16

Industrial Hygiene Exposure Assessment—Data Analysis and Interpretation

PAUL HEWETT

As discussed in Hewett[1] (see Chapter 15), industrial hygiene exposure assessments come in three varieties: qualitative, semiquantitative, and quantitative. Quantitative surveys, which involve the measurement of current worker exposure using personal sampling equipment or direct reading instruments, are often necessary for initial or baseline evaluations. Furthermore, periodic sampling and occasional audits are necessary for validating earlier assessments and for detecting upward trends in exposure. Consequently, an industrial hygienist is often faced with questions regarding the collection, analysis, interpretation, and management of occupational exposure data. Hewett[1] described the rationale behind exposure monitoring and covered data collection and data management. The purpose of this chapter is to suggest appropriate procedures for analyzing and criteria for interpreting exposure data.

As discussed in Hewett,[1] exposure-monitoring programs must be designed and tailored for a wide variety of work environments. But first, as noted by Roach[2] in 1967, "[it] is important that hygienic standards should not be given widely different interpretations." We should also agree that the goal of an effective exposure-monitoring program is to routinely and accurately characterize the exposure profile* of each worker. It is less critical that we adopt identical or similar data analysis and interpretation procedures. There are numerous data analysis techniques—parametric and nonparametric—that will yield similar *decisions* regarding the acceptability of the work environment. Regardless of the number of measurements collected or the sophistication of the analysis technique, there is always a role for professional judgment and common sense.[1]

*If we measured the full-shift TWA exposure of a single worker for each of the approximately 250 working days per year and plotted these measurements in a histogram, then the shape of the histogram would be an estimate of the true exposure distribution for that worker during that year. The term "exposure profile" is used to refer to this distribution. One can also conceptualize within-shift exposure profiles (for short-term exposures) or exposure profiles for periods shorter than or longer than a year.

The procedures presented here emphasize the calculation of the point estimate of a relevant exposure parameter, and in addition the 95% lower confidence limit (LCL) and the 95% upper confidence limit (UCL). Taken together, these confidence limits comprise a 90% confidence interval for the true parameter. The advantage of this interval estimate is that it can readily be used to gauge (a) the accuracy of our point estimate of the true parameter, and (b) the acceptability of the work environment by comparing the LCL or UCL to an exposure limit or other relevant criterion. These procedures are consistent with those recommended by the Exposure Assessment Strategies Committee (EASC)[3] of the American Industrial Hygiene Association (AIHA) and the Comité Européen de Normalisation (CEN).[4]

Each worker should expect a work environment devoid of unreasonable risks. Our goal is to protect *each* individual worker, but limited resources usually compel industrial hygienists to (a) aggregate workers into exposure groups; (b) determine which exposure groups warrant priority attention; and (c) evaluate the "exposure profile" of each exposure group in order of priority. Consequently, our data collection strategies (see Hewett[1]), and data analysis and interpretation procedures must be designed so that our *conclusions* regarding the exposure group are reasonably *predictive* of the exposures experienced by each member of the group.

Data Analysis—Goodness-of-Fit and Other Issues

We need to assure ourselves that our exposure measurements have *predictive value.* If we can demonstrate that (a) the exposure distribution is reasonably stable; (b) the data are reasonably independent and uncorrelated; and (c) the data are described reasonably well by a lognormal distribution, then it is logical to assume that they have predictive value for the exposure group. Furthermore, we often require that the exposure group be reasonably homogeneous with regard to the source and

conditions of exposure for each group member. Otherwise, the exposure profile for the group may not be predictive of the exposure profiles for one or more of the individuals within the group.

Is the Exposure Distribution Stationary?

When evaluating exposure measurements one always assumes that the underlying processes and work practices that generate and influence exposures will remain *stationary*, or relatively constant, until the next survey or evaluation. If exposures are currently acceptable, but trending upward, then a decision based on current measurements will have little or no predictive value. Symanski et al.[5] found that most exposure distributions will remain reasonably stationary for periods up to a year, after which systematic changes are more likely.

The "x-bar, r-charts" used in quality control are sometimes adapted for use in industrial hygiene.[6,7] Such charts are useful for demonstrating that an industrial process is stationary. Generally speaking, however, the number of measurements necessary to first establish statistical control limits and the number collected thereafter during each survey are usually beyond the means of most exposure-monitoring programs.[8] However, there are control charting techniques[9] designed for small sample sizes that could be considered.*

More useful, perhaps, are time series graphs such as recommended by the CEN,[4] Roach et al.,[2] and Roach.[10] The CEN recommended plotting moving or weighted averages for visually detecting trends. George et al.[11] discussed the use of formal statistical tests of stationarity, but cautioned that "visual investigation" always be part of any analysis.

Production line processes are often fairly stable. In such work environments conclusions based on current or recent exposure measurements may be reasonably predictive for periods up to a year or longer. However, the stability of a particular work environment may not be known until an exposure history of several years has been developed. Other work environments, such as mining, are inherently unstable and subject to considerable change within a short time span. In such environments regular monitoring is necessary to detect trends toward higher exposures.

Are the Exposure Data Independent?

In principle, exposure data should be independent; that is, there is no linear relationship between measurements collected on successive days. However, exposure data collected during unusually elevated or depressed production rates may be correlated and unrepresentative of exposures in general. In such a situation, measurements collected randomly throughout the observation period would be more appropriate than the usual campaign type survey.

Analysis of autocorrelation requires large datasets of consecutive measurements. However, since many researchers[11,12] have found little evidence of significant autocorrelation in measurements collected on successive days, it is reasonable to assume, in the absence of compelling information or data to conclude otherwise, that full-shift exposure measurements are reasonably independent. However, because there are often patterns of exposures within a shift, it is logical to expect that short-term measurements collected sequentially within a shift will be somewhat correlated.

Is the Lognormal Distribution Assumption Valid?

Goodness-of-fit testing involves both graphical and analytical evaluations.[13] The data should be plotted so that (a) inconsistent data points can be identified and (b) the goodness-of-fit can be subjectively evaluated. This should be followed by an objective analytical test of the hypothesis of lognormality. The lognormal distribution model is often used when zero is the physical lower limit for possible values, large values occasionally occur, and the processes that generate or control exposures tend to interact in a multiplicative manner.[14] Experience has shown that multiple exposure measurements collected either from a single individual or from multiple individuals within an exposure group tend toward a lognormal distribution.[15] Therefore, it is reasonable to assume that the underlying distribution for workplace exposure data is the lognormal distribution unless there is a compelling reason to conclude otherwise. There are, however, instances where the normal distribution may be more appropriate, for example, when multiple measurements are simultaneously collected at the same location.

*Traditional x bar and r control charts can be adapted for use with the lognormal distribution. Such charts are not used to determine compliance with federal or authoritative OELs, but can be used to assess whether or not the process remains in "statistical control" (e.g., stable, stationary). The control limits are calculated using the log-transformed exposure measurements and then exponentiated. This results in asymmetric upper and lower control limits. Trends can then be evaluated using the resulting "geometric mean" control chart. Changes in variability can be detected using the resulting "ratio" control chart (referring to the ratio of the largest to smallest measurement in each dataset).

Log-Probability Plotting Techniques

Log-probability plotting is the traditional method for qualitatively assessing the adequacy of the lognormal distribution model assumption. Odd patterns in the data and inconsistent data points can be readily identified.[6] However, the procedure is subjective and, when done by hand, can be tedious, particularly for large datasets.

Procedure:

(1) Sort or rank order the data: $\mathbf{x} = \{x_1, \ldots, x_n\}$ where x_1 is the smallest value and x_n is the largest.
(2) Assign a rank (r) to each sorted value where r ranges from 1 to n, starting with x_1.
(3) Calculate a plotting position, p, (basically a pseudo-cumulative frequency) using Blom's formula:*

$$p_i = \frac{i - \frac{3}{8}}{n + \frac{1}{4}} = \frac{i - 0.375}{n + 0.25}, \quad 1 \le i \le n \qquad (1)$$

(4) *Log-probability plot:* Using log-probability paper, plot p versus x. (Normal-probability paper can be used if the normal distribution assumption is of interest.)
Log-probit plot: Using regular graph paper or a spreadsheet program, plot m_i versus y_i, where $y = \ln(x)$ and m represents the probit (probability unit; also called "normal order statistic" or z-value) corresponding to p:†

$$m_i = \phi^{-1}[p_i]$$

For both the log-probability and log-probit plots, a straight line can be drawn emphasizing the influence of the central 80% of the data. That is, measurements in the tails should be given less weight when fitting the straight line. If *most* of the data fall along or near the straight line, then one can qualitatively state that the data appears to be lognormal.

Personal computer statistics or spreadsheet programs can be used to produce both log-probability and log-probit plot. It is also a relatively simple matter to have the program display a linear regression line along with the data to assist in the visual evaluation.‡ An alternative plot is the cumulative distribution function (CDF) plot. The plotting position, p, is graphed versus $\ln(x)$. Ideally, the data should appear to fall along a sigmoidal, or S-shaped curve. If *most* of the data fall along or near a sigmoidal curve, then one can qualitatively state that the data appears to be lognormal.

Formal Goodness-of-Fit Tests

There are numerous statistical tests for determining whether or not a particular set of data departs significantly from the normal distribution assumption. The one recommended here, Filliben's test[16] (as modified by Looney and Gulledge[17]), is easily implemented using a computer spreadsheet program or programmable calculator. This procedure is complementary to the graphical technique in that it incorporates the concept of the probability plot. The purpose of this test is to determine if a particular set of data departs significantly from the lognormal distribution assumption by evaluating whether or not the log-transformed values depart significantly from normal.

Filliben's test can be applied to sample sizes ranging from 3 to 100. Although not presented here, a similar procedure by Royston[13] can be applied to any sample size between 5 and 5000.§

Filliben's Test Filliben[16] developed a goodness-of-fit test based on normal order statistics. Looney and Gulledge[17] recalculated Filliben's critical values and recommended substituting Blom's formulae (see Step 3 below) for the plotting position formulae developed by Filliben. The CEN,[4] in their general guidance regarding exposure assessment, listed Filliben's test as one method for formally assessing the lognormal distribution assumption.

Procedure

(1) Sort or rank order the data: $\mathbf{x} = \{x_1, \ldots, x_n\}$ where x_1 is the smallest value and x_n is the largest.

*There are several formulae for estimating the plotting position, for example, $p_i = (i + 0.5)/n$ or $p_i = i/(n + 1)$. Blom's formula is preferred.

†For example, for a p_i of 0.95, $m_i = 1.645$; for a p_i of 0.50, $m_i = 0.000$. These values can be obtained from the cumulative normal distribution table found in texts on statistics or calculated using the statistical functions in a spreadsheet program.

‡Simple linear regression using $\ln(x)$ versus m can be used to produce a straight line through the datapoints.

§It can be shown that Filliben's and Royston's tests are essentially identical and yield consistent results for sample sizes ranging from 5 to 100, the range of overlap for the two procedures.

(2) Assign a rank (r) to each sorted value where r ranges from 1 to n, starting with x_1.
(3) For each x_i, calculate a plotting position, p_i, using Blom's formula (Eq. 1).
(4) Determine the corresponding normal order statistic, m, for each plotting position.
(5) Using simple linear regression, regress y versus m, where $y = \ln(x)$.
(6) Calculate the correlation coefficient r.
(7) Evaluate the following hypotheses by comparing r to a table of critical values (Table A1):

H_o: y is from a normal distribution.
H_a: y is not from a normal distribution.

(8) If r is less than or equal to the critical value, reject H_o with 95% confidence. Otherwise, accept H_o and conclude that y is normally distributed. Therefore, x is lognormally distributed.

How It Works If y is truly normally distributed, the calculated correlation coefficient r will tend to be near unity, or 1.0. Under the normal distribution assumption, Filliben determined the lower 5th percentile r value for sample sizes ranging from 3 to 100. If the calculated r value is less than or equal to this critical r value, then one can state, with 95% confidence, that the distribution from which these data were drawn is not normal. Consequently, if r is less than the critical value, reject H_o with a $(1 - \alpha)$ 100% confidence level. Otherwise, there is not enough evidence to reject, H_o. If H_o cannot be rejected, then the conclusion is that y is normally distributed or at least approximately normally distributed. This being the case, one can infer that x is lognormally distributed.

Filliben's test can be applied to both the actual values and the log-transformed values. One could then use the magnitude of the correlation coefficient to select between the normal or lognormal distribution assumptions. Sometimes both the lognormal and normal distribution assumptions can be rejected. It also should be noted that an observed correlation coefficient, or r value, that is less than the critical value does not necessarily mean that the underlying data do not stem from lognormal distributions. It may be that the data reflect two or more underlying lognormal distributions, reflecting perhaps the inadvertent combination of two or more distinctly different exposure groups.

Is the Exposure Group Reasonably Homogeneous?

Ideally, each member of an exposure group should have an identical exposure profile, although on any single day the exposures will vary. In practice, identical exposure profiles are unlikely to be observed. If an exposure group is reasonably homogeneous with respect to the conditions of exposure (agent/jobs/task/controls), then we expect that exposures primarily reflect the influence of process and work environment and can be reduced or controlled through direct control of the process and/or by general ventilation. If the exposure group is decidedly heterogeneous, then we expect that exposures tend to reflect the effectiveness of individual ventilation controls; differences in the number and duration of assigned tasks; and/or the influence of individual work practices, in which case exposures are reduced by focusing on individual work environments and individual work practices.

The process of grouping workers using observational skills has been described as the "observational approach" and criticized for being subjective and prone to classification errors.[15,18] How then does one objectively determine if a particular exposure group is sufficiently homogeneous so that decisions based on an analysis of group exposures are relevant to each member of the group, whether measured or not? There are no recognized criteria for objectively grouping workers or for distinguishing between a reasonably homogeneous exposure group and a clearly heterogeneous exposure group, although at least one has been proposed.[15] The CEN[4] offered the following "rule of thumb":

> [I]f an individual exposure is less than half or greater than twice the arithmetic mean [of the n measurements collected from the homogeneous group], the relevant work factors should be closely re-examined to determine whether the assumption of homogeneity was correct.

As a practical measure, many industrial hygienists use a process of continuous improvement to increase worker similarity within an exposure group. Basically, the observational approach is used to devise initial, logical exposure groups that are homogeneous with respect to process, agent, job/task, and type of controls. If, after the baseline survey, the exposure profile for the group appears acceptable, then periodic follow-up surveys are planned to evaluate the individual work practices of randomly selected workers or workers who are suspected, based upon previous measurements or professional judgment, to experience generally higher exposures. This continual, cyclic evaluation and modification of individual work practices, habits, and controls is expected to result in exposure groups that become more homogeneous over time.

Within-group homogeneity is less of an issue in those situations where the group exposures can be rated

highly controlled or well-controlled (see the Hewett[1] scheme for rating exposures as minimal, well-controlled, controlled, poorly controlled, and uncontrolled). Considerable heterogeneity within an exposure group may not matter if it is highly likely that all individual exposure profiles are appropriately controlled.[3] A similar reasoning applies to the exposure group where exposures are rated poorly controlled or uncontrolled: remedial action is necessary regardless.

For those in-between situations, where the group exposure profile appears to be controlled or borderline poorly controlled, differences in individual exposure profiles may be profound for some workers. It may be that only a fraction of the workers experience the majority of the overexposures. Identification of maximum-risk employees within the group, through personal sampling or the use of direct reading instruments, may lead to splitting the group into two or more groups, each containing workers with similar exposure profiles. ANOVA-based procedures, such as those used by Woskie et al.[19] and recommended by Rappaport,[15] should be considered when experiencing difficulty establishing reasonably similar exposure groups (see subsection titled Analysis of Repeat Measurements below).

Data Analysis—Descriptive and Compliance Statistics

When characterizing workplace exposures industrial hygienists are interested in accurately estimating the population parameters for the exposure profile associated with a particular work environment. These exposures may be specific to an individual employee or an exposure group.

Population parameters are almost always unknown and must be estimated from a sample of n measurements. Estimates of the population parameters are called "sample statistics" or "point estimates." Table 16.1 contains those population parameters and associated sample statistics that usually are of interest to industrial hygienists. Most of the statistics in Table 16.1 are considered "descriptive statistics," useful for characterizing the location and shape of the underlying distribution of exposures. Several statistics identified as "compliance" statistics are useful for determining whether a particular work environment is acceptable, unacceptable, or in need of further evaluation. (Here the term "compliance" is used to apply to the determination that an *exposure profile* complies with or conforms to some accepted definition of "acceptable." This use is different from the concept that compliance implies that each and every exposure must be less than the applicable federal OEL.)

Statistics can either be parametric or nonparametric. Parametric statistics are based on the assumption that the underlying distribution of exposures can be reasonably described by a known probability distribution function, such as the lognormal or normal distribution. Nonparametric statistics are not based on any distributional assumption and for this reason are sometimes called distribution-free statistics.

Parametric Statistics

Descriptive Statistics

Arithmetic Mean and Standard Deviation The sample mean ($\bar{x}$) is the most commonly used measure of central tendency. It is an unbiased estimate of the true population mean, regardless of the underlying distribution.

Geometric Mean and Geometric Standard Deviation Authorities are in general agreement that exposure profiles are often best described by the lognormal distribution.[3,4,6,14,15] The sample geometric mean represents an estimate of the median, or 50th percentile, of the expo-

TABLE 16.1 Exposure Profile Parameters and Sample Statistics of Interest to Industrial Hygienists

	Population Parameter	Sample Statistic
Descriptive	mean (μ)	sample mean ($\bar{x}$)
	standard deviation (σ)	sample standard deviation (s)
	geometric mean (GM)	sample geometric mean (gm)
	geometric standard deviation (GSD)	sample geometric standard deviation (gsd)
	median ($\tilde{\mu}$)	sample median ($\tilde{x}$)
Compliance	95th percentile ($X_{0.95}$)	sample 95th percentile ($x_{0.95}$)
	exceedance fraction (θ)	sample exceedance fraction (f)
	mean (μ)*	sample mean ($\bar{x}$) or MVUE*

*The 95% upper or lower confidence limit for the true mean of the exposure profile is compared to a long-term average OEL (LTA OEL), not the TWA OEL.

sure distribution. The sample geometric mean will always be less than the sample arithmetic mean. The sample geometric standard deviation is a measure of the spread or degree of dispersion in the data. It can be interpreted as the ratio of the 84th percentile to the 50th percentile (geometric mean) exposure, or the 50th percentile to the 16th percentile:

$$gm = \exp\left[\frac{1}{n}\sum_{i=1}^{n}\ln x_i\right] = \exp\left[\frac{1}{n}\sum_{i=1}^{n} y_i\right]$$

$$gsd = \exp\ (s_y)$$

where

$$s_y = \sqrt{\frac{\sum_{i=1}^{n}(\ln x_i - \ln gm)^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}}$$

The lowest GSD is a theoretical 1.0, indicating absolutely no variability in the log-transformed values (the exponential of zero is one). Exposure variability can be classified according to the following rules-of-thumb:

low-exposure variability	GSD ≤ 1.5
moderate-exposure variability	1.5 < GSD ≤ 2.5
high-exposure variability	GSD > 2.5.

An exposure profile for an exposure group may have a large GSD because it contains a number of dissimilar workers. True GSDs greater than 4 are unusual, particularly for individual workers. As a rule-of-thumb, sample GSDs of 3 or more should be checked to see if dissimilar workers or activities have been combined (e.g., indoor and outdoor activities), if there is seasonal variation, or simply too few data.

Minimum Variance Unbiased Estimator If the underlying distribution for the data is approximately lognormal, which is assumed to be the case with most exposure data, then the minimum variance unbiased estimator (MVUE) is the *preferred* point estimate of the true mean, particularly when the sample size is small and/or the sample geometric standard deviation is large.[20,21] The MVUE is calculated using the following formula:

$$\bar{x}_m = \exp(\bar{y})\ \psi\left(\frac{s_y^2}{2}\right) = gm \cdot \psi\left(\frac{s_y^2}{2}\right)$$

The ψ function is defined for any argument g as:

$$\psi(g) = \left[1 + \frac{(n-1)g}{n} + \frac{(n-1)^3 g^2}{n^2(n+1)2!} + \frac{(n-1)^5 g^3}{n^3(n+1)(n+3)3!} + \frac{(n-1)^7 g^4}{n^4(n+1)(n+3)(n+5)4!} + \ldots\right]$$

The above equation is easily calculated using a programmable calculator or personal computer. Calculation to at least five terms is accurate to the third decimal place.[21] (There is often little difference between the simple arithmetic mean and the MVUE. Because of its familiarity the arithmetic mean may be preferred for presentation and reporting purposes.)

Compliance Statistics

The exceedance fraction and 95th percentile statistics are useful for evaluating whether or not an exposure profile is acceptable, relative to some evaluation criterion, or OEL.

Exceedance Fraction A point estimate of the exceedance fraction (f), or fraction of exposures greater than the OEL, for lognormal distributed data can be calculated using the following equation:

$$f = P(c > OEL) = P(Z > z)$$

where

$$z = \frac{\ln OEL - \bar{y}}{s_y} \qquad (2)$$

and $\bar{y}$ and s_y are the sample mean and sample standard deviation of the log-transformed data calculated from a sample of n measurements. Equation 2 can be read as "the (sample) exceedance fraction equals the probability that a future concentration exceeds the OEL." This probability can be determined in the usual fashion by consulting a Z-value table found in any statistics textbook or by using the inverse z-function found in most computer statistics or spreadsheet programs.

95th Percentile Exposure The point estimate of the ith percentile of the underlying distribution for a sample of n measurements is estimated by

$$x_{1-\alpha} = \exp[\ln gm + Z_{1-\alpha} \cdot \ln gsd]$$
$$= \exp[\bar{y} + Z_{1-\alpha} \cdot s_y]$$

where α is the area under the distribution curve to the right of the ith percentile. When estimating the 95th percentile $\alpha = 0.05$ and the Z-value is replaced by $Z_{1-0.05}$, or 1.645.

Confidence Intervals

Sample statistics are rarely identical to the true population parameters. One can gain insight into the precision of the statistic or point estimate by calculating the 90% confidence interval around the sample statistic. The 90%

confidence interval can be thought of as the interval in which we will, 90% of the time, find the true value or population parameter. If the sample size is large and/or the variability in the exposures is low, then the confidence interval can be narrow, providing assurance that the point estimate is not far from the true value. However, if the sample size is small and/or the variability is large, then the confidence interval can be quite broad, suggesting that the true value may be considerably different from our point estimate.

The bounds of the 90% interval are the 95%LCL and 95%UCL. The 95%LCL and 95%UCL are useful in that they can be directly compared to target acceptable values. For example, if the 95%UCL is less than a target value, then one can be at least 95% confident that the true value of the statistic (the population parameter) is less than the target value. Conversely, if the 95%LCL is greater than a target value, then one can be at least 95% confident that the true value of the statistic is greater than the target value.

Arithmetic Mean Two procedures are presented. The first procedure can be used for sample data where it is assumed that the underlying distribution is normal. The second is preferred when it is assumed that the underlying distribution is lognormal.

Normal Distribution Assumption This procedure is fairly robust, that is, it works well for many non-normal distributions, especially as the sample size increases, producing reasonably accurate confidence limits.

Procedure:

(1) Calculate the sample mean ($\bar{y}$) and sample standard deviation (s).
(2) Calculate the 95% upper or lower confidence limit:

$$CL = \bar{x} + t \cdot \frac{s}{\sqrt{n}}$$

where $t = t_{0.95,n-1}$ for the 95%UCL and $t = t_{0.05,n-1}$ for the 95%LCL.

Taken together, the 95%LCL and 95%UCL form a 90% confidence interval for the true arithmetic mean when the data are normally distributed and an approximate 90% confidence interval when the distribution departs from normality.

Lognormal Distribution Assumption The following procedure was adapted from Land[22] and is preferred when the underlying distribution is assumed to be reasonably lognormal. (See Hewett[23] for a review of alternative procedures.)

Procedure:

(1) Calculate an estimate of the mean of the lognormal distribution:*

$$\bar{x}_m = \exp\left(\bar{y} + \frac{1}{2} s_y^2\right)$$

(2) Obtain the appropriate C-factor. Table A2a lists the C-factors necessary to estimate the 95%UCL: $C(s_y; n, 0.95)$. Table A2b lists the C-factors used to estimate the 95%LCL: $C(s_y; n, 0.05)$. Linear or Lagrange interpolation within and between sample sizes may be required.
(3) Calculate the 95% lower or upper confidence limit:

$$CL = \exp\left[\ln(\bar{x}_m) + C\frac{s_y}{\sqrt{n-1}}\right]$$

where

CL = the upper or lower confidence limit, depending upon the choice of C
$C = C(s_y; n, 1-\alpha)$ for the UCL where $\alpha = 0.05$
$= C(s_y; n, \alpha)$ for the LCL where $\alpha = 0.05$

Taken together, the 95%LCL and 95%UCL form a 90% confidence interval for the true arithmetic mean of lognormal distribution data. See Hewett and Ganser[24] for approximation formulae for calculating the appropriate C-factor. Also, interpolation can be avoided by using a table s_y value that is slightly greater than the calculated value. This will result in a confidence interval that is slightly wider than 90%.†

Exceedance Fraction The following procedure for calculating the 90% confidence interval around the exceedance fraction (for lognormally distributed data) was adapted from Odeh and Owen.[25]

*There are several formulae for estimating the mean, or average, of a lognormal distribution. See Reference 23 for a related discussion.

†Occasionally when the sample *GSD* is large and/or the sample size is small the 95%UCL for μ will be quite large, sometimes greater than the 95%UCL for the 95th percentile. While seemingly illogical, the 95% upper confidence limit for the true mean is still valid: 95% of the time the true mean will be less than the calculated 95%UCL. The best advice here is to reexamine the data and the exposure group definition. Perhaps high- and low-exposure jobs or tasks are represented in the dataset, resulting in an unusually large sample GSD. Otherwise, collect more data and recalculate.

Procedure:

(1) Calculate z using Equation 2
(2) Using z and the sample size, n, the 95% LCL for f can be read from Table A3. However, interpolation (linear or Lagrange) will usually be necessary to obtain reasonable accuracy. The 95%UCL can also be determined from Table A3. Obtain the table value using n and the negative of z. The 95%UCL for f is the complement of this value (complement = 1 − value).

Interpolation can be avoided by (a) using a simplified procedure[24] or (b) using a table z-value that is slightly greater than the calculated z-value for the LCL and UCL calculations, respectively. The latter option will result in a confidence interval that is slightly wider than 90%, but simplifies the calculations.

95th Percentile Exposure The 95%LCL and 95%UCL for the $x_{0.95}$ (assuming the exposure data are lognormally distributed) are easily estimated using K-factors developed by Odeh and Owen.[25] Unlike the exceedance fraction, the 95th percentile is not specific for any particular OEL. Consequently, it is useful when there are multiple exposure limits that can be applied (e.g., a substance may have an OSHA PEL, a NIOSH REL, and an ACGIH TLV).

Procedure:

(1) Calculate the sample mean ($\bar{y}$) and sample standard deviation (s_y) of the log-transformed data where $y = \ln(x)$.
(2) Calculate the 95% lower or upper confidence limit:

$$95\%\text{LCL} = \exp\lfloor \bar{y} + K_{0.05,0.95,n} \cdot s_y \rfloor$$

$$95\%\text{UCL} = \exp\lfloor \bar{y} + K_{0.95,0.95,n} \cdot s_y \rfloor$$

See Table A4 for the appropriate 95%LCL and 95%UCL K values. (The K values for the 95%LCL are nearly identical to the K' (K prime) values described by Tuggle.[26]) Taken together, the 95%LCL and 95%UCL form a 90% confidence interval for the true 95th percentile ($X_{0.95}$), assuming a lognormal distribution. Note that the 95% UCL for the sample 95th percentile exposure is identical to the 95% upper tolerance limit recommended by several authorities.[3,6,27,28]

Nonparametric Statistics

There may be situations where there is compelling evidence that the lognormal distribution assumption does not apply, for example, when the log-probability plot is far from linear and the data fails a formal goodness-of-fit test such as Filliben's test presented above. In this situation the preferred approach would be to evaluate the data and supporting documentation for indications that the exposure group definition needs to be reevaluated. Perhaps several distinctly different exposure groups are represented in one larger group. Another approach would be to apply nonparametric procedures for estimating the median, 95th percentile, and exceedance fraction. One disadvantage to using nonparametric, or distribution-free, statistics is that the confidence intervals are wider than those estimated assuming a particular distribution, such as the lognormal or normal distribution. Perhaps this is why nonparametric statistics are not often reported in the industrial hygiene literature. Readers interested in nonparametric statistics should review the recommendations by Rock,[29] Esmen,[30] and the EASC.[3,6] A good general reference is that of Conover.[31]

Descriptive Statistics

The first step in a nonparametric analysis is to sort the data from low to high values. The ith ordered observation will be referred to as x_i.

Mean The simple arithmetic mean is an unbiased estimate of the center of mass of a distribution, regardless of the shape of the distribution.

Median The 50th percentile of any distribution is known as the median. If the sample size is odd, then the sample median is simply the middle value. If the sample size is even, the sample median is the average of the two middle values. Specifically:

$$\text{for } n = 2k+1 \text{ (odd } n\text{): } \tilde{x} = x_{k+1}$$

$$\text{for } n = 2k \quad \text{(even } n\text{): } \tilde{x} = \frac{1}{2}(x_k + x_{k+1})$$

The median exposure is one measure of central tendency of the exposure profile. It has no current use in determining compliance.

Compliance Statistics

Exceedance Fraction The observed, or nonparametric, exceedance fraction ($\tilde{f}$) is simply the ratio of the number of overexposures (i.e., measurements greater than the OEL) to the sample size: $\tilde{f} = m/n$, where m = number of overexposures and n = sample size. The observed exceedance fraction can often be zero for small sample sizes,

even when the actual probability of an overexposure is quite large.*

95th Percentile Exposure The nonparametric 95th percentile can be estimated for $n \geq 20$. Linear interpolation between rank-ordered data is usually necessary to estimate the 95th percentile exposure.[32,33]

$$\tilde{x}_{0.95} = x_i + (0.95n - i)(x_{i+1} - x_i)$$

where i = integer portion of $0.95 \cdot n$

Confidence Intervals

Median For $n \leq 30$ confidence intervals for the median (50th percentile) exposure can be calculated using the following procedure.[31] Let q equal the quantile (percentile) of interest, 0.5 in this case. The ordered rank of the 95%LCL, l, is determined by finding l such that

$$\sum_{x=0}^{l-1} b(x; n, q) = 0.05 \qquad (3)$$

where $b(x; n, q)$ is the binomial function: $b(x; n, q) = \binom{n}{x} q^x (1-q)^{n-x}$, $x = 0, 1, 2, \ldots, n$. The ordered rank of the 95%UCL, u, is determined by setting q to 0.5 and finding u such that

$$\sum_{x=0}^{u-1} b(x; n, q) = 0.95$$

Table A5 contains the rank of the 95%LCL and 95%UCL for the sample median for sample sizes ranging from 5 to 30. If $n > 30$, the following formulae can be used to estimate the rank of the 95%LCL or 95%UCL value, where q equals 0.5, and $Z = 1.645$[32,34]:

$$l = q(n+1) - Z\sqrt{nq(1-q)} \qquad (4)$$

$$u = q(n+1) + Z\sqrt{nq(1-q)} \qquad (5)$$

The resulting rank—l or u—should be rounded to the next smaller or larger integer, respectively. Alternatively, linear interpolation between the ordered values can be used to determine the 95%LCL or 95%UCL. Equations 4 and 5 are conservative, thus resulting in a confidence interval slightly wider than 90%.

Exceedance Fraction Confidence limits for the nonparametric exceedance fraction can be determined from Table A6 for $n \leq 30$. Confidence intervals for larger sample sizes are determined by finding θ (pronounced "theta"; $0 < \theta < 1$) such that the following equations are true[35]:

$$\sum_{x=m}^{n} b(x; n, \theta_{LCL}) = 0.05$$

$$\sum_{x=0}^{m} b(x; n, \theta_{UCL}) = 0.05$$

For $n > 30$ the following equations can be used to calculate reasonably accurate confidence intervals.[32,36] The 95%LCL and 95%UCL is estimated by substituting 1.645 for Z in the following equations:

$$\theta_{LCL} = \frac{1}{n+Z^2} \cdot \left\{ (m-0.5) + \frac{Z^2}{2} - Z \cdot \sqrt{(m-0.5) - \frac{(m-0.5)^2}{n} + \frac{Z^2}{4}} \right\}$$

$$\theta_{UCL} = \frac{1}{n+Z^2} \cdot \left\{ (m+0.5) + \frac{Z^2}{2} + Z \cdot \sqrt{(m+0.5) - \frac{(m+0.5)^2}{2} + \frac{Z^2}{4}} \right\}$$

except if

$m = 0$, then $\theta_{LCL} = 0$

$m = 1$, then $\theta_{LCL} = 1 - (1-\alpha)^{1/n}$ where $\alpha = 0.05$ for the 95%LCL(θ)

$m = n$, then $\theta_{UCL} = 1$.

95th Percentile The 95%LCL can be calculated for sample sizes greater than 4 (although 19 measurements are necessary before the point estimate can be calculated). In contrast, the sample size has to exceed 58 before the more useful 95%UCL can be calculated. The 95%LCL can be determined using Equation 3: let $q = 0.95$ and determine l such that the equation is true. This procedure was used to determine the values in Table A5. This table contains the rank (r) of the 95%LCL for the sample 95th percentile for sample sizes ranging from 5 to 30. If $n > 30$, Equation 4 is adequate for estimating the rank of the 95%LCL[32]: let $q = 0.95$ and $Z = 1.645$ and calculate l. If $n > 58$, Equation 5 may be used in a similar manner to estimate the rank of the 95%UCL.

Comments Confidence intervals for the nonparametric 95th percentile require large sample sizes, particularly for the 95%UCL, thus reducing their usefulness for risk-management decision making. In contrast, 90% confidence intervals for the exceedance fraction can be calculated for virtually any sample size. Note that these confidence intervals will often be considerably broader

*For large sample sizes, say $n > 30$, the calculated and observed exceedance fraction should be similar if the data are approximately lognormal.

than those estimated using the parametric lognormal assumption. Consequently, nonparametric statistics should only be used when one is confident in rejecting the lognormal (or normal) distribution assumption. Reevaluating the exposure group definitions and the range of tasks or activities within the exposure group may lead to new exposure groups where the lognormal assumption is reasonable.

Data Analysis Issues

Censored Data

Measurements less than the limit of detection (LOD) (or limit-of-quantification) often occur, particularly for well-controlled work environments and when the exposure limit is close to the LOD. Simply ignoring the LOD values will bias the sample GM upward and the sample GSD downward. There are various substitution techniques for estimating distribution parameters from datasets containing censored data: each censored datum is simply replaced with the (1) LOD, (2) half of the LOD, or (3) LOD divided by $\sqrt{2}$. The CEN[4] recommended using half of the LOD. Simple substitution works well when the percentage of LOD values is small. More sophisticated techniques generally produce more accurate sample estimates, but require tables and intermediate calculations. The reader is advised to consult the references before adopting any particular scheme.[32,37,38]

Inconsistent Data Points

Inconsistent data points, or what some might call outliers, should be carefully considered. A true outlier represents a gross error in either sampling or analysis, or an aberrant condition in the work environment that is clearly unlikely to be repeated. Unless strongly justified, inconsistent data should be not be eliminated from a dataset. Such values most likely are part of the actual exposure profile and reflect exposure conditions that should be investigated.[3,7]

Analysis of Repeat Measurements

Exposure variability within an exposure group can be divided into within-worker and between-worker components.[3,15] Within-worker factors that influence exposure variability include differences in assigned tasks or time at task, work practices, and individual exposure controls. Between-worker factors include production level variation and general ventilation. There is growing interest in the use of analysis-of-variance (ANOVA) or components-of-variance techniques to estimate these within- and between-worker components of overall variability. Rappaport[15] illustrates the calculation techniques and advocates their use in assessing the degree of heterogeneity within exposure groups. Woskie et al.[19] described the application of similar techniques to short-term (within-shift) and full-shift measurements. Such analyses can guide decisions to focus on modifying general engineering/ventilation controls, or to focus on modifying individual engineering/ventilation controls or work practices. However, these techniques require the collection of repeat measurements for each of n randomly selected workers in each exposure group. Others have described ANOVA-based techniques for assessing compliance.[39,40] These proposals merit consideration, but apply only to true long-term average OELs. In summary, when repeat measurements are available for workers, the techniques described in these and similar papers may be useful for gaining insight into the relative contribution of the various sources of workplace variability and help direct intervention efforts.

Data Interpretation

The underlying goal of any data interpretation scheme is to *reliably* determine, given the available data, that the exposure profile of *each* worker in *each* exposure group is either acceptable or unacceptable for the current observation period.[1] Often an industrial hygienist may defer a final decision until additional information or data can be collected. However, unless action is taken to reduce exposures, a decision to defer is by default a decision that the work environment is acceptable. It is generally accepted that decisions in the presence of uncertainty should always be made in favor of the workers.[3]

For those exposure profiles judged acceptable, the decision should be made with a high degree of confidence. How much confidence is needed depends on the numerous factors involved. For example, if resampling intervals are close, then less confidence can be tolerated during each survey, such as when using a control chart approach or a simple decision scheme. If resampling occurs infrequently, then a high degree of confidence is needed such as provided by a formal statistical test using the 95%UCL. If the OEL is uncertain, such as when a working or provisional[1] OEL is used or the legal OEL is dated and no longer considered protective, then a high degree of confidence is needed. The same can be said for situations where the substance is particularly toxic or the toxicological information is uncertain.

Data interpretation schemes generally fall into one of three categories: nonparametric decision logics, parametric decision logics, and control chart techniques.

Nonparametric Decision Logics

Nonparametric decision logics are based on simple decision rules or the use of non-parametric statistics. Rule-based decision logics require few, if any, statistical calculations, and no underlying distribution is assumed.

Rule-Based Decision Logics

Historically, a common procedure for determining the acceptability of a "work environment" is to collect one or several measurements from one or more maximum-risk employees and apply simple decision rules. As the quantity of exposure data has increased, industrial hygienists have adopted ever more sophisticated and statistically defensible data analysis procedures. However, simple decision rules have a legitimate place in federal regulations[41] and when designing exposure surveillance programs (where one or a few exposures over the OEL or an action limit triggers a more comprehensive evaluation).

The 1977 NIOSH[42] sampling strategy and decision logic is summarized in Table 16.2. Notice that a decision can be rapidly reached with one or a few measurements. NIOSH introduced the concept of an "Action Level," or half of the OEL. Measurements above the Action Level, but below the OEL, triggered additional sampling. This scheme permitted rapid and efficient decision making. However, as noted by Tuggle,[27] the decision may not be the right decision, even when the true exceedance fraction is 0.25 or greater.

More recently the CEN[4] described the use of simple decision rules. For datasets containing less than six measurements, collected from a specific "homogeneous exposure group," *each* measurement should be compared to the OEL. A single overexposure should trigger an investigation and remedial action, if necessary. The CEN provided several examples for determining when to resample and how many measurements to collect. One example is summarized in Table 16.3.

Nonparametric Statistical Analysis

Nonparametric statistical analysis has been recommended for those instances where one cannot assume that the data are derived from a lognormal distribution.[3,6,29,30] In general, larger sample sizes are needed before statistical significance is reached. Because of this limitation, judgment is often needed when balancing the expense of further sampling against the expense of installing additional controls.[3,30]

If the nonparametric 95th percentile exposure is less than the OEL, then one has evidence that the true 95th percentile is less than the OEL, but the confidence level is unknown. At least 20 measurements are needed just to estimate the 95th percentile, while 58 measurements are necessary to estimate the corresponding 95%UCL. Is this a reasonable use of resources? Perhaps attention should be given to defining exposure groups where parametric statistics can be used.

If the nonparametric exceedance fraction ($\bar{f} = m/n$) is less than 0.05, then again one has evidence that the true exceedance fraction is less than 0.05. However, nonparametric exceedance fractions can be extremely misleading when the sample size is small. For example, consider a dataset where $n = 2$ and both measurements are below the OEL. The point estimate of the exceedance

TABLE 16.2 Example Decision Logic for Small Sample Sizes

Sampling Strategy:

Collect one or more measurements (C) from one or more maximum risk employees (MRE). Compare each measurement to the Permissible Exposure Limit (PEL) and the Action Level (AL = 0.5 · PEL).

Decision Rules	Decison/Action
1. If C < AL	then conclude the work environment is *acceptable* for the exposure group presented by the MRE.
2. If C > PEL	then conclude that the work environment is *unacceptable*, take corrective action, and re-sample at least monthly.
3. If AL ≤ C ≤ PEL	then collect an additional measurement at least every two months until either two consecutive measurements are less than half the PEL (and conclude that the exposures are acceptable) or any single measurement is above the PEL (and conclude that exposures are *unacceptable* and take appropriate actions to reduce exposures).

Note. Adapted from NIOSH.[42]

TABLE 16.3 Example Decision Logic for Small Sizes

Sampling strategy:

Collect a single measurement every two months from each exposure group. Compare *each* measurement to four action limits:
N1 = 0.40 · OEL
N2 = 0.70 · OEL
N3 = 1.00 · OEL
N4 = 1.50 · OEL

Decision Rules	Decision/Action
1. If $C \leq N1$ twice consecutively	then collect one measurement every 6 months.
2. If $C \leq N2$	then continue collecting one measurement every two months.
3. If $N2 < C \leq N4$	then collect two measurements every two months.*
4. If $N2 < C \leq N4$ twice consecutively	then collect two measurements every two months for eight months.†
5. If $N3 \leq C \leq N4$ twice consecutively	then immediate action is warranted to reduce exposures.
6. If $C > N4$	then immediate action is warranted to reduce exposures.

Note. Adapted from CEN.[4]
*Every overexposure (C > N3) should be investigated, the reasons for the overexposure identified, and appropriate measures taken.
†If the two measurements were collected on the same survey, then immediate action is warranted to reduce exposures.

fraction is 0.0, but because the 95%UCL is 0.777 (see Table A6) we have little confidence that the true exceedance fraction is less than 0.05.

Parametric Decision Logics

In 1977 NIOSH[42] stated that the goal of an effective exposure-monitoring program is to "attain 95% confidence that no more than 5% of employee days are over the standard." NIOSH was not implying that it is permissible to overexpose each employee once every twenty shifts, but was merely providing a "statistical" goal for designing an exposure monitoring program. A similar goal has been suggested by numerous organizations,[3,4,6] authoritative individuals,[10,29,43,44,45] and corporations.[46,47]

According to the AIHA[3,6] the exposure profile for a similar exposure group is usually deemed acceptable if it is highly likely that only a small percentage of the measurements exceed the OEL. This is the case if the parametric or nonparametric upper confidence limit on the 95th percentile is less than the OEL, or the upper confidence limit on the exceedance fraction is less than 0.05.

If six or measurements are randomly collected from an exposure group, the CEN[4] recommends that one use statistics to estimate the probability of overexposure for individuals within the HEG. If this probability exceeds 5%, then corrective action should take place. Otherwise, periodic monitoring should be used to confirm that the point estimate of the probability of overexposure remains less than 5%.

The above recommendations are pertinent to TWA OELs. When dealing with long-term average OELs (LTA OEL; see Chapter 15) some researchers advocate the use of ANOVA techniques when analyzing exposure measurements.[39,40] Such techniques require the characterization of the distribution of individual long-term mean exposures *within* an exposure group. This necessitates collecting multiple measurements from each of n randomly selected workers per exposure group. ANOVA techniques are then used for determining the probability that the long-term mean of any single worker in the exposure group exceeds LTA OEL.[15] Rappaport et al.[39] suggest that the exposure profile of the exposure group may be judged acceptable if this probability is 0.10 or less. Tests for comparing the average exposure for an exposure group or individual worker have also been described.[48]

Consistent with these authoritative recommendations, Table 16.4 contains generic criteria for determining whether a dataset suggests that the exposure profile for a particular exposure group or work environment is currently acceptable or unacceptable relative to a TWA OEL or LTA OEL. The evaluation statistic utilized can be either the point estimate of the exposure parameter of interest or its associated upper or lower confidence limit. For example, if the point estimate of the 95th percentile exposure is less than the TWA OEL, then one has evidence that the work environment is acceptable, but not necessarily compelling evidence. Compelling evidence exists when the 95% upper confidence limit for the 95th percentile is less than the TWA OEL. A similar logic ap-

TABLE 16.4 Evaluation Criteria for Testing Whether a Dataset Represents "Acceptable" or "Unacceptable" Exposure Conditions

Evaluation Criteria			
TWA OEL or STEL OEL	LTA OEL	Exposure Profile	Appropriate Action
95%UCL($x_{0.95}$) ≤ OEL or 95%UCL(f) ≤ 0.05	95%UCL($\bar{x}$) ≤ LTA OEL	(clearly) acceptable	Periodically re-sample.
$x_{0.95}$ ≤ OEL or f ≤ 0.05	$\bar{x}$ ≤ LTA OEL	acceptable	Periodically re-sample.
$x_{0.95}$ > OEL or f > 0.05	$\bar{x}$ > LTA OEL	unacceptable	Take steps to reduce exposures. Re-sample.
95%LCL($x_{0.95}$) > OEL or 95%LCL(f) > 0.05	95%LCL($\bar{x}$) > LTA OEL	(clearly) unacceptable	Take *immediate* steps to reduce exposures. Re-sample.

plies to the exceedance fraction and to the long-term mean (when a LTA OEL applies).

Note that concluding that an exposure profile is currently "acceptable" relative to an OEL permits one to *state* that the exposure profile "appears to be controlled." If the 95%UCL is compared to the OEL one could state that the exposure profile "is controlled, with at least 95% confidence." The OEL referred to in Table 16.4 should be replaced by either half or a tenth of the actual OEL in order to evaluate whether or not the true exposure profile can be rated "well-controlled" or "highly controlled" (see Table 15.3, Chapter 15[1]).

Control Chart Techniques

Control charts or time series plots can be used for recognizing trends[4,32] and cycles, to assess the stability of the work environment, and for visually comparing measurements to OELs and Action Limits. Roach et al.[2] and Roach[10] recommended the use of time series plots coupled with simple decision rules. For example, a warning line could be established at or below the OEL. Measurements above the warning line should elicit a repeat visit. Two measurements in a row above the warning line should result in immediate action. The CEN[4] also recommended a time-series plot for routinely collected measurements. Both the individual measurements and the moving average is plotted against time. In this manner the acceptability of individual values can be assessed, as well as any trends toward higher exposures.

The combination of time-series plots and simple decision rules should be considered when dealing with dynamic work environments where significant change, as part of the normal work process, is expected. Such techniques are also suited to work environments that are rated "controlled" (see Table 15.3, Chapter 15[1]). This is because continual surveillance and continuous improvement are necessary until such time that statistical tests permit the exposure profile to be rated "well-controlled" or "minimal."

The application of statistical process control techniques to the control of exposures in the work environment is encouraged. The TWA OEL could serve as an upper specification limit that, in principle, should not be exceeded. For practical purposes, the upper specification limit could be defined as the 95th percentile of a "controlled" exposure profile. As measurements accumulate the annual (or less) long-term mean should be one-third or less of the TWA OEL, and not be permitted to exceed half the single-shift limit. It is even conceivable for stable processes to calculate traditional control limits for gauging drift and trends. As exposures are controlled, the upper specification limit could be progressively reduced, encouraging further reductions in exposure, thereby minimizing risk to the workers and the future liability risk of the employer.

Data Interpretation Issues

Interpretation of a Single Exposure Measurement

Authoritative sources are unanimous in recommending that *each* overexposure be *investigated*, regardless of the level of sophistication of the exposure-monitoring program or the past exposure history[4,6] (see related discussion in Ref. 41). If there is compelling or convincing past exposure data to suggest that the overexposure is most likely a random occurrence in an otherwise controlled exposure profile, then it is reasonable to take no action beyond merely documenting the investigation. However, if no rational explanation can be found for the overexposures, one is compelled to conclude that a systematic change of some sort *may* have occurred; after all, in a

"controlled" work environment overexposures should be infrequent to rare. Follow-up actions may consist of fine tuning existing controls, installation or modification of controls, or an evaluation of individual work practices. Regardless, additional measurements are usually warranted in order to verify the need for additional controls or to evaluate the effectiveness of any intervention.

As a rule-of-thumb, one should always be suspicious when one or more measurements exceeds 50% of the OEL in a small dataset, particularly when $n < 6$. This suggests that the exceedance fraction of the current exposure profile may be unacceptable (i.e., greater than 0.05).

Dual Limits

Many substances have dual limits; that is, both a TWA OEL and an STEL OEL. OSHA[49] noted that if the full-shift, TWA exposure is less than the PEL, then the short-term exposures are also likely controlled. Along similar lines, Spear et al.[50] concluded, using mathematical analysis, that the exceedance fraction for short-term, 15-minute average exposures will be limited to 5% or less if the exceedance fraction for full-shift, TWA exposure is limited to 5% or less. However, one cannot always rely upon a satisfactory TWA measurement to indicate that within-shift excursions are also acceptable. Short-term exposures should always be evaluated when there are predictable or recognizable within-shift cycles and episodes of high exposure.

In Great Britain vinyl chloride is subject to a 7 ppm single shift limit and a 3 ppm annual average limit.[51] In the United States there are no examples of explicit dual limits that consist of a TWA OEL and a LTA OEL. However, in a "controlled" work environment the long-term average exposure for each individual worker should be much less than the OEL, particularly for chronic disease agents, and the single shift limit should be infrequently exceeded, if at all. As discussed in Hewett[1] (Chapter 15), it is reasonable to devise a provisional LTA OEL using the AIHA recommendation of one third the single shift limit (or TWA OEL). The long-term average should be less than the LTA OEL and at no time exceed one half of the single shift limit.

Examples

Example 1

Table 16.5 contains total welding fume data collected in three welding departments at an automobile frame manufacturing plant. The data represent measurements collected on a single day in 1987 from randomly selected welders. Each measurement represents the full-shift, 8-hour TWA exposure to medium-steel welding fumes for each of n welders in the three departments. The applicable OEL is the 1987 ACGIH TLV of 5 mg/m^3 for welding fumes (total particulate,* not otherwise classified). Measurements were also collected on other days during this week as part of a research study,[52] but the data from a single day will be used to illustrate the calculations presented in this chapter. Assume that these data were collected as part of an initial baseline survey of these welding operations. It was not possible to monitor all welders, so a reasonable number were selected from each department. Since all the welders in each department were engaged in similar tasks, used similar welding consumables and equipment, and were subject to similar ventilation controls, it was logical to initially consider each "department" as an exposure group. What then can be said from an analysis of these data?

Goodness-of-Fit

Figure 16.1 shows the log-probit curves for these data. This graph was created using a standard computer spreadsheet program. The x-axis can be displayed in "normal" or "log" terms. The "log" x-axis is usually selected under the usual assumption that the data are lognormal distributed. The y-axis is in terms of the z-value equivalent of the plotting position (see Goodness-of-Fit section). In all three cases, the data appear "reasonably" linear with no obvious outliers or data points that need individual investigation. All three datasets pass the revised Filliben's goodness-of-fit test: the calculated correlation coefficient, r, exceeded the appropriate critical value found in Table A1. Since neither the subjective graphical evaluation nor the objective goodness-of-fit test suggest that the data do not come from a lognormal distribution, we may calculate descriptive and compliance statistics with reasonable assurance that our point estimates will be valid. Of course, with limited data, say n less than 10, it is difficult to reject any distributional assumption.

Descriptive Statistics

All three departments exhibited moderate exposure variability as indicated by the sample GSDs, which ranged from 1.56 to 1.67 (see Table 16.6). The sample GMs

*Since welding fumes are typically smaller than 1 μm in diameter, the "total particulate" specified in 1987 is essentially equivalent to the "inhalable particulate" specified in the 1997 TLV booklet.

TABLE 16.5 Data Used in the Examples (Sorted in Ascending Order)

Automobile Frame Manufacturing (welding fumes, mg/m³)[52]			Chemical Plant (inorganic lead, μg/m³)[53]		
Department B	Department C	Department E	Worker A		
0.21	1.63	6.39	3.9	12.4	21.5
0.42	2.02	6.89	7.9	12.9	21.9
0.49	2.04	9.59	8.6	13.0	22.2
0.58	2.32	10.89	9.0	14.4	24.6
	4.28	19.97	9.0	15.0	25.4
	6.04		9.5	15.9	25.6
			10.0	17.1	25.7
			10.0	18.6	28.9
			10.2	19.1	30.4
			10.4	19.5	34.0
			11.3	19.6	46.9
			11.4	20.2	56.4

are useful for estimating the median, or 50th percentile exposure, of a lognormal exposure distribution. The sample arithmetic mean is useful for indicating whether or not the long-term average exposure will be controlled to the ideal range of less than one-third of the OEL. In this example only Department B had a sample mean near this range. The confidence intervals about the sample mean were quite broad, reflecting the small sample size. Consequently, we have to interpret the estimates of the sample mean with some caution, as the true values may be considerably different from the sample estimates. It will require periodic monitoring over a period of a year or more to determine conclusively whether or not the long-term mean for Department B is routinely maintained in this range.

If the underlying distributions are reasonably lognor-

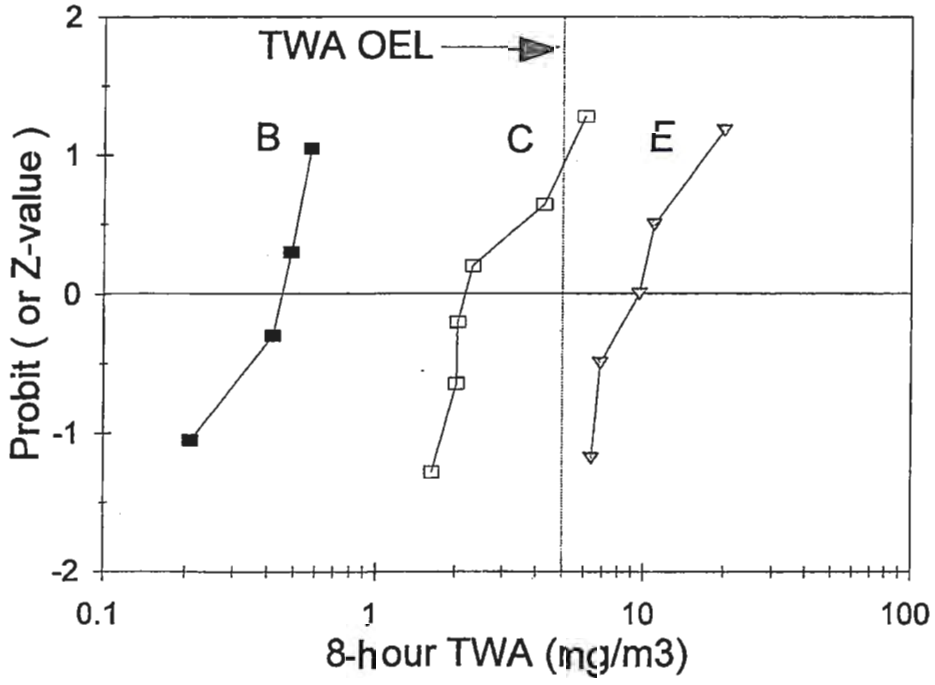

FIGURE 16.1 Log-probit curves for Example 1.

TABLE 16.6 Parametric Descriptive and Compliance Statistics for Examples 1 and 2 (the 90% confidence intervals are given in parentheses)

	Auto Frame Manufacturing Plant[52] (welding fumes, mg/m³)			Chemical Plant[53] (inorganic lead, μg/m³)
	Department B	Department C	Department E	Worker A
n	4	6	5	36
min	0.21	1.63	6.39	3.9
max	0.58	6.04	19.97	56.4
Descriptive statistics				
$\bar{x}$	0.43	3.06	10.75	18.7
	(0.29–1.06)	(2.15–5.66)	(7.58–20.60)	(16.1–22.4)
gm	0.40	2.72	9.83	16.2
gsd	1.56	1.67	1.57	1.73
mvue	0.43	3.02	10.66	18.7
Compliance statistics				
f_{OEL}	<0.01	0.12	0.93	0.02
	(<<0.01–0.03)	(0.02–0.40)	(0.63–0.99)	(<0.01–0.06)
$x_{0.95}$	0.83	6.29	20.75	39.7
	(0.55–3.95)	(4.25–18.01)	(14.25–66.33)	(32.5–52.5)
Goodness-of-fit				
log-probit graph	reasonably lognormal	reasonably lognormal	reasonably lognormal	lognormal
Filliben's test	0.933 ≥ 0.868*	0.937 ≥ 0.889*	0.956 ≥ 0.880*	0.987 ≥ 0.969*
r (lognormal)	∴ lognormal	∴ lognormal	∴ lognormal	∴ lognormal

*Critical value. If *r* is less than the critical value, then there is evidence that the underlying distribution is not lognormal.

mal, then the MVUE is the best point estimate of the true mean. In most cases, it will be similar to the simple arithmetic sample mean, but may vary substantially when the sample size is small and the sample GSD is large.

Compliance Statistics

The sampled welders working in Department B experienced exposures all of which were less than 12% of the OEL. In contrast, the sampled welders in Department E all experienced exposures greater than the OEL. In Department C all of the measurements were greater than 10% of the OEL, with one out of the six exceeding the OEL.

For Department B the compliance statistics suggest the point estimate of the exceedance fraction is virtually zero, with a 95%UCL of 0.03. Even with this small sample size we can be more than 95% confident that the true exceedance fraction is less than 0.05. Similarly, both the point estimate of the 95th percentile and its 95%UCL were less the OEL. For Department C the point estimate of the exceedance fraction is 0.12 while the point estimate of the 95th percentile exceeds the OEL. For Department E the 95%LCLs for the 95th percentile and exceedance fraction exceeded the TWA OEL and 0.05, respectively.

Conclusions and Recommendations

With small sample sizes such as these all of the measurements should be less than the OEL for one to declare that the exposure profile appears "controlled" and that the work environment is currently "acceptable" (see Table 16.4). Even so, the confidence intervals will often be broad with 95%UCL values well into the unacceptable range. This necessitates occasional remonitoring to verify or validate the initial assessment and to determine if higher risk employees were missed in the baseline survey.

Exposures in Department B *appear* to be "well-controlled" (see Hewett,[1] Chapter 15, Table 15.3, for the statistical interpretations of these terms). Because the 95%UCL for the 95th percentile is less than the TWA OEL, we can *conclude*, with at least 95% confidence, that the exposure profile for this exposure group is "controlled." Occasional remonitoring and reassessment should take place in order to confirm this assessment.

preferably with a mix of repeat sampling and sampling of other welders.

Regarding Department C, exposures routinely were substantial (i.e., greater than 10% of the OEL), with a point estimate of 12% overexposures. Steps should be taken to reduce exposures and to evaluate others in this exposure group. It is possible that adjusting existing controls and modifying individual work practices will be sufficient to permit, after follow-up sampling, a rating of "controlled" or better. However, periodic remonitoring will probably be necessary until the stability of the process can be ascertained and an exposure history is developed over a period of several years.

Exposures in Department E are obviously unacceptable without any calculations. We can expect approximately 93% of the exposures in this department to exceed the OEL. Immediate action is warranted, to include the use of interim respiratory protection. No further monitoring is necessary until improvements have been made in either the ventilation system and/or work practices.

Observations

As a general rule, it is recommended that one have at least 6–10 measurements before using statistical techniques for characterizing an exposure profile.[4,6] While fewer than six measurements often lead to highly variable *estimates* (i.e., statistics) of the true exposure profile parameters, it is often possible to reach a highly accurate *decision* regarding the question "Is the exposure profile acceptable?" Here we see that for Departments B and E sound conclusions can be reached, despite the fact that less than six measurements were used. If exposure profiles can be controlled to the point that the exposures can be rated "well-controlled" or "highly controlled" and the process is relatively stable, then the future sampling burden will be minimal.

The nonparametric descriptive and compliance statistics in Table 16.7 provide several interesting comparisons to Table 16.6. For Department B we calculated a parametric 95%UCL for the exceedance fraction of 0.03. If we used nonparametric statistics, we would be compelled to accept the proposition that the true exceedance fraction might be as great as 0.53, despite the fact that the maximum value was only 12% of the TWA OEL. Furthermore, the nonparametric 95th percentile cannot be estimated with such small sample sizes.

Example 2

Table 16.5 also contains inorganic lead personal exposure data collected at an alkyl lead manufacturing plant.[53] The measurements were collected from a single worker over the course of 6 weeks. These data were collected as part of a research project, but are useful as an example of the analysis of a larger dataset collected over a longer span of time. We will use the 1978 OSHA PEL of 50 $\mu g/m^3$ as the relevant OEL.

Goodness-of-Fit

Figure 16.2 shows the log-probit plot of the data. The regression line was determined by simple linear regression of the log-transformed concentration values and the corresponding z-value. The data appear to follow the regression line suggesting that the lognormal distribution assumption is reasonable. This assessment is corroborated by Filliben's test where the correlation coefficient for the lognormal assumption was greater than the critical value. Consequently, we can conclude that there is no reason to reject the lognormal distribution and proceed to calculate parametric descriptive and compliance statistics.

Descriptive Statistics

Here we have a substantial number of measurements, permitting us to calculate reasonably accurate estimates of distribution parameters (Table 16.6). If the process is reasonably stable, these estimates should have considerable predictive value. The sample GM suggests that the true median will be approximately 16 $\mu g/m^3$. The sample GSD of 1.73 suggests a moderately variable work environment. The average exposure is 37% of the OEL, close to our goal of one-third of the TWA OEL or less.

Compliance Statistics

The point estimate of the exceedance fraction is 0.02, but the 95%UCL is slightly above 0.05. Similarly, the point estimate of the 95th percentile is less than the OEL, but the 95%UCL slightly exceeds the OEL. In comparison, the observed, or nonparametric, exceedance fraction (see Table 16.7) was 0.03 (one of 36 measurements exceeded the OEL). If the underlying distribution is truly lognormal, the observed and calculated exceedance fractions should, with increasing sample size, converge to the same value.

Conclusions and Recommendations

This exposure profile *appears* to be "controlled"; that is, the point estimates of the parametric 95th percentile and exceedance fraction are less than the target values. However, the 95%UCLs are slightly greater than the target values. Consequently, regular monitoring is necessary to confirm the designation of "controlled" exposure

TABLE 16.7 Nonparametric Description and Compliance Statistics for Examples 1 and 2

	Auto Frame Manufacturing Plant[52] (welding fumes, mg/m³)			Chemical Plant[53] (inorganic lead, µg/m³)
	Department B	Department C	Department E	Worker A
Descriptive statistics				
$\tilde{x}$ (median)	0.455 (NA–NA)	2.18 (1.63–6.04)	9.59 (6.39–19.97)	16.5 (12.4–20.2)
Compliance statistics				
$\tilde{f}_{OEL} = m/n$	0.00 (0.00–0.53)	0.17 (0.01–0.58)	1.00 (0.54–1.00)	0.03 (<0.01–0.13)
$\tilde{x}_{0.95}$ (95th percentile)	NA $n < 20$	NA $n < 20$	NA $n < 20$	34.1 (28.9–NA)*

The 90% Confidence Intervals are given in parentheses. In addition, the sample size, minimum and maximum value, and arithmetic mean should also be reported.
*Because $n < 58$, only the 95%LCL can be estimated.

profile. Notice that even with 36 measurements the 95%UCL for the observed (i.e., nonparametric) exceedance fraction (see Table 16.7) is much larger than the parametric 95%UCL.

Analysis of work practices may reduce peak exposures within each shift, thus reducing the full-shift TWA exposures. In this manner, the long-term mean could be reduced to below the target value of one-third the TWA OEL. Furthermore, the 95%UCL for the either the 95th percentile or exceedance fraction may then fall below the target values—the TWA OEL and 0.05, respectively—permitting one to *conclude*, with at least 95% confidence, that the exposure profile is "controlled," or even "well-controlled."

Observations

These measurements were collected sequentially from a single worker; therefore, without additional information, our conclusions apply only to this worker. Let's assume

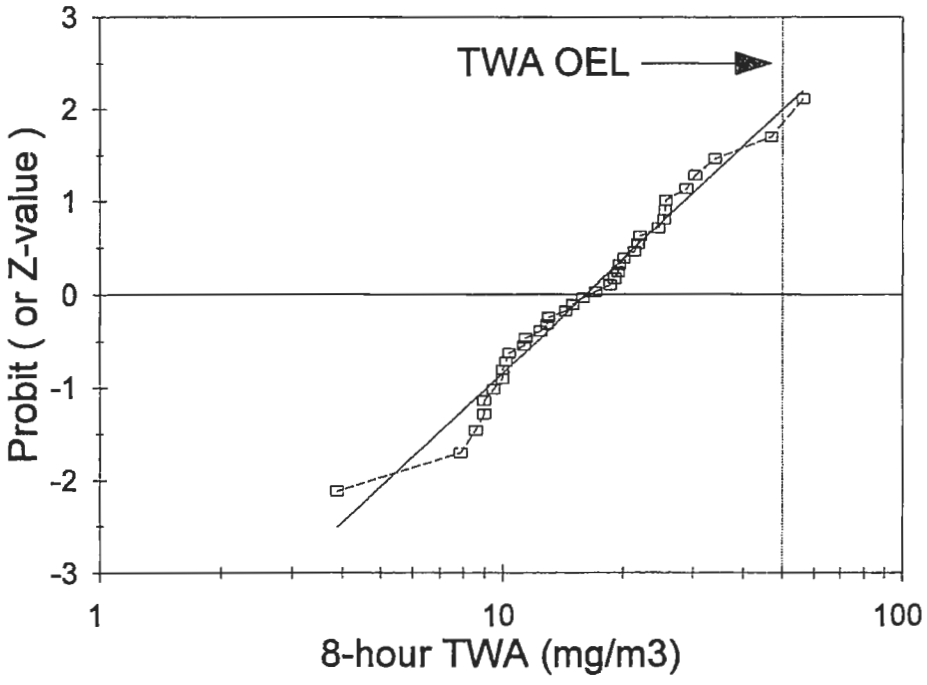

FIGURE 16.2 Log-probit curve for Example 2.

that the data represent measurements collected from one or more presumably maximum-risk employees. In this scenario we could state that the exposure profiles of these members of the exposure group, and by extension, all members of the exposure group, appear to be "controlled." Our selection of maximum-exposed workers should be verified by collecting measurements from randomly selected members of the exposure group during future surveys. Otherwise, our conclusion may not apply to all of the unmeasured members of the exposure group.

Alternatively, let's assume that the data represent measurements collected from randomly selected workers drawn from an exposure group where it is difficult to select maximum risk employees. Somewhat more caution is warranted. We could conclude from our statistical analysis that the exposure profile for the group appears to be "controlled." However, we are ultimately interested in each individual exposure profile. Whether or not this conclusion applies to all exposure profiles within the group depends on how homogeneous the exposure group is with respect to exposure controls and work practices. Ideally, we should reserve judgment until an exposure history has been developed. In the meantime, we can state that the exposure profiles appear to be "controlled," but that periodic surveillance is warranted to confirm our tentative rating.

In either case—selection of presumed maximum exposure employees or random selection—task analysis and assessment of individual work practices should increase the degree of homogeneity within each exposure group, thus increasing the predictive value of our measurements. If more than one measurement is available per employee, then an analysis of within-worker and between-worker variability may be useful for determining whether exposures are primarily affected by factors common to all workers in the exposure group or by factors specific to each worker.[15,19]

APPENDIX

TABLE A1 Filliben's Test Critical *r* Values ($\alpha = 0.05$) for Determining Goodness-of-Fit for the Normal or Lognormal Distribution Assumption (Ref. 17)

n	*r*	*n*	*r*	*n*	*r*
3	0.879	23	0.956	42	0.973
4	0.868	24	0.957	43	0.974
5	0.880	25	0.959	44	0.974
6	0.888	26	0.960	45	0.974
7	0.898	27	0.961	46	0.975
8	0.906	28	0.962	47	0.976
9	0.912	29	0.963	48	0.976
10	0.918	30	0.964	49	0.976
11	0.923	31	0.965	50	0.977
12	0.928	32	0.966	55	0.979
13	0.932	33	0.967	60	0.980
14	0.935	34	0.968	65	0.981
15	0.939	35	0.969	70	0.983
16	0.941	36	0.969	75	0.984
17	0.944	37	0.970	80	0.985
18	0.946	38	0.971	85	0.985
19	0.949	39	0.971	90	0.986
20	0.951	40	0.972	95	0.987
21	0.952	41	0.973	100	0.987
22	0.954				

TABLE A2a C Factors for Estimating the 95%UCL ($\alpha = 0.05$) for the Mean of Lognormally Distributed Data (Ref. 22)

	Sample Size (n)								
s_y*	3	4	5	6	7	8	9	10	15
0.01	2.415	2.054	1.918	1.849	1.807	1.779	1.759	1.745	1.706
0.1	2.750	2.222	2.035	1.942	1.886	1.849	1.822	1.802	1.749
0.2	3.295	2.463	2.198	2.069	1.992	1.943	1.908	1.881	1.809
0.3	4.109	2.777	2.402	2.226	2.125	2.058	2.011	1.977	1.882
0.4	5.220	3.175	2.651	2.415	2.282	2.195	2.134	2.089	1.968
0.5	6.495	3.658	2.947	2.638	2.465	2.354	2.277	2.220	2.068
0.6	7.807	4.209	3.287	2.892	2.673	2.534	2.439	2.368	2.181
0.7	9.120	4.801	3.662	3.173	2.904	2.735	2.618	2.532	2.306
0.8	10.43	5.414	4.062	3.477	3.155	2.952	2.813	2.710	2.443
0.9	11.74	6.038	4.478	3.796	3.420	3.184	3.021	2.902	2.589
1.0	13.05	6.669	4.905	4.127	3.698	3.426	3.239	3.103	2.744
1.25	16.33	8.265	6.001	4.990	4.426	4.068	3.820	3.639	3.163
1.5	19.60	9.874	7.120	5.880	5.184	4.741	4.433	4.207	3.612
1.75	22.87	11.49	8.250	6.786	5.960	5.432	5.065	4.795	4.081
2.0	26.14	13.11	9.387	7.701	6.747	6.135	5.710	5.396	4.564
2.5	32.69	16.35	11.67	9.546	8.339	7.563	7.021	6.621	5.557
3.0	39.23	19.60	13.97	11.40	9.945	9.006	8.350	7.864	6.570
3.5	45.77	22.85	16.27	13.27	11.56	10.46	9.688	9.118	7.596
4.0	52.31	26.11	18.58	15.14	13.18	11.92	11.03	10.38	8.630

	Sample Size (n)								
s_y*	20	30	40	60	101	201	401	601	1001
0.01	1.689	1.673	1.666	1.659	1.653	1.649	1.647	1.647	1.646
0.1	1.725	1.702	1.6911	1.680	1.670	1.662	1.658	1.656	1.654
0.2	1.776	1.744	1.728	1.712	1.697	1.685	1.677	1.674	1.671
0.3	1.838	1.796	1.775	1.753	1.733	1.716	1.705	1.700	1.696
0.4	1.922	1.859	1.832	1.803	1.777	1.755	1.740	1.734	1.728
0.5	1.999	1.932	1.898	1.862	1.830	1.802	1.784	1.776	1.769
0.6	2.097	2.015	1.974	1.930	1.891	1.857	1.835	1.825	1.816
0.7	2.205	2.108	2.058	2.007	1.960	1.919	1.892	1.881	1.870
0.8	2.324	2.209	2.151	2.090	2.035	1.988	1.957	1.944	1.931
0.9	2.451	2.318	2.251	2.181	2.117	2.062	2.027	2.012	1.997
1.0	2.586	2.434	2.357	2.277	2.205	2.143	2.102	2.085	2.068
1.25	2.952	2.750	2.648	2.542	2.447	2.364	2.310	2.288	2.266
1.5	3.347	3.094	2.966	2.832	2.713	2.609	2.542	2.514	2.486
1.75	3.763	3.457	3.303	3.142	2.997	2.872	2.791	2.757	2.723
2.0	4.193	3.835	3.654	3.465	3.295	3.148	3.053	3.013	2.974
2.5	5.079	4.617	4.384	4.139	3.920	3.729	3.605	3.553	3.503
3.0	5.988	5.424	5.138	4.838	4.569	4.334	4.183	4.119	4.057
3.5	6.910	6.244	5.907	5.552	5.233	4.956	4.776	4.700	4.627
4.0	7.841	7.074	6.685	6.276	5.908	5.588	5.380	5.293	5.208

*$s_y = \ln(gsd)$

TABLE A2b C Factors for Estimating the 95% LCL ($\alpha = 0.95$) for the Mean of Lognormally Distributed Data (Ref. 22)

	Sample Size (n)								
s_y*	3	4	5	6	7	8	9	10	15
0.01	−2.355	−2.022	−1.896	−1.831	−1.791	−1.766	−1.747	−1.734	−1.697
0.1	−2.130	−1.898	−1.806	−1.759	−1.731	−1.712	−1.699	−1.690	−1.666
0.2	−1.949	−1.791	−1.729	−1.697	−1.678	−1.667	−1.658	−1.653	−1.640
0.3	−1.816	−1.710	−1.669	−1.650	−1.639	−1.633	−1.629	−1.627	−1.625
0.4	−1.717	−1.650	−1.625	−1.615	−1.611	−1.610	−1.610	−1.611	−1.617
0.5	−1.644	−1.605	−1.594	−1.592	−1.594	−1.596	−1.599	−1.603	−1.618
0.6	−1.589	−1.572	−1.573	−1.578	−1.584	−1.591	−1.597	−1.602	−1.625
0.7	−1.549	−1.550	−1.560	−1.572	−1.582	−1.592	−1.600	−1.608	−1.638
0.8	−1.521	−1.537	−1.555	−1.572	−1.586	−1.599	−1.610	−1.620	−1.656
0.9	−1.502	−1.530	−1.556	−1.577	−1.595	−1.611	−1.625	−1.637	−1.680
1.0	−1.490	−1.530	−1.562	−1.588	−1.610	−1.628	−1.644	−1.658	−1.707
1.25	−1.486	−1.549	−1.596	−1.632	−1.662	−1.687	−1.708	−1.727	−1.793
1.5	−1.508	−1.590	−1.650	−1.696	−1.733	−1.764	−1.791	−1.814	−1.896
1.75	−1.547	−1.647	−1.719	−1.774	−1.819	−1.857	−1.889	−1.916	−2.015
2.0	−1.598	−1.714	−1.799	−1.864	−1.917	−1.960	−1.998	−2.029	−2.144
2.5	−1.727	−1.877	−1.986	−2.070	−2.138	−2.193	−2.241	−2.283	−2.430
3.0	−1.880	−2.065	−2.199	−2.301	−2.384	−2.452	−2.510	−2.560	−2.740
3.5	−2.051	−2.272	−2.429	−2.550	−2.647	−2.727	−2.795	−2.855	−3.067
4.0	−2.237	−2.491	−2.672	−2.810	−2.922	−3.015	−3.093	−3.161	−3.406

	Sample Size (n)								
s_y*	20	30	40	60	101	201	401	601	1001
0.01	−1.682	−1.668	−1.661	−1.655	−1.651	−1.647	−1.646	−1.646	−1.645
0.1	−1.656	−1.648	−1.646	−1.643	−1.642	−1.643	−1.644	−1.645	−1.645
0.2	−1.637	−1.636	−1.636	−1.638	−1.641	−1.646	−1.649	−1.651	−1.653
0.3	−1.626	−1.631	−1.635	−1.641	−1.648	−1.656	−1.663	−1.666	−1.669
0.4	−1.624	−1.634	−1.641	−1.651	−1.662	−1.674	−1.684	−1.688	−1.693
0.5	−1.629	−1.644	−1.655	−1.668	−1.683	−1.699	−1.711	−1.717	−1.723
0.6	−1.640	−1.661	−1.675	−1.692	−1.711	−1.731	−1.746	−1.753	−1.760
0.7	−1.658	−1.684	−1.700	−1.721	−1.744	−1.768	−1.786	−1.795	−1.804
0.8	−1.680	−1.711	−1.731	−1.756	−1.783	−1.811	−1.832	−1.842	−1.853
0.9	−1.708	−1.744	−1.767	−1.795	−1.826	−1.859	−1.884	−1.895	−1.907
1.0	−1.740	−1.781	−1.807	−1.839	−1.874	−1.912	−1.940	−1.953	−1.966
1.25	−1.835	−1.889	−1.923	−1.965	−2.012	−2.060	−2.097	−2.114	−2.131
1.5	−1.949	−2.016	−2.058	−2.111	−2.169	−2.229	−2.275	−2.296	−2.318
1.75	−2.078	−2.158	−2.208	−2.272	−2.341	−2.414	−2.469	−2.494	−2.514
2.0	−2.218	−2.311	−2.370	−2.445	−2.526	−2.611	−2.675	−2.705	−2.736
2.5	−2.525	−2.645	−2.721	−2.817	−2.921	−3.032	−3.115	−3.154	−3.194
3.0	−2.856	−3.003	−3.096	−3.214	−3.342	−3.478	−3.581	−3.628	−3.677
3.5	−3.204	−3.377	−3.488	−3.627	−3.780	−3.940	−4.062	−4.119	−4.177
4.0	−3.563	−3.764	−3.892	−4.052	−4.228	−4.414	−4.555	−4.620	−4.688

*$s_y = \ln(gsd)$

TABLE A3 Confidence Limit Values for Estimating the 95%LCL and 95%UCL for the Proportion in the Tail of a Normal (or Lognormal) Distribution (Ref. 25; Table 7.4)

	Sample Size (n)								
z	2	3	4	5	6	7	8	9	10
3.0	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00001	0.00001	0.00001
2.8	0.00000	0.00000	0.00000	0.00000	0.00001	0.00001	0.00002	0.00003	0.00004
2.6	0.00000	0.00000	0.00001	0.00002	0.00003	0.00005	0.00007	0.00009	0.00011
2.4	0.00000	0.00001	0.00003	0.00006	0.00010	0.00014	0.00020	0.00026	0.00031
2.2	0.00000	0.00003	0.00009	0.00018	0.00029	0.00041	0.00054	0.00067	0.00080
2.0	0.00002	0.00011	0.00029	0.00052	0.00079	0.00107	0.00135	0.00163	0.00191
1.8	0.00008	0.00038	0.00085	0.00140	0.00199	0.00257	0.00315	0.00371	0.00424
1.6	0.00030	0.00115	0.00226	0.00344	0.00461	0.00574	0.00682	0.00783	0.00879
1.4	0.00106	0.00311	0.00543	0.00772	0.00987	0.01187	0.01372	0.01543	0.01701
1.2	0.00317	0.00753	0.01188	0.01589	0.01951	0.02277	0.02571	0.02837	0.03079
1.0	0.00820	0.01629	0.02364	0.03006	0.03565	0.04055	0.04488	0.04874	0.05220
0.8	0.01831	0.03164	0.04292	0.05236	0.06034	0.06719	0.07315	0.07839	0.08305
0.6	0.03544	0.05534	0.07131	0.08424	0.09493	0.10394	0.11168	0.11840	0.12433
0.4	0.05997	0.08771	0.10903	0.12581	0.13941	0.15071	0.16028	0.16853	0.17574
0.2	0.09006	0.12725	0.15462	0.17559	0.19228	0.20594	0.21740	0.22718	0.23566
0.0	0.12240	0.17114	0.20542	0.23099	0.25095	0.26707	0.28044	0.29175	0.30148
−0.2	0.15403	0.21648	0.25856	0.28913	0.31255	0.33120	0.34649	0.35932	0.37027
−0.4	0.18342	0.26118	0.31177	0.34762	0.37460	0.39580	0.41299	0.42728	0.43938
−0.6	0.21024	0.30413	0.36360	0.40483	0.43534	0.45898	0.47795	0.49357	0.50671
−0.8	0.23469	0.34491	0.41329	0.45976	0.49359	0.51947	0.54001	0.55677	0.57076
−1.0	0.25712	0.38345	0.46048	0.51188	0.54869	0.57648	0.59829	0.61592	0.63052
−1.2	0.27784	0.41982	0.50506	0.56090	0.60025	0.62954	0.65226	0.67045	0.68537
−1.4	0.29713	0.45416	0.54703	0.60671	0.64806	0.67839	0.70162	0.72002	0.73498
−1.6	0.31518	0.48661	0.58642	0.64929	0.69205	0.72292	0.74625	0.76451	0.77921
−1.8	0.33219	0.51732	0.62332	0.68867	0.73223	0.76315	0.78616	0.80395	0.81812
−2.0	0.34828	0.54640	0.65781	0.72489	0.76866	0.79913	0.82146	0.83849	0.85187
−2.2	0.36356	0.57398	0.68997	0.75806	0.80144	0.83104	0.85234	0.86834	0.88077
−2.4	0.37814	0.60013	0.71989	0.78826	0.83072	0.85905	0.87906	0.89384	0.90516
−2.6	0.39207	0.62495	0.74766	0.81563	0.85668	0.88341	0.90192	0.91535	0.92547
−2.8	0.40543	0.64851	0.77335	0.84028	0.87950	0.90440	0.92125	0.93325	0.94215
−3.0	0.41828	0.67087	0.79706	0.86237	0.89942	0.92229	0.93741	0.94796	0.95565
−3.2	0.43064	0.69210	0.81886	0.88204	0.91666	0.93740	0.95077	0.95990	0.96643
−3.4	0.44258	0.71224	0.83886	0.89946	0.93146	0.95003	0.96168	0.96946	0.97491
−3.6	0.45411	0.73134	0.85713	0.91478	0.94405	0.96048	0.97049	0.97701	0.98149
−3.8	0.46527	0.74945	0.87376	0.92818	0.95468	0.96903	0.97752	0.98291	0.98652
−4.0	0.47609	0.76661	0.88885	0.93983	0.96357	0.97596	0.98306	0.98744	0.99031
−4.2	0.48658	0.78285	0.90249	0.94988	0.97095	0.98151	0.98737	0.99088	0.99313
−4.4	0.49677	0.79821	0.91476	0.95849	0.97701	0.98592	0.99069	0.99346	0.99519
−4.6	0.50668	0.81272	0.92577	0.96584	0.98196	0.98938	0.99321	0.99537	0.99668
−4.8	0.51633	0.82641	0.93560	0.97205	0.98595	0.99207	0.99510	0.99676	0.99774
−5.0	0.52572	0.83932	0.94434	0.97727	0.98915	0.99413	0.99651	0.99776	0.99848
−5.2	0.53488	0.85147	0.95208	0.98164	0.99169	0.99570	0.99754	0.99847	0.99899
−5.4	0.54381	0.86290	0.95890	0.98525	0.99368	0.99688	0.99829	0.99897	0.99934
−5.6	0.55252	0.87363	0.96489	0.98823	0.99524	0.99776	0.99882	0.99932	0.99958
−5.8	0.56103	0.88369	0.97013	0.99067	0.99644	0.99841	0.99920	0.99955	0.99973
−6.0	0.56935	0.89311	0.97468	0.99265	0.99736	0.99888	0.99946	0.99971	0.99983

(*continued*)

TABLE A3 *Continued*

	Sample Size (n)								
n	11	15	21	30	40	60	80	100	120
3.0	0.00002	0.00003	0.00006	0.00011	0.00016	0.00024	0.00030	0.00036	0.00040
2.8	0.00005	0.00009	0.00017	0.00027	0.00037	0.00054	0.00068	0.00078	0.00087
2.6	0.00014	0.00025	0.00041	0.00064	0.00085	0.00118	0.00143	0.00163	0.00180
2.4	0.00038	0.00062	0.00097	0.00142	0.00182	0.00244	0.00290	0.00325	0.00353
2.2	0.00094	0.00145	0.00214	0.00299	0.00373	0.00481	0.00559	0.00618	0.00665
2.0	0.00218	0.00320	0.00447	0.00597	0.00723	0.00904	0.01029	0.01123	0.01197
1.8	0.00476	0.00660	0.00881	0.01130	0.01334	0.01617	0.01809	0.01951	0.02062
1.6	0.00969	0.01281	0.01641	0.02031	0.02341	0.02761	0.03040	0.03243	0.03401
1.4	0.01848	0.02341	0.02889	0.03463	0.03909	0.04499	0.04884	0.05162	0.05375
1.2	0.03301	0.04030	0.04814	0.05612	0.06219	0.07005	0.07511	0.07872	0.08147
1.0	0.05534	0.06545	0.07602	0.08653	0.09436	0.10434	0.11066	0.11513	0.11852
0.8	0.08723	0.10046	0.11396	0.12710	0.13674	0.14884	0.15640	0.16172	0.16572
0.6	0.12959	0.14604	0.16250	0.17822	0.18958	0.20367	0.21238	0.21846	0.22302
0.4	0.18211	0.20173	0.22101	0.23913	0.25206	0.26790	0.27760	0.28433	0.28935
0.2	0.24310	0.26577	0.28768	0.30795	0.32224	0.33956	0.35007	0.35732	0.36270
0.0	0.30997	0.33553	0.35982	0.38197	0.39740	0.41592	0.42705	0.43467	0.44032
−0.2	0.37977	0.40802	0.43444	0.45816	0.47450	0.49389	0.50544	0.51331	0.51911
−0.4	0.44981	0.48048	0.50867	0.53361	0.55056	0.57048	0.58223	0.59020	0.59604
−0.6	0.51795	0.55063	0.58014	0.60583	0.62308	0.64311	0.65482	0.66270	0.66846
−0.8	0.58264	0.61677	0.64703	0.67294	0.69010	0.70978	0.72116	0.72877	0.73431
−1.0	0.64284	0.67773	0.70809	0.73361	0.75027	0.76913	0.77990	0.78706	0.79224
−1.2	0.69787	0.73280	0.76255	0.78708	0.80283	0.82041	0.83033	0.83686	0.84156
−1.4	0.74741	0.78161	0.81008	0.83305	0.84755	0.86347	0.87233	0.87811	0.88225
−1.6	0.79132	0.82407	0.85068	0.87164	0.88461	0.89861	0.90628	0.91123	0.91475
−1.8	0.82967	0.86037	0.88463	0.90325	0.91454	0.92648	0.93291	0.93702	0.93991
−2.0	0.86268	0.89084	0.91243	0.92855	0.93809	0.94797	0.95319	0.95649	0.95879
−2.2	0.89067	0.91596	0.93473	0.94831	0.95614	0.96407	0.96818	0.97074	0.97251
−2.4	0.91407	0.93630	0.95224	0.96338	0.96963	0.97581	0.97893	0.98085	0.98216
−2.6	0.93334	0.95248	0.96569	0.97460	0.97946	0.98412	0.98642	0.98781	0.98875
−2.8	0.94896	0.96512	0.97582	0.98276	0.98642	0.98984	0.99148	0.99246	0.99311
−3.0	0.96145	0.97481	0.98328	0.98855	0.99123	0.99366	0.99480	0.99546	0.99590
−3.2	0.97127	0.98210	0.98866	0.99256	0.99447	0.99615	0.99691	0.99735	0.99763
−3.4	0.97888	0.98750	0.99246	0.99527	0.99660	0.99772	0.99821	0.99849	0.99867
−3.6	0.98469	0.99141	0.99508	0.99706	0.99796	0.99869	0.99900	0.99917	0.99927
−3.8	0.98906	0.99420	0.99685	0.99821	0.99880	0.99926	0.99945	0.99955	0.99962
−4.0	0.99229	0.99615	0.99803	0.99894	0.99932	0.99960	0.99971	0.99977	0.99980
−4.2	0.99464	0.99748	0.99879	0.99938	0.99962	0.99979	0.99985	0.99988	0.99990
−4.4	0.99633	0.99839	0.99927	0.99965	0.99979	0.99989	0.99992	0.99994	0.99995
−4.6	0.99752	0.99898	0.99957	0.99981	0.99989	0.99994	0.99996	0.99997	0.99998
−4.8	0.99835	0.99937	0.99975	0.99989	0.99994	0.99997	0.99998	0.99999	0.99999
−5.0	0.99892	0.99962	0.99986	0.99994	0.99997	0.99999	0.99999	0.99999	1.00000
−5.2	0.99930	0.99977	0.99992	0.99997	0.99999	0.99999	1.00000	1.00000	1.00000
−5.4	0.99955	0.99987	0.99996	0.99999	0.99999	1.00000	1.00000	1.00000	1.00000
−5.6	0.99972	0.99992	0.99998	0.99999	1.00000	1.00000	1.00000	1.00000	1.00000
−5.8	0.99983	0.99996	0.99999	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
−6.0	0.99989	0.99998	0.99999	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000

(*continued*)

TABLE A3 *Continued*

	Sample Size (n)		
z	240	600	1000
3.0	0.00058	0.00080	0.00090
2.8	0.00121	0.00160	0.00178
2.6	0.00239	0.00307	0.00338
2.4	0.00455	0.00567	0.00617
2.2	0.00831	0.01007	0.01084
2.0	0.01453	0.01718	0.01832
1.8	0.02437	0.02818	0.02979
1.6	0.03926	0.04447	0.04665
1.4	0.06075	0.06757	0.07038
1.2	0.09039	0.09892	0.10240
1.0	0.12939	0.13964	0.14378
0.8	0.17844	0.19028	0.19501
0.6	0.23737	0.25056	0.25579
0.4	0.30504	0.31931	0.32493
0.2	0.37941	0.39445	0.40033
0.0	0.45772	0.47323	0.47926
−0.2	0.53687	0.55255	0.55860
−0.4	0.61380	0.62932	0.63527
−0.6	0.68583	0.70084	0.70656
−0.8	0.75088	0.76502	0.77037
−1.0	0.80759	0.82053	0.82539
−1.2	0.85536	0.86683	0.87109
−1.4	0.89425	0.90406	0.90767
−1.6	0.92484	0.93295	0.93589
−1.8	0.94811	0.95456	0.95687
−2.0	0.96521	0.97016	0.97191
−2.2	0.97736	0.98102	0.98229
−2.4	0.98571	0.98831	0.98920
−2.6	0.99125	0.99303	0.99363
−2.8	0.99481	0.99598	0.99636
−3.0	0.99701	0.99776	0.99800
−3.2	0.99833	0.99879	0.99893
−3.4	0.99910	0.99937	0.99945
−3.6	0.99953	0.99968	0.99973
−3.8	0.99976	0.99984	0.99987
−4.0	0.99988	0.99993	0.99994
−4.2	0.99994	0.99997	0.99997
−4.4	0.99997	0.99999	0.99999
−4.6	0.99999	0.99999	1.00000
−4.8	1.00000	1.00000	1.00000
−5.0	1.00000	1.00000	1.00000
−5.2	1.00000	1.00000	1.00000
−5.4	1.00000	1.00000	1.00000
−5.6	1.00000	1.00000	1.00000
−5.8	1.00000	1.00000	1.00000
−6.0	1.00000	1.00000	1.00000

TABLE A4 K Factors for Calculating the 95%LCL and 95%UCL for the 95th Percentile of Normally or Lognormally Distributed Data (Ref. 25; Tables 1.10.1–1.10.4 and 1.4.1–1.4.4)

$K_{0.05, 0.95, n}$				$K_{0.95, 0.95, n}$			
n	*K*	*n*	*K*	*n*	*K*	*n*	*K*
2	0.475	40	1.297	2	26.260	40	2.125
3	0.639	41	1.300	3	7.656	41	2.118
4	0.743	42	1.304	4	5.144	42	2.111
5	0.818	43	1.308	5	4.203	43	2.105
6	0.875	44	1.311	6	3.708	44	2.098
7	0.920	45	1.314	7	3.399	45	2.092
8	0.958	46	1.317	8	3.187	46	2.086
9	0.990	47	1.321	9	3.031	47	2.081
10	1.017	48	1.324	10	2.911	48	2.075
11	1.041	49	1.327	11	2.815	49	2.070
12	1.062	50	1.329	12	2.736	50	2.065
13	1.081	55	1.343	13	2.671	55	2.042
14	1.098	60	1.354	14	2.614	60	2.022
15	1.114	65	1.364	15	2.566	65	2.005
16	1.128	70	1.374	16	2.524	70	1.990
17	1.141	75	1.382	17	2.486	75	1.976
18	1.153	80	1.390	18	2.453	80	1.964
19	1.164	85	1.397	19	2.423	85	1.954
20	1.175	90	1.403	20	2.396	90	1.944
21	1.184	95	1.409	21	2.371	95	1.935
22	1.193	100	1.414	22	2.349	100	1.927
23	1.202	120	1.433	23	2.328	120	1.899
24	1.210	140	1.447	24	2.309	140	1.879
25	1.217	160	1.459	25	2.292	160	1.862
26	1.225	180	1.469	26	2.275	180	1.849
27	1.231	200	1.478	27	2.260	200	1.837
28	1.238	300	1.507	28	2.246	300	1.800
29	1.244	400	1.525	29	2.232	400	1.778
30	1.250	500	1.537	30	2.220	500	1.763
31	1.255	600	1.546	31	2.208	600	1.752
32	1.261	700	1.553	32	2.197	700	1.744
33	1.266	800	1.559	33	2.186	800	1.737
34	1.271	900	1.563	34	2.176	900	1.732
35	1.276	1000	1.567	35	2.167	1000	1.727
36	1.280	1500	1.581	36	2.158	1500	1.712
37	1.284	2000	1.590	37	2.149	2000	1.703
38	1.289	3000	1.600	38	2.141	3000	1.692
39	1.293	5000	1.610	39	2.133	5000	1.681
		10000	1.620			10000	1.670

TABLE A5 Ranks (r) for the Nonparametric 95%LCL and 95%UCL for the 50th Percentile and the 95%LCL for the 95th Percentile (taken together, the LCL and UCL form a 90% confidence interval for the true median)

	Median (50th Percentile)		95th Percentile		Median (50th Percentile)		95th Percentile
n	r LCL	r UCL	r LCL	n	r LCL	r UCL	r LCL
5	1	5	4	18	6	13	15
6	1	6	5	19	6	14	16
7	1	7	6	20	6	15	17
8	2	7	6	21	7	15	18
9	2	8	7	22	7	16	19
10	2	9	8	23	8	16	20
11	3	9	9	24	8	17	21
12	3	10	10	25	8	18	22
13	4	10	11	26	9	18	23
14	4	11	12	27	9	19	24
15	4	12	13	28	10	19	25
16	5	12	14	29	10	20	25
17	5	13	14	30	11	20	26

Note. The sample size has to exceed 58 before a 95%UCL rank can be calculated for the 95th percentile.

TABLE A6 Nonparametric 95% Lower and Upper Confidence Limit on the Fraction Exceeding an OEL [The left and right values represent the 95%LCL and 95%UCL, respectively. Taken together, they form the nonparametric 90% confidence interval for the true exceedance fraction (m refers to the number of overexposures out of n measurements).]

m	$n=1$		$n=2$		$n=3$		$n=4$		$n=5$	
0	0.000	0.950	0.000	0.777	0.000	0.632	0.000	0.527	0.000	0.451
1	—	—	0.025	0.975	0.016	0.865	0.012	0.752	0.010	0.658
2	—	—	—	—	0.135	0.984	0.097	0.903	0.076	0.811
3	—	—	—	—	—	—	—	—	0.189	0.924
m	$n=6$		$n=7$		$n=8$		$n=9$		$n=10$	
0	0.000	0.393	0.000	0.348	0.000	0.313	0.000	0.283	0.000	0.259
1	0.008	0.582	0.007	0.521	0.006	0.471	0.005	0.429	0.005	0.395
2	0.062	0.729	0.053	0.659	0.046	0.600	0.041	0.550	0.036	0.507
3	0.153	0.847	0.128	0.775	0.111	0.711	0.097	0.655	0.087	0.607
4	—	—	0.225	0.872	0.193	0.807	0.168	0.749	0.150	0.697
5	—	—	—	—	—	—	0.251	0.832	0.222	0.778
m	$n=11$		$n=12$		$n=13$		$n=14$		$n=15$	
0	0.000	0.239	0.000	0.221	0.000	0.206	0.000	0.193	0.000	0.181
1	0.004	0.365	0.004	0.339	0.003	0.317	0.003	0.297	0.003	0.280
2	0.033	0.470	0.030	0.438	0.028	0.410	0.026	0.386	0.024	0.364
3	0.078	0.565	0.071	0.528	0.066	0.495	0.061	0.466	0.056	0.440
4	0.135	0.651	0.122	0.610	0.112	0.573	0.104	0.540	0.096	0.511
5	0.199	0.729	0.181	0.685	0.165	0.646	0.152	0.610	0.141	0.578
6	0.271	0.801	0.245	0.755	0.224	0.713	0.206	0.675	0.190	0.641
7	—	—	—	—	0.287	0.776	0.263	0.737	0.243	0.700
8	—	—	—	—	—	—	—	—	0.300	0.757

(*continued*)

TABLE A6 *Continued*

m	$n = 16$		$n = 17$		$n = 18$		$n = 19$		$n = 20$	
0	0.000	0.171	0.000	0.162	0.000	0.154	0.000	0.146	0.000	0.140
1	0.003	0.264	0.003	0.251	0.002	0.238	0.003	0.227	0.002	0.217
2	0.022	0.344	0.021	0.327	0.020	0.311	0.019	0.296	0.018	0.283
3	0.053	0.417	0.049	0.396	0.047	0.377	0.044	0.360	0.042	0.344
4	0.090	0.485	0.084	0.461	0.079	0.439	0.075	0.420	0.071	0.401
5	0.132	0.549	0.123	0.522	0.116	0.498	0.109	0.476	0.104	0.456
6	0.177	0.609	0.166	0.581	0.156	0.554	0.147	0.530	0.139	0.508
7	0.226	0.667	0.211	0.636	0.199	0.608	0.187	0.582	0.177	0.558
8	0.278	0.722	0.260	0.690	0.244	0.660	0.229	0.632	0.217	0.607
9	—	—	0.310	0.740	0.291	0.709	0.274	0.680	0.258	0.653
10	—	—	—	—	—	—	0.320	0.726	0.302	0.698

m	$n = 21$		$n = 22$		$n = 23$		$n = 24$		$n = 25$	
0	0.000	0.133	0.000	0.128	0.000	0.123	0.000	0.118	0.000	0.113
1	0.002	0.207	0.002	0.198	0.002	0.191	0.002	0.183	0.002	0.177
2	0.017	0.271	0.016	0.260	0.015	0.250	0.015	0.240	0.014	0.231
3	0.040	0.330	0.038	0.316	0.036	0.304	0.034	0.293	0.033	0.282
4	0.067	0.385	0.064	0.369	0.061	0.355	0.059	0.342	0.056	0.330
5	0.098	0.437	0.094	0.420	0.089	0.404	0.085	0.390	0.082	0.376
6	0.132	0.488	0.126	0.469	0.120	0.451	0.114	0.435	0.110	0.420
7	0.168	0.536	0.160	0.516	0.152	0.497	0.145	0.479	0.139	0.463
8	0.205	0.583	0.195	0.561	0.186	0.541	0.178	0.522	0.170	0.504
9	0.245	0.628	0.232	0.605	0.221	0.584	0.211	0.563	0.202	0.544
10	0.285	0.672	0.271	0.648	0.258	0.625	0.246	0.604	0.235	0.584
11	0.328	0.715	0.311	0.689	0.296	0.665	0.282	0.643	0.269	0.622
12	—	—	—	—	0.335	0.704	0.319	0.681	0.305	0.659
13	—	—	—	—	—	—	—	—	0.341	0.695

m	$n = 26$		$n = 27$		$n = 28$		$n = 29$		$n = 30$	
0	0.000	0.109	0.000	0.105	0.000	0.102	0.000	0.099	0.000	0.095
1	0.001	0.170	0.001	0.164	0.001	0.159	0.001	0.154	0.001	0.149
2	0.013	0.223	0.013	0.216	0.012	0.209	0.012	0.202	0.011	0.196
3	0.032	0.272	0.031	0.263	0.029	0.255	0.028	0.247	0.027	0.239
4	0.054	0.319	0.052	0.308	0.050	0.298	0.048	0.289	0.046	0.280
5	0.079	0.363	0.075	0.351	0.073	0.340	0.070	0.329	0.068	0.319
6	0.105	0.406	0.101	0.393	0.097	0.380	0.094	0.368	0.090	0.357
7	0.133	0.447	0.128	0.433	0.123	0.419	0.119	0.406	0.115	0.394
8	0.163	0.487	0.156	0.472	0.150	0.457	0.145	0.443	0.140	0.430
9	0.194	0.527	0.186	0.510	0.179	0.494	0.172	0.479	0.166	0.465
10	0.225	0.565	0.216	0.547	0.208	0.530	0.200	0.515	0.193	0.500
11	0.258	0.602	0.248	0.583	0.238	0.566	0.229	0.549	0.221	0.533
12	0.292	0.638	0.280	0.619	0.269	0.601	0.259	0.583	0.249	0.566
13	0.326	0.674	0.313	0.653	0.300	0.634	0.289	0.616	0.278	0.599
14	—	—	—	—	0.333	0.667	0.320	0.648	0.308	0.630
15	—	—	—	—	—	—	0.352	0.680	0.338	0.662

References

1. Hewett, P. Industrial Hygiene Exposure Assessment—Data Collection and Management. In *Handbook of Chemical Health and Safety*; American Chemical Society: Washington, DC, 2001; Chapter 15.
2. Roach, S. A.; Baier, E. J.; Ayer, H. E.; Harris, R. L. Testing Compliance with Threshold Limit Values for Respirable Dusts. *Am. Ind. Hyg. Assoc. J.* **1967**, *28*, 543–553
3. Mulhausen, J., Damiano, J., Eds. *A Strategy for Assessing and Managing Occupational Exposures*, 2nd ed.; American Industrial Hygiene Association: Fairfax, VA, 1998.
4. CEN (Comité Européen de Normalisation). Workplace atmospheres—Guidance for the assessment of exposure by inhalation of chemical agents for comparison with limit values and measurement strategy. *European Standard EN 689*, effective no later than Aug. 1995 (English version) (Feb. 1995).
5. Symanski, E.; Kupper, L. L.; Kromhout, H.; Rappaport, S. M. An Investigation of Systematic Changes in Occupational Exposure. *Am. Ind. Hyg. Assoc. J.* **1996**, *57*, 724–735.
6. Hawkins, N. C., Norwood, S. K., Rock, J. C., Eds. *A Strategy for Occupational Exposure Assessment*; American Industrial Hygiene Association: Fairview, VA, 1991.
7. Oudyk, J. D. Review of an Extensive Ferrous Foundry Silica Sampling Program. *Appl. Occup. Environ. Hyg.* **1995**, *10*, 331–340.
8. Hawkins, N. C.; Landenberger, B. D. Statistical Control Charts: A Technique for Analyzing Industrial Hygiene Data. *Appl. Occup. Environ. Hyg.* **1991**, *6*, 689–695.
9. Montgomery, D. C. *Introduction to Statistical Quality Control*, 3rd ed.; Wiley & Sons: New York, 1996.
10. Roach, S. *Health Risks from Hazardous Substances at Work—Assessment, Evaluation, and Control*; Pergamon Press: New York, 1992.
11. George, D. K.; Flynn, M. R.; Harris, R. L. Autocorrelation of Interday Exposures at an Automobile Assembly Plant. *Am. Ind. Hyg. Assoc. J.* **1995**, *56*, 1187–1194.
12. Francis, M.; Selvin, S.; Spear, R.; Rappaport, S. The Effect of Autocorrelation on the Estimation of Workers' Daily Exposures. *Am. Ind. Hyg. Assoc. J.* **1989**, *50*, 37–43.
13. Royston, P. A Pocket-Calculator Algorithm for the Shapiro-Francia Test for Non-normality: An Application to Medicine. *Stat. Med.* **1993**, *12*, 181–184.
14. Esmen, N. A.; Hammad, Y. Y. Log-normality of Environmental Sampling Data. *J. Environ. Sci. Health* **1977**, *A12*, 29–41.
15. Rappaport, S. M. Interpreting Levels of Exposures to Chemical Agents. In *Patty's Industrial Hygiene and Toxicology*, 3rd ed., Vol. 3, Part A; Harris, R. L., Cralley, L. J., Cralley, L. V., Eds.; Wiley & Sons, Inc.: New York, 1994.
16. Filliben, J. J. The Probability Plot Correlation Coefficient Test for Normality. *Technometrics* **1975**, *17*, 111–117.
17. Looney, S. W.; Gulledge, T. R. Use of the Correlation Coefficient with Normal Probability Plots. *The American Statistician* **1985**, *39*, 75–79.
18. Rappaport, S. M. Assessment of Long-term Exposures to Toxic Substances in Air. *Ann. Occup. Hyg.* **1991**, *35*, 61–121.
19. Woskie, S. R. et al. The Real-Time Dust Exposures of Sodium Borate Workers: Examination of Exposure Variability. *Am. Ind. Hyg. Assoc. J.* **1994**, *55*, 207–217.
20. Aitchison, J.; Brown, J. A. C. *The Lognormal Distribution with Special Reference to Its Uses in Economics*; Cambridge University Press: New York, 1957.
21. Attfield, M. D.; Hewett, P. Exact Expressions for the Bias and Variance of Estimators of the Mean of a Lognormal Distribution. *Am. Ind. Hyg. Assoc. J.* **1992**, *53*, 432–435.
22. Land, C. E. Tables of Confidence Limits for Linear Functions of the Normal Mean and Variance. In *Selected Tables in Mathematical Statistics*, Vol. III; Harter, H., Owen, D., Eds.; 1975; pp. 385–419.
23. Hewett, P. Mean Testing: II. Comparison of Several Alternative Procedures. *Appl. Occup. Environ. Hyg.* **1997**, *12*, 347–355.
24. Hewett, P.; Ganser, G. H. Simple Procedures for Calculating Confidence Intervals Around the Sample Mean and Exceedance Fraction Derived from Lognormally Distributed Data. *Appl. Occup. Environ. Hyg.* **1997**, *12*, 132–142.
25. Odeh, R. E.; Owen, D. B. *Statistics: Textbooks and Monographs Series;* Vol. 32—Tables for Normal Tolerance Limits, Sampling Plans, and Screening, 1980.
26. Tuggle, R. M. Assessment of Occupational Exposure Using One-Sided Tolerance Limits. *Am. Ind. Hyg. Assoc. J.* **1982**, *43*, 338–346.
27. Tuggle, R. M. The NIOSH Decision Scheme. *Am. Ind. Hyg. Assoc. J.* **1981**, *42*, 493–498.
28. Leidel, N. A.; Busch, K. A. Statistical Design and Data Analysis requirements. In *Patty's Industrial Hygiene and Toxicology*, 3rd ed., Vol. 3, Part A; Harris, R. L., Cralley, L. J., Cralley, L. V., Eds.; Wiley & Sons, Inc.: New York, 1994.
29. Rock, J. C. Occupational Air Sampling Strategies. In *Air Sampling Instruments for Evaluation of Atmospheric Contaminants*, 8th ed.; Cohen, B. S., Hering, S. V., Eds. American Conference of Governmental Industrial Hygienists: Cincinnati, OH, 1995.
30. Esmen, N. A. A Distribution-free Double-sampling Method for Exposure Assessment. *Appl. Occup. Environ. Hyg.* **1992**, *7*, 613–621.
31. Conover, W. J. *Practical Nonparametric Statistics, Second Edition*; Wiley & Sons: New York, 1980; pp. 111–112.
32. Gilbert, R. O. *Statistical Methods for Environmental Pollution Monitoring*; Van Nostrand Reinhold: New York, 1987.
33. Mood, A. M.; Graybill, F. A. *Introduction to the Theory of Statistics*, 2nd ed.; McGraw-Hill Book Company, Inc.: New York, 1963.
34. Snedecor, G. W.; Cochran, W. G. *Statistical Methods*, 7th ed.; Iowa State University Press: Ames, IA, 1980.
35. Beyer, W. H. *Handbook of Tables for Probability and Statistics*. Chemical Rubber Company (CRC): 1968.
36. Blyth, C. R.; Still, H. A. Binomial Confidence Intervals. *J. Am. Stat. Assoc.* **1983**, *78*, 108–116.
37. Hornung, R. W.; Reed, L. D. Estimation of Average Concentration in the Presence of Nondetectable Values. *Appl. Occup. Environ. Hyg.* **1990**, *5*, 46–51.
38. Perkins, J. L.; Cutter, G. N.; Cleveland, M. S. Estimating the Mean, Variance, and Confidence Limits from Censored (<Limit of Detection), Lognormally-distributed Exposure Data. *Am. Ind. Hyg. Assoc. J.* **1990**, *51*, 416–419.
39. Rappaport, S. M.; Lyles, R. H.; Kupper, L. L. An Exposure-assessment Strategy Accounting for Within- and Between-worker Sources of Variability. *Ann. Occup. Hyg.* **1995**, *39*, 469–495.
40. Lyles, R. H.; Kupper, L. L.; Rappaport, S. M. A Lognormal Distribution-based Exposure Assessment Method for Unbalanced Data. *Ann. Occup. Hyg.* **1997**, *41*, 63–76.

41. Hewett, P. Interpretation and Use of Occupational Exposure Limits for Chronic Disease Agents. *Occupational Medicine: State of the Art Reviews* **1996**, *11*(3).
42. Leidel, N. A.; Busch, K. A.; Lynch, J. R. *Occupational Exposure Sampling Strategy Manual*; National Institute for Occupational Safety and Health (NIOSH) Publication No. 77-173 (available from the National Technical Information Service (NTIS), Publication No. PB274792, or the NIOSH Web site as a .pdf file), 1977.
43. Corn, M.; Esmen, N. A. Workplace Exposure Zones for Classification of Employee Exposures to Physical and Chemical Agents. *Am. Ind. Hyg. Assoc. J.* **1979**, *40*, 47–57.
44. Still, K. R.; Wells, B. Quantitative Industrial Hygiene Programs: Workplace Monitoring. (Industrial Hygiene Program Management series, Part VIII). *Appl. Ind. Hyg.* **1989**, *4*, F14–F17.
45. Ayer, H. E. Occupational Air Sampling Strategies. In *Air Sampling Instruments for Evaluation of Atmospheric Contaminants*, 7th ed.; Hering, S. V., Ed.; American Conference of Governmental Industrial Hygienists: Cincinnati, OH, 1989.
46. Damiano, J. Quantitative Exposure Assessment Strategies and Data in the Aluminum Company of America. *Appl. Occup. Environ. Hyg.* **1995**, *10*, 289–298.
47. McHattie, G. V.; Rackham, M.; Teasdale, E. L. The Derivation of Occupational Exposure Limits in the Pharmaceutical Industry. *J. Soc. Occup. Med.* **1988**, *38*, 105–108.
48. Hewett, P. Mean Testing: II. Comparison of Several Alternative Procedures. *Appl. Occup. Environ. Hyg.* **1997**, *12*, 347–355.
49. OSHA (Occupational Safety and Health Administration). *Code of Federal Regulations 29*; Part 1910.1048, Formaldehyde (Appendix B), 1994.
50. Spear, R. C.; Selvin, S.; Francis, M. The Influence of Averaging Time on the Distribution of Exposures. *Am. Ind. Hyg. Assoc. J.* **1986**, *47*, 365–368.
51. HSC (Health and Safety Executive). *EH40/97 Occupational Exposure Limits 1997*; HSE Books: Suffolk, Great Britain, 1997.
52. Hewett, P. Characterization of Exposures to Welding Fumes and Gases during Production Line Welding. Presented at the 1989 American Industrial Hygiene Conference, 1989.
53. Cope, R. F.; Pancamo, B. P.; Rinehart, W. E.; Haar, G. L. Personnel Monitoring for Tetraalkyl Lead in the Workplace. *Am. Ind. Hyg. Assoc. J.* **1979**, *40*, 372–379.

17

Reproductive Hazards in the Workplace

ELIZABETH ANNE JENNISON

Reproductive disorders rank among the 10 leading work-related illnesses and injuries in the United States, with an estimated 14 million workers having potential occupational exposure to known or suspected reproductive hazards.[1] Unfortunately, efforts to prevent these disorders face major gaps in knowledge. Disorders of reproduction represent an interaction between individual genetic makeup, environmental conditions, and the intensity, duration, and timing of exposure to those conditions. A single toxicant can produce a variety of adverse outcomes depending on the specific conditions of exposure. Conversely, each class of reproductive outcomes can result from a variety of different agents, acting through several biologic mechanisms. Workers are often exposed to more than one agent, so there may be interactive effects from complex mixtures in the workplace or environmental agents outside the workplace. Finally, disorders of reproduction arising from occupational factors may be difficult to distinguish from those with nonoccupational etiologies.[2]

The term *reproductive hazard* is properly restricted to hazards that interfere with or prevent conception. Reproductive hazards may have adverse effects on libido, sexual behavior, any aspect of spermatogenesis or oogenesis, hormonal activity or physiological response that would interfere with the capacity to fertilize, fertilization itself, or the development of the fertilized ovum up to and including implantation.[3] They are distinct from *developmental hazards*, which produce structural abnormalities, functional deficits, pathological alterations to growth, or death. Using this distinction, reproductive hazards are those that affect the worker and their effects may be reversible. The effects of developmental hazards will be confined to the fetus or offspring and are almost invariably permanent. A given agent, for example, ionizing radiation, may present both a reproductive and a developmental hazard.[4]

Despite the OSHA Hazard Communication Standard,[5] which requires employers to provide workers with Material Safety Data Sheets (MSDSs) and training pertaining to hazardous substances used on the job, worker education regarding potential reproductive effects appears to be inadequate.[6] This may be related, in part, to the observation that a significant percentage of MSDSs contain

HANDBOOK OF
CHEMICAL HEALTH AND SAFETY

ROBERT J. ALAIMO
Editor

American Chemical Society
Washington, D.C.

OXFORD
UNIVERSITY PRESS

2001

OXFORD
UNIVERSITY PRESS

Oxford New York
Athens Auckland Bangkok Bogotá Buenos Aires Calcutta
Cape Town Chennai Dar es Salaam Delhi Florence Hong Kong Istanbul
Karachi Kuala Lumpur Madrid Melbourne Mexico City Mumbai
Nairobi Paris São Paulo Shanghai Singapore Taipei Tokyo Toronto Warsaw

and associated companies in
Berlin Ibadan

Developed and distributed in partnership by the
American Chemical Society and Oxford University Press

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016

Oxford is a registered trademark of Oxford University Press.

Library of Congress Cataloging-in-Publication Data
Handbook of chemical health and safety / Robert J. Alaimo, editor.
p. cm.
Includes bibliographic references and index.
ISBN 0-8412-3670-4
1. Hazardous substances—Safety measures—Handbooks, manuals, etc.
I. Alaimo, Robert J., 1940–
T55.3.H3 H344 2001
660'.2804—dc21 99-52958

1 3 5 7 9 8 6 4 2
Printed in the United States of America
on acid-free paper