Occupational Hazards

Industrial Hygiene Exposure Assessments

"Worst-case" versus random sampling

By Jerome E. Spear

Abstract: *Industrial hygiene exposure assessments are performed to evaluate a person's exposure to a chemical or physical agent. Two types of sampling strategies should be considered when planning an exposure assessment study: worst-case sampling and random sampling. This article compares the two strategies and provides an eight-step process of profiling a similar exposure group using a random-sampling approach.*

AN INDUSTRIAL HYGIENE exposure assessment is the process used to evaluate a person's exposure to a chemical or physical agent. To measure a worker's actual exposure, a monitor would need to be placed in his/her breathing zone each day, something that is cost-prohibitive and time-consuming. Since it is often not possible (or practical) to measure each person's actual exposures to the chemical or physical agent, judgments regarding the acceptability of exposure are made based on an estimate of the person's dose via representative sampling.

The exposure assessment strategy employed depends on the purpose and goal of the monitoring and what the sample(s) should represent. The two types of sampling strategies to consider when planning an exposure assessment study are "worst-case" sampling and random sampling. The broad difference is that worst-case sampling involves more subjectivity than a random-sampling approach.

In worst-case sampling, workers who are subjectively believed to have the highest exposures are nonrandomly selected. If no worst-case sample exceeds the occupational exposure limit(s), one can be subjectively satisfied (but not statistically confident) that the exposure profile is acceptable. For example, operators who have similar job duties within a plant's production unit may be identified as a similar exposure group (SEG). For such workers, a worst-case exposure might occur on a day that their process unit generates the highest production output. Such a day would subjectively be considered the worst-case exposure period and would be targeted for evaluation.

A random-sampling strategy requires workers within an SEG and the sample periods to be randomly selected; exposure data is then subjected to statistical analysis. Decisions on the acceptability of the exposure profile can then be determined with a known level of confidence based on the central tendency and spread (or dispersion) of the sample distribution. Applications of a random sampling strategy include:

- Describe the eight-hour time-weighted average (TWA) concentrations over different days for a single worker or SEG.
- Describe the 15-minute TWA concentrations over one workshift for a single worker or SEG.
- Estimate the full-shift TWA concentration over one workshift based on short-term (or grab) samples for a single worker or SEG.

Both types of sampling strategies have advantages and limitations. The primary advantage of using a worst-case strategy is that fewer samples must be collected, making it more cost-efficient and less time-consuming than a random-sampling approach. A limitation of worst-case sampling is that it requires the industrial hygienist to recognize the worst-case conditions, which may include professional judgment about the specific task and/or a specific work practices unique to an individual worker.

However, because a person's judgment may not be as good as perceived, worst-case (i.e., judgmental) samples taken will likely be based on inherent personal biases. As a result, it is not possible to measure the accuracy of worst-case samples (Albright, et al 329). The conclusions about exposure will also include potential biases of judgments about job conditions, work practices and/or other conditions believed to have an impact on exposure.

Random sampling eliminates such biases since samples of the population are selected randomly. As a result, the

Jerome E. Spear, CSP, CIH, *is managing member of J.E. Spear Consulting LLC in Magnolia, TX. He has more than 15 years' experience helping organizations prevent and control losses through the application of incident prevention and exposure control techniques. Spear has a B.S. in Industrial Engineering with a specialization in System Safety Engineering from Texas A&M University. He is a professional member of ASSE's Gulf Coast Chapter, and is a member of System Safety Society and American Industrial Hygiene Assn.*

Table 1

Industrial Hygiene Assessment Strategies: Worst-Case Versus Random Sampling

Sampling Strategy	Advantages	Limitations
Worst-Case Sampling	•Fewer samples are typically collected to make a decision. •Statistical skills are not required.	•Relies on subjective judgments on worst-case exposure conditions. •Difficult to capture the actual worst-case exposure period(s). •Requires the industrial hygienist to recognize the environmental conditions and employee work practices that have the most significant affect on employees' exposures. •Decisions on exposure acceptability are made without a known level of confidence.
Random Sampling	•Exposure groups can be characterized with a known level of certainty. •Considered to be a more defensible strategy since the outcome is based on an objective analysis.	•More samples are required in order to profile the exposure group. •Takes more time to interpret and analyze the sampling results.

variability in the data (due to work practice variations between workers, day-to-day variations, process variations, etc.) is measured and used to estimate the exposure parameters. As a result, a random sampling strategy is based on an objective analysis of the SEG. Table 1 provides a comparison of a the two types of sampling strategies.

Exposure Profiling: An Eight-Step Process

Statistically speaking, it is always possible to make the wrong decision no matter how confident one is. However, quantifying the level of certainty through a random-sampling strategy maximizes the chances of making the right decision. With the availability of computers and statistical software applications, much of the statistical grunt work is automated, making it easier to employ a random-sampling strategy. Several statistical terms and considerations must be understood when using this approach. These are described in the following eight-step process to exposure profiling.

Step 1: Identify the SEG to Profile

The key consideration in categorizing SEGs is to select the exposure group in order to minimize the amount of variation between the samples; otherwise, the resulting confidence intervals (calculated from the mean and variance) will be too wide to be useful. An SEG may be a single worker performing a single task; however, it is often impractical to perform random sampling for each worker. A more practical approach is to include multiple employees who have similar exposures. For example, the employees assigned to operate propane-powered lift trucks in a warehouse may be grouped as having similar potential exposures to carbon monoxide. If multiple SEGs are identified (as most facilities have), one must devise a method for ranking data collection needs. AIHA's *A Strategy for Assessing and Managing Occupational Exposures* provides such a method for study based on the toxicity of the material, conditions of the workplace environment and their likelihood to influence exposures, and individual work practices that are likely to influence exposures (Mulhausen and Damiano 89-101).

Step 2: Randomly Select Workers & Exposure Periods within the Selected SEG

A random sample is one where each worker and time period has an equal probability of being selected for sample collection. It is important to collect samples as randomly as possible; otherwise, the resulting statistics will be biased. A random number table and/or the random number function in a spreadsheet computer program (such as Microsoft Excel) are useful tools in this process. Another consideration is the number of samples that should be collected in order for the exposure profile to be useful. The answer depends on several factors, including the variability of the sample. However, AIHA generally recommends that six to 10 samples (Mulhausen and Damiano 106) are needed to perform a baseline exposure profile.

Step 3: Collect Breathing Zone Samples of the Selected Workers at Randomly Selected Intervals

Breathing zone samples should be collected near the employee's nose and mouth. The breathing zone can be visualized as a hemisphere about six to nine inches around the employee's face. Those conducting the monitoring should understand the nature of the job or process in which the agent(s) is used or generated as well as the basics of field sampling methods and techniques.

Although formal training and/or hands-on training on exposure monitoring techniques and analytical methods are useful, the mechanics of conducting personal air sampling can be self-taught. Several sources of information are helpful in identifying the appropriate sampling and analytical methodology and equipment. These include the *OSHA Technical Manual* and the *NIOSH Manual of Analytical Methods.* An AIHA-accredited laboratory can provide guidance and instructions on specific sampling methods as well.

Step 4: Calculate the Descriptive Statistics for the Data Set

Descriptive statistics include the mean, median, percent above occupational exposure limit (OEL), range and standard deviation that characterize the sample's distribution such as the central tendency and the variability in the data. The mean and median are used to measure the central tendency of the data, whereas the range and standard deviation are measures of variability. By looking at the data from several perspectives, information and patterns in the data may be discovered. Many data sets can be interpreted simply by comparing the OEL with descriptive statistics. When most of the data are clustered well below or well above the OEL, a decision can generally be made on workplace acceptability by using descriptive statistics (Mulhausen and Damiano 235).

The geometric mean (GM) and geometric standard deviation (GSD) are descriptive statistics used to estimate parameters of a lognormal distribution (see Step 5). The GM is the antilog of the arithmetic mean of the logtransformed values (sidebar, above). It is the value below and above which lie 50 percent of the elements in the population (i.e., the population median). The GSD is the antilog of the standard deviation of the logtransformed values (sidebar, above). It is unitless and reflects variability in the population around the GM; therefore, confidence intervals will have a larger spread as the GSD increases.

Calculating the Geometric Mean & Geometric Standard Deviation

1) Transform the original data to the natural log ($\log_e$).

$$y_i = \log_e(x_i)$$

where x_i = original values and y_i = logtransformed values

2) Calculate the mean ($\bar{y}$) and standard deviation (s) of the logtransformed data.

3) Take the antilog of the mean and the antilog of standard deviation of the logtransformed data.

$$GM = \text{antilog}(\bar{y}) = e^{\bar{y}}$$

$$GSD = \text{antilog}(s) = e^{s}$$

Step 5: Determine Whether the Data Fit a Lognormal and/or Normal Distribution

Upper- and lower-confidence limits (UCL and LCL, respectively) and upper-tolerance limits (UTL) are calculated based on knowing (or assuming) a certain underlying distribution of the data set. The type of distribution (i.e., normal or lognormal) will generate different confidence intervals and tolerance limits. A random variable is called normally distributed if the distribution (as plotted on a histogram) looks like a bell-shaped curve (Figure 1).

However, industrial hygiene sampling data is often "skewed to the right" (Figure 2) since occupational exposure values have a lower boundary (i.e., the measured exposure value cannot be less than zero). Taking the log of the variable often mitigates such skewness. In such cases, the distribution is then considered lognormally distributed—or lognormal—if the log of the variable is normally distributed. The lognormal distribution is often applied to occupational exposures, yet the assumption of lognormality is seldom verified (Waters, et al 493). If the data follow neither a lognormal nor a normal distribution, the data may not represent a single SEG. In such cases, the data may need to be divided into two or more SEGs and analyzed separately (Mulhausen and Damiano 242).

If the sample size is large with lots of values, the data can be visualized by creating a histogram of the data (i.e., plotting the relative frequency of elements falling in specified intervals). However, industrial hygiene sampling data often consist of small data sets with fewer than 10 samples due to cost and/or other constraints. One way to qualitatively determine whether the underlying distribution follows a lognormal and/or normal distribution for small sample sizes is to plot the data on probability paper. A lognormal distribution is suggested if the data plotted on lognormal probability paper form a straight line. Likewise, a normal distribution is suggested if the data plotted on normal probability paper form a straight line.

A major advantage of probability plotting is the amount of information that can be displayed in a compact form; however, it requires subjectivity in deciding how well the model fits the data (Waters, et al 493) since probability plotting relies on whether the plotted data forms a straight line. Identifying large deviations from linearity is based on the subjective valuation of viewing the probability plot. Probability paper is available for various types of sample distributions and plotting procedures are described in Technical Appendix I of NIOSH's *Occupational Exposure Sampling Strategy Manual* (Leidel, et al, 97-105). If a statistical program is available, a more quantitative approach should be used to evaluate the goodness-of-fit of the distribution. If both a lognormal and normal distribution are indicated, the confidence limits and upper-tolerance limit should be calculated assuming a lognormal distribution.

Step 6: Calculate the Estimated Arithmetic Mean, One-Sided UCL & LCL of the Arithmetic Mean, 95th Percentile & UTL for the Data Set

• **Estimated arithmetic mean.** For a normal distribution, the estimated arithmetic mean is the same as the sample mean. However, if the data are lognormally distributed, several methods are available for estimating the arithmetic mean and for calculating confidence limits. The sample mean and t distribution confidence limits ("Estimating" sidebar, pg. 43) have the advantage of being easy to calculate, although they can be more variable than other estimates (Mulhausen and Damiano 254). Other preferred methods for estimating the arithmetic mean and com-

Figure 1

Normal Distribution

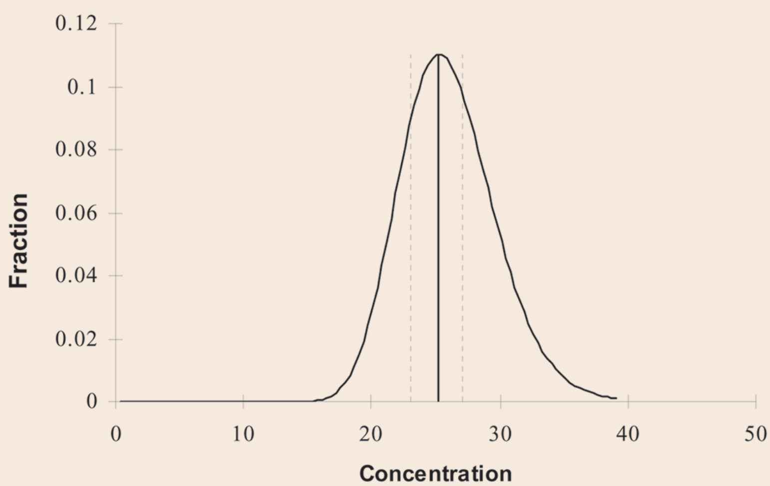
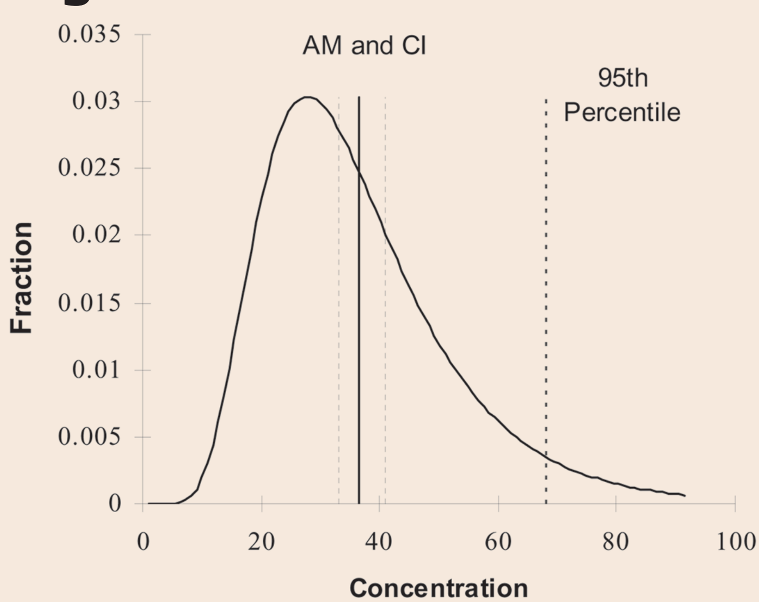


Figure 2

Lognormal Distribution



puting the confidence limits are described in AIHA's *A Strategy for Assessing and Managing Occupational Exposures,* (Mulhausen and Damiano 253-264). However, estimating the arithmetic mean and computing the confidence limits using such preferred methods are difficult without the use of a computer and/or specialized software.

•**$UCL_{1,95\%}$ and $LCL_{1,95\%}$ of the arithmetic mean.** When evaluating toxicants that produce chronic diseases, the mean exposure should be examined (Rappaport and Selvin 378). However, to evaluate acute (short-term) exposures, the UTL of the 95th percentile should be examined. The $UCL_{1,95\%}$ is the one-sided upper value at a 95-percent confidence level. Likewise, the $LCL_{1,95\%}$ is the one-sided lower value at a 95-percent confidence level. If the $UCL_{1,95\%}$ is below the OEL, there is a 95-percent confidence level that the long-term average exposure is below the OEL.

From a compliance perspective, the burden of proof is on the compliance officer to show that the worker was overexposed with at least 95-percent certainty. Therefore, an OSHA compliance officer must demonstrate that the $LCL_{1,95\%}$ exceeds OSHA's permissible exposure limit (PEL). If the data are lognormally distributed, different methods are available to calculate the confidence limits for the arithmetic mean. The sample mean and t distribution confidence limits have the advantage of being easy to calculate, but they can be more variable than other estimates (Mulhausen and Damiano 254).

•**95th percentile.** The 95th percentile is the value in which 95 percent of the population will be included. For example, the median is the 50th percentile.

•**$UTL_{95\%,95\%}$.** This is the UTL of the 95th percentile and is typically used to examine acute (short-term) exposures (e.g., fast-acting contaminants). For a lognormal distribution, the $UTL_{95\%,95\%}$ is calculated using this equation: $UTL_{95\%,95\%} = e^{(\bar{y}+Ks)}$ (Mulhausen and Damiano 270-272) where, $\bar{y}$ and s are the mean and standard deviation, respectively, of the logtransformed data. K is a factor for tolerance limits that is obtained from a table given the confidence level, percentile and number of samples.

Step 7: Make a Decision on the Acceptability of the Exposure Profile

Generally, a $UCL_{1,95\%}$ that results in a value greater than the long-term OEL suggests that the exposure profile is unacceptable, whereas a $UCL_{1,95\%}$ which results in a value below the long-term OEL suggests that the exposure profile is acceptable. For chemicals with acute (or short-term) effects, the UTL of the 95th percentile should be examined.

However, calculating the $UTL_{95\%,95\%}$ with few data points tends to produce a wide tolerance interval, which limits its usefulness. A $UTL_{95\%,95\%}$ which results in a value below the short-term exposure level and ceiling limit suggests that the exposure profile is acceptable, but large numbers of samples are needed to identify "acceptable" environments (Selvin, et al 89).

Step 8: Refine the SEG

Results of the exposure profile may indicate that the exposure group needs to be further refined. For example, it may appear that the sampling for certain individuals results in higher exposures. To statistically test the significance of this variation, an analysis of variance (ANOVA) may be performed. An ANOVA is an inferential statistical test that compares two or more means to determine whether the means are significantly different. If the means are statistically different, the SEG may need to be further refined.

Estimating the Arithmetic Mean & Calculating Confidence Limits

Multiple techniques can be used to estimate the arithmetic mean and compute confidence limits. The sample mean and t distribution confidence limits has the advantage of being easy to calculate but can be more variable than other estimates (Mulhausen and Damiano 254). The sample mean and t distribution confidence limits are calculated as follows:

Step 1: Calculate the sample mean ($\bar{y}$) and sample standard deviation(s).

Step 2: Calculate the confidence limits (CL).

$CL = \bar{y} \pm t(s/\sqrt{n})$

Where:

- The value of t can be looked up in a "student-t" table available in most statistics books.
- $UCL_{1,95\%} = \bar{y} + t_{0.95}(s/\sqrt{n})$
- $LCL_{1,95\%} = \bar{y} - t_{0.95}(s/\sqrt{n})$

Step 3: Compare the $UCL_{1,95\%}$ and $LCL_{1,95\%}$ to the long-term OEL.

Case Studies: Applying Random Sampling Strategies

To illustrate the random-sampling approach, two case studies are provided. One describes how to evaluate the exposure of a single worker using a random-sampling method. The other example involves estimating a full-shift TWA based on random short-term samples.

Example 1: Estimating the Long-Term Exposure of a Single Employee

An employer requested an evaluation of a particular employee's eight-hour TWA exposure (in relation to OSHA's PEL) to formaldehyde during a manufacturing task. The first step in assessing exposures to environmental agents is to have a thorough understanding of the processes, tasks and contaminants to be studied. Information may be obtained through observations and possibly the use of direct-reading devices. Interviews with workers, managers, maintenance personnel and other relevant personnel (such as technical experts) provide an additional source of information and knowledge. In addition, a review of records and documents (including past exposure monitoring data), relevant industry standards, and/or other literature can provide some insights on the magnitude of exposures for given processes and tasks performed at the worksite.

In this case, the employee operated equipment that forms fiberglass-matting material which is used in building materials (such as backing for shingles). Potential exposure to formaldehyde was identified since the MSDS for the resin used to bind the fiberglass matting lists formaldehyde as a component. To assess this employee's exposure using a worst-case sampling strategy, breathing zone air samples would be collected on days believed to produce the highest exposure to formaldehyde. Sampling results would then be used to derive conclusions based on professional judgment (and potential biases) of worst-case conditions.

To eliminate potential biases or incorrect worst-case judgments, a random-sampling strategy was designed. A total of five workdays were randomly selected, which represented the eight-hour sample periods. A formaldehyde monitor was placed in the employee's breathing zone for the duration of the randomly selected workdays to measure the TWA concentration of formaldehyde. Descriptive statistics were calculated, which resulted in the following:

- Maximum: 0.85 ppm
- Minimum: 0.43 ppm
- Range: 0.42 ppm
- Mean: 0.624 ppm
- Median: 0.60 ppm
- Standard deviation: 0.15 ppm
- Geometric mean: 0.610 ppm
- Geometric standard deviation: 1.274 ppm

These statistics indicate no substantial outliers in the data since the mean and median are relatively close in value.

Next, to ensure that the data followed a normal and/or lognormal distribution, the data were plotted on probability paper and a statistical goodness-of-fit test was performed, which indicated both a lognormal and normal distribution. As a result, the arithmetic mean, $UCL_{1,95\%}$ and $LCL_{1,95\%}$ were estimated assuming a lognormal distribution (see below):

- Estimated arithmetic mean: 0.624 ppm
- $UCL_{1,95\%}$: 0.823 ppm
- $LCL_{1,95\%}$: 0.512 ppm

In this example, it can be concluded, with 95-percent certainty, that this employee's exposure to formaldehyde will be less than 0.823 ppm on any given workday. Although the mean concentration of formaldehyde (0.624 ppm) for this employee is less than OSHA's PEL of 0.75 ppm, the most conservative approach would be to reduce exposures to a level less than the $UCL_{1,95\%}$. Therefore, for the employer concerned with ensuring a safe workplace, additional interventions to reduce this employee's exposure to formaldehyde are warranted since the employer cannot conclude with 95-percent certainty that his exposure is below the PEL. However, from a compliance perspective, it would not be possible to demonstrate with 95-percent certainty that the employee's exposure exceeds the PEL since the $LCL_{1,95\%}$ is less than the PEL.

The exposure assessment strategy employed depends on the purpose and goal of the monitoring and what the sample(s) should represent.

Example 2: Estimating TWA Concentration Based on Random Short-Term Samples

An employer identified a potential exposure to 1,6-hexamethylene diisocyanate (HDI) among employees who were spray painting the exterior shell of an aboveground storage tank. Due to limitations with the selected field sampling and analytical method for HDI, the sample duration had to be limited to approximately 15 minutes.

The spray-painting task was performed by employees working from a boom-supported elevated platform. At times, two painters worked from the same platform; at other times, only one painter performed the task. As a result, a sampling strategy was developed that employed both worst-case and random-sampling techniques. Due to the sampling duration limitations, the strategy included collecting random, short-term samples from the breathing zone of each painter and estimating TWA concentration for the task based on those samples. The worst-case exposure condition was also targeted. Random short-term samples were collected when both painters worked side-by-side from the same elevated platform since this condition was assumed to represent the highest exposures to the painters.

Four random, short-term samples were collected from the breathing zone of each painter (i.e., a total of eight random samples), which resulted in the following descriptive statistics:

- Maximum: 0.0042 mg/m^3
- Minimum: <0.00037 mg/m^3
- Range: 0.00394 mg/m^3
- Mean: 0.002 mg/m^3
- Median: 0.002 mg/m^3
- Standard deviation: 0.001 mg/m^3
- Geometric mean: 0.002 mg/m^3
- Geometric standard deviation: 2.545 mg/m^3

These statistics indicate no significant outliers in the data since the mean and median are equivalent.

To ensure that the data followed a normal and/or lognormal distribution, the data were plotted on probability paper and a statistical goodness-of-fit test was performed, which indicated both a lognormal and normal distribution. As a result, the arithmetic mean, $LCL_{1,95\%}$, and $UCL_{1,95\%}$ were estimated assuming a lognormal distribution:

- Estimated arithmetic mean: 0.002 mg/m^3
- $UCL_{1,95\%}$: 0.007 mg/m^3
- $LCL_{1,95\%}$: 0.001 mg/m^3

In this example, one can conclude with 95-percent certainty that the TWA concentration of HDI for the task is less than 0.007 mg/m^3. Since the $UCL_{1,95\%}$ is below the PEL of 0.034 mg/m^3, the exposure was determined to be acceptable from both the employer's perspective and the OSHA compliance officer's perspective.

It is important to note that since the sampling was limited to periods within a single workshift, this case study does not account for day-to-day variation which may exist. Therefore, the results represent the TWA concentration for the task that was performed on the date of the sampling. However, this concentration represents the worst-case condition for this task since breathing zone samples were collected during worst-case conditions—when two painters were spray painting from the same platform.

Beyond Airborne Exposures

Historically, more attention has been given to airborne exposures than to exposures to physical agents and dermal exposures. However, in many situations, random-sampling strategies may be applied to data other than airborne samples. For some chemicals, skin absorption may be the predominant route of exposure and airborne samples would not be the most appropriate variable to study such exposures. Biological monitoring may be more appropriate in such circumstances.

Random-sampling approaches may also be applied to physical agents. However, for noise data and/or other variables measured in a logarithmic scale, analyzing the allowable dose (i.e., percent of dose), rather than decibels, should be considered so that statistical tools can be applied to such exposure measurements.

Both worst-case sampling and random-sampling strategies are useful in assessing exposures. It is important to understand the limitations of each and to correctly apply the selected sampling strategy. A primary benefit of a random-sampling strategy is that it allows SEGs to be profiled with a known level of certainty, which makes it a more defensible and objective sampling strategy. The primary benefit of a worst-case sampling approach is that fewer samples are needed (making it a less costly and less time-consuming strategy) to make an exposure judgment. In some cases, a combination of both approaches may be beneficial. ■

References

Albright, S., et al. *Data Analysis and Decision Making with Microsoft Excel.* Pacific Grove, CA: Brooks/Cole Publishing Co., 1999.

Leidel, N., et al. *Occupational Exposure Sampling Strategy Manual.* DHEW/NIOSH Pub. 77-173. Cincinnati, OH: NIOSH, 1977.

Mulhausen, J. and J. Damiano. *A Strategy for Assessing and Managing Occupational Exposures.* 2nd ed. Fairfax, VA: AIHA Press, 1998.

NIOSH. *NIOSH Manual of Analytical Methods.* Washington, DC: U.S. Dept. of Health and Human Services, CDC, NIOSH. <http://www.cdc.gov/niosh/nmam>.

OSHA. *OSHA Technical Manual.* Washington, DC: U.S. Dept. of Labor, OSHA. <http://www.osha.gov/dts/osta/otm/otm_toc.html>.

Rappaport, S. and S. Selvin. "A Method for Evaluating the Mean Exposure from a Lognormal Distribution." *AIHA Journal.* 48(1987): 374-379.

Selvin, S., et al. "A Note on the Assessment of Exposure Using One-Sided Tolerance Limits." *AIHA Journal.* 48(1987): 89-93.

Waters, M., et al. "A Measure of Goodness-of-Fit for the Lognormal Model Applied to Occupational Exposures." *AIHA Journal.* 52(1991): 493-502.