When the True Mean is 0.25 of the Exemption level

$(1-a)$	Coefficient of Variation				
	0.1 (see Figure C.6)	0.5 (see Figure C.7)	1.0 (see Figure C.8)	2.0	3.0
.90	4*	4	23	**	**
.95	4*	5	35	**	**
.99	4*	8	70	**	**

* Indicates the calculated sample size is less than four.
** Indicates that the interval cannot be achieved at the selected confidence level with any number of samples.

Table 4. Estimated Number of Samples Needed to Demonstrate That the Upper 99th Percentile Is Less Than or Equal to the HWIR Exemption level When the True Mean is 0.50 of the Exemption level

$(1-a)$	Coefficient of Variation				
	0.1 (see Figure C.9)	0.5	1.0	2.0	3.0
.90	4*	**	**	**	**
.95	4	**	**	**	**
.99	5	**	**	**	**

* Indicates the calculated sample size is less than four.
** Indicates that the interval cannot be achieved at the selected confidence level with any number of samples.

The parametric approach presented above provides a statistical basis for minimum sample sizes that could be used to demonstrate compliance with the HWIR exemption level. The sample sizes should be adequate as long as the following assumptions can be made: (1) the data exhibit an approximately normal distribution, (2) the mean concentration lies considerably below the exemption level (e.g., 1/4 to 1/10 the exemption level), and (3) there is very little variability in the constituent concentrations (e.g., CV should be about 1 or less).

Non-Parametric approach: In the non-parametric setting, a balance must be struck between having a sufficiently high number of samples to conclude the waste qualifies for the exemption and having a sufficiently low number to minimize the chance of a Type II error (that is, we want to minimize the chance of obtaining a sample greater than the exemption level when in fact the true 99th percentile is less than or equal to the exemption level).

The non-parametric approach is simple. A set of samples is obtained and if one or more measurements exceeds the exemption level, then the waste does not qualify for the exemption. The approach does have statistical properties, as shown in Table 5.

Table 5. Estimated Number of Samples Needed When Using a "No Exceedance" Decision Rule* (all samples must be less than or equal to HWIR exemption level, otherwise the standard is not met).

Proportion (p) of waste that must be less than or equal to the HWIR exemption level	100 ($1-\alpha$)% Confidence (or the level of assurance we have that the waste meets the established decision rule).						
	.50	.60	.75	.80	.90	.95	.99
0.50	1	2	2	3	4	5	7
0.60	2	2	3	4	5	6	10
0.75	3	4	5	6	9	11	17
0.80	4	5	7	8	11	14	21
0.90	7	9	14	16	22	29	44
0.95	14	18	28	32	45	59	90
0.99	69	92	138	161	230	299	459

* The approach is based on a non-parametric (distribution-free) one-sided confidence bound on a percentile.

To use Table 5, select the level of confidence required and the percent of the waste expected to be less than or equal to the HWIR exemption level. For example, if we want to be 90 percent confident that 90 percent of all possible samples are less than or equal to the HWIR exemption level, we need 22 random samples and <u>all 22 samples must be less than or equal to the HWIR exemption level</u>. Otherwise, we can not conclude that 90 percent of all possible samples are less than the standard unless we take more samples or we are willing to accept a lower level of confidence.

As mentioned earlier, one drawback of the no exceedance rule is that with an increasing number of samples, there is a increasing chance that a rare but high concentration sample will be obtained. This can occur even if the true 99th percentile is exactly equal to the exemption level. In fact, if 22 samples are obtained from a distribution in which the true 99th percentile is equal to the exemption level, then there is about a 20 percent chance of obtaining an "unlucky" sample from the tail of the distribution (see further explanation in Appendix D).

3.1.6 Can Composite Sampling Be Used with the "No Exceedance" Rule?

Yes. Composite sampling is a strategy in which multiple individual or "grab" samples (from different locations or times) are physically combined and mixed into a single sample. Even though composite sampling results in a physical rather than a mathematical averaging of individual samples, it is still possible to estimate the maximum value that could be found in any individual sample used to form a composite. A detailed explanation of the technique is presented in Section 9.4.3.6 of the revised Chapter Nine SW-846, however, the basic approach is summarized below.

Let *EL* be the HWIR Exemption level that can not be exceeded. Note that *EL* must be large relative to the detection limit for the constituent of concern. For a measurement x_i from a composite sample formed from g individual samples, the following rules apply assuming analytical and sampling error are negligible:

- If $x_i < EL / g$, then no single individual sample can be > *EL*.
- If $x_i > EL$, then at least one *must*, and as many as all individual samples *may* be > *EL*.
- If $x_i > EL / g$, then at least one of the *g* individual samples must be > *EL*.

If all individual (i.e., grab) samples are found to be less than the Exemption level, the statistical performance can be determined from Table 5 or from the following relationship:

$$(1-a) = 1-P^n .$$

For example, if 9 composite samples are analyzed and each composite consists of 5 grab samples (for a total of 45 grab samples), and all grab samples are less than the exemption level, you can conclude with at least 90 percent confidence that at least 95 percent of the waste represented by the sampling is less than the exemption level:

$$(1-a) = 1-P^n = 1-0.95^{45} = 0.90$$

3.1.7 How is Subsequent Testing Performed After the Initial Exemption is Demonstrated?

The appropriate approach for subsequent testing after the initial demonstration depends upon how the null and alternative hypotheses are framed. For the initial exemption demonstration, the hypotheses are framed as follows:

Null: the upper 99^{th} percentile concentration of the constituent exceeds the HWIR exemption level.
Alt.: the upper 99^{th} percentile concentration in the waste is less than or equal to HWIR exemption level.

If a waste is highly variable, it is most “protective” to maintain the above hypothesis structure for each round of testing. However, for a relatively homogeneous waste stream, it may be appropriate to reverse the hypotheses. If the hypothesis structure is reversed, then the *lower* confidence bound on the upper percentile would be compared to the exit level. Formulas for calculating lower and upper confidence bounds on an upper percentile *both* involve adding a multiple of the standard deviation to the sample mean. The only difference is that a smaller “kappa” factor is used for the lower limit while a larger k is used for the upper limit.

One of the key questions regarding subsequent testing is: Can fewer samples be used for subsequent testing? If the original hypothesis structure remains unchanged, then the number of samples should continue to be determined based on the distributional characteristics of the waste concentration data and the desired level of confidence in the decision.

If the hypothesis structure is reversed, fewer samples would results in a wider confidence limit (assuming all other parameters remain the same) such that the lower confidence bound would get lower and decrease the chance of detecting a statistically significant increase above the exemption level. If the non-parametric bound on a percentile is used (see Table 5), then decreasing the sample size will result in either a deceased level of confidence associated with the decision, a decreased proportion of the waste that can be judged below the standard, or both.

In summary, we recommend that the number of samples for subsequent testing under a “no exceedance” rule be based on data quality objectives and the distributional characteristics of the underlying population (i.e., the estimated mean and variance).

3.2 Option No. 2: Using an Upper Confidence Limit on the Mean to Qualify a Waste for the HWIR Exemption

3.2.1 Description

Under this option, the mean must remain below the exemption level for the HWIR claimant to retain the exemption. Because the true mean is not known, it must be estimated from samples drawn from the waste and tested. Due to uncertainty associated with the information obtained from samples, a confidence interval on the mean is constructed and the upper confidence limit (UCL) can then be used to check compliance. If the UCL is less than or equal to the standard, the waste is judged in compliance with the standard. Otherwise, the opposite conclusion is reached.

3.2.2 Advantages

- The mean is usually used as the parameter of interest to compare to standards that protect again chronic long-term health effects (such a the HWIR exemption levels). For example, the UCL on the mean is used in Superfund to establish risk-based cleanup levels.

- Requires considerably fewer samples than the “not to exceed” decision rule.

- The UCL on the mean is consistent with the method generators use for the Toxicity Characteristics and Comparable Fuels.

- It is enforceable through the use of composite (physical average) samples.

- Calculation of the number of samples required is simple.

- Calculation of the confidence limit is simple.

- It is consistent with PBMS and SW-846 Chapter 9.

- Generators with highly variable waste close to the exemption level pay a “penalty” because a large number of samples would be required. On the other hand, generators with low variability waste that is well below the exemption level require much fewer samples to demonstrate compliance.

- Commenters on the 1995 proposed rule generally favored the UCL on the mean over the “not to exceed” approach (though most favored use of a rolling average).

3.2.3 Limitations

- Perceived complexity of statistical methods by some people.

- Number of samples needed to demonstrate compliance could be very large for generators with highly variable wastes with mean constituent concentrations close to their respective exemption levels (this also could be viewed as an advantage).

- If HWIR exemption levels replace certain LDR standards, then testing compliance becomes more complicated because the standards represent two different parameters: HWIR exemption levels protect against chronic (long-term average) risks, whereas LDR treatment standards represent concentrations achievable by the "best" treatment technologies.

3.2.4 Approach

A confidence interval on the mean is used to define the upper and lower limits that bracket the true mean with a known level of confidence. If the upper confidence limit (UCL) or confidence bound is less than the standard, then we can conclude the true value is below the standard with a level of confidence given by at least $1-a$ (see Figure 2).

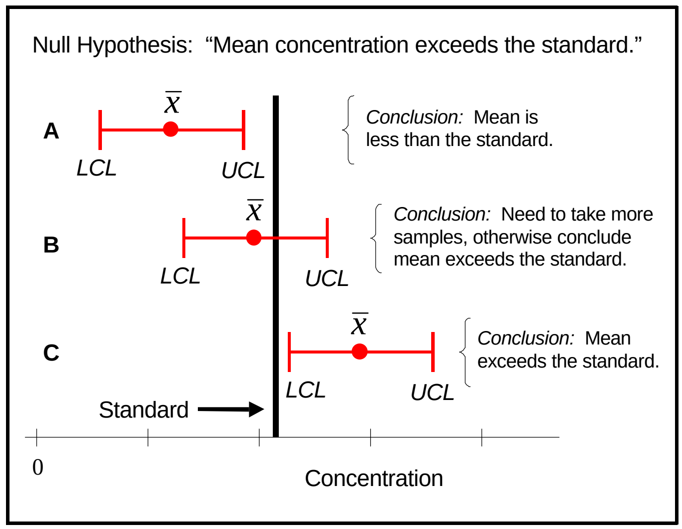


Figure 2. Using confidence limits on the mean to compare waste concentrations to a fixed standard.

3.2.5 Number of Samples Required

The number of samples required to construct a confidence interval for a normal distribution mean can be computed from a simple approximation formula such as:

$$n = \left[t_{1-a,n-1} s / d\right]^2$$

In the equation, $t_{1-a,n-1}$ is the standard normal distribution percentile for a confidence level of $1-a$. The length of the interval is denoted by *d* and is defined by the difference between the true (but unknown) mean and the upper bound of the "grey region" (set equal to the HWIR exemption level). The approach does not specify a Type II error rate (it is implicitly set at 50 percent).

The formula cannot be solved directly; rather an iterative method is required because the "*t*" value depends upon *n*, and *n* is not known. The TINV function (the inverse of the Student's t-distribution) in MS Excel was used to perform the necessary iterations.

Tables 6, 7, 8, and 9 present sample size estimates for several values of the coefficient of variation at various levels of confidence with the mean set at 0.1, 0.25, 0.5, and 0.75 of the exemption level, respectively.

Table 6. Estimated Number of Samples Needed to Demonstrate That the Mean Is Less Than or Equal to the HWIR Exemption Level When the True Mean is 0.1 of the Exemption level

$(1-a)$	Coefficient of Variation				
	0.1	0.5	1.0	2.0	3.0
.90	4*	4*	4*	4*	4*
.95	4*	4*	4*	4*	4*
.99	4*	4*	4*	4*	4*

* Indicates the calculated sample size is less than four.

Table 7. Estimated Number of Samples Needed to Demonstrate That the Mean Is Less Than or Equal to the HWIR Exemption Level When the True Mean is 0.25 of the Exemption level

$(1-a)$	Coefficient of Variation				
	0.1	0.5	1.0	2.0	3.0
.90	4*	4*	4*	4*	4
.95	4*	4*	4*	4	5
.99	4*	4*	4*	6	9

* Indicates the calculated sample size is less than four.

Table 8. Estimated Number of Samples Needed to Demonstrate That the Mean Is Less Than or Equal to the HWIR Exemption Level When the True Mean is 0.50 of the Exemption level

$(1-a)$	Coefficient of Variation				
	0.1	0.5	1.0	2.0	3.0
.90	4*	4*	4	9	17
.95	4*	4*	5	13	27
.99	4*	5	9	25	53

* Indicates the calculated sample size is less than four.

Table 9. Estimated Number of Samples Required to Demonstrate That the Mean Is Less Than or Equal to the HWIR Exemption Level When the True Mean is 0.75 of the Exemption level

$(1-a)$	Coefficient of Variation				
	0.1	0.5	1.0	2.0	3.0
.90	4*	6	17	61	135
.95	4*	9	27	100	222
.99	4	16	53	198	439

* Indicates the calculated sample size is less than four.

3.2.6 Can Composite Sampling Be Used to Calculate the UCL on the Mean?

Yes. Composite sampling can be used to obtain a more precise estimate of the mean. Confidence limits for the mean can be calculated from the analytical results using the same procedure that is used for non-composite sampling, except that n represents the number of composite samples and s represents the standard deviation of the n composite samples. There are a number of limitations to the use of composite sampling such as the loss of volatile constituents during the physical mixing of individual samples to form each composite.

3.2.7 How is Subsequent Testing Performed After the Initial Exemption is Demonstrated?

Once an HWIR exemption is granted, the exemption can be "maintained" through subsequent testing. In general, there are two approaches that can be used:

(1) ***Periodically repeat the sampling exercise that was used initially to qualify the waste for exemption***. This approach is best suited for highly variable wastes or wastes generated or treated in batches such that the characteristics from batch to batch are not known in advance. If this approach is used, then the number of samples required for subsequent testing would be about the same as that required for the initial exemption.

(2) ***Use a prediction limit constructed from the data generated from the qualifying study to determine if there is statistically significant evidence that the mean has changed.*** A *prediction limit* constructed with specified confidence on the initial sample set would allow a much smaller number of samples to be analyzed in subsequent testing than required for the initial exit demonstration. Basic prediction limits are no more difficult to compute than a basic confidence interval, but are designed to set an upper bound on the range of *individual* measurements one is likely to see in the next *k* sampling rounds. That is, if quarterly re-testing is to be conducted once the initial HWIR exemption is granted, so that 4 new samples will be available each year, a prediction limit could be constructed to set an upper bound on the range of individual concentrations one would expect to see over the next four sampling events.

Prediction intervals of this sort are commonly used in the RCRA groundwater monitoring program and are codified not only in CFR 264.97 but also explained in detail in existing USEPA guidance documents. The most significant advantage to using a prediction limit as opposed to computing another confidence interval around the mean is that much less data are required as well as the fact that a mean confidence interval tries to estimate the location of the true mean and not the likely range of individual measurements. The statistical hypothesis being tested by either technique, however, is the same. If any of the individual samples tested in subsequent testing exceeded the prediction limit, there would be statistically significant evidence that the mean level of the waste had changed and now exceeded the exemption level.

The prediction limit (*PL*) could be calculated using the following formula:

$$PL = EL + t_{n-1,1-a/k}\, s\sqrt{1+\frac{1}{n}}$$

where *EL* = HWIR exemption level, *s* = standard deviation of the initial phase sample measurements, n = number of initial phase sample measurements, and t = upper percentile from a Student's t-table based on the choice of confidence level and the number of future values, k, to be predicted. Of course, it is immediately apparent from this formula that the upper prediction limit so calculated will itself exceed the HWIR exemption level criteria. But this is again to account for possible variability in the waste stream and the fact that individual sample values may exceed the exemption level even though the true average does not. Note also that no additional data are required to compute the upper prediction limit. Once enough sampling information has been collected to compute the initial confidence interval used to establish the exemption, these same data can be used to compute the prediction limit.

In spite of the simplicity and low cost of the prediction limit method, there is one potential disadvantage to EPA: Because prediction limits would be specific to a particular waste at an individual facility, enforcement and compliance determinations would be more difficult for EPA. An out-of-compliance situation could only be determined after a review and analysis of historical waste analysis data; but this situation is not much different than that currently used by EPA in the RCRA ground-water monitoring program in which each individual facility has a site-specific ground-water monitoring system and monitoring plan developed based on knowledge of local hydrogeology, constituents of concern, and other factors.

The number of samples required to use prediction limits would depend upon the frequency of sampling; however, if quarterly sampling is used, as few as <u>one sample per quarter</u> or per a given volume of waste could be used.

3.3 Option No. 3: Using a Sample Mean that is Less than the Exemption level with the Maximum Sample Value Less than a Multiple of the Exemption level to Qualify a Waste for Exemption

3.3.1 Description

Under this approach, a waste is judged in compliance with the requirements of the HWIR exemption if two conditions are met: (1) the sample mean (the average of all samples collected to represent a given decision unit) must be less than or equal to the HWIR exemption level, AND (2) no single sample result can exceed a multiple of the exemption level.

Development of a Multiple of the Exemption Level:

For the purpose of this analysis, it was first necessary to establish a value above the mean that will serve as a maximum value not to be exceeded. Because the mean remains below the exemption level, there will be no increase in risk to human health and the environment if occasional and rare exceedances of the standard occur -- even if those exceedances include concentrations several times the exemption levels. The question then becomes: what is the maximum value above the exit level that can be tolerated as long as the mean remains below the exemption level?

First, an assumption must be made about the underlying distribution. If rare but high concentration exceedances are acceptable, a lognormal distribution is a reasonable default model to use because in a lognormal distribution a relatively small proportion of the distribution includes relatively large values.

Second, an assumption must be made about the range (i.e., the difference between the maximum and the minimum) of concentration values that might be found in a waste. For the purpose of this analysis, it is assumed that range of concentration values will fall within one order of magnitude with the "maximum" value set equal to the 99th percentile of the distribution (C_{99}).

Third, an assumption must be made about the shape of the lognormal distribution. The shape of the distribution is defined by s_y^2 (the variance of the logged values) and m_y (the mean of the logged values). A value for s_y^2 can be derived from an approximation of s_y. An approximation for s_y may be computed by dividing the range by 4.652 (based on the assumption that about 99 percent of all values in the distribution will fall within plus or minus 2.326 standard deviations of the mean), as follows:

$$s_y = \frac{\ln(\min) - \ln(\max)}{4.652} = \frac{\ln(10)}{4.652} = 0.495$$

The maximum allowed value above an HWIR exemption level can then be defined using the same approach EPA uses to develop "variability factors" for Land Disposal Restriction concentration based treatment standards (USEPA 1991). The variability factor (*VF*) is defined as the ratio of the 99th percentile (C_{99}) of the lognormal distribution to its arithmetic mean where:

$$C_{99} = \exp(m_y + 2.326 s_y)$$

$$Mean = \exp(m_y + 0.5 s_y^2)$$

By substitution, a multiple (*M*) of the HWIR exemption level that defines the maximum allowed concentration value can be defined in the same manner as the variability factor:

$$VF = M = \exp(2.326 s_y - 0.5 s_y^2)$$

Substitution of $s_y = 0.495$ (derived above) gives a multiple (*M*) of 2.8.

3.3.2 Advantages

- Ensures that the mean is always below the exemption level.
- Setting a maximum value above the exemption level limits the amount a variability that can be present in a qualifying waste.
- Easy to calculate.
- Easy to enforce because it provides an unambiguous standard (such as 2.8 X HWIR exemption level) that enforcement can use to demonstrate non-compliance.

3.3.3 Limitations

- Calculation of appropriate number of samples is more complicated than for Options 1 and 2.

- Because the decision rule sets a maximum value for any single sample, there is an increasing risk of an exceedance of the standard as more samples are collected due simply to random chance or due to sampling and analytical error.

- The physical size, shape, and orientation of each sample will determine the “maximum” analytical result (this is also an issue of concern for Option No. 1).

3.3.4 Approach

The basic approach requires application of two decision rules: (1) the sample mean must not exceed the applicable HWIR exemption level, AND (2) no single sample value may exceed 2.8 times the applicable exemption level. While statistical method can be used to estimate sample sizes, compliance determinations can be made simply by calculating the sample mean and the maximum.

3.3.5 Number of Samples Required

The decision rule requires that the mean be maintained at or below the exemption level with the maximum value in any sample no greater than the 2.8 times the exemption level.

If we set the mean and the exemption level equal to one, assume that the maximum sample is equal to 2.8 times the exemption level, and assume the UCL on the 99th percentile is a reasonable approximation of the maximum in a set of samples, we can calculate sample sizes using a similar approach to that used above in Section 3.1. The key to solving the sample size problem is to establish a set of distribution models with the following parameters:

- The mean must less than or equal to 1;

- The upper 99th percentile must be equal to or less than 2.8; and

- No values may be less than zero.

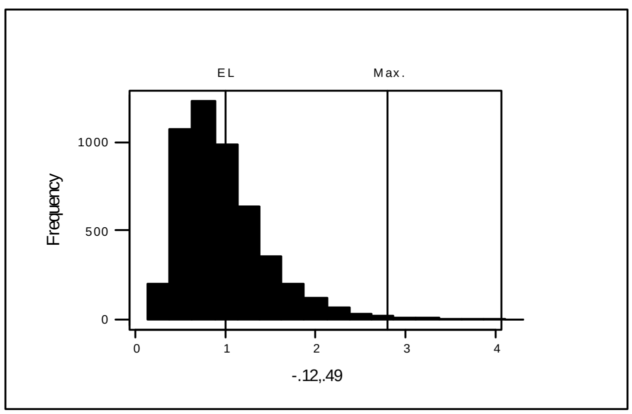


Figure 3. Example of a lognormal distribution with a true mean equal to the HWIR exemption level (1.0) and the 99th percentile equal to 2.8 times the exemption level.

Figure 3 is an example of such a distribution. The figure shows a lognormal distribution with an untransformed true mean ($\hat{m}$) of 1.0 and the 99th percentile equal to 2.8. A key feature of the distribution is that it is skewed to the right, with most of the values falling below the exemption level. Of course, a normal distribution also could be fit using the criteria noted above, but the CV would need to be relatively small (e.g., about 0.3 or less) to minimize the proportion of the distribution the would fall below zero (see Figure E-7 in Appendix E).

Sample size estimates can be made using the parametric procedures described in Section 3.1, with some modifications. Because the assumption of lognormality is reasonable, the number of samples required to estimate an upper confidence limit on an upper percentile can be determined by backsolving the following equation for k :

$$UL_{1-a}(x_p) = \exp\left[\bar{y} + s_y k_{1-a,p}\right]$$

The k factor is determined by rearranging the above as follows:

$$\frac{\ln(UL)\ -\ \bar{y}}{s_y} = k_{1-a\,,\,p}$$

where:

$k_{1-a,p}$	is the upper $1-a$ factor for the *P*th percentile with n sample measurements.
UL	represents a value not to be exceeded (we'll assume it is equal to the 99th percentile),
$\bar{y}$	arithmetic mean of the log-transformed values, and
s_y	standard deviation of the log-transformed values.

By inputing assumed values for the mean ($\bar{y}$) and standard deviations (s_y), a k factor can be estimated. The k factor can then matched with a sample size (n) for the desired $1-a$. Note that the approach does not specify of a Type II error rate (b).

Table 10 below provides sample sizes required to estimate the UCL on the 99th percentile for various lognormal distributions. For all distributions, the mean is less than the exemption level and the 99th percentile is less than or equal to 2.8 times the exemption level (see histograms in Appendix E).

Table 10. Number of Samples Required When the Decision Rule Requires the Mean to Be Less than the Exemption level with the Maximum Sample Value Less than 2.8 Times the Exemption Level

$1-a$	Parameters of Lognormal Distributions						Normal Distribution
	m_y = -0.1 CV$_y$ = 0.10 (Fig. E-1)	m_y = -0.3 CV$_y$ = 0.30 (Fig. E-2)	m_y = -0.5 CV$_y$ = 0.40 (Fig. E-3)	m_y = -0.8 CV$_y$ = 0.53 (Fig. E-4)	m_y = -0.50 CV$_y$ = 0.59 (Fig. E-5)	m_y = -0.60 CV$_y$ = 0.66 (Fig. E-6)	m = 1.0 CV = 0.33 (Fig. E-7)
0.90	4*	6	7	9	40	55	4
0.95	4*	8	12	13	60	87	6
0.99	4	13	18	24	120	184	9
where: $CV_y = \sqrt{\exp\left(s_y^2\right)-1}$							$CV = \frac{s}{m}$

*Indicates that the calculated sample size is less than four.

Non-Parametric Approach. A non-parametric approach can be used to estimate sample sizes. The approach is the same as that presented in Section 3.1.5 (see Table 5). In addition to no single sample exceeding a multiple of the exemption level, the mean must remain below the exemption level. Accordingly, the HWIR claimant must calculate the mean as well as compare each value to the multiple of the exemption level. If either parameter is exceeded, the waste does not qualify for the exemption.

3.3.6 Can Composite Sampling Be Used?

Yes. Composite sampling can be used. The information obtained from composite samples must be used in two ways to check compliance with the HWIR exemption levels: (1) to estimate the mean, and (2) to estimate the maximum. Composite samples provide and efficient method for estimating the mean (the average of the composite samples provides a point estimate of the mean), and the technique described above in Section 3.1.6 can be used to determine if any individual sample exceeds the maximum value allowed.

3.3.7 How is Subsequent Testing Performed After the Initial Exemption is Demonstrated?

Once an HWIR exemption is granted, the exemption can be "maintained" through subsequent testing. Similar to Option 2, one of two approaches might be used:

(1) ***Periodically repeat the sampling exercise that was used initially to qualify the waste for exemption***. This approach is best suited for highly variable wastes or wastes generated or treated in batches such that the characteristics from batch to batch are not known in advance. If this approach is used, then the number of samples required for subsequent testing would be about the same as that required for the initial exemption. The decision rule is simple: the sample mean must be less than or equal to the exemption level AND no single sample value may exceed 2.8 time the applicable exemption level.

(2) ***Use a prediction limit constructed from the data generated from the qualifying study to determine if there is statistically significant evidence that the mean has changed with the added constraint that no single value may exceed 2.8 times the exemption level.*** As with Option 2, subsequent testing could include the use of a prediction limit on the mean with the additional constraint that no single sample value can exceed 2.8 times the applicable exemption level. As noted in Section 3.2.7, the primary advantage of prediction limits as opposed to calculating a new mean after each round of sampling is that much less data are required.

4. Summary and Recommendations

This paper presents estimates of the number of samples a waste generator would need to obtain to demonstrate compliance with applicable HWIR exemption levels under each of three decision rule options under consideration by EPA. Regardless of the decision rule ultimately selected, it is recommended that the generator determine the appropriate number of samples using the data quality objective process or a similar planning process to define:

- the parameter of interest and the decision rule;
- the desired degree of precision;
- the level of confidence required in the estimate; and
- the statistical distribution, if any, to be assumed including estimates of parameters such as the

mean and standard deviation.

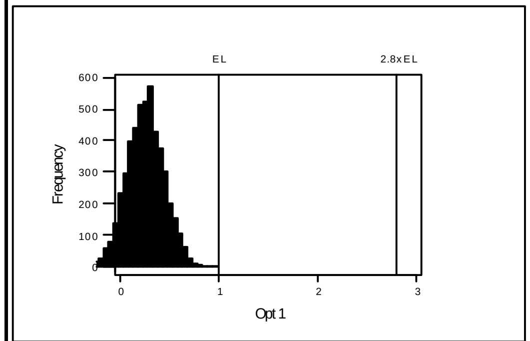
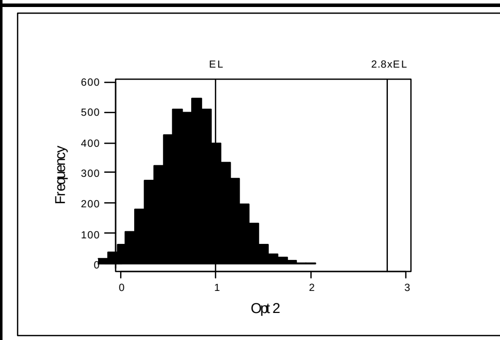
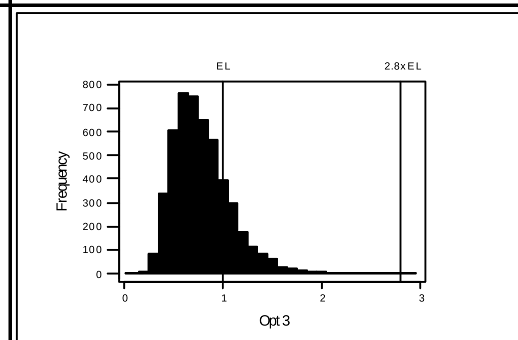
To derive estimates of samples sizes required under each of the three decision rule options, it was necessary to establish baseline data quality objectives (DQOs) to specify the parameters of interest, the decision rules, and desired confidence levels. A range of distributional models was created and sample sizes were estimated for each distributional assumption.

Each decision rule option uses a different parameter or combination of parameters to measure attainment of the HWIR exemption levels. As such, there is not a simple way to compare “apples to apples” for the purpose of comparing sample size requirements. However, given a relatively small number of samples (e.g., *between six and nine*) and a specified level of confidence, it is possible to make general statement about the distributional characteristics required for a generator’s waste to “pass” the exemption.

Table 11 lists the range of sample sizes that might be required to make an exemption demonstration. The table provides a narrative description of distribution models required for a waste to pass the initial exemption. The last column of the table includes histograms of large samples drawn from a population expected to pass the HWIR exemption. An infinite variety of such samples and histograms is possible, and those shown in the table are for illustrative purposes only.

The sample size estimates are only valid under the assumed distributional characteristics specified in the table coupled with other assumptions stated in this report. (For example, the number of samples would only be valid for making a decision about particular mass of waste that is sampled -- the DQO “decision unit”.) In practice, the actual number of samples required may be lower or significantly higher. In fact, some generators will find that to attain the desired degree of precision, the number of samples required will be so high as to make the exemption demonstration impractical. This will occur in situations where a constituent concentration is either too close to the exemption level, the concentrations are highly variable, or both.

Table 11. Summary Table of Number of Samples Needed to Demonstrate That A Waste Qualifies for an HWIR Exemption.

Decision Rule Option	**Number of Samples**	**Confidence Level**	**Distribution Model and Parameters Required to Achieve Compliance With The Exemption**	**Example Histograms** (*n* = 5000 drawn from a population expected to pass an HWIR exemption)
Option 1: "No Exceedance"	4 to >23	90%	• True mean must be approximately 1/4 to 1/10 of the exemption level. • Very little variability can be tolerated (e.g., the CV must be approximately 1 or less).	 Mean = 0.25, SD = 0.125. About 7 samples needed to demonstrate 99th percentile ≤ EL with 1-α=0.90.
	4 to >35	95%		
	4 to >70	99%		
Option 2: UCL on Mean Less Than Exemption Level	4 to >135	90%	• The mean can be higher than in Option 1 (but also must be less than the exemption level) • High variability and skewness can be tolerated, though a larger number of samples would be required for a highly variably waste with a mean close to the exemption level.	 Mean = 0.75, SD = 0.375. About 6 samples needed to demonstrate UCL on mean ≤ EL at 1-α=0.90.
	4 to >222	95%		
	4 to >439	99%		
Option 3: Sample Mean Less Than or Equal to Exemption Level, and No Sample Greater Than 2.8 x EL.	4 to >55	90%	• The sample mean can be higher than allowed under Options 1 and 2 (The mean can be equal to the exemption level). • Moderate positive skewness can be tolerated.	 Mean = 0.80, GSD = 1.43. 9 samples needed to demonstrate 99th perc. ≤ 2.8 x EL with 1-α=0.90.
	4 to >87	95%		
	4 to >184	99%		

Appendix A

Implementation Language from the November 1995 Proposed Rule

Data Evaluation

i. Compliance with the Exemption levels

The Agency is requesting comment on three approaches of data evaluation.

First, the Agency is proposing that, for exemptions under today's proposed rule, generators would be required to evaluate their waste based on the maximum detected concentrations of the exemption constituents. If any constituent concentration is greater than its specified exemption level, then the waste would be ineligible for exemption under today's proposed rule. One advantage of this approach is that facilities can use process and waste knowledge to determine the appropriate number of representative samples without relying on a complex, potentially costly statistical approach to determine an appropriate number of samples. However, generators will need to be sufficiently knowledgeable about their waste and process to make an unbiased determination regarding the appropriate number of samples. Actual sample representativeness might be difficult to verify or otherwise assess (on a statistical basis). Finally, the level of uncertainty associated with the results cannot be defined. Because of this, under this approach, a single composite sample that validly exceeds the HWIR exemption levels would indicate that the waste is hazardous and must be handled in Subtitle C.

Second, the Agency requests comment on also allowing a second data evaluation method whereby the analytical results are evaluated in terms of an upper confidence limit around an average concentration. An example of one method for determining an upper confidence limit is presented in the statistical approach found in Chapter Nine of SW-846 (Third Edition, as amended by Updates I, II, IIA, and IIB), where, for the purpose of evaluating solid wastes, the probability level (confidence interval) of 80 percent is used. Sample measurements for which the upper limit of the 80 percent confidence interval about the sample mean is below the regulatory level for the chemical contaminant are not considered to be present at levels of regulatory concern. One main advantage of this approach is that the number of samples is statistically determined and thus it eliminates any bias that might otherwise be introduced when using knowledge to determine the appropriate number of samples. In addition, the level of uncertainty associated with the results can be determined. However, the main disadvantage of this approach is that it could be more costly for some facilities than the proposed approach. For example, it might require multiple rounds of sampling to determine the mean and variance. Highly variable wastes may require the collection of many more additional samples than might otherwise be determined to be necessary using the first approach. However, this statistical approach allows occasional samples to be above exemption level, as long as the upper confidence limit of the data overall is below the exemption level.

The Agency also requests comment on a third data evaluation method that would allow facilities to use long-term average data to demonstrate compliance without consideration of the upper confidence limit. A rolling average of samples would be taken over the course of a year on a schedule determined by the initial sampling and analysis plan. As long as the average of the samples was below the HWIR exemption level, the waste stream would be considered non-hazardous. This approach would have the advantage of being simpler than the second option, while allowing occasional exceedences of the exemption levels by single samples, as long as the average concentration is below exemption levels.

EPA has modelled risk with the assumption that the constituents of concern are uniformly distributed within the waste at the exit concentrations. In discussion with the Hazardous Waste

Identification Dialogue Group, some representatives noted that actual levels might need to average significantly below the exemption levels if the exit criteria are to be consistently met. The second and third data evaluation methods discussed above help address this issue.

However, EPA and the States have noted that the only practical approach for enforcement purposes is to independently collect samples for analysis (which may represent a composite of materials spatially or over a short time span) and to set up the regulation so that an exceedence by any single composite sample during an inspection could constitute a violation. It would then be the responsibility of the generator to refute this, using historic sampling data and possibly additional samples to show that the sample exceedence does not constitute an overall violation of the HWIR levels.

EPA believes it is important to retain the practical approach whereby a single composite sample of a waste at some arbitrary point in time or space during a short visit is considered sufficient for enforcement purposes. However, because the exit numbers were modeled based on long-term average concentrations, the Agency requests comment on allowing occasional exceedences as long as the average concentration meets the exemption level.

In addition to the concern about enforceability, however, EPA has identified two additional concerns about using average concentration to determine compliance. First, not all waste streams would be disposed of in the same place. Thus the wastes may on average be in compliance when they are generated, but the wastes arriving at the disposal site (possibly from multiple sources) may not be, on average, below the exemption levels. Second, EPA has not modeled the constituents for acute risk. While the average concentration of constituents may be below the exemption levels, the occasional "high" concentration may be of concern due to acute health or ecological effects.

One possible way to address some of these concerns is, in addition to requiring that the average meet the exemption levels (as in the second and third data evaluation methods), EPA could require that all samples be below some "peak" concentration.

Under this approach, if the average concentrations are below the exemption levels, and all individual samples are below the higher peak level, then the generator would be in compliance and need take no further action to support the exemption. EPA or a State would then be able to confirm waste status without total reliance on the generator's data and without the expense of periodic sampling by EPA or the State. EPA requests comment on this issue, including any information on setting peak levels.

For any of the three data evaluation approaches, representative samples must be collected in support of exemption under today's proposed rule, consisting of a sufficient number of samples to represent the spatial and temporal variability of the waste characteristics, regardless of how the sample number is determined.

For the identification and handling of "outliers", the Agency is recommending that testing for outliers should be done if an observation seems particularly high or low compared to the rest of the data set. If an outlier is identified, the result should not be treated as such until a specific reason for the abnormal measurement can be determined (e.g. contaminated sampling equipment, laboratory contamination, data transcription error). If a specific reason is documented, the result should be excluded from further data evaluation. If a plausible reason cannot be found, the observations should be treated as a true, albeit extreme, value and not excluded from the data evaluation, as waste composition can vary. The Agency solicits comments on implementable techniques for the identification of analytical outliers.

The results of the tests of all of the constituents on the exemption list would be required to show the constituent concentration to be at or below the exemption level in order for the claimant to be eligible for an exemption. In the case where a constituent's exemption level is based on the quantitation criteria (EQC, as described in section IV.E.), in addition to showing a non-detect at the exemption level, the waste would be required to meet applicable requirements set forth at 40 CFR 268. Certain facilities may have difficulty quantifying a constituent at the exemption level due to matrix interference effects, but the Agency expects exempted wastes to have relatively clean matrices such that exemption levels should be able to be achieved. The Agency believes that the exemption level must be met in order for a waste to exit Subtitle C; therefore, waste streams that cannot meet exemption levels would not exit under today's rule. The Agency asks for comment on this approach.

Appendix B

Table B-1. Factors $K_{1-a,0.99}$ for Estimating an Upper $100(1-a)\%$ Confidence Limit on the 99th Percentile of a Normal Distribution

n	$1-a$		
	0.90	**0.95**	**0.99**
2	18.500	37.094	185.62
3	7.340	10.553	23.896
4	5.438	7.042	12.387
5	4.668	5.749	8.939
6	4.243	5.065	7.335
7	3.972	4.643	6.412
8	3.783	4.355	5.812
9	3.641	4.144	5.389
10	3.532	3.981	5.074
11	3.444	3.852	4.829
12	3.371	3.747	4.633
13	3.310	3.659	4.472
14	3.257	3.585	4.337
15	3.212	3.520	4.222
16	3.172	3.463	4.123
17	3.137	3.414	4.037
18	3.106	3.370	3.960
19	3.078	3.331	3.892
20	3.052	3.295	3.832
21	3.028	3.262	3.777
22	3.007	3.233	3.727
23	2.987	3.206	3.681
24	2.969	3.181	3.640
25	2.952	3.158	3.601
30	2.884	3.064	3.447
35	2.833	2.994	3.334
40	2.793	2.941	3.245
45	2.762	2.897	3.181
50	2.735	2.863	3.125
60	2.694	2.807	3.038
70	2.663	2.766	2.974*
80	2.638	2.733	2.924*
90	2.618	2.706	2.883*
100	2.601	2.684	2.850*
120	2.574	2.649	2.797
∞	2.326	2.326	2.326

(Compiled from Gilbert 1987 (Table A3), Hahn and Meeker 1991 (Table A.12d), and Natrella 1966 (Table A-7). * calculated by K. Cameron.)

Appendix C

Normal distributions with various CVs and Means relative to an HWIR exemption level.

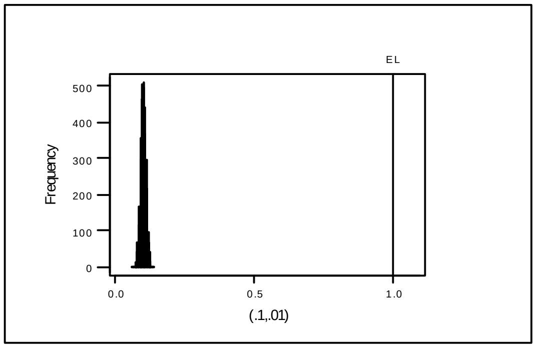


Figure C-1. Random sample, n=5000, drawn from a population with true mean 0.1 of the Exemption level (exemption level= 1.0) and CV=0.1.

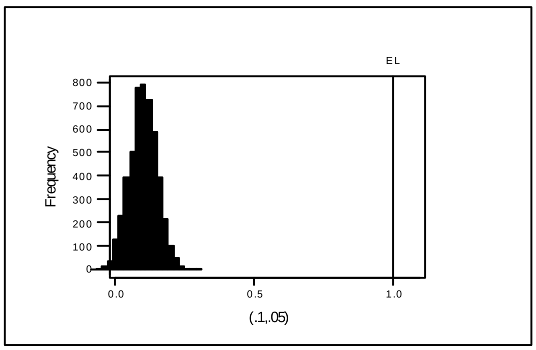


Figure C-2. Random sample, n=5000, drawn from a population with true mean 0.1 of the Exemption level (exemption level= 1.0) and CV=0.5.

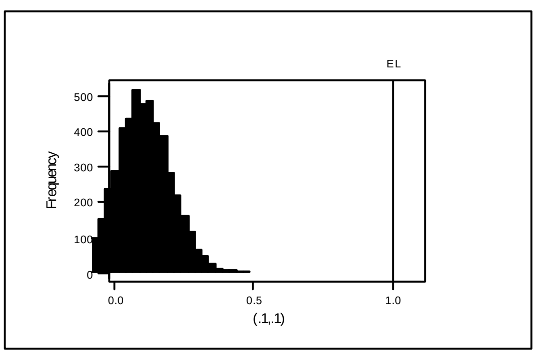


Figure C-3. Random sample, n=5000, drawn from a population with true mean 0.1 of the Exemption Level (exemption level= 1.0) and CV=1.0.

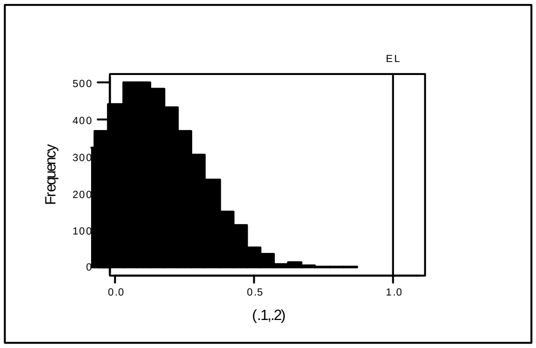


Figure C-4. Random sample, n=5000, drawn from a population with true mean 0.1 of the Exemption level (exemption level= 1.0) and CV=2.0.

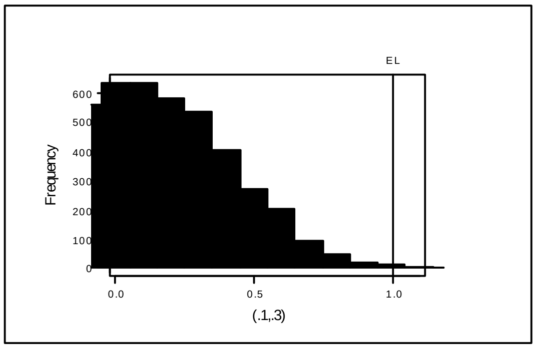


Figure C-5. Random sample, n=5000, drawn from a population with true mean 0.1 of the Exemption Level (exemption level= 1.0) and CV=3.0.

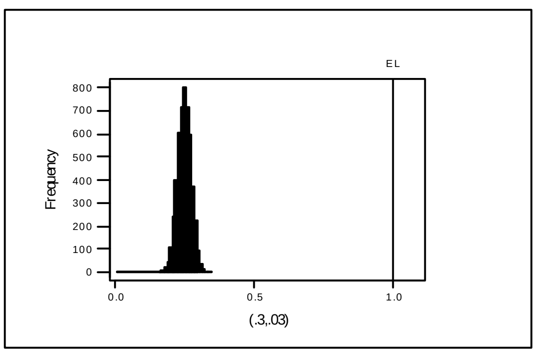


Figure C-6. Random sample, n=5000, drawn from a population with true mean 0.25 of the Exemption Level (exemption level= 1.0) and CV=0.1.

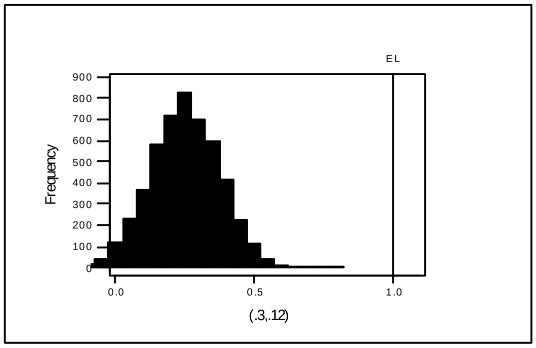


Figure C-7. Random sample, n=5000, drawn from a population with true mean 0.25 of the Exemption level (exemption level= 1.0) and CV=0.5.

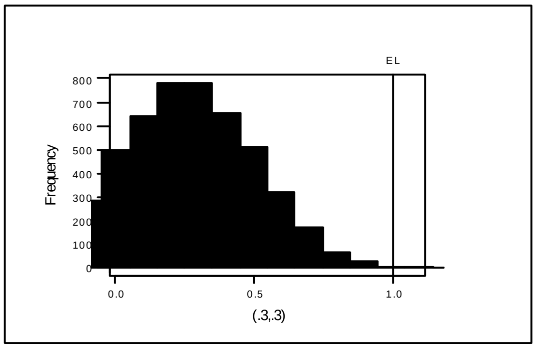


Figure C-8. Random sample, n=5000, drawn from a population with true mean 0.25 of the Exemption level (exemption level= 1.0) and CV=1.0.

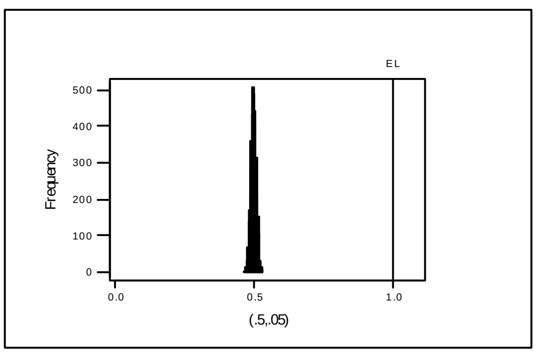


Figure C-9. Random sample, n=5000, drawn from a population with true mean 0.5 of the Exemption level (exemption level= 1.0) and CV=0.1.

Appendix D

A Limitation of the "No Exceedance" Rule for Granting an HWIR Exemption

Under the "no exceedance" decision rule, no single sample can exceed the HWIR exemption level if the waste is to retain an HWIR exemption. However, the more samples obtained, the greater the probability that an exceedance will be found.

This decision rule implies that the entire distribution must lie below the exemption level **(see Figure D-1**). It also implies that the maximum value in any single sample must be less than or equal to the exemption level.

Statistically, we cannot determine the maximum value, however, we can use an upper percentile (such as the 99th percentile) as a reasonable approximation of the maximum.

Suppose we have a waste population in which the true (but unknown) 99th percentile is equal to the HWIR exemption level. We wish to use a decision rule to determine the waste should be exempted under HWIR.

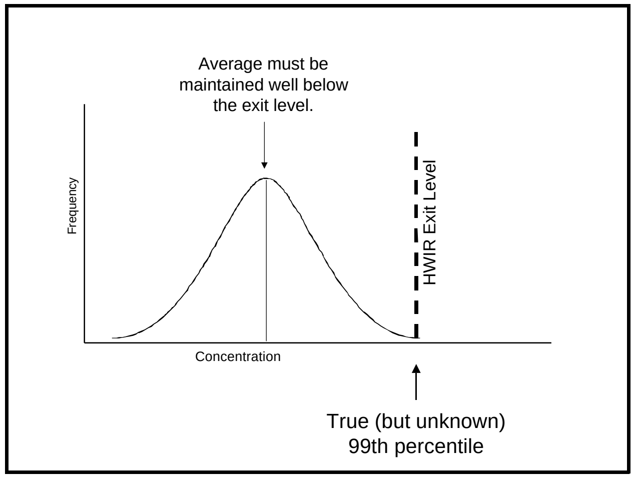


Figure D-1. If no samples can exceed the exemption level, then the entire distribution must lie below the exemption level.

Suppose n representative samples are obtained and the data are used to estimate the 99th percentile of the underlying distribution. We denote this estimate as $X_{0.99}$. Suppose our estimate is in fact equal to the true 99th percentile. Then by the definition of a percentile, the probability (q) is 0.99 (or 99 chances out of 100) that a single newly obtained representative sample drawn from the waste will be less than or equal to the exemption level.

If a decision rule is used that allows for zero exceedances of the exemption level, what is the probability that at least one sample out of n samples will exceed the exemption level even though the true 99th percentile is less than or equal to the exemption level?

The probability that at least one of the n measurements will be greater than the exemption level is $1-q^n$. The probability of a waste not qualifying for the exemption under the "no exceedance" decision rule for various numbers of samples is given as:

n	$1-0.99^n$ (or probability of not qualifying for the exemption)
1	0.01
4	0.04
8	0.08
12	0.11
22	0.20
44	0.36
70	0.50
100	0.63

Conclusion:

The more samples you take, the greater the chance that you will conclude the waste does not qualify for an HWIR exemption, even when the true 99th percentile is equal to the exemption level.

For example, if a waste generator takes 22 samples from a waste in which the true (but unknown) 99th percentile is equal to the exemption level, there is about a 20 percent chance of concluding the waste does not meet the exemption level when in fact it does.

Furthermore, the problem is compounded over time with subsequent testing. For example, if only 4 samples are obtained for the initial exemption, and 4 more samples are obtained each quarter for the next five years (for a total of 80 samples), there is a 55 percent chance of losing the exemption even though the waste characteristics have not changed!

The “no exceedance” rule works well from an enforcement perspective, however, it may be a major limitation for the implementability of the rule: eventually, a generator’s waste will be disqualified and returned to the RCRA Subtitle C jurisdiction. One possible solution is to reverse the null hypothesis one the waste qualifies for the exemption, and then require that the lower confidence limit on the upper percentile be used to check non-compliance.

Appendix E

Lognormal distribution histograms for various values of m_y and s_y.

"EL" = HWIR exemption level.
"Max" = maximum value allowed for any single grab sample under decision rule Option No. 3.

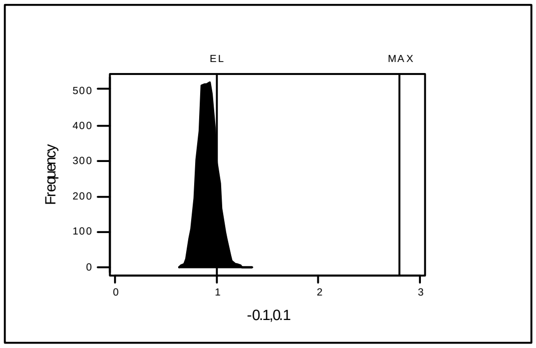


Figure E-1. Lognormal distribution with true mean =0.9 of the exemption level. GSD=1.11.

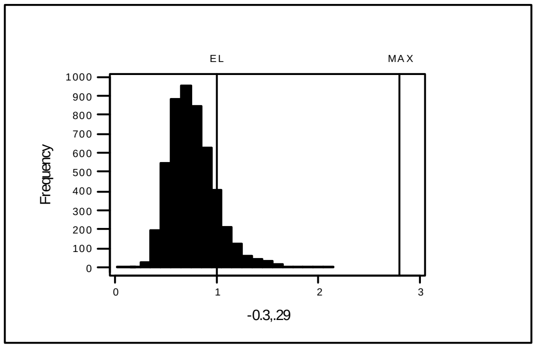


Figure E-2. Lognormal distribution with true mean =0.77 of the exemption level. GSD=1.34.

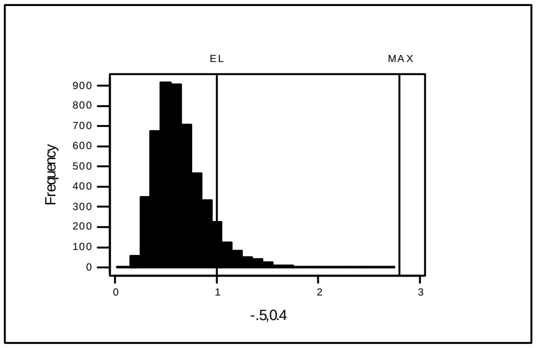


Figure E-3. Lognormal distribution with true mean =0.65 of the exemption level. GSD=1.47.

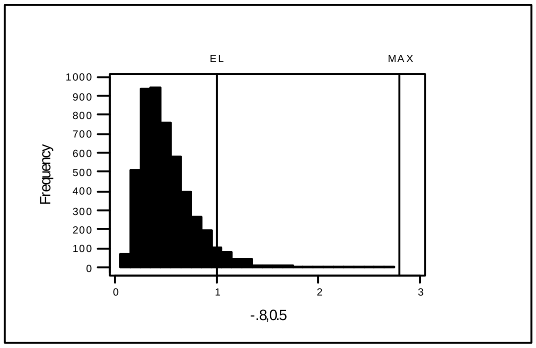


Figure E-4. Lognormal distribution with true mean =0.51 of the exemption level. GSD=1.65.

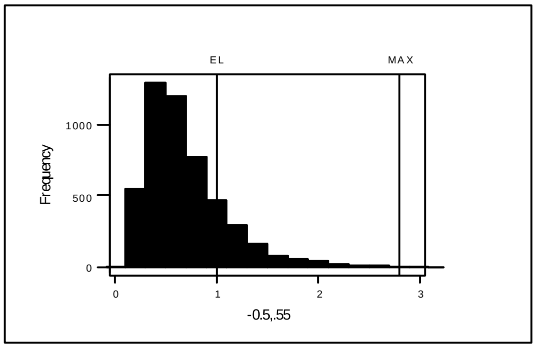


Figure E-5. Lognormal distribution with true mean =0.70 of the exemption level. GSD=1.73.

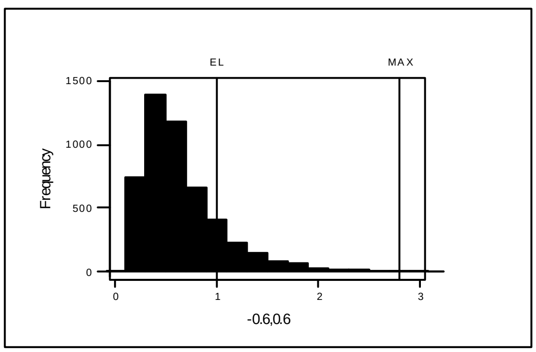


Figure E-6. Lognormal distribution with true mean =0.65 of the exemption level. GSD=1.82.

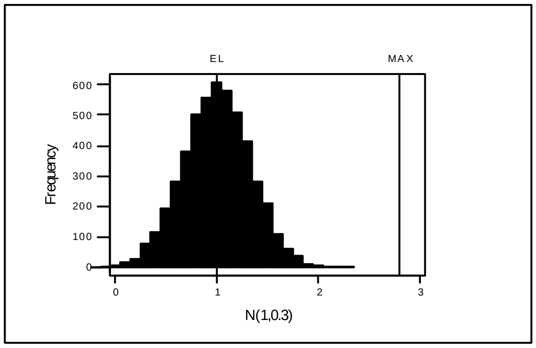


Figure E-7. Normal distribution with true mean =1.0 of the exemption level with a standard deviation of 0.333.

References

Gilbert, R.O. 1987. Statistical Methods for Environmental Pollution Monitoring. Van Nostrand Reinhold (now John Wiley), New York. 320 pp.

Hahn, G.J. and W.Q. Meeker. 1991. Statistical Intervals: A Guide for Practitioners. John Wiley & Sons, Inc. 392 p.

Natrella, M.G. 1966. Experimental Statistics, National Bureau of Standards Handbook 91, United States Department of Commerce, Washington DC.

USEPA. 1991c. Final Best Demonstrated Available Technology (BDAT) Background Document For Quality Assurance/Quality Control Procedures and Methodology (NTIS PB95-230926). Office of Solid Waste. October 23.

USEPA. 1994. Guidance for the Data Quality Objectives Process, EPA QA/G-4 <http://es.epa.gov/ncerqa/qa/qa_docs.html#G-4> (EPA/600/R-96/055), Office of Research and Development, Washington, DC 20460, September 1994.