OSWER 9285.6-10

December 2002

CALCULATING UPPER CONFIDENCE LIMITS FOR EXPOSURE POINT CONCENTRATIONS AT HAZARDOUS WASTE SITES

Office of Emergency and Remedial Response
U.S. Environmental Protection Agency
Washington, D.C. 20460

ADMIN RECORD

SW-A-005687

Disclaimer

This document provides guidance to EPA Regions concerning how the Agency intends to exercise its discretion in implementing one aspect of the CERCLA remedy selection process. The guidance is designed to implement national policy on these issues.

The statutory provisions and EPA regulations described in this document contain legally binding requirements. However, this document does not substitute for those provisions or regulations, nor is it a regulation itself. Thus, it cannot impose legally-binding requirements on EPA, States, or the regulated community, and may not apply to a particular situation based upon the circumstances. Any decisions regarding a particular remedy selection decision will be made based on the statute and regulations, and EPA decisionmakers retain the discretion to adopt approaches on a case-by-case basis that differ from this guidance where appropriate. EPA may change this guidance in the future.

TABLE OF CONTENTS

1.0 INTRODUCTION 1

2.0 APPLICABILITY OF THIS GUIDANCE 2

3.0 DATA EVALUATION 2

3.1 Outliers 3
3.2 Non-detects 4

4.0 UCL CALCULATION METHODS 6

4.1 UCL Calculation with Methods for Specific Distributions 8

UCLs for Normal Distributions 8
UCLs for Lognormal Distributions 10

Land Method 10
Chebyshev Inequality Method 11

UCLs for Other Specific Distribution Types 14

4.2 UCL Calculation with Nonparametric or Distribution-Free Methods 14

Central Limit Theorem (Adjusted) 15
Bootstrap Resampling 16
Jackknife Procedure 18
Chebyshev Inequality Method 18

5.0 OPTIONAL USE OF MAXIMUM OBSERVED CONCENTRATION. 20

6.0 UCLs AND THE RISK ASSESSMENT 20

7.0 PROBABILISTIC RISK ASSESSMENT 22

8.0 CLEANUP GOALS 22

9.0 REFERENCES 23

APPENDIX A: USING BOUNDING METHODS TO ACCOUNT FOR NON-DETECTS 26

APPENDIX B: COMPUTER CODE FOR COMPUTING A UCL WITH THE BOOTSTRAP SAMPLING METHOD 28

1.0 INTRODUCTION

This document updates a 1992 guidance originally developed to supplement EPA's *Risk Assessment Guidance for Superfund (RAGS), Volume 1 – Human Health Evaluation Manual* (RAGS/HHEM, EPA 1989), which describes a general approach for estimating exposure of individuals to chemicals of potential concern at hazardous waste sites. It addresses a key element of the risk assessment process for hazardous waste sites: estimation of the concentration of a chemical in the environment. This concentration, commonly termed the exposure point concentration (EPC), is a conservative estimate of the average chemical concentration in an environmental medium. The EPC is determined for each individual exposure unit within a site. An exposure unit is the area throughout which a receptor moves and encounters an environmental medium for the duration of the exposure. Unless there is site-specific evidence to the contrary, an individual receptor is assumed to be equally exposed to media within all portions of the exposure unit over the time frame of the risk assessment.

EPA recommends using the average concentration to represent "a reasonable estimate of the concentration likely to be contacted over time" (EPA 1989). The guidance previously issued by EPA in 1992, *Supplemental Guidance to RAGS: Calculating the Concentration Term* (EPA 1992), states that, "because of the uncertainty associated with estimating the true average concentration at a site, the 95 percent upper confidence limit (UCL) of the arithmetic mean should be used for this variable." The 1992 guidance addresses two kinds of data distributions: normal and lognormal. For normal data, EPA recommends an upper confidence limit (UCL) on the mean based on the Student's *t*-statistic. For lognormal data, EPA recommends the Land method using the *H*-statistic. EPA describes approaches for testing distribution assumptions in *Guidance for Data Quality Assessment: Practical Methods for Data Analysis* (EPA 2000b, section 4.2).

The 1992 guidance has been helpful for EPC calculation, but it does not address data distributions that are neither normal nor lognormal. Moreover, as has been widely acknowledged, the Land method can sometimes produce extremely high values for the UCL when the data exhibit high variance and the sample size is small (Singh et al. 1997; Schulz and Griffin 1999). EPA's 1992 guidance recognizes the problem of extremely high UCLs, and recommends that the maximum detected concentration become the default when the calculated UCL exceeds this value. Singh et al. (1997) and Schulz and Griffin (1999) suggest several alternate methods for calculating a UCL for non-normal data distributions. This guidance provides additional tools that risk assessors can use for UCL calculation, and assists in applying these methods at hazardous waste sites. It begins with a discussion of issues related to evaluating the available site data and then presents brief discussions of alternative methods for UCL calculation, with recommendations for their use at hazardous waste sites. In addition, EPA has worked with its contractor, Lockheed Martin to develop a software package, ProUCL, to perform many of the calculations described in this guidance (EPA 2001a). Both ProUCL and this guidance make recommendations for calculating UCLs, and are intended as tools to support risk assessment.

To obtain a copy of the ProUCL software or receive technical assistance in using it, risk assessors should contact:

Director of the Technical Support Center
USEPA Office of Research and Development
National Exposure Research Laboratory
Environmental Sciences Division
Las Vegas, Nevada
702-798-2270.

The ultimate responsibility for deciding how best to represent the concentration data for a site lies with the project team.[1] Simply choosing a statistical method that yields a lower UCL is not always the best representation of the concentration data at a site. The project team may elect to use a method that yields a higher (i.e., more conservative) UCL based on its understanding of site-specific conditions, including the representativeness of the data collection process, and the limits of the available statistical methods for calculating a UCL.

2.0 APPLICABILITY OF THIS GUIDANCE

This document updates 1992 guidance developed by EPA's Office of Emergency and Remedial Response; yet it can be applied to any hazardous waste site. It provides alternative methods for calculating the 95 percent upper confidence limit of the mean concentration, which can be used at sites subject to the discretion of the regulatory agencies and programs involved. The approaches described in this document are not specific to a particular medium (e.g., soil, groundwater), or receptor (e.g., human ecological), but apply to any media or receptor for which the UCL would be calculated.[2]

This document does not substitute for any statutory provisions or regulations, nor is it a regulation itself. Thus, it cannot impose legally-binding requirements on EPA, States, or the regulatory community, and may not apply to a particular situation based upon the circumstances. Any decision regarding cleanup of a particular site will be made based on the statutes and regulations, and EPA decisionmakers retain the discretion to adopt approaches on a case-by-case basis that differ from this guidance to a particular situation. The Agency accepts public input on this document at any time.

This guidance is based on the state of knowledge at present. The practices discussed herein may be refined, updated, or superseded by future advances in science and mathematics.

---

[1] The project team typically consists of a site manager (e.g., the Remedial Project Manger) and a multidisciplinary team of technical experts, including human health and ecological risk assessors, hydrogeologists, chemists, toxicologists, and quality assurance specialists.

[2] Note that this guidance does not apply to lead-contaminated sites. The Technical Review Working Group for Lead recommends that the average concentration is used in evaluating lead exposures (see http://www.epa.gov/superfund/programs/ lead/trwhome.htm).

3.0 DATA EVALUATION

In the risk assessment process, data evaluation precedes exposure assessment. Because this guidance deals with a component of exposure assessment, it therefore assumes that data have already undergone validation and evaluation and that the data have been determined to meet data quality objectives (DQOs) for the project in question. DQOs are important for any project where environmental data are used to support decision-making, as at hazardous waste sites.

One factor to consider in data evaluation is whether the number of sample measurements is sufficient to characterize the site or exposure unit. The minimum number of samples to conduct any of the statistical tests described in this document should be determined using the DQO process (EPA 2000a). Use of the methods described in this guidance is not a substitute for obtaining an adequate number of samples. Sample size is especially important when there is large variability in the underlying distribution of concentrations. However, defaulting to the maximum value of small data sets may still be the last resort when the UCL appears to exceed the range of concentrations detected.

Another important issue to consider is the method of sampling. All the statistical methods described in this guidance for calculating UCLs are based on the assumption of random sampling. At many hazardous waste sites, however, sampling is focused on areas of suspected contamination. In such cases, it is important to avoid introducing bias into statistical analyses. This can be achieved through stratified random sampling, i.e., random sampling within specified targeted areas. So long as the statistical analysis is constructed properly (i.e., there is no mixing of samples across different populations) bias can be minimized. The risk assessor should always note any potential bias in EPC estimates.

The risk assessor should also consider the duration of exposure and the time scale of the toxicity. For example, a chronic exposure may warrant the use of different concentrations or sample locations from an acute exposure. The time periods over which data were collected should also be considered. See EPA 1989, Chapters 5.1 and 6.4.2, for further details.

Once a set of data from a site has been evaluated and validated, it is appropriate to conduct exploratory analysis to determine whether there are outliers or a substantial number of non-detect values that can adversely affect the outcome of statistical analyses. The following sections describe the potential impact of outliers and non-detect values on the calculation of UCLs and approaches for addressing these types of values.

3.1 Outliers

Outliers are values in a data set that are not representative of the set as a whole, usually because they are very large relative to the rest of the data. There are a variety of statistical tests for determining whether one or more observations are outliers (EPA 2000b, section 4.4). These tests should be used judiciously, however. It is common that the distribution of concentration data at a site is strongly skewed so that it contains a few very high values corresponding to local hot spots of contamination. The receptor could be exposed to these hot spots, and to estimate the EPC correctly it is important to take account of these values. Therefore, one should be careful not to exclude values merely because they are large relative to the rest of the data set.

Extreme values in the data set may represent true spatial variation in concentrations. If an observation or group of observations is suspected to be part of a different contamination source or exposure unit, then regrouping of the data may be most appropriate. In this case, it may be necessary to evaluate these data as a separate hot spot or to resample. The behavior of the receptor and the size and location of the exposure unit will determine which sample locations to include. Such decisions depend on project-specific assessments based on the conceptual site model.

EPA guidance suggests that, when outliers are suspected of being unreliable and statistical tests show them to be unrepresentative of the underlying data set, any subsequent statistical analyses should be conducted both with and without the outlier(s) (EPA 2000b). In addition, the entire process, including identification, statistical testing and review of outliers, should be fully documented in the risk characterization.

3.2 Non-detects

Chemical analyses of contaminant concentrations often result in some samples being reported as below the sample detection limit (DL). Such values are called non-detects. Non-detects may correspond to concentrations that are actually or virtually zero, or they may correspond to values that are considerably larger than zero but which are below the laboratory's ability to provide a reliable measurement. Elevated detection limits need to be investigated, especially if there are high percentages of non-detects. It is not appropriate to simply account for elevated detection limits with statistical techniques; improvements in sampling and analysis methods may be needed to lower detection limits.

In this guidance, the term "detection limit" is used to represent the reported limit of the non-detect. In reality, this could be any of a number of detection or quantitation limits. For further discussion of detection and quantitation limits in the risk assessment, see text box and Chapter 5 of EPA 1989.

Alternative Quantitation Limits

Method Detection Limit (MDL): The lowest concentration of a hazardous substance that a method can detect reliably in either a sample or blank.

Contract-Required Quantitation Limit (CRQL): The substance-specific level that a CLP laboratory must be able to routinely and reliably detect in specific sample matrices. The CRQL is not the lowest detectable level achievable, but rather the level that a CLP laboratory must reliably quantify. The CRQL may or may not be equal to the quantitation limit of a given substance in a given sample.

Source: Superfund Glossary of Terms and Acronyms (http://www.epa.gov/superfund/resources/hrstrain/htmain/glossal.htm

In the statistical literature, data sets containing non-detects are called censored or left-censored. The detection limit achieved for a particular sample depends on the sensitivity of the measuring method used, the instrument quantitation limit, and the nature of dilutions and other preparations employed for the sample. In addition, there may be different degrees of censoring. For instance, some laboratories use the letter code "J" to indicate that a value was below the quantitation limit and the letter "U" to indicate that a value was below the detection limit. These code systems vary among laboratories, however, and it is essential to understand what the laboratory notations indicate about the reliability of its measurements.[3] Censoring can cause problems in calculating the UCL. There are several common options for handling non-detects.

Reexamining the conceptual site model may suggest that the data be partitioned. For instance, it may be clear from the spatial pattern of non-detects in the data that the region sampled can be subdivided into contaminated and non-contaminated areas. Evidence for this depends on the observed pattern of contamination, how the contamination came to be located in the medium, and how the receptors will come in contact with the medium. It may be necessary to collect more samples to obtain an adequate site characterization.

Simple Substitution methods assign a constant value or constant fraction of the detection limit (DL) to the non-detects. Three common conventions are: (1) assume non-detects are equal to zero; (2) assume non-detects are equal to the DL; or (3) assume non-detects are equal to one-half the DL. Whatever proxy value is assigned, it is then used as though it were the reliably estimated value for that measurement. Because of the complicated formulas used to compute UCLs, there is no general rule about which substitution rule will yield an appropriate UCL. The uncertainty associated with the substitution method increases, and its appropriateness decreases, as the detection limit becomes larger and as the number of non-detects in the data set increases.

Bounding methods estimate limits on the UCL in a distribution-free way. This method involves determining the lower and upper bounds of the UCL based on the full range of possible values for non-detects. If the uncertainty arising from censoring is relatively small, then the difference between the lower and upper bound estimates will be small. It is not possible to bound the UCL by using simple substitution methods such as computing the UCL once with the non-detects replaced by zeros and once with the non-detects replaced by their respective detection limits. Sometimes using all zeros will inflate the estimate of the standard deviation of the concentration values to such a degree that the resulting value for the UCL is larger than the value from using the detection limits (Ferson et al. 2002, Rowe 1988, Smith 1995). See Appendix A for an example of how to compute bounds on the UCL.

Distributional methods rely on applying an assumption that the shape of the distribution of non-detect values is similar to that of measured concentrations above the detection limit. EPA provides guidance on handling non-detects using several distributional methods, including Cohen's method (EPA 2000b, section 4.7). In addition, Helsel (1990) reviews a variety of distributional methods (see also Hass and Scheff 1990; Gleit 1985; Kushner 1976; Singh and Nocerino 2001). EnvironmentalStats for S-PLUS (Millard 1997) offers an array of methods for estimating parameters from censored data sets.

---

[3] Information concerning the quantitation limits also should be incorporated into the appropriate supplemental tables in the framework for risk assessment planning, reporting, and review described in the *Risk Assessment Guidance for Superfund Volume 1: Human Health Evaluation Part D (RAGS, Part D)* (EPA 1998.)

The appropriate method to use depends on the severity of the censoring, the size of the data set, and what distributional assumptions are reasonable. There are five recommendations about how to treat censoring in the estimation of UCLs.

1) Detection limits should always be reported for non-detects. Non-detects should also be reported with observed values where possible.
2) It is inappropriate to convert non-detects into zeros without specific justification (e.g., the analyte was not detected above the detection limit in any sample at the site).
3) If a bounding analysis reveals that the quantitative effects of censoring are negligible, then no further analysis may be required.
4) If further analysis is desired, consider using a distribution-specific method.
5) If the proportion of non-detects is high ($\geq$75%) or the number of samples is small (n<5), no method will work well. In this case, it is reasonable to report the percentage of data below the detection limit, and resort again to a bounding approach in which non-detects are replaced by the detection limit and used to compute a UCL value that is reported as a number likely to be considerably larger than the true mean.

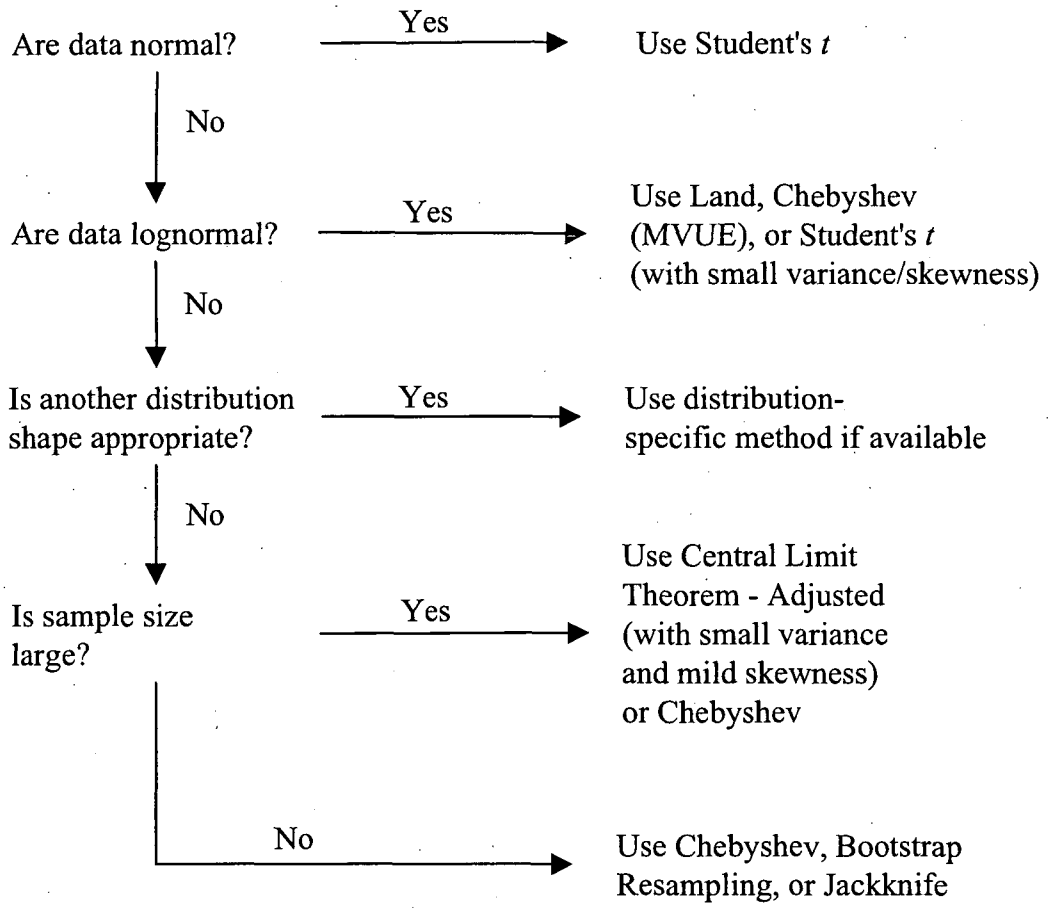
4.0 UCL CALCULATION METHODS

There are a number of different methods for calculating UCLs. Before an appropriate method can be selected the site data must be characterized through exploratory analysis. Fitting distributions to the data is a crucial part of this exploratory data analysis (Schulz and Griffin 1999). As recommended by EPA (1992), "where there is a question about the distribution of the data set, a statistical test should be used to identify the best distributional assumption for the data set." This is necessary because no single distribution type fits all environmental data sets. Risk assessors deal with some environmental data sets that appear normally distributed, and with others that appear lognormally distributed. They also encounter data sets that do not fit either normal or lognormal distributions. Distributions can be analyzed by a variety of methods, many of which are described in Gilbert (1987) and EPA (2000b). Data plotting can also help identify a useful distributional assumption. Some of these methods have been incorporated in the ProUCL software. Whatever method is used, it should be chosen in consultation with the EPA regional risk assessor and other project team members as appropriate. The assistance of a statistician may also be helpful in some cases.

The two most commonly used methods for computing UCLs are distributional methods. When the concentration distribution is normal, the classical approach based on the Student's t-statistic has typically been used. When the distribution is lognormal, the Land method based on the H-statistic has been used. Distribution-free or nonparametric methods are available if the risk assessor cannot reasonably make assumptions about the distributional type. EPA describes several methods (EPA, 2000c). For large data sets, an approach based on the Central Limit Theorem with a correction for positive skewness may be used. For data sets that are not large enough for this approach, there is more than one approach available, although none is ideal in all circumstances. General methods include an approach based on the Chebyshev inequality and an approach based on the bootstrap resampling procedure. These are described in EPA (2000c) and in Schulz and Griffin (1999). Both papers give examples and comparisons of the UCLs calculated by various methods. The flow chart shown in Figure 1 summarizes the recommendations in this guidance.

It should be noted that the "variance" in Figure 1 represents the variance of the log-transformed data. For detailed definitions of skewness, refer to the User's Guide for the ProUCL software.

Figure 1: UCL Method Flow Chart



Risk assessors are encouraged to use the most appropriate estimate for the EPC given the available data. The flow chart in Figure 1 provides general guidelines for selecting a UCL calculation method. This guidance presents descriptions of these methods, including their applicability, advantages and disadvantages. It also includes examples of how to calculate UCLs using the methods. While the methods identified in this guidance may be useful in many situations, they will probably not be appropriate for all hazardous waste sites. Moreover, other methods not specifically described in this guidance may be most appropriate for particular sites. The EPA risk assessor should be involved in the decision of which method(s) to use.

4.1 UCL Calculation with Methods for Specific Distributions

This section of the guidance presents methods for calculating UCLs when data can be shown to fit a specific distribution. Directions for using methods to calculate UCL for normal, lognormal, and other specific distributions are included, as are example calculations.

UCLs for Normal Distributions

If the data are normally distributed, then the one-sided (1-α) upper confidence limit $UCL_{1-\alpha}$ on the mean should be computed in the classical way using the Student's *t*-statistic (EPA 1992; Gilbert 1987, page 139; Student 1908). There is no change in EPA's prior recommendations for this type of data set (EPA 1992). Exhibit 1 gives the procedure for computing the UCL of the mean when the underlying distribution is normal. Exhibit 2 gives a numerical example of an application of the method.

Exhibit 1: Directions for Computing UCL for the Mean of a Normal Distribution — Student's *t*

Let $X_1, X_2, \ldots, X_n$ represent the n randomly sampled concentrations.

STEP 1: Compute the sample mean $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

STEP 2: Compute the sample standard deviation $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$.

STEP 3: Use a table of quantiles of the Student's *t* distribution to find the $(1-\alpha)^{th}$ quantile of the Student's *t* distribution with n-1 degrees of freedom. For example, the value at the 0.05 level with 40 degrees of freedom is 1.684. A table of Student's *t* values can be found in Gilbert (1987, page 255, where the values are indexed by p=1-α, rather than α level). The *t* value appropriate for computing the 95% UCL can be obtained in Microsoft Excel® with the formula TINV((1-0.95)*2, n-1).

STEP 4: Compute the one-sided (1-α) upper confidence limit on the mean

$$UCL_{1-\alpha} = \overline{X} + t_{\alpha,n-1}\, s / \sqrt{n}$$

Exhibit 2: An Example Computation of UCL for a Normal Distribution — Student's *t*

25 samples were collected at random from an exposure unit. The values observed are 228, 552, 645, 208, 755, 553, 674, 151, 251, 315, 731, 466, 261, 240, 411, 368, 492, 302, 438, 751, 304, 368, 376, 634, and 810 µg/L. It seems reasonable that the data are normally distributed, and the Shapiro-Wilk W test for normality fails to reject the hypothesis that they are ($W = 0.937$). The UCL based on Student's t is computed as follows.

STEP 1: The sample mean of the n=25 values is $\overline{X} = 451$.

STEP 2: The sample standard deviation of the values is $s = 198$.

STEP 3: The t-value at the 0.05 level for 25-1 degrees of freedom is $t_{0.05,25\text{-}1} = 1.710$.

STEP 4: The one-sided 95% upper confidence limit on the mean is therefore

$$UCL_{95\%} = 451 + 1.710 \times 198 / \sqrt{25} = 519$$

Testing for normality. For mildly skewed data sets, the student's t-statistic approach may be used to compute the UCL of the mean. But for moderate to highly skewed data sets, the t-statistic-based UCL can fail to provide the specific coverage for the population mean. This is especially true for small n. For instance, the 95% UCL based on 10 random samples from a lognormal distribution with mean 4.48 and standard deviation 5.87 will underestimate the true mean about 20% of the time, rather than the nominal rate of 5%. Therefore it is important to test the data for normality. EPA (2000b, section 4.2) gives guidance for several approaches for testing normality. The tests described therein are available in DataQUEST and ProUCL, which are convenient software tools (EPA 1997 and 2001a).

Accounting for non-detects. The use of substitution methods to account for non-detects is recommended only when a very small percentage of the data is censored (e.g., ≤ 15%), under the presumption that the numerical consequences of censoring will be minor in this case. As the percentage of the data censored increases, substitution methods tend to alter the distribution and violate the assumption of normality. Moreover, the effect of the various substitution rules on UCL estimation is difficult to predict. Replacing non-detects with half the detection limit can underestimate the UCL, and replacing them with zeros may overestimate the UCL (because doing so inflates the estimate of the standard deviation).

When censoring is moderate (e.g., >15% and ≤ 50%), it is preferable to account for non-detects with Cohen's method (Gilbert 1987). EPA provides guidance on the use of Cohen's method, which is a maximum likelihood method for correcting the estimates of the sample mean and the sample variance to account for the presence of non-detects among the data (EPA 2000b, beginning on page 4-43). This method requires that the detection limit be the same for all the data and that the underlying data are normally distributed.

UCLs for Lognormal Distributions

It is inappropriate to extend the methods of the previous section to lognormally distributed samples by log-transforming the data, computing a UCL and then back-transforming the results. For

concentration data sets that appear to be lognormally distributed, it may instead be preferable to use one of several methods available that are specifically well-suited to this type of distribution. These methods are described in the following sections.

Land Method

In past guidance, EPA had recommended using the Land method to compute the upper confidence limit on the mean for lognormally distributed data (Land 1971, 1975; Gilbert 1987; EPA 1992; Singh et al. 1997). This method requires the use of the *H*-statistic, tables for which were published by Land (1975) and Gilbert (1987, Tables A10 and A12). Exhibit 3 gives step-by-step directions for this method and Exhibit 4 gives a numerical example of its application.

Caveats about this method. Land's approach is known to be sensitive to deviations from lognormality. The formula may commonly yield estimated UCLs substantially larger than necessary when distributions are not truly lognormal if variance or skewness is large (Gilbert 1987). When sample sizes are small (less than 30), the method can be impractical even when the underlying distribution is lognormal (Singh et al. 1997).

Exhibit 3: Directions for Computing UCL for the Mean of a Lognormal Distribution— Land Method

Let $X_1, X_2, \ldots, X_n$ represent the n randomly sampled concentrations.

STEP 1: Compute the arithmetic mean of the log-transformed data $\overline{\ln X} = \frac{1}{n}\sum_{i=1}^{n} \ln(X_i)$.

STEP 2: Compute the associated standard deviation $s_{\ln X} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\ln(X_i) - \overline{\ln X}\right)^2}$.

STEP 3: Look up the $H_{1-\alpha}$ statistic for sample size n and the observed standard deviation of the log-transformed data. Tables of these values are given by Gilbert (1987, Tables A-10 and A-12) and Land (1975).

STEP 4: Compute the one-sided (1-α) upper confidence limit on the mean

$$UCL_{1-\alpha} = \exp\left(\overline{\ln X} + s_{\ln X}^2 / 2 + H_{1-\alpha} s_{\ln X} / \sqrt{n-1}\right)$$

Testing for lognormality. Because the Land method assumes lognormality, it is very important to test this assumption. EPA gives guidance for several approaches to testing distribution assumptions (EPA 2000b, section 4.2). The tests are also available in the DataQUEST and ProUCL software tools (EPA 1997 and 2001a).

Exhibit 4: An Example Computation of UCL for a Lognormal Distribution — Land Method

31 samples were collected at random from an exposure unit. The observed values are 2.8, 22.9, 3.3, 4.6, 8.7, 30.4, 12.2, 2.5, 5.7, 26.3, 5.4, 6.1, 5.2, 1.8, 7.2, 3.4, 12.4, 0.8, 10.3, 11.4, 38.2, 5.6, 14.1, 12.3, 6.8, 3.3, 5.2, 2.1, 19.7, 3.9, and 2.8 mg/kg. Because of their skewness, the data may be lognormally distributed. The Shapiro-Wilk *W* test for normality rejects the hypothesis, at both the 0.05 and 0.01 levels, that the distribution is normal. The same test fails to reject at either level the hypothesis that the distribution is lognormal. The UCL on the mean based on Land's *H* statistic is computed as follows.

STEP 1: Compute the arithmetic average of the log-transformed data $\overline{\ln X} = 1.8797$.

STEP 2. Compute the standard deviation of the log-transformed data $s_{\ln X} = 0.8995$.

STEP 3. The *H* statistic for $n = 31$ and $s_{\ln X}=0.90$ is 2.31.

STEP 4: The one-sided 95% upper confidence limit on the mean is therefore

$$UCL_{95\%} = \exp\left(1.8797 + 0.8995^2/2 + 2.31 \times 0.8995/\sqrt{31-1}\right) = 14.4$$

Accounting for non-detects. Gilbert (1987, page 182) suggests extending Cohen's method to account for non-detect values in lognormally distributed concentrations. Cohen's method (EPA 2000b, page 4-43) assumes the data are normally distributed, so it must be applied to the log-transformed concentration values. If $\hat{\mu}_y$ and $\hat{\sigma}_y$ are the corrected sample mean and standard deviation, respectively, of the log-transformed concentrations, then the corrected estimates of the mean and standard deviation of the underlying lognormal distribution can be obtained from the following expressions:

$$\hat{\mu} = \exp(\hat{\mu}_y + \hat{\sigma}_y^2/2)$$

$$\hat{\sigma} = \hat{\mu}\sqrt{\exp(\hat{\sigma}_y^2) - 1}$$

This method requires there be a single detection level for all the data values.

Chebyshev Inequality Method

Singh et al. (1997) and EPA (2000c) suggest the use of the Chebyshev inequality to estimate UCLs which should be appropriate for a variety of distributions so long as the skewness is not very large. The one-sided version of the Chebyshev inequality (Allen 1990, page 79; Savage 1961, page 216) is appropriate in this context (cf. Singh et al. 1997, EPA 2000c). It can be applied to the sample mean to obtain a distribution-free estimate of the UCL for the population mean when the population variance or standard deviation are known. In practice, however, these values are not known and must be estimated from data. For lognormally distributed data sets, Singh et al. (1997) and EPA (2000c) suggest using the minimum-variance unbiased estimators (MVUE) for the mean and variance to obtain an UCL of the mean. (See also Gilbert 1987, for discussion of the MVUE). This

approach may yield an estimated UCL that is more useful than that obtained from the Land method (when the underlying distribution of concentrations is lognormal). This alternative approach for a lognormal distribution is described in Exhibit 5 and is available in the ProUCL software tool (EPA 2001a). A numerical illustration of the Chebyshev inequality method using the sample mean and standard deviation appears in Exhibit 6. In this example the estimate of the UCL based on the Chebyshev inequality is less than that based on the Land method. The Chebyshev inequality estimate of the UCL is 1,965 mg/kg; while applying the Land method to this same data set yields a higher UCL estimate of 2,658 mg/kg.

Exhibit 5: Steps for UCL Calculation Based on the Chebyshev Inequality — MVUE Approach for Lognormal Distributions

Let $X_1, X_2, \ldots, X_n$ represent the n randomly sampled concentrations.

STEP 1: Compute the arithmetic mean of the log-transformed data $\overline{\ln X} = \frac{1}{n}\sum_{i=1}^{n} \ln(X_i)$.

STEP 2: Compute the associated variance $s^2_{\ln X} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\ln(X_i) - \bar{y}\right)^2$.

STEP 3: Compute the minimum-variance unbiased estimator (MVUE) of the population mean for a lognormal distribution $\hat{\mu}_{LN} = \exp\left(\overline{\ln X}\right) g_n\left(s^2_{\ln X}/2\right)$, where g_n denotes a function for which tables are available (Aitchison and Brown 1969, Table A2; Koch and Link 1980, Table A7).

STEP 4: Compute the MVUE of the associated variance of this mean

$$\sigma^2_{\mu} = \exp(2\ln X)\left(\left(g_n(s^2_{\ln X}/2)\right)^2 - g_n\left(\frac{n-2}{n-1}s^2_{\ln X}\right)\right)$$

STEP 5: Compute the one-sided (1-α) upper confidence limit on the mean

$$UCL_{1-\alpha} = \hat{\mu}_{LN} + \sqrt{\left(\frac{1}{\alpha} - 1\right)\sigma^2_{\mu}}$$

Caveats about the Chebyshev method. EPA (2000c) points out that for highly skewed lognormal data with small sample size and large standard deviation, the Chebyshev 99% UCL may be more appropriate than the 95% UCL, because the Chebyshev 95% UCL may not provide adequate coverage of the mean. As skewness increases further, the Chebyshev method is not recommended. See the ProUCL User's Guide (2001a) for specific recommendations on use of these two UCL estimates.

Exhibit 6: An Example Computation of UCL Based on the Chebyshev Inequality

29 samples were collected at random from an exposure unit. The observed values are 107, 175, 1796, 2002, 109, 30, 273, 83, 127, 254, 466, 12, 403, 31, 1042, 923, 24, 537, 5667, 59, 158, 59, 353, 10, 8, 33, 1129, 3 and 279 mg/kg. The observed skewness of this data set is 3.8, and these data may be lognormally distributed. The assumption of normality is rejected at the 0.05 level by a Shapiro-Wilk W test, but the same test fails to reject a test of lognormality even at the 0.1 level. The UCL on the mean can be computed based on the Chebyshev Inequality as follows.

STEP 1: The arithmetic mean of the log-transformed data $\overline{\ln X}$ is 4.9690.

STEP 2: The associated variance $s^2_{\ln X} = 3.3389$.

STEP 3: The MVUE of the mean for a lognormal distribution $\hat{\mu}_{LN} = 666.95$.

STEP 4: The MVUE of the variance of the mean $\sigma^2_{\mu} = 88552$.

STEP 5: The resulting one-sided 95% upper confidence limit on the mean of the concentration

$$UCL_{95\%} = 666.95 + \sqrt{(19)88552} = 1{,}965$$

The 95% UCL based on the Land method for these data would be 2,658.

EPA (2000c, Table 7) suggests that the Chebyshev inequality method for computing the UCL may be preferred over the Land method, even for lognormal distributions, in certain situations. Exhibit 7 describes the conditions, in terms of the sample size and the standard deviation of the log-transformed data, under which the Chebyshev inequality method will probably yield more useful results than the Land method.

Exhibit 7

Conditions Likely to Favor Use of Chebyshev Inequality (MVUE) over Land Method

Standard deviation of log-transformed data	Sample Size	Recommendation
1 - 1.5	<25	95% Chebyshev (MVUE) UCL
1.5 - 2	<20	99% Chebyshev (MVUE) UCL
	20 - <50	95% Chebyshev (MVUE) UCL
2 - 2.5	<25	99% Chebyshev (MVUE) UCL
	25 - 70	95% Chebyshev (MVUE) UCL
2.5 - 3.0	<30	99% Chebyshev (MVUE) UCL
	30 - <70	95% Chebyshev (MVUE) UCL

UCLs for Other Specific Distribution Types

Methods for computing UCLs on the mean of other types of distributions have appeared in the statistical literature. For example, Johnson (1978) describe a method for computing the UCL for asymmetrical distributions such as the exponential. Schulz and Griffin (1999) described Wong's (1993) method for obtaining confidence limits on the mean of a gamma distribution. In general, if there are arguments that suggest a population of concentrations should fit a particular distribution shape, and if statistical testing confirms the expected shape reasonably conforms with available data, then the UCL computed by a method developed specifically for the distribution shape, if one exists, is likely to be appropriate for the data set. An analyst should consider using a distribution-specific method if possible because it is likely to produce more valid statistical results. The advice and support of a statistician may be invaluable in such cases, both for characterizing the distribution and for identifying and evaluating possible ways to derive confidence limits.

4.2 UCL Calculation With Nonparametric or Distribution-free Methods

There are also distribution-free approaches to computing UCLs on the mean that do not make specific assumptions about the shape of the underlying distribution of concentrations. While these methods assume the samples are representative of the underlying distribution of concentrations, they require no assumptions about the shape of that distribution and are applicable to a variety of situations. Although parametric statistical methods that depend on a distributional assumption are usually more efficient and powerful than nonparametric methods, it can be difficult to justify their use through empirical testing of the shape of the distribution. In such cases, one of the following nonparametric, or distribution-free techniques are often preferred. For information on how to account for non-detects, see the earlier discussion under "Data Evaluation" above.

Central Limit Theorem (Adjusted)

If sample size is sufficiently large, the Central Limit Theorem (CLT) implies that the mean will be normally distributed, no matter how complex the underlying distribution of concentrations might be. This is the case even if the underlying distribution is strongly skewed, has outliers, or is a mixture of different populations, so long as it is stationary (not changing over time), has finite variance, and the samples are collected independently and randomly. However, the theorem does not say how many samples are sufficient for normality to hold. When sample size is moderate or small the means will not generally be normally distributed, and this non-normality is intensified by the skewness of the underlying distribution. Chen (1995) suggested an approach that accounts for positive skewness. Singh et al. (1997) and EPA (2000c) call this approach the "adjusted CLT" method. They suggest it is an appropriate alternative to the distribution-specific Land's method even if the distribution is lognormal when the standard deviation is less than one and sample size is larger than 100. Exhibit 8 describes the steps for this method, and Exhibit 9 gives a numerical example.

Exhibit 8: Directions for Computing UCL Using the Central Limit Theorem (Adjusted)

Let $X_1, X_2, \ldots, X_n$ represent the n randomly sampled concentrations.

STEP 1: Compute the sample mean $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

STEP 2: Compute the sample standard deviation $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$.

STEP 3: Compute the sample skewness $\beta = \frac{n}{(n-1)(n-2)}\sum_{i=1}^{n}\left(\frac{x_i - \overline{x}}{s}\right)^3$. This can be calculated in Microsoft® Excel with the SKEW function.

STEP 4: Let z_α be the $(1-\alpha)^{th}$ quantile of the standard normal distribution. For the 95% confidence level, $z_\alpha = 1.645$.

STEP 5: Compute the one-sided $(1-\alpha)$ upper confidence limit on the mean

$$UCL_{1-\alpha} = \overline{X} + \left(z_\alpha + \frac{\beta}{6\sqrt{n}}\left(1 + 2z_\alpha^2\right)\right)s/\sqrt{n} .$$

Exhibit 9: Example UCL Computation Based on the Central Limit Theorem (Adjusted)

60 samples were collected at random from an exposure unit. The values observed are 35, 111, 105, 27, 25, 20, 17, 21, 32, 32, 23, 17, 35, 32, 29, 25, 97, 20, 26, 18, 17, 18, 26, 25, 16, 28, 29, 28, 21, 119, 23, 98, 20, 21, 24, 21, 22, 117, 27, 25, 22, 21, 26, 24, 33, 33, 21, 24, 30, 31, 23, 30, 28, 25, 22, 23, 25, 28, 26, and 107 mg/L. Filliben's test shows that this distribution is significantly different (at the 1% level) from both a normal and a lognormal distribution. The UCL based on the Central Limit Theorem is computed as follows.

STEP 1: The sample mean of the n=60 values is $\overline{X} = 34.57$.

STEP 2: The sample standard deviation of the values is $s = 27.33$.

STEP 3: The sample skewness $\beta = 2.366$.

STEP 4: The z statistic is 1.645.

STEP 5: The one-sided 95% upper confidence limit on the mean is

$$UCL_{95\%} = 34.57 + \left(1.645 + \frac{2.366}{6\sqrt{60}}\left(1 + 2 \times 1.645^2\right)\right) 27.33 / \sqrt{60} = 42$$

Caveats about this method. A sample size of 30 is sometimes prescribed as sufficient for using an approach based on the Central Limit Theorem, but when using this CLT or adjusted CLT method and the data are skewed (as many concentration data sets are), larger samples may be needed to approximate normality. EPA's ProUCL User's Guide (2001) suggests that a sample size of 100 or more may be needed, based on Monte Carlo studies by EPA (2000c).

Bootstrap Resampling

Bootstrap procedures (Efron 1982) are robust nonparametric statistical methods that can be used to construct approximate confidence limits for the population mean. In these procedures, repeated samples of size n are drawn with replacement from a given set of observations. The process is repeated a large number of times (e.g., thousands), and each time an estimate of the desired unknown parameter (e.g., the sample mean) is computed. There are different variations of the bootstrap procedure available. One of these, the bootstrap t procedure, is described in the ProUCL User's Guide (EPA 2001a). An elaborated bootstrap procedure that takes bias and skewness into account is described in Exhibit 10 (Hall 1988 and 1992; Manly 1997; Schulz and Griffin 1999; Zhou and Gao 2000).

Caveats about resampling. Bootstrap procedures assume only that the sample data are representative of the underlying population. However, since they involve extensive resampling of the data and, thus, exploit more of the information in a sample, that sample must be a statistically accurate characterization of the underlying population in all respects (not just in its mean and standard deviation). In practice, it is random sampling that satisfies the representativeness assumption. Therefore the data must be random samples of the underlying population. Bootstrapping procedures are inappropriate for use with data that were idiosyncratically collected or focused especially on contamination hot spots.

Exhibit 10: Steps for Calculating a Hall's Bootstrap Estimate of UCL

Let $X_1, X_2, \ldots, X_n$ represent the n randomly sampled concentrations.

STEP 1: Compute the sample mean $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

STEP 2: Compute the sample standard deviation $s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$.

STEP 3: Compute the sample skewness $k = \frac{1}{ns^3}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^3$.

STEP 4: For $b = 1$ to B (a very large number) do the following:
4.1: Generate a bootstrap sample data set; i.e., for $i = 1$ to n let j be a random integer between 1 and n and add observation X_j to the bootstrap sample data set.
4.2: Compute the arithmetic mean $\overline{X}_b$ of the data set constructed in step 4.1.
4.3: Compute the associated standard deviation s_b of the constructed data set.
4.4: Compute the skewness k_b of the constructed data using the formula in Step 3.
4.5: Compute the studentized mean $W = (\overline{X}_b - \overline{X})/s_b$.
4.6: Compute Hall's statistic $Q = W + k_b W^2/3 + k_b^2 W^3/27 + k_b/(6n)$.

STEP 5: Sort all the Q values computed in Step 4 and select the lower α^{th} quantile of these B values. It is the $(\alpha B)^{th}$ value in an ascending list of Q's. This value is from the *left* tail of the distribution.

STEP 6: Compute $W(Q) = \frac{3}{k}\left(\left(1 + k\left(Q_\alpha - \frac{k}{6n}\right)\right)^{1/3} - 1\right)$.

STEP 7: Compute the one-sided (1-α) confidence limit on the mean.

$$UCL_{1-\alpha} = \overline{X} - W(Q_\alpha)s$$

Exhibit 11: An Example Computation of Bootstrap Estimate of UCL

Using the same concentration values given in Exhibit 4, the UCL can also be computed based on the Bootstrap Resampling method.

STEP 1: The sample mean of the n =31 values is $\overline{X}$= 9.59.

STEP 2: The standard deviation (using n as divisor) of the values is s = 8.946.

STEP 3: The skewness k = 1.648.

The Pascal-language software shown in Appendix B estimates the UCL with 100,000 bootstrap iterations. The one-sided 95% UCL on the mean is 13.3. Because this value depends on random deviates, it can vary slightly on recalculation.

Jackknife Procedure

Like bootstrap, the jackknife technique is a robust procedure based on resampling (Tukey 1977). In this procedure repeated samples are drawn from a given set of observations by omitting each observation in turn, yielding n data sets of size n-1. An estimate of the desired unknown parameter (e.g., sample mean) is then computed for each sample. When the standard estimators are used for the mean and standard deviation, this procedure reduces to the UCL based on Student's t. However, when other estimators (such as MVUE) are used this jackknife procedure does not reduce to the UCL based on Student's t. Singh et al. (1997) suggest that this method could be used with other estimators for the population mean and standard deviation to yield UCLs that may be appropriate for a variety of distributions.

Chebyshev Inequality Method

As described previously, Singh et al. (1997) and EPA (2000c) suggested the use of the Chebyshev inequality to estimate UCLs which should be appropriate for a variety of distributions as long as the skewness is not very large. The one-sided version of the Chebyshev inequality (Allen 1990, page 79; Savage 1961, page 216) is appropriate in this context (cf. Singh et al. 1997, EPA 2000c). It can be applied to the sample mean to obtain a distribution-free estimate of the UCL for the population mean when the population variance or standard deviation are known. In practice, however, these values are not known and must be estimated from data. Singh et al. (1997) and EPA (2000c) suggest that the population mean and standard deviation can be estimated by the sample mean and sample standard deviation. This approach is described in Exhibit 12 and is available in the ProUCL software tool (EPA 2001a). A numerical illustration of the Chebyshev inequality method using the sample mean and standard deviation appears in Exhibit 13.

Caveats about the Chebyshev method. Although the Chebyshev inequality method makes no distributional assumptions, it does assume that the parametric standard deviation of the underlying distribution is known. As Singh et al. (1997) acknowledge, when this parameter must be estimated from data, the estimate of the UCL is not guaranteed to be larger than the true mean with the prescribed frequency implied by the α level. In fact, using only an estimate of the standard deviation can substantially underestimate the UCL when the variance or skewness is large, especially for small sample sizes. In such cases, a Chebyshev UCL with a higher confidence coefficient such as 0.99 may be used, according to Singh, et al.

Exhibit 12: Steps for Computing UCL Based on the Chebyshev Inequality — Nonparametric

Let $X_1, X_2, \ldots, X_n$ represent the n randomly sampled concentrations.

STEP 1: Compute the arithmetic mean of the data $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

STEP 2: Compute the sample standard deviation $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$.

STEP 3: Compute the one-sided (1-α) upper confidence limit on the mean

$$UCL_{1-\alpha} = \overline{X} + \sqrt{\frac{1}{\alpha} - 1}\,(s/\sqrt{n})$$

Exhibit 13: An Example Computation of UCL Based on Chebyshev Inequality — Nonparametric

Using the same concentration values given in Exhibit 4 and used in Exhibit 11, the UCL on the mean can also be computed based on the Chebyshev inequality.

STEP 1: The sample mean of the n=31 values is $\overline{X}$= 9.59.

STEP 2: The sample standard deviation of the values is s = 9.094

STEP 3: The one-sided 95% upper confidence limit on the mean is therefore

$$UCL_{95\%} = 9.59 + 4.3589 \times 9.094/\sqrt{31} = 16.7$$

5.0 OPTIONAL USE OF MAXIMUM OBSERVED CONCENTRATION

Because some of the methods outlined above (particularly the Land method) can produce very high estimates of the UCL, EPA (1992) allows the maximum observed concentration to be used as the exposure point concentration rather than the calculated UCL in cases where the UCL exceeds the maximum concentration.

It is important to note, however, that defaulting to the maximum observed concentration may not be protective when sample sizes are very small because the observed maximum may be smaller than the population mean. Thus, it is important to collect sufficient samples in accordance with the DQOs for a site. The use of the maximum as the default exposure point concentration is reasonable only when the data samples have been collected at random from the exposure unit and the sample size is large.

6.0 UCLs AND THE RISK ASSESSMENT

Risk assessors are encouraged to use the most appropriate estimate for the EPC given the available data. The flow chart in Figure 1 provides general guidelines for selecting a UCL calculation method. Exhibit 14 summarizes the methods described in this guidance, including their applicability, advantages and disadvantages. While the methods identified in this guidance may be useful in many situations, they will probably not be appropriate for all hazardous waste sites. Moreover, other methods not specifically described in this guidance may be most appropriate for particular sites. The EPA risk assessor and, potentially, a trained statistician should be involved in the decision of which method(s) to use.

When presenting UCL estimates, the risk assessor should identify:

- how the shape of the underlying distribution was identified (or, if it was not identified, what methods were used in trying to identify it),
- the chosen UCL method,
- reasons that this UCL method is appropriate for the site data, and
- assumptions inherent in the UCL method.

It may also be appropriate to include information such as advantages and disadvantages of the distribution-fitting method, advantages and disadvantages of the UCL method, and how the risk characterization would change if other assumptions were used.

Exhibit 14
Summary of UCL Calculation Methods

Method	Applicability	Advantages	Disadvantages	Reference
For Normal or Lognormal Distributions				
Student's *t*	means normally distributed, samples random	simple, robust if *n* is large	distribution of means must be normal	Gilbert 1987; EPA 1992
Land's *H*	lognormal data, small variance, large *n*, samples random	good coverage[1]	sensitive to deviations from lognormality, produces very high values for large variance or small *n*	Gilbert 1987; EPA 1992
Chebyshev Inequality (MVUE)	skewness and variance small or moderate, samples random	often smaller than Land	may need to resort to higher confidence levels for adequate coverage	Singh et al. 1997
Wong	gamma distribution	second order accuracy[2]	requires numerical solution of an improper integral	Schulz and Griffin 1999; Wong 1993
Nonparametric/Distribution-free Methods				
Central Limit Theorem - Adjusted	large *n*, samples random	simple, robust	sample size may not be sufficient	Gilbert 1987; Singh et al. 1997
Bootstrap t Resampling	sampling is random and representative	useful when distribution cannot be identified	inadequate coverage for some distributions; computationally intensive	Singh et al. 1997; Efron 1982
Hall's Bootstrap Procedure	sampling is random and representative	useful when distribution cannot be identified; takes bias and skewness into account	inadequate coverage for some distributions; computationally intensive	Hall 1988; Hall 1992; Manly 1997; Schultz and Griffin 1999
Jackknife Procedure	sampling is random and representative	useful when distribution cannot be identified	inadequate coverage for some distributions; computationally intensive	Singh et al. 1997
Chebyshev Inequality	skewness and variance small or moderate, samples random	useful when distribution cannot be identified	inappropriate for small sample sizes when skewness or variance is large	Singh et al. 1997; EPA 2000c

[1] Coverage refers to whether a UCL method performs in accordance with its definition.
[2] As opposed to maximum likelihood estimation, which offers first order accuracy.

7.0 PROBABILISTIC RISK ASSESSMENT

The estimates of the UCL described in this guidance can be used as point estimates for the EPC in deterministic risk assessments. In probabilistic risk assessments, a more complete characterization of the underlying distribution of concentrations may be important as well. Risk assessors should consult *Risk Assessment Guidance for Superfund, Volume 3 - Part A, Process for Conducting a Probabilistic Risk Assessment* (EPA 2001b) for specific guidance with respect to probabilistic risk assessments.

8.0 CLEANUP GOALS

Cleanup goals are commonly derived using the risk estimates established during the risk assessment. Often, a cleanup goal directly proportional to the EPC will be used, based on the relationship between the site risk and the target risk as defined in the National Contingency Plan. In such cases, the attainment of the cleanup goal should be measured with consideration of the method by which the EPC was derived. For more details, see *Surface Soil Cleanup Strategies for Hazardous Waste Sites* (EPA, to be published).

9.0 REFERENCES

Aitchison, J. and J.A.C. Brown (1969). *The Lognormal Distribution*. Cambridge University Press, Cambridge.

Allen, A.O. (1990). *Probability, Statistics and Queueing Theory with Computer Science Applications*, second edition. Academic Press, Boston.

Chen, L. (1995). Testing the mean of skewed distributions. *Journal of the American Statistical Association* 90: 767-772.

Efron, B. (1982) *The Jackknife, the Bootstrap and Other Resampling Plans*. SIAM, Philadelphia, Pennsylvania.

EPA (1989). *Risk Assessment Guidance for Superfund, Volume I - Human Health Evaluation Manual (Part A). Interim Final.* http://www.epa.gov/superfund/programs/risk/ragsa/, EPA/540/1-89/002. Office of Emergency and Remedial Response, U.S. Environmental Protection Agency, Washington, D.C.

EPA (1992). *A Supplemental Guidance to RAGS: Calculating the Concentration Term.* http://www.deq.state.ms.us/newweb/opchome.nsf/pages/HWDivisionFiles/$file/uclmean.pdf. Publication 9285.7-081. Office of Solid Waste and Emergency Response, U.S. Environmental Protection Agency, Washington, D.C.

EPA (1997). *Data Quality Evaluation Statistical Toolbox (DataQUEST) User's Guide.* http://www.epa.gov/quality/qs-docs/g9d-final.pdf, EPA QA/G-9D QA96 Version. Office of Research and Development, U.S. Environmental Protection Agency, Washington, D.C. [The software is available at http://www.epa.gov/quality/qs-docs/dquest96.exe]

EPA(1998). *Risk Assessment Guidance for Superfund, Volume 1 - Human Health Evaluation Manual Part D.* http://www.epa.gov/superfund/programs/risk/ragsd/ Publication 9285:7-01D. Office of Solid Waste and Emergency Response, U.S. Environmental Protection Agency, Washington D.C.

EPA (2000a). *Data Quality Objectives Process for Hazardous Waste Site Investigations.* http://www.epa.gov/quality/qs-docs/g4hw-final.pdf, EPA QA/G-4HW, Final. Office of Environmental Information, U.S. Environmental Protection Agency, Washington, D.C.

EPA (2000b). *Guidance for Data Quality Assessment: Practical Methods for Data Analysis.* http://www.epa.gov/r10earth/offices/oea/epaqag9b.pdf, EPA QA/G-9, QA00 Update. Office of Environmental Information, U.S. Environmental Protection Agency, Washington, D.C.

EPA (2000c). *On the Computation of the Upper Confidence Limit of the Mean of Contaminant Data Distributions, Draft.* Prepared for EPA by Singh, A.K., A. Singh, M. Engelhardt, and J. M. Nocerino. Copies of the article can be obtained from Office of Research and Development, Technical Support Center, U.S. Environmental Protection Agency, Las Vegas, Nevada.

EPA (2001a). ProUCL- Version 2. [software for Windows 95, accompanied by "ProUCL User's Guide."] Prepared for EPA by Lockheed Martin.

EPA (2001b). *Risk Assessment Guidance for Superfund, Volume 3 - Part A, Process for Conducting a Probabilistic Risk Assessment Draft.* http://www.epa.gov/superfund/programs/risk/ags3adt/. Office of Solid Waste and Emergency Response, U.S. Environmental Protection Agency, Washington D.C.

EPA (To be published). *Guidance or Surface Soil Cleanup at Hazardous Waste Sites: Implementing Cleanup Levels, Draft.* Office of Solid Waste and Emergency Response, U.S. Environmental Protection Agency, Washington D.C.

Ferson, S. et al. 2002. Bounding uncertainty analyses. Accepted for Publication in *Proceedings of the Workshop on the Application of Uncertainty Analysis to Ecological Risks of Pesticides.* edited by A. Hart. SETAC Press, Pensacola, Florida.

Gilbert, R.O. (1987). *Statistical Methods for Environmental Pollution Monitoring.* Van Nostrand Reinhold, New York.

Gleit, A. (1985). Estimation for small normal data sets with detection limits. *Environmental Science and Technology* 19: 1201-1206.

Haas, C.N. and P.A. Scheff (1990). Estimation of averages in truncated samples. *Environmental Science and Technology* 24: 912-919.

Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals. *Annals of Statistics* 16: 927-953.

Hall, P. (1992). On the removal of skewness by transformation. *Journal of the Royal Statistical Society B* 54: 221-228.

Helsel, D.R. (1990). Less than obvious: Statistical treatment of data below the detection limit. *Environmental Science and Technology* 24: 1766-1774.

Johnson, N.J. (1978). Modified t-tests and confidence intervals for asymmetrical populations. *The American Statistician* 73:536-544.

Koch, G.S., Jr., and R.F. Link (1980). *Statistical Analyses of Geological Data.* Volumes I and II. Dover, New York.

Kushner, E.J. (1976). On determining the statistical parameters for pollution concentration from a truncated data set. *Atmospheric Environ.* 10: 975-979.

Land, C.E. (1971). Confidence intervals for linear functions of the normal mean and variance. *Annuals of Mathematical Statistics* 42: 1187-1205.

Land, C.E. (1975). Tables of confidence limits for linear functions of the normal mean and variance. *Selected Tables in Mathematical Statistics* Vol III p 385-419.

Manly, B.F.J. (1997). *Randomization, Bootstrap, and Monte Carlo Methods in Biology* (2nd edition). Chapman and Hall, London.

Millard, S.P. (1997). EnvironmentalStats for S-Plus user's manual, Version 1.0. Probability, Statistics, and Information, Seattle, WA.

Rowe, N.C. (1988). Absolute bounds on set intersection and union sizes from distribution information. *IEEE Transactions on Software Engineering.* SE-14: 1033-1048.

Savage, I.R. (1961). Probability inequalities of the Tchebycheff type. *Journal of Research of the National Bureau of Standards-B. Mathematics and Mathematical Physics* 65B: 211-222.

Schulz, T.W. and S. Griffin (1999). Estimating risk assessment exposure point concentrations when the data are not normal or lognormal. *Risk Analysis* 19:577-584.

Singh, A.K., A. Singh, and M. Engelhardt (1997). The lognormal distribution in environmental applications. EPA/600/R-97/006

Singh, A., and J.M. Nocerino (2001). Robust Estimation of Mean and Variance Using Environmental Datasets with Below Detection Limit Observations. Accepted for Publication in *Journal of Chemometrics and Intelligent Laboratory Systems.*

Smith, J.E. (1995). Generalized Chebychev inequality: theory and applications in decision analysis. *Operations Research.* 43: 807-825.

Student [W.S. Gossett] (1908).On the probable error of the mean. *Biometrika* 6: 1-25.

Tukey, J.W. (1977). *Exploratory Data Analysis.* Addison Wesley, Reading, MA.

Wong, A. (1993). A note on inference for the mean parameter of the gamma distribution. *Stat. Prob. Lett.* 17: 61-66.

Zhou, X.-H. and S. Gao (2000). One-sided confidence intervals for means of positively skewed distributions. *The American Statistician* 54: 100-104.

Appendix A: Using Bounding Methods to Account for Non-detects

This appendix presents an iterative procedure that can be used to account for non-detects in data when estimating a UCL. It provides a step-by-step approach for computing an upper bound on the UCL using the "Solver" feature in Microsoft ® Excel spreadsheets.

STEP 1. Enter all the detected values in a column.

STEP 2. At the bottom of the same column, append as place holders as many copies of the formula

=RAND( )*DL

as there were non-detects. In these formulas, *DL* should be replaced by the detection limit.

STEP 3. Copy all the cells you have entered in steps 1 and 2 to a second column.

STEP 4. In another cell, enter the formula for the UCL that you wish to use. For instance, to use the 95% UCL based on Student's *t*, enter the formula

=AVERAGE(*range*)+TINV((1-0.95)*2, *n*-1)*SQRT(VAR(range)/n)

where *range* denotes the array of cell references in the second column you just created and *n* denotes the number of measurements (both detected values and non-detects).

STEP 5. From the Excel menu, select Tools / Solver.

STEP 6. In the "Solver Parameters" dialog box, specify the cell in which you entered the UCL formula as the Target Cell.

STEP 7. To find the upper bound of the UCL click on the Max indicator; to find the lower bound of the UCL click on the Min indicator.

STEP 8. Enter references to the cells containing the place holders for the non-detects in the field under the label "By Changing Cells." (Do not click the "Guess" button.)

STEP 9. For each cell that represents a non-detect, add a constraint specifying that the cell is to be greater than or equal to (">=") the detection limit *DL*.

STEP 10. Click on the Options button and check the box labeled "Assume Non-Negative."

STEP 11. Then click OK and then the Solver button. The program will automatically locate a local extreme value (i.e., maximum or minimum) for the UCL.

STEP 12. Record this value. You can use the Save Scenario button and Excel's scenario manager to do this.

STEP 13. Again copy all the detected values and randomized place holders for the non-detects from the first column to the same spot in the second column.

STEP 14. Select Tools / Solver and click the Solve button.

STEP 15. If calculating the upper bound, record the resulting value of the UCL if it is larger than previously computed. If calculating the lower bound, record the resulting value of the UCL if it is smaller than previously computed.

STEP 16. Repeat steps 13 through 15 to search for the global maximum or minimum value for the UCL.

Appendix B: Computer Code for Computing a UCL with the Hall's Bootstrap Sampling Method

This appendix presents Pascal code that can be used to compute the bootstrap estimate of a UCL. To use it, place data in the vector *x*. Then specify the sample size *n*, the vector *x* and the alpha-level, and call the procedure bootstrap. When the procedure finishes, the estimated value will be in the variable UCL. To obtain a 95% UCL, let alpha be 0.05. Up to 100 data values and up to 10,000 bootstrap iterations are supported, but these limits may be changed.

```
const
  max = 100;
  bmax = 10000;

type
  index = 1..max;
  bindex = 1..bmax;
  float = extended;{could just be real}
  vector = array[index] of float;
  bvector = array[bindex] of float;

var
  qq : bvector;

function getmean(n : integer; x : vector) : float;
  var s : float; i : integer;
  begin
  s := 0.0;
  for i := 1 to n do s := s + x[i];
  getmean := s / n;
  end;

function getstddev(n:integer; xbar:float; x:vector) : float;
  var s : float; i : integer;
  begin
  s := 0.0;
  for i := 1 to n do s := s + (x[i] - xbar) * (x[i] - xbar);
  getstddev := sqrt(s / n); {not n-1}
  end;

function getskew(n:integer; xbar:float; stddev:float; x:vector) :
float;
  var s,s3 : float; i : integer;
  begin
  s := 0.0;
  s3 := stddev * stddev * stddev;
  for i:=1 to n do s:=s+(x[i]-xbar)*(x[i]-xbar)*(x[i]-xbar)/s3;
  getskew := s / n;
  end;

procedure qsort(var a: bvector; lo,hi: integer);
  procedure sort(l,r: integer);
  var i,j : integer; x,y: float;
  begin
  i:=l; j:=r; x:=a[(l+r) div 2];
  repeat
    while a[i]<x do i:=i+1;
    while x<a[j] do j:=j-1;
    if i<=j then
      begin
      y:=a[i]; a[i]:=a[j]; a[j]:=y;
      i:=i+1; j:=j-1;
      end;
```

```
    until i>j;
  if l<j then sort(l,j);
  if i<r then sort(i,r);
  end;
  begin {qsort}
  sort(lo,hi);
  end;

procedure bootsample(n : integer; x : vector; var y : vector);
  var i,j : integer;
  begin
  for i := 1 to n do
    begin
    j := random(n) + 1;
    y[i] := x[j];
    end;
  end;

procedure bootstrap(n:integer; x:vector; alpha:float; var
ucl:float);
{let alpha be 0.05 to compute a 95% UCL}
var
  i,b,bb : integer;
  xbar, stddev, skew, bxbar, bstddev, bskew, k, w, q, a : float;
  bx : vector;
begin
bb := bmax;
for b:=1 to bmax do qq[b] := 0.0;
xbar := getmean(n,x);
stddev := getstddev(n,xbar,x);
skew := getskew(n,xbar,stddev,x);
for b := 1 to bb do
  begin
  bootsample(n,x,bx);
  bxbar := getmean(n,bx);
  bstddev := getstddev(n,bxbar,bx);
  k := getskew(n,bxbar,bstddev,bx);
  w := (bxbar - xbar) / bstddev;
  q := w + skew * w*w / 3 + k*k * w*w*w / 27 + k / (6 * n);
  qq[b] := q;
  end;
qsort(qq,1,bb);
q := qq[round(alpha * bb)];
a := 1 + skew * (q-skew / (6 * n));
if a = 0.0 then w := -3 / skew
           else w := (3 / skew) * (exp((1/3) * ln(a)) - 1);
ucl := xbar - w * stddev;
end;
```